IDENTIFYING HADRONIC TAU DECAYS USING CALORIMETER TOPOLOGICAL CLUSTERS AT ATLAS

by

Michel Trottier-McDonald

BSc., Université du Québec à Trois-Rivières, 2008

THESIS SUBMITTED IN PARTIAL FULFILLMENT

OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

IN THE DEPARTMENT

OF

PHYSICS

SIMON FRASER UNIVERSITY

Summer 2010

APPROVAL

Name: Michel Trottier-McDonald

Degree: Master of Science

Title of Thesis: Identifying hadronic tau decays using calorimeter topological clusters at ATLAS

Examining Committee: Dr. Simon Watkins (Chair)

Dr. Dugan O'Neil, Simon Fraser University
Senior Supervisor

Dr. Michel Vetterli, Simon Fraser University
Supervisor

Dr. Bernd Stelzer, Simon Fraser University
Supervisor

Dr. Howard Trottier, Simon Fraser University
Internal Examiner

Date Approved: August 23$^{\text{th}}$, 2010

Abstract

The ATLAS experiment collects data on the world's highest energy proton-proton collisions. Its primary goals are to search for direct evidence of physics beyond the Standard Model of particle physics, and to provide a final verdict on the existence of the Standard Model Higgs boson. A robust τ identification procedure is an important ingredient in the search for Standard Model and supersymmetric Higgs bosons. In this study, new τ identification variables based on calorimeter topological clusters are investigated. These variables complement existing calorimeter-based variables by bringing additional discriminating power against hadronic jets. They have a straightforward physics interpretation which allows them to be calculated at particle-level. Some of these new variables are shown to be robust against background noise from the underlying event and in-time pileup, two effects which negatively impact τ identification. Several of these new variables are now part of the official τ reconstruction software at ATLAS.

À ceux et celles qui m'encouragent et qui n'ont jamais douté de moi malgré mes propres doutes: Papa, Maman, Mirka, Maxime, mes amis de toujours et Elisabeth.

Acknowledgments

As long as I can remember, I wanted to be a scientist. I can also remember many instances when I could have given up that dream, thinking that I did not have what it takes. If it hasn't been for the support of a lot of people, I wonder where I would be today.

My parents, Lise and Jimmy, always valued creativity, curiosity, inquiry, patience and hard work. Every single time I manifested these, I have been encouraged.

I had a lot of wonderful teachers during secondary school. Benoît DeCarufel, Michel Ferron, Martin Thiffeault, François Garceau and Jean-Noël Pronovost are all very gifted teachers who made a strong impression on me. They are the main reason I chose natural sciences for my post-secondary studies.

In CÉGEP, Renée Bourassa opened my eyes to critical thinking in her wonderful philosophy class. I had my third physics course with Simon Langlois, who was teaching his own class for the first time. He is the reason I decided to keep going in physics after CÉGEP.

Studying at the physics department of Université du Québec à Trois-Rivières has been a wonderful experience. I have many thanks particularly to Prof. Louis Marchildon and Prof. Adel Antippa for teaching me precisely the nature of science and taking my passion for physics to an unprecedented level. I would also like to thank Prof. Jean Hamelin for giving me the opportunity to experience the world of research. This is what decided me to do graduate studies.

I would like to thank Catherine Russell, who introduced me to Dugan. This has been a tipping point in my life. Before, I could not see any opportunity to work in particle physics. I want to thank my closest friends, Guillaume, Étienne, Gabriel R., Jérémie, David, Gabriel S., Maxime, Léanne, Renaud and all the others. Your friendship is forever, which is the reason I was not afraid to leave for British Columbia.

I keep this last paragraph for Dugan. I have difficulty imagining a better supervisor for a student like me. He knows how to get me to surpass my own limits. He is always on the lookout for opportunities for his students. It gave me a chance to finish this degree in two years. Thank you so much for taking me as your graduate student and providing me with detailed constructive feedback on all me research activities. You are making me a better scientist.

Contents

Approval **ii**

Abstract **iii**

Dedication **iv**

Acknowledgments **v**

Contents **vii**

List of Tables **x**

List of Figures **xii**

1 General introduction **1**

1.1 Units and conventions . . . 2

2 Theory **4**

2.1 The Standard Model of particle physics . . . 4

2.2 The Minimally Supersymmetric Standard Model . . . 7

2.3 τ production as a signature of new phenomena . . . 9

2.3.1 The Standard Model Higgs boson . . . 9

2.3.2 The MSSM Higgs bosons . . . 11

3 Experimental setup **13**

3.1 The Large Hadron Collider . . . 13

3.2 The ATLAS experiment 16
3.2.1 Hardware 16
3.2.2 Software 19

4 τ reconstruction and identification at ATLAS 21
4.1 τ leptons 21
4.2 Hadronic jets 23
4.3 τ reconstruction in the ATLAS inner detector 25
4.4 τ reconstruction in the ATLAS Calorimeter 26
4.4.1 Shower physics 27
4.4.2 The ATLAS Calorimeter system 31
4.4.3 Shower reconstruction 38
4.4.4 Jet reconstruction 41
4.4.5 τ calorimeter seed 42

5 cluster-based τ identification variables 46
5.1 Parameters of the study 46
5.2 Variables constructed from the 4-momenta of topological clusters 47
5.2.1 Number of clusters 49
5.2.2 Effective number of clusters 51
5.2.3 Invariant mass of all clusters 53
5.2.4 Invariant mass of the effective clusters 55
5.2.5 Average ΔR to the geometrical center of all clusters 57
5.2.6 Average ΔR to the geometrical center of the effective clusters 57
5.2.7 Linear correlations 61
5.3 Variables constructed from cluster moments 63
5.3.1 Number of cells 63
5.3.2 Average over effective clusters of the second moment in radial distance to the shower axis 64
5.3.3 Average of the electromagnetic energy fraction of the effective clusters 65
5.3.4 Isolation of the effective clusters 67
5.3.5 Linear correlations 68

6 Performance **71**
6.1 Boosted decision trees 71
6.2 Selection of the cluster-based variables 73
6.2.1 Selection of the 4-momentum variables 73
6.2.2 Selection of the cluster moment variables 73
6.3 Performance of the cluster-based variables 74
6.3.1 Performance of the 4-momentum variables 74
6.3.2 Performance of the cluster moment variables 82
6.3.3 Combining all Calorimeter-based variables 86
6.4 Effect of pileup 89
6.4.1 Effect on the distributions 90
6.4.2 Effect on the performance 98

7 Summary, conclusions and future directions **103**
7.1 Summary and conclusions 103
7.2 Future directions 104

Bibliography **107**

List of Tables

2.1 Basic properties of known bosons in the Standard Model. 5

2.2 Basic properties of known fermions in the Standard Model. The quark masses are estimated using the $\overline{\mathrm{MS}}$ scheme at a scale $\mu = 2\mathrm{GeV}$. 5

4.1 Most important decay modes for τ [2]. 22

4.2 Granularity and segmentation of the ATLAS electromagnetic and hadronic calorimeters [22]. The forward calorimeter is not shown, since it is not used in τ reconstruction. 36

6.1 Gain in background rejection at 30% signal efficiency when adding the 4-momentum variables to the cell-based variables. 76

6.2 Variable rankings for 4-momentum variables and cell-based variables . . . 77

6.3 Variable rankings for effective-cluster and cell-based variables. 79

6.4 Variable rankings for all-cluster and cell-based variables. 80

6.5 Ratio of the in background rejection at 30% signal efficiency based on the 4-momentum variables between an ideal calorimeter and the ATLAS calorimeter. 82

6.6 Gain in background rejection at 30% signal efficiency when adding the cluster moment variables to the cell-based variables. 84

6.7 Variable rankings for cluster moment variables and cell-based variables . . 84

6.8 Gain in background rejection at 30% signal efficiency when adding the cluster moment variables to the 4-momentum variables. 86

6.9 Variable rankings for cluster moment variables with 4-momentum variables 86

6.10 Gain in background rejection at 30% signal efficiency when adding the cluster moment variables to the cell-based and 4-momentum variables. . . . 88
6.11 Variable rankings for all variables 88
6.12 Relative mean and RMS shifts of the cell-based and 4-momenta variables caused by the presence of pileup. Smaller values are consistent with pileup resistance. 93
6.13 Variable rankings for 4-momentum variables and original variables, pileup vs. no pileup 102

List of Figures

2.1 First order processes of Higgs production at the LHC. (a) Gluon fusion and (b) vector boson fusion. 10
2.2 (a) Cross-sections for the main Higgs production channels at the LHC with a center of mass energy of 14TeV. (b) Branching Ratios for the Higgs decay modes of interest at the ATLAS experiment. ATLAS Experiment Images: ©CERN [8]. 11
2.3 Area of the supersymmetry parameter space already excluded by the CDF experiment at Fermilab and the area that will be probed by the ATLAS experiment for potential discovery. ATLAS Experiment Image: ©CERN [10]. 12

3.1 Partial schematic of the CERN accelerator complex relevant to LHC proton-proton collisions. The maximum proton energy is listed at the end of each acceleration stage. The positions of the four major experiments are also shown. 14
3.2 Representation of the ATLAS detector showing the main subdetectors. ATLAS Experiment Image: ©CERN [12]. 17

4.1 Hadronization process. The black dots are quarks, the line connecting them are gluon fields. 1) The initial quarks have momenta going in opposite directions. 2) Part of their kinetic energy is transferred to the gluon field. 3) The gluon field has enough energy to conjure a quark/antiquark pair, breaking it in two. 4) There is energy left in the gluon fields. 5) The remaining energy is dissipated through the creation of more quark/antiquark pairs and fragmentation continues. 24
4.2 Development of an electromagnetic shower according to the simplified model. A wiggly line represents a photon, while a straight line represents an electron or a positron. 28
4.3 Real electromagnetic shower from a cosmic ray observed in the ICARUS T600 detector during a technical run in 2001. ICARUS Experiment Image: ©ETH Institute for Particle Physics. 30
4.4 Real hadronic shower from a cosmic ray observed in the ICARUS T600 detector during a technical run in 2001. ICARUS Experiment Image: ©ETH Institute for Particle Physics. 32
4.5 Longitudinal section of the inner detector and calorimeter system of the ATLAS detector. The sections highlighted in blue are the barrel and end-cap electromagnetic calorimeters. The sections highlighted in orange are the barrel, extended barrel and end-cap hadronic calorimeters. Not highlighted but also shown are the forward calorimeters, between the beam axis and the end-cap modules. ATLAS Experiment Image: ©CERN. [22] . . . 33
4.6 Cross-section view of the accordion geometry of the lead sheets and copper electrodes in the electromagnetic barrel. The total space taken by the sheets is seen to increase proportionally with the radius (from left to right). ATLAS Experiment Image: ©CERN. [22] 34
4.7 Granularity in η of the barrel of the electromagnetic calorimeter, and depth in radiation lengths. ATLAS Experiment Image: ©CERN. [22] 35
4.8 Cell division of the tile calorimeter in η. ATLAS Experiment Image: ©CERN. [22] 38
4.9 Amount of material composing the ATLAS calorimeters as a function of η, given in absorption lengths. ATLAS Experiment Image: ©CERN. [22] . . 39

4.10 Calorimeter cell-based τ identification variables: (a) Electromagnetic radius , (b) hadronic radius, (c) centrality fraction, (d) isolation fraction, (e) η width in the strip layer, (f) number of hits in the strip layer. 45

5.1 Number of clusters/particles. (a), (b) and (c) show the integrated distributions for $10\text{GeV} < E_T < 70\text{GeV}$ and $-2.5 < \eta < 2.5$ for visible final state τ decay products only (signal) and final state particles including underlying event (background), final state particles including underlying event (signal and background) and reconstructed clusters respectively. (d), (e) and (f) show the E_T profiles for the same selections as for (a), (b) and (c) respectively. 50

5.2 Effective number of clusters/particles. (a), (b) and (c) show the integrated distributions for $10\text{GeV} < E_T < 70\text{GeV}$ and $-2.5 < \eta < 2.5$ for visible final state τ decay products only (signal) and final state particles including underlying event (background), final state particles including underlying event (signal and background) and reconstructed clusters respectively. (d), (e) and (f) show the E_T profiles for the same selections as for (a), (b) and (c) respectively. 52

5.3 Invariant mass of all clusters/particles. (a), (b) and (c) show the integrated distributions for $10\text{GeV} < E_T < 70\text{GeV}$ and $-2.5 < \eta < 2.5$ for visible final state τ decay products only (signal) and final state particles including underlying event (background), final state particles including underlying event (signal and background) and reconstructed clusters respectively. (d), (e) and (f) show the E_T profiles for the same selections as for (a), (b) and (c) respectively. 54

5.4 Invariant mass of effective clusters/particles only. (a), (b) and (c) show the integrated distributions for $10\text{GeV} < E_T < 70\text{GeV}$ and $-2.5 < \eta < 2.5$ for visible final state τ decay products only (signal) and final state particles including underlying event (background), final state particles including underlying event (signal and background) and reconstructed clusters respectively. (d), (e) and (f) show the E_T profiles for the same selections as for (a), (b) and (c) respectively. 56

5.5 Average ΔR to the geometrical center for all clusters/particles. (a), (b) and (c) show the integrated distributions for 10GeV $< E_T <$ 70GeV and $-2.5 < \eta < 2.5$ for visible final state τ decay products only (signal) and final state particles including underlying event (background), final state particles including underlying event (signal and background) and reconstructed clusters respectively. (d), (e) and (f) show the E_T profiles for the same selections as for (a), (b) and (c) respectively. 58
5.6 Average ΔR to the geometrical center for effective clusters/particles only. (a), (b) and (c) show the integrated distributions for 10GeV $< E_T <$ 70GeV and $-2.5 < \eta < 2.5$ for visible final state τ decay products only (signal) and final state particles including underlying event (background), final state particles including underlying event (signal and background) and reconstructed clusters respectively. (d), (e) and (f) show the E_T profiles for the same selections as for (a), (b) and (c) respectively. 60
5.7 Linear correlation coefficients between cell-based and 4-momentum variables for (a) 1prong signal, (b) 3 prongs signal, (c) background. An empty square represents a correlation factor of 0. 62
5.8 Number of cells. (a) shows the integrated distribution for 10GeV $< E_T <$ 70GeV and $-2.5 < \eta < 2.5$ for reconstructed clusters and (b) shows the E_T profile for the same τ candidates. 63
5.9 Average over effective clusters of the second moment in radial distance to the shower axis. (a) shows the integrated distribution for 10GeV $< E_T <$ 70GeV and $-2.5 < \eta < 2.5$ for reconstructed clusters and (b) shows the E_T profile for the same τ candidates. 65
5.10 Average over effective clusters of the fraction of energy in the EM calorimeter. (a) shows the integrated distribution for 10GeV $< E_T <$ 70GeV and $-2.5 < \eta < 2.5$ for reconstructed clusters and (b) shows the E_T profile for the same τ candidates. 66
5.11 Average over effective clusters of the isolation of the clusters. (a) shows the integrated distribution for 10GeV $< E_T <$ 70GeV and $-2.5 < \eta < 2.5$ for reconstructed clusters and (b) shows the E_T profile for the same τ candidates. 67

5.12 Linear correlation coefficients between cell-based and cluster moment variables for (a) 1prong signal, (b) 3 prongs signal, (c) background. An empty square represents a correlation factor of 0. 69
5.13 Linear correlation coefficients between 4-momentum and cluster moment variables for (a) 1prong signal, (b) 3 prongs signal, (c) background. An empty square represents a correlation factor of 0. 70

6.1 Comparison of the performance of the 4-momentum variables and the cell-based variables. (a) shows 1 prong candidates only, (b) shows 3 prong candidates only and (c) shows all candidates, each for $E_T < 35$GeV. (d), (e) and (f) show the same but for $E_T > 35$GeV. 75
6.2 Comparison of the performance of the effective-cluster variables and the all-cluster variables when added to the cell-based variables. (a) shows 1 prong candidates only, (b) shows 3 prong candidates only and (c) shows all candidates, each for $E_T < 35$GeV. (d), (e) and (f) show the same but for $E_T > 35$GeV. 78
6.3 Comparison of the performance of the 4-momentum variables when calculated from truth particles and from reconstructed clusters. (a) shows 1 prong candidates only, (b) shows 3 prong candidates only and (c) shows all candidates, each for $E_T < 35$GeV. (d), (e) and (f) show the same but for $E_T > 35$GeV. 81
6.4 Comparison of the performance of cluster moment variables and the cell-based variables. (a) shows 1 prong candidates only, (b) shows 3 prong candidates only and (c) shows all candidates, each for $E_T < 35$GeV. (d), (e) and (f) show the same but for $E_T > 35$GeV. 83
6.5 Comparison of the performance of the cluster moment variables and the 4-momentum variables. (a) shows 1 prong candidates only, (b) shows 3 prong candidates only and (c) shows all candidates, each for $E_T < 35$GeV. (d), (e) and (f) show the same but for $E_T > 35$GeV. 85

6.6 Comparison of the performance of the cell-based variables, the 4-momentum variables and the cluster moment variables when put together. (a) shows 1 prong candidates only, (b) shows 3 prong candidates only and (c) shows all candidates, each for $E_T < 35$GeV. (d), (e) and (f) show the same but for $E_T > 35$GeV. 87

6.7 Number of clusters. (a) shows effect of pileup on signal, (b) shows effect of pileup on background, (c) shows E_T profiles when pileup is present. . . . 91

6.8 Effective number of clusters. (a) shows effect of pileup on signal, (b) shows effect of pileup on background, (c) shows E_T profiles when pileup is present. 92

6.9 Invariant mass of all clusters. (a) shows effect of pileup on signal, (b) shows effect of pileup on background, (c) shows E_T profiles when pileup is present. 94

6.10 Invariant of the effective clusters. (a) shows effect of pileup on signal, (b) shows effect of pileup on background, (c) shows E_T profiles when pileup is present. 95

6.11 Average ΔR to the geometrical center of all clusters. (a) shows effect of pileup on signal, (b) shows effect of pileup on background, (c) shows E_T profiles when pileup is present. 96

6.12 Average ΔR to the geometrical center of the effective clusters. (a) shows effect of pileup on signal, (b) shows effect of pileup on background, (c) shows E_T profiles when pileup is present. 97

6.13 Fraction of background rejection retained when pileup is taken into account with respect to signal efficiency for different sets of variables. (a) 1 prong candidates only, (b) 3 prong candidates only, (c) all candidates. Statistical error bars on the background rejection are plotted. Above a signal efficiency of 0.8, they become larger than the difference between the points. 99

6.14 Ratio of background rejection for a subset of variable to the complete set of variables (cell-based and 4-momenta based) with no pileup conditions. (a) 1 prong candidates only, (b) 3 prong candidates only, (c) all candidates. Statistical error bars on the background rejection are plotted. Above a signal efficiency of 0.8, they become larger than the difference between the points. 100

6.15 Ratio of background rejection for a subset of variable to the complete set of variables (cell-based and 4-momenta based) with pileup conditions. (a) 1 prong candidates only, (b) 3 prong candidates only, (c) all candidates. Statistical error bars on the background rejection are plotted. Above a signal efficiency of 0.8, they become larger than the difference between the points. 101

Chapter 1

General introduction

The reductionist approach in science consists of determining the components of a system, and their interactions, to understand the behavior of the complete system. Pushed to the extreme, this approach gives us particle physics, the description of the most fundamental constituents of matter and their interactions. The pinnacle of our understanding of particle physics is the so-called Standard Model, which describes successfully a large body of data accumulated by dozens of experiments. However, a growing number of observations lead physicists to think that the Standard Model is incomplete, motivating direct approaches to find what is missing.

The Large Hadron Collider (LHC) is a particle collider at CERN, Geneva, Switzerland. Its main purpose is to provide proton-proton collisions to four massive detectors, one of them being the ATLAS detector. The LHC brought the energy frontier to a center of mass energy of 7 TeV on March 30$^{\text{th}}$ 2010. ATLAS has been recording data on these collisions ever since. Both the LHC and the ATLAS detector have been designed to provide evidence of whatever lies beyond the Standard Model.

The Higgs boson is the only elementary particle in the Standard Model which has not yet been observed. If its mass is below 140 GeV$/c^2$, the presence of τ leptons (or τ) provides one of its most important decay signatures. Furthermore, τ are an essential signature for supersymmetric Higgs bosons. Identifying τ is thus a crucial step in looking for these new particles.

Since τ are short-lived, they have to be recognized via their decay products. Unfortunately, the scattering of proton constituents produces sprays of particles called hadronic jets which are composed of the same type of particles as τ decay products. Pairs of hadronic jets are produced between six and seven orders of magnitude more often than W bosons decaying to τ, which is the most abundant Standard Model source of τ at the LHC. The discrimination of τ decay products versus hadronic jets is a primary goal of τ identification at ATLAS.

There are a few basic differences between τ decays and hadronic jets that can be used for discrimination. τ identification is performed by first selecting τ candidates and calculating variables quantifying their appearance. The discrimination power of these variables is then consolidated using multi-variate techniques. τ identification at ATLAS uses information coming from the calorimeter system and the tracking system.

The design of the baseline calorimeter variables is based on extracting information from different geometric regions of the calorimeter. A different and complementary approach is presented in this thesis. By taking advantage of the fine spatial granularity of the ATLAS calorimeter, it is possible to reconstruct individual decay products. The topological clustering algorithm connects neighbouring calorimeter cells having a significant energy, in an attempt to reconstruct the showers produced by the particles. By using the resulting clusters, it is possible to quantify directly the physical differences between τ decays and hadronic jets.

The content of the current thesis is organized as follows. Chapter 2 goes more in depth in the theoretical motivations for identifying τ at ATLAS. Chapter 3 introduces the LHC and the ATLAS experiment. Chapter 4 explains how τ are reconstructed using the ATLAS detector. Chapter 5 introduces a new class of identification variables called the cluster-based variables. Chapter 6 discusses the discriminating information content of these variables. Chapter 6.4 addresses the impact of pileup on calorimeter τ identification variables, and how it can be minimized. Finally, Chapter 7 closes the discussion with a summary, conclusion and some suggestions for future directions of investigation.

1.1 Units and conventions

Throughout this thesis, the following conventions are observed. Firstly, the electric charge is given as a multiple q of the magnitude of e, the electron charge. Whenever a particle X possesses a non-zero electric charge q, there is also an antiparticle $\bar{X}$ which has the opposite sign electric charge $-q$. When the electric charge of a particle is listed as $\pm q$ ($\mp q$), it means that the particle has charge $+q$ ($-q$) and the associated antiparticle has charge $-q$ ($+q$). Unless explicitly stated, the designation of a particle as electron, quark or τ will refer to both the particle and the antiparticle.

Secondly, the energy is expressed in electronvolts (eV). An electron acquires 1 eV of energy when it is accelerated by an electric potential of 1 Volt. However, this unit is very small for high energy physics applications, so the GeV (10^9 eV) and the TeV (10^{12} eV) will be the most frequently used units of energy in this thesis.

Thirdly, the speed of light is taken to be a dimensionless quantity of magnitude 1, which allows mass (eV/c^2), momentum (eV/c) and energy (eV) to be expressed in terms of eV.

Finally, the cross-section which is used to quantify the probability of any scattering event in particle physics is given in barns (b) where $1\mathrm{b} \equiv 1 \times 10^{-28}\mathrm{m}^2$ [1].

Chapter 2

Theory

2.1 The Standard Model of particle physics

The underpinning of the Standard Model of particle physics is quantum field theory. Quantum field theory is a powerful framework resulting from the combination of the fundamental principles of quantum mechanics and special relativity. So far, the interactions we have successfully described using this framework are the electromagnetic, the weak and the strong interactions. Electromagnetism accounts for all atomic, molecular, optical and condensed matter physics. The strong and the weak interactions account for all known nuclear processes. The range of phenomena ultimately explained by the Standard Model is astonishingly large. Gravity is the only known fundamental interaction for which a sound quantum theory has yet to be formulated and confirmed.

The Standard Model includes a description of the strong interaction called quantum chromodynamics (QCD). It also includes the electroweak theory, which describes both electromagnetic and weak interactions. All the interactions manifest themselves as spin-1[1] force carrier particles commonly called bosons. They are listed in Table 2.1 [2].

The particles subject to these interactions are also a part of the Standard Model. They are spin-$\frac{1}{2}$ particles, referred to as fermions, and they are listed in Table 2.2 [2]. Fermions are further classified into leptons and quarks. The main differences between these two

[1]The spin of a fundamental particle is an internal discrete degree of freedom which is conserved along with angular momentum.

Table 2.1: Basic properties of known bosons in the Standard Model.

interaction	carrier	electric charge	mass (GeV)
electromagnetic	γ (photon)	0	0
weak	$W^{\pm}$	± 1	80.4
	Z	0	91.2
strong	g (gluon)	0	0

Table 2.2: Basic properties of known fermions in the Standard Model. The quark masses are estimated using the $\overline{\text{MS}}$ scheme at a scale $\mu = 2$ GeV.

generation	leptons			quarks		
	symbol	electric charge	mass (GeV)	symbol	electric charge	mass (GeV)
1	e	± 1	0.511×10^{-3}	u	$\pm\frac{2}{3}$	$1.5 - 3.3 \times 10^{-3}$
	ν_e	0	0	d	$\mp\frac{1}{3}$	$3.5 - 6.0 \times 10^{-3}$
2	μ	± 1	0.106	c	$\pm\frac{2}{3}$	1.27
	ν_μ	0	0	s	$\mp\frac{1}{3}$	0.104
3	τ	± 1	1.78	t	$\pm\frac{2}{3}$	171
	ν_τ	0	0	b	$\mp\frac{1}{3}$	4.2

classes are that leptons do not interact via the strong interaction since they lack colour charge[2]. Also, quarks have fractional electric charge. Because the strong interaction does not allow for asymptotic freedom, quarks are always found in bound states. The colour charge must always be neutralized by binding two or three quarks together (colour confinement). The resulting objects are called mesons (a quark and an anti-quark) or baryons (three quarks) and collectively called hadrons. Hadrons always have an electric charge which is an integer multiple of the electron charge.

[2]The colour charge is another degree of freedom of fundamental particles which is the conserved quantity shaping strong interactions.

A peculiar feature of the fermions is that they can be organized in three generations. The first generation is comprised of the up quark u, the down quark d, the electron e and the electron neutrino ν_e. The other two generations are copies of the first in all respects except that the quarks and the charged leptons are more massive.

Neutrinos interact very little with matter, which makes them very difficult to study. For most practical purposes in high energy experimental physics, the neutrinos can be approximated to be massless, which is why the neutrinos masses in Table 2.2 are given as 0.

The bare quark masses are also given as approximate values in Table 2.2. There are many subtleties involved in the determination of the quark masses, the most obvious being that it must be done indirectly since quarks are always bound to other quarks[3]. Deducing the quark masses by observing the properties of these bound states is an enterprise that is dependent on the theoretical framework used [2].

Compared to neutrinos and quarks, charged leptons are simple. Since they do not interact strongly, they can only form bound states with other particles via the electromagnetic interaction. The binding energies involved are weak enough so that free charged leptons are simple to obtain and study. Furthermore, their mass can be easily defined and measured and no evidence exists that would indicate that they have substructure. They are the archetypical "point particles".

There is one more particle that is necessary to make the Standard Model a coherent whole. The unification of the electromagnetic and the weak interactions is done by first making a non-Abelian gauge theory based on the $\mathrm{SU}(2) \times \mathrm{U}(1)$[4] symmetry group. This results in a theory of interactions between fermions mediated by four massless bosons [4]. Then, the $\mathrm{SU}(2) \times \mathrm{U}(1)$ gauge symmetry is broken with the help of an additional scalar field. The symmetry breaking transforms the four massless bosons into three massive (W^+, W^- and Z) and one massless boson (γ). The additional scalar field is the famous Higgs field. The mechanism by which it is exploited to break the electroweak gauge symmetry is

---

[3]The top quark is an exception. Theoretically, it decays fast enough after production so that it does not form any bound state via the strong interaction [2].

[4]$\mathrm{SU}(2)$ refers to a group of transformations represented by 2×2 unitary matrices, while $\mathrm{U}(1)$ refers to a group of transformations represented by a single complex scalar. For an insightful introduction to group theory in the context of quantum mechanics see [3].

the Higgs mechanism.

The Higgs field manifests itself as a massive neutral boson in the Standard Model. It can be used to give mass to the fermions. It is rather unsettling that a particle playing such a fundamental role in our best description of the smallest constituents of matter has yet to be observed. The determination of the existence (or non-existence) of the Higgs boson is arguably the most urgent experimental goal in contemporary particle physics.

As comprehensive as it is, the Standard Model of particles physics is not a complete description of matter. By itself, it cannot account for gravitational interactions and thus it can only play a partial role in cosmology. Furthermore, compelling astronomical evidence indicates that 22% of the universe is made of an unknown form of matter, referred to as dark matter [5]. In addition, everywhere we look in the universe, we find no evidence of large quantities of anti-matter. The Standard Model predicts that there should be almost[5] as much matter as anti-matter in the universe.

There are also a number of epistemological reasons to think that the Standard Model is not the whole story. For example, despite the fact that there is a mechanism to explain the provenance of mass, it cannot explain the exact spectrum of masses observed in fermions [1]. It also cannot explain the strengths of the different interactions, nor the strengths of the couplings responsible for cross-generation transmutations of quarks. All this is usually summarized by saying that there are more than 20 free parameters in the Standard Model. These parameters must be evaluated through experimentation if the theory is to have any predictive power. Also, no reason is found in the Standard Model to explain why the fundamental particles can be organized in three generations.

2.2 The Minimally Supersymmetric Standard Model

The discovery of supersymmetry is strongly anticipated during the exploration of TeV-scale physics. Supersymmetry is the hypothesis that the exchange of bosons and fermions leaves the dynamics of fundamental particles unchanged. Adding this as a constraint on the Standard Model has a number of interesting consequences.

In a viable supersymmetric theory, each of the Standard Model particles introduced in

[5] The Standard Model provides sources of violation of the matter/anti-matter asymmetry, but they are not abundant enough to explain astronomical observations.

section 2.1 must be a member of a "supermultiplet". Every supermultiplet has fermionic and bosonic degrees of freedom, which can be arranged to obtain a fermion field and a boson field. These fields manifest themselves as two particles with the same mass, electric charge and colour degrees of freedom [6]. In other words, every Standard Model particle must have a "superpartner". This implies that there must exist a boson with exactly the mass and electric charge of the electron which has not been observed. It is hard to imagine how such a particle would have escaped experimental scrutiny. Supersymmetry is still plausible if it undergoes spontaneous symmetry breaking. This minimal inclusion of supersymmetry in the Standard Model is referred to as the Minimally Supersymmetric Standard Model (MSSM).

One of the biggest challenges for physicists working on supersymmetric models is to find out how to break supersymmetry. A general supersymmetry breaking scheme requires the introduction of a large number of new parameters. In this respect, supersymmetry does not improve the Standard Model.

A first interesting prediction of the MSSM is that the lightest superpartner has to be a stable, colourless and neutral particle. It is a natural candidate to explain dark matter [1]. A second compelling argument in favour of supersymmetry relies on the assumption that the Higgs mechanism is real. In the Standard Model, the mass of the Higgs boson is sensitive to corrections from its coupling to massive particles. These corrections can easily propel the Higgs mass into the TeV range unless some fine tuning prevents it. This possibility is hard to reconcile with current experimental constraints, which limit the Higgs mass to be of the order of 100 GeV. This difficulty of the Standard Model is usually referred to as the "hierarchy problem". In the MSSM, the corrections responsible for the Higgs mass instability are exactly cancelled by the corrections brought by the superpartners, completely eliminating the hierarchy problem.

Unifying the different interactions found in nature is the goal of any physicist who believes that all phenomena can ultimately be described by a single theory. The unification of electromagnetic and weak interactions is the latest success in this enterprise, although it is not the first. Another salient example is the unification of electricity and magnetism, done by Maxwell in the 19th century. An even older example is the unification of gravity on Earth with gravity in the solar system, achieved by Newton in the 16th century. There are also plenty of unification examples which are not directly about interactions. Good examples are the phenomena explained by the motion of atoms: temperature, pressure and sound. Such unifications are landmarks in our understanding of the laws of nature. However, unifying the strong and electroweak interactions is not possible within the current formulation of the Standard Model. The current formulation does not allow the relevant coupling constants to converge at any energy scale. Supersymmetry makes the convergence possible by changing the energy dependence of these coupling constants.

In the MSSM, a single Higgs multiplet is not enough for the Higgs mechanism to give mass to particles. The simplest alternative is to consider two $SU(2)$ multiplets yielding eight degrees of freedom. When electroweak symmetry is broken, three of these degrees of freedom become part of the three weak bosons. The five remaining degrees of freedom are interpreted to be five different Higgs scalar potentials, usually referred to as the neutral h^0, H^0 and A^0, and the charged H^+ and H^-. This prediction of the MSSM opens a door for experimental exploration.

2.3 τ production as a signature of new phenomena

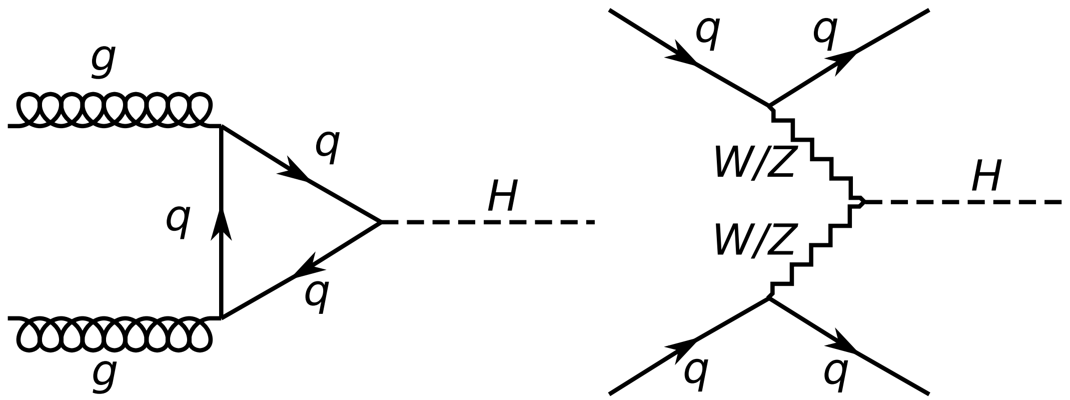
All Higgs bosons discussed previously are short-lived particles. If they are produced in the LHC collisions, they will be observed through the signature of their decay products. Something they have in common is that they all have decay modes to τ. In some cases, these decay modes may play a major part in the discovery and/or exclusion of Higgs bosons. This is sufficient motivation for developing a τ identification procedure at the ATLAS experiment[6].

2.3.1 The Standard Model Higgs boson

The most important processes for Standard Model Higgs production at the LHC are gluon fusion and vector boson fusion. The lowest order Feynman diagrams for these processes are illustrated in Fig. 2.1. The cross-section for these processes are given as a function of the mass of the Higgs boson in Fig. 2.2(a). Of greater relevance are the branching ratios of the main decay modes as a function of the Higgs mass which are given in Fig. 2.2(b). The

[6] For a more comprehensive survey of other sources of τ that are beyond the Standard Model for which the ATLAS experiment has discovery and exclusion potential, read Ref. [7].

success of the ATLAS experiment in discovering or excluding the Higgs boson depends on probing several of these decays modes.



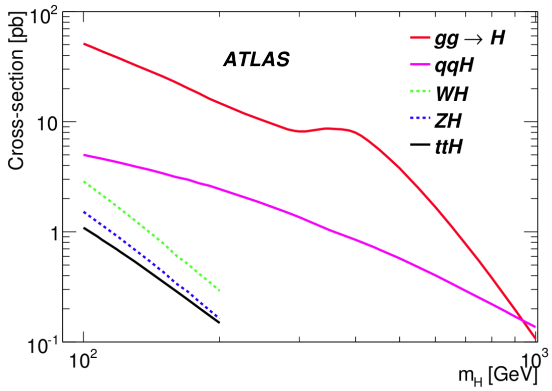
(a) (b)

Figure 2.1: First order processes of Higgs production at the LHC. (a) Gluon fusion and (b) vector boson fusion.

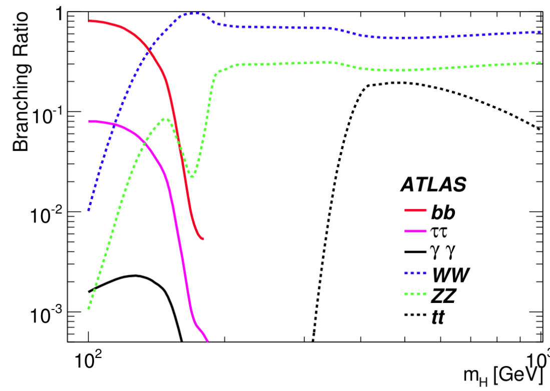
The search for the Higgs boson at ATLAS begins for Higgs masses of $\simeq$ 115 GeV[7]. Up to a mass of ~140 GeV, the dominant Higgs boson decay channel is a pair of b quarks. However, this channel is difficult to exploit due to its high sensitivity to systematic uncertainties and large backgrounds [8]. The second decay mode in importance is $H \rightarrow \tau\tau$. This channel presents its own challenges but they are more easily met than those of $H \rightarrow bb$. These challenges will be exposed in Chapter 4. The $H \rightarrow \gamma\gamma$ channel is not very abundant, but it offers a clean signal. When combined, the $H \rightarrow \tau\tau$ and $H \rightarrow \gamma\gamma$ channels can lead to a 95% confidence level exclusion after gathering $2fb^{-1}$ of collision data[8] at 14 TeV [10]. However, it will take much more data to be able to claim discovery for a low mass Higgs boson.

[7] Results from the LEP experiment exclude the existence of a Higgs boson with a mass smaller than 114.4 GeV [9] at 95% confidence level.

[8] fb^{-1} or inverse femtobarns are units of integrated luminosity, which will be explained in Chapter 3.



(a)



(b)

Figure 2.2: (a) Cross-sections for the main Higgs production channels at the LHC with a center of mass energy of 14 TeV. (b) Branching Ratios for the Higgs decay modes of interest at the ATLAS experiment. ATLAS Experiment Images: ©CERN [8].

2.3.2 The MSSM Higgs bosons

The five Higgs bosons of the MSSM have two principal free parameters usually denoted as $\tan\beta$ and m_A (which can also be the mass of any of the five Higgs bosons). The parameter space which must be explored by the ATLAS experiment to find MSSM Higgs bosons is two-dimensional. The area of this two-dimensional space that will be probed at the LHC is shown in Fig. 2.3.

Neutral MSSM Higgs bosons have a branching ratio to $\tau\tau$ of the order of 10% in different scenarios considered in Ref. [8]. Since these particles may have masses of a few hundred GeV, they can produce τ with energies difficult to obtain through Standard Model processes alone, like $Z \to \tau\tau$. This makes $A/H/h \to \tau\tau$ a very interesting channel to find evidence for supersymmetry.

The role that τ play in the search for charged Higgs bosons is even more important. The decay process $H^{\pm} \to \tau^{\pm}\nu_{\tau}$ has the largest branching ratio in many regions of the MSSM parameter space. Because of how charged MSSM Higgs are produced, the τ coming directly from the Higgs decay is accompanied by an assortment of other objects [10]. To correctly identify τ decays is necessary to flag these complex final states as a signature of $H^{\pm}$ production.

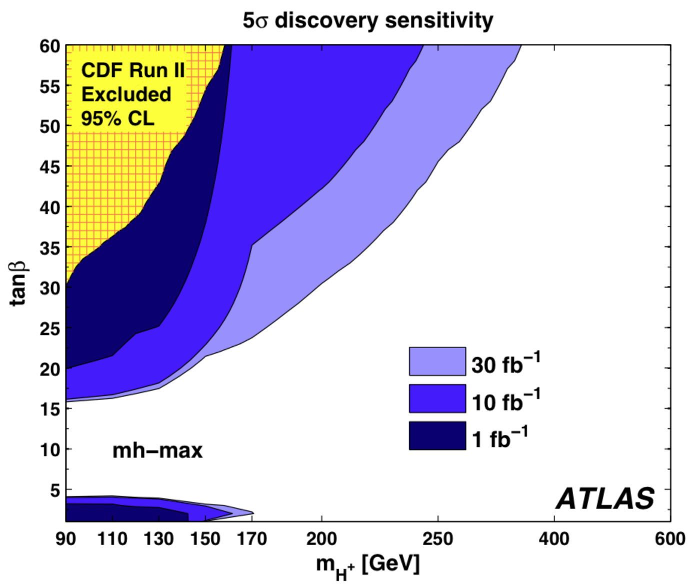


Figure 2.3: Area of the supersymmetry parameter space already excluded by the CDF experiment at Fermilab and the area that will be probed by the ATLAS experiment for potential discovery. ATLAS Experiment Image: ©CERN [10].

Chapter 3

Experimental setup

Situated near Geneva, Switzerland, CERN currently hosts the largest particle physics accelerator complex in the world. The final stage of the complex, called the Large Hadron Collider (LHC) [11], is responsible for accelerating two counter-rotating beams of protons to an energy of 7 TeV each and making them cross at four precise locations. The ATLAS detector takes data at one of these four locations. ATLAS is a highly complex machine designed to provide as much information as possible about the LHC collisions. By analyzing the large data sample collected by the ATLAS detector, it will be possible to make scientific claims about whatever lies beyond the Standard Model. The purpose of this chapter is to introduce the LHC and the ATLAS experiment, as well as some key vocabulary that will be used throughout the rest of the thesis.

3.1 The Large Hadron Collider

The protons which eventually collide at the design energy of 14 TeV at the LHC[1] are originally nuclei of hydrogen atoms. After stripping the electrons from the protons, the protons are accelerated to 50 MeV by a linear accelerator called LINAC2. Fig. 3.1 shows a schematic of the systems of the CERN accelerator complex needed to produce the 14 TeV collisions of the LHC.

[1] At the time of writing, the LHC is accelerating both beams to 3.5 TeV, producing collisions of 7 TeV in the center of mass reference frame. The LHC requires further commissioning before it can be operated at its design energy.

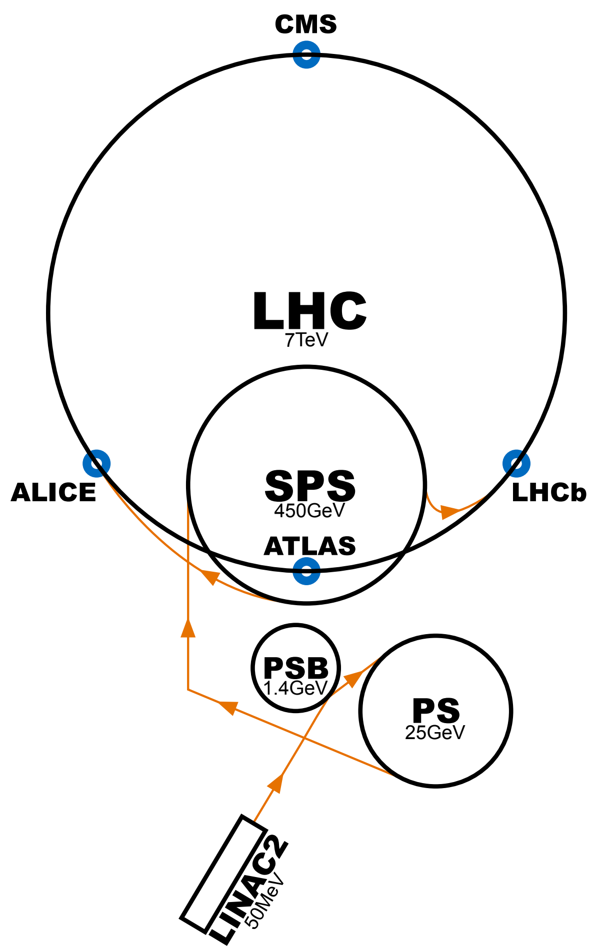


Figure 3.1: Partial schematic of the CERN accelerator complex relevant to LHC proton-proton collisions. The maximum proton energy is listed at the end of each acceleration stage. The positions of the four major experiments are also shown.

LINAC2 and all subsequent accelerators are based on radio-frequency (RF) cavities. An RF cavity is a conducting enclosure in which an electromagnetic standing wave is maintained, much like a microwave oven. The oscillation of the standing wave is controlled such that proton bunches entering the cavity will be consistently accelerated in the desired direction. An important fact about RF cavities is that they cannot maintain a continuous beam of particles. The beam has to be split into a train of bunches.

The next step is to inject the 50 MeV protons into the Proton Synchrotron Booster (PSB), the first synchrotron of the CERN accelerator complex. The advantage of a synchrotron over a linear accelerator is that the same RF cavities can be used many times on the same protons. This is why all following stages of acceleration are also done in synchrotrons. The protons have to be kept on a circular trajectory, which necessitates an additional magnet system. Also, the energy loss of the protons due to synchrotron radiation must be compensated. The PSB takes the protons to an energy of 1.4 GeV before sending them to the Proton Synchrotron (PS). The PS accelerates the protons to 25 GeV and sends them to the Super Proton Synchrotron (SPS), which takes them to an energy of 450 GeV before passing them to the LHC.

The LHC is a complex device. While a high quality vacuum is maintained in its 27km long beam-pipes, quadrupole magnets prevent the injected proton bunches from dispersing. Dipole magnets are responsible for keeping the protons on a precise circular orbit while RF cavities take the bunches up to 7 TeV of energy. Other complex magnet systems are responsible for making the two counter-rotating beams cross each other at well-defined collision points. Also, most LHC magnets are based on superconductors, which requires the largest cryogenic installation in the world.

The LHC is not only at the energy frontier of particle physics, it is also the highest luminosity hadron collider yet built. The luminosity is simply the number of particles per unit area and unit time which can be given in $\mathrm{b}^{-1}\mathrm{s}^{-1}$ or in $\mathrm{cm}^{-2}\mathrm{s}^{-1}$. When the luminosity is multiplied by the cross-section of any given process, one gets the rate at which this process occurs. A high luminosity is necessary to study processes with very small cross-sections. The LHC is designed to achieve a luminosity of $10\mathrm{nb}^{-1}\mathrm{s}^{-1}$ ($10^{34}\mathrm{cm}^{-2}\mathrm{s}^{-1}$). The luminosity integrated over time is referred to as the integrated luminosity, given in b^{-1}. The integrated luminosity is a convenient measure of how much data have been collected.

3.2 The ATLAS experiment

The ATLAS detector is designed for a very broad physics programme. It must be capable of characterizing many types of particles, which no single technology can do. The design of a multi-purpose detector must involve a number of complementary subsystems and compromises must be made. The quantity of data and the geographical distribution of the physicists analyzing the data also raise challenges that require state-of-the-art computing solutions. The aim of this section is to introduce the components of the ATLAS system which make it possible to search for τ signatures.

3.2.1 Hardware

The ATLAS detector can be divided in three main subsystems. The inner detector is the closest to the interaction point. It specializes in non-destructive measurements of charged particles. The calorimeter system surrounds the inner detector and it conducts destructive measurements on a wider range of particle types, since it can also detect some neutral particles. The outermost system is the muon spectrometer, which is entirely dedicated to tracking muons. This section will survey these three systems. All other systems are directly or indirectly assisting the operation of the inner detector, the calorimeter system and the muon spectrometer. These supporting systems include powerful magnets to bend the trajectories of charged particles, cooling systems, alignment sensors, physical support structures, etc. Fig. 3.2 shows a digital representation of the ATLAS detector.

All of these systems must be able to support the stress imposed by the LHC collisions. At design luminosity, proton bunches will collide every 25ns, and there will be 23 collisions per bunch crossing on average [12]. Systems close to the beam line must be radiation resistant or easy to replace. Also, the detector must have enough time to recover before it can respond to a new collision. The fine segmentation of many subsystems of the ATLAS detector helps to avoid high occupancy.

The ATLAS coordinate system

The conventional ATLAS coordinate system is used throughout this thesis. It is defined as follows. First, the z-axis of a cartesian frame is defined to be along the beam axis. The

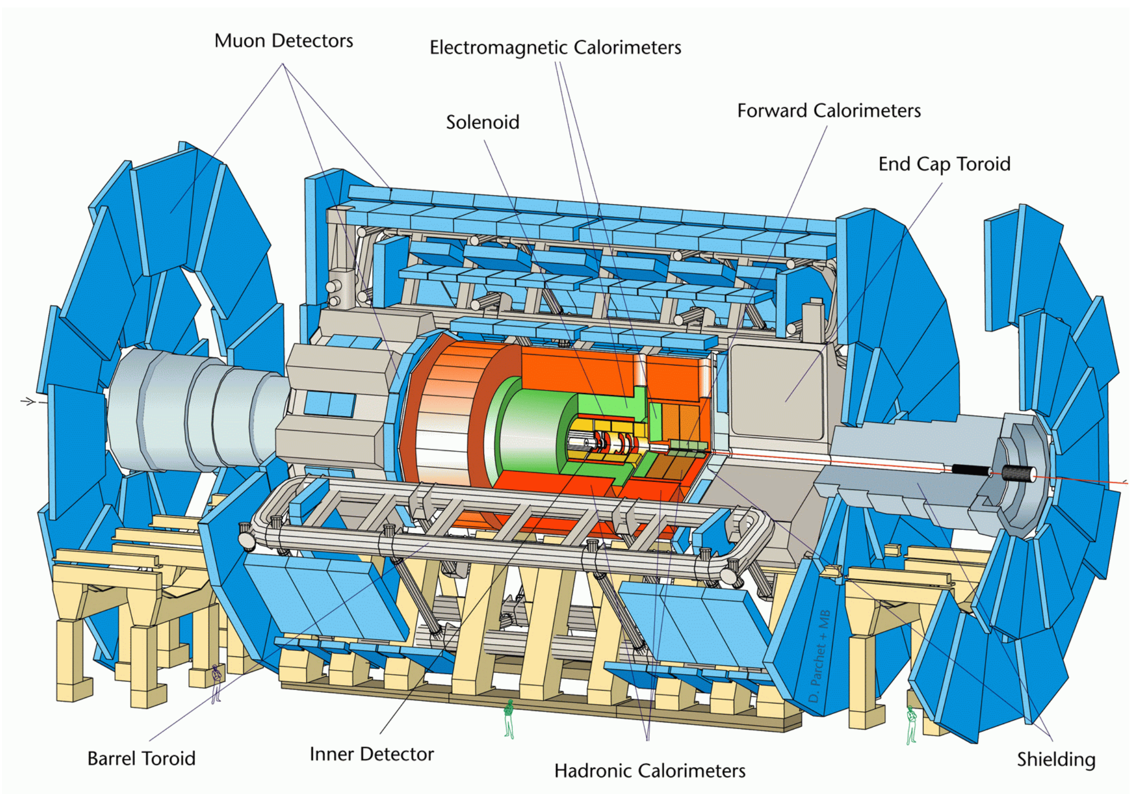


Figure 3.2: Representation of the ATLAS detector showing the main subdetectors. ATLAS Experiment Image: ©CERN [12].

perpendicular x-y plane is oriented such that the x-axis points toward the center of the LHC ring and the y-axis points upward. The azimuthal angle in the x-y plane is represented by ϕ and the polar angle taken with respect to the z-axis is θ. For physics analysis, θ is rarely used since particles coming out of the collisions will be more or less boosted (in the special relativistic sense) along the z-axis. The pseusdorapidity $\eta = -\ln\tan(\theta/2)$ is an equivalent measure of the polar angle which has the benefit of having Lorentz invariant intervals. The two-dimensional space defined by η and ϕ is where most of the work in this thesis takes place.

The inner detector

The main purpose of the inner detector is to precisely measure the direction and momentum of charged particles, while interacting with them as little as possible. There are three different technologies used in the inner detector: the pixel detector, the silicon tracker (SCT) and the transition radiation tracker (TRT). The pixel detector and the SCT exploit the same physical principle: the creation of electron-hole pairs in semiconductors by passing charged particles. The TRT relies on the ionization trail left by a charged particle in the gas to perform tracking. In both types of technology, the freed charges are collected and transformed into usable signal. The TRT can also detect transition radiation, which is mostly used for electron identification. The response to a charged particle returned by the inner detector is a chain of point-like "hits" which are used to interpolate a trajectory. The complete inner detector is bathed in a strong magnetic field which bends the particle's path according to their momentum. The spatial resolution of the pixel detector and the silicon tracker makes it possible to reconstruct points of origin of multiple tracks, or vertices, either coming from primary interactions or from the decay of particles like B mesons and τ. The TRT is there to provide additional hits and electron/hadron separation.

The calorimeter system

The purpose of a particle physics calorimeter is to present a volume of sensitive material sufficiently dense to completely stop most neutral and charged particles in order to measure their energy. The only particles that the ATLAS calorimeter system is not intended to stop are neutrinos and muons. The particles causing a calorimeter response are usually separated in two categories: electromagnetic and hadronic. Electrons and photons create relatively dense energy deposits which are generally measured in the electromagnetic calorimeter. Hadronic particles include long-lived hadrons like protons, neutrons and other mesons. These particles can only deposit most of their energy in the calorimeter by interacting strongly with atomic nuclei. Since these interactions are less likely, hadronic showers penetrate more deeply in the material. The hadronic calorimeter provides more depth of sensitive material to contain hadronic energy deposits. The calorimeter system is of utmost relevance to this thesis and will be covered in detail in chapter 4.

The muon spectrometer

The muon spectrometer takes advantage of the large quantity of matter between itself and the interaction point to filter out most long-lived particles. Muons are usually the only visible particles passing through. The muon spectrometer possesses its own dedicated magnet system which consists of eight large toroidal magnets. It is capable of identifying muons and measuring their momentum with high efficiency. Since it is of no relevance to the work presented in this thesis, it will be discussed no further.

3.2.2 Software

Data acquisition and triggers

Recording the data coming out of the ATLAS detector is a technical challenge. At design luminosity, the rate of bunch crossings is ~40MHz [12] and for each bunch crossing, there is some data coming from each of the millions of channels of the ATLAS detector. It is not practical to record all of these data. The solution adopted by the ATLAS collaboration consists of three levels of filters, or triggers, which reject events (bunch crossings). The first level has to deliver the fastest decisions based on very little data processing. Subsequent levels have to deal with reduced rates of incoming data, allowing for further processing and more informed decisions. A trigger menu is designed to select events relevant to the large variety of analyses done at ATLAS. As the luminosity conditions of the LHC change, the triggers also change. At design luminosity, the trigger selection should be capable of reducing the recorded data rate to ~200Hz.

Data distribution and analysis

ATLAS will collect several petabytes of raw data per year at design luminosity. A world-wide network of computing facilities, the LHC Computing Grid (LCG), is used by ATLAS and the other LHC experiments to store and process this data. The LCG is structured in Tiers [13]. The Tier-0 facility at CERN is responsible for processing and storing the raw data. The data files it produces are forwarded to the Tier-1 facilities where they are made available to the research community for longer periods of time. ATLAS currently has 10 Tier-1 facilities around the world, including a Canadian Tier-1 located at TRIUMF. Ideally, data files should be available from two different facilities to ensure availability and security. There is another level of smaller, but more numerous, centers called Tier-2 facilities which are dedicated to providing computing resources for data analysis. This is also where the centralized production of simulated data is taking place.

Simulated ATLAS data

In any experiment, a model which embodies our understanding and predictions about the experiment must be compared to the data. At ATLAS, the understanding of the physics of the collisions is expressed using a range of Monte Carlo particle physics event generators. Collision events are generated, and then, the response of the ATLAS detector to these events is simulated. The result of this simulation is a data sample in the exact same format as events from real collisions. It is thus possible to run any analysis on both simulated and real data for the purpose of direct comparison. Until late November 2009, no collision data were available. Simulated data were instrumental in preparing the ATLAS reconstruction software, which is responsible for building physics objects from the raw detector output.

Chapter 4

τ reconstruction and identification at ATLAS

The term "reconstruction" is commonly used in the ATLAS collaboration to designate all the computational steps from the raw data of the detector to what is used for physics data analysis. Reconstruction at ATLAS can be roughly divided into low-level and high-level. Low-level reconstruction refers to the construction of basic objects like interpolated tracks or clusters of calorimeter cells. High-level reconstruction uses these basic objects to build physics object *candidates*, like jets, muons, electrons or τ candidates. A sample of τ candidates may contain a large fraction of fake τ. Identification is required on top of reconstruction to remove as many fakes as possible. The purpose of the work presented in this thesis is to facilitate τ identification.

The aim of the first two sections of this chapter is to build an understanding of the τ lepton and its principal background: hadronic jets. Then, a summary of τ reconstruction from the inner detector is given, followed by a detailed description of τ reconstruction in the calorimeter system.

4.1 τ leptons

τ are the charged leptons of the third generation. They are more massive than a proton ($m_\tau = 1.777$ GeV [2]) and they are short-lived. At ATLAS, their average lifetime of

Table 4.1: Most important decay modes for τ [2].

decay modes	branching fraction (%)
leptonic modes	
$\tau^{\pm} \rightarrow e^{\pm}\nu_e\nu_\tau$	17.4
$\tau^{\pm} \rightarrow \mu^{\pm}\nu_\mu\nu_\tau$	17.9
hadronic modes	
$\tau^{\pm} \rightarrow \pi^{\pm}\nu_\tau$	10.9
$\tau^{\pm} \rightarrow \pi^{\pm}\pi^{0}\nu_\tau$	25.5
$\tau^{\pm} \rightarrow \pi^{\pm}\pi^{0}\pi^{0}\nu_\tau$	9.3
$\tau^{\pm} \rightarrow \pi^{\pm}\pi^{0}\pi^{0}\pi^{0}\nu_\tau$	1.0
$\tau^{\pm} \rightarrow \pi^{\pm}\pi^{\pm}\pi^{\mp}\nu_\tau$	9.3
$\tau^{\pm} \rightarrow \pi^{\pm}\pi^{\pm}\pi^{\mp}\pi^{0}\nu_\tau$	4.6

291×10^{-15}s is long enough so that their decay products can sometimes be seen to originate some distance from the interaction point. However, τ never leave the beam pipe of the LHC before decaying. Only their decay products are seen in the detector.

τ decay via the weak interaction and all decay modes include neutrinos. Their different decay modes are usually categorized in purely leptonic decays or semi-hadronic (hadronic) decays. The most important decay modes are listed in Table 4.1 with their branching fractions.

Since neutrinos cannot be seen at ATLAS, the only way to distinguish a purely leptonic τ decay from prompt electrons or muons is by using the decay vertex. Purely leptonic decay modes are very difficult to exploit for τ identification in a busy environment like that of the LHC. They are not considered here.

Hadronic modes feature pions (π mesons). Neutral pions (π^0) are bound states of a light quark and its anti-quark which decay almost instantly (for all practical purposes at ATLAS) 99% of the time to two photons [2]. Charged pions ($\pi^{\pm}$) are combinations of a u and a d quark which are longer-lived. In fact, they live long enough to be considered stable particles at ATLAS. τ hadronic decays are commonly classified according to the number of charged decay products, conventionally called "prongs". 1 prong τ decays make up about

50% of all decay modes, while 3 prong decays make up 15%.

While $\tau^{\pm} \rightarrow \pi^{\pm}\nu_{\tau}$ is sometimes difficult to distinguish from a prompt electron, τ decays with many decay products are harder to distinguish from hadronic jets. Fortunately, when there are multiple decay products, quantities like the transverse width of the decay or the invariant mass of the decay products can be used to discriminate against these jets. The width of a τ decay is controlled by the boost of the parent τ and its mass. Even at energies as low as 10 GeV, a τ decay is narrower than most hadronic jets.

4.2 Hadronic jets

When probed with very high energy particles, protons are seen to have a more complex substructure than just three bare quarks bound by a few gluons. Virtual processes inside create a complicated network of gluons and particle/antiparticle pairs. When two protons collide at the LHC, any of the bare quarks, gluons or virtual quarks, collectively called partons, may interact. In fact, the momentum of the whole proton is divided among its constituents. The energy of the interactions taking place between the partons of one proton and those of another is known empirically in the form of statistical distributions called parton distribution functions [14]. It follows that a broad energy spectrum can be investigated without changing the configuration of a hadron collider, which is convenient in the search for new phenomena. The price to pay is that the protons are practically ripped apart when they collide. Because of a process called hadronization, large numbers of hadrons are sprayed all over the detector. It can be difficult to recognize specific processes in such a messy environment.

In a so-called hard-scattering event, two partons interact and produce significant momentum transverse to the beam direction. A common process of this type is the production of a pair of gluons or quarks going in opposite directions in the center-of-mass frame of the collision. At first sight, this should be impossible because of colour confinement. Hadronization, for which one of the best descriptions is found in the Lund string fragmentation model [15], is the process by which the colour field of the final state is broken in two by the creation of a quark/antiquark pair. The result is two colourless bound states flying in opposite directions in the center of mass frame. A common visualization of the process is to pull on both ends of an elastic band until it breaks in two. Ends of the elastic bands represent quarks while the band itself represents the gluon field. Hadronization does not stop there. There may still be large quantities of energy in the gluon fields. Newly formed hadrons will keep on dividing in the same way until there is not enough energy left [16]. The process is illustrated in Fig. 4.1. In the end, the result of a dijet hard-scattering event is two collimated sprays of hadrons which are called hadronic jets, or QCD jets[1].

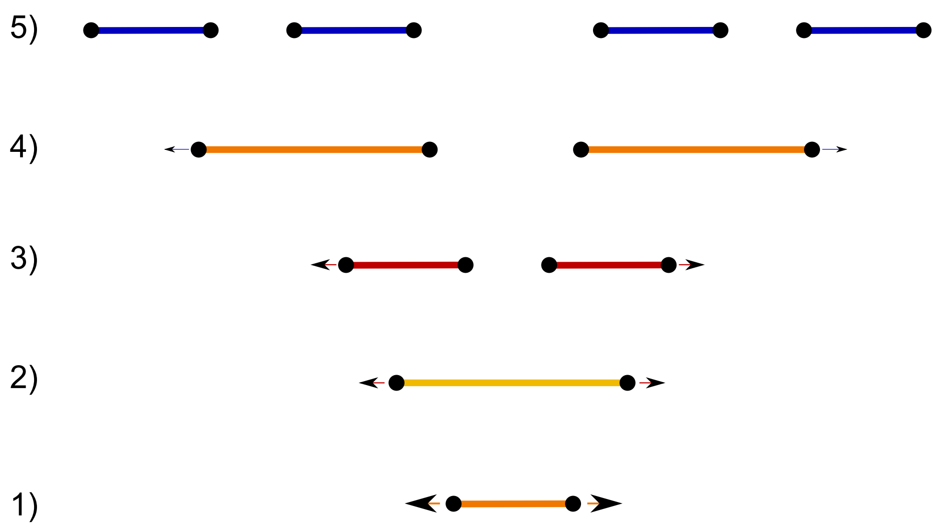


Figure 4.1: Hadronization process. The black dots are quarks, the line connecting them are gluon fields. 1) The initial quarks have momenta going in opposite directions. 2) Part of their kinetic energy is transferred to the gluon field. 3) The gluon field has enough energy to conjure a quark/antiquark pair, breaking it in two. 4) There is energy left in the gluon fields. 5) The remaining energy is dissipated through the creation of more quark/antiquark pairs and fragmentation continues.

These hadronic jets are mostly made of very common hadrons: pions. In addition, hadronic jets are produced in copious quantities at the LHC, orders of magnitude more frequently than τ. Reducing the hadronic jet background is a top-priority in τ identification.

Fortunately, hadronic jets are different from τ decays in a few respects. The transverse

[1]What actually happens is complicated by higher-order processes. Initial and final state gluon radiation will produce multi-jet final states.

width of hadronic jets is not the result of the conversion of the mass of a decaying particle to kinetic energy. It comes from the energy stored in the gluon fields of the fragmenting hadrons. This makes hadronic jets significantly wider than hadronic τ decays. Also, it makes the invariant mass calculated from the final state hadrons strongly correlated with the jet energy. Furthermore, hadronic jets usually have more final state particles than hadronic τ decays.

4.3 τ reconstruction in the ATLAS inner detector

The reconstruction of τ is an example of high-level reconstruction. The basic idea is to find τ candidates using loose selection criteria to accept as many real τ as possible. Then, for each candidate, a number of quantities are calculated which can be used to discriminate against the background. τ reconstruction at ATLAS uses two sets of criteria to find τ candidates. The first is based on the information from the inner detector, while the second uses the calorimeter. Any track or jet satisfying either set of criteria is commonly referred to as a seed.

As mentioned in section 3.2.1, the inner detector is only capable of detecting charged particles. 1 prong τ decays are seen as a single track, while 3 prong τ decays are perceived as a tight cluster of three tracks. The 4-momentum of the charged particles is reconstructed with great accuracy. Hadronic jets will usually produce many tracks.

Charged particle reconstruction via the tracks is not perfect. It depends on which type of particle produces the track, whether the track is isolated or not, noise in the electronic readouts, etc. The tracking system at ATLAS does not cover $|\eta| > 2.5$. Since tracking information is very important to achieve good τ identification, no τ candidate is reconstructed beyond $|\eta| > 2.5$.

A τ seed found in the inner detector, or track seed, has to be a track which satisfies the following criteria:

- a transverse momentum[2] $p_T \geq 6$ GeV;
- at least 7 hits in the silicon tracker;

[2] p_T is related to the momentum p by $p_T = p \sin\theta$ or $p = p_T \cosh\eta$. Momentum in the transverse plane to the beam axis in a collider experiment can only be generated by a scattering event.

- a transverse impact parameter[3] $d_0 > 2$mm.

Once such a track seed is found, additional tracks in the immediate vicinity of the seed are collected as part of the τ candidate. The information contained in the resulting set of tracks is used to build identification variables. The identification variables based on the tracking information include the invariant mass and the width of the track system, which are useful only if there is more than one track. It also includes the impact parameter in 3-dimensional space.

There are also a number of variables which use both the tracks and the associated energy deposits in the calorimeter. These variables are essentially energy/momentum ratios where the momentum considered is that of the leading track or that of all tracks, while the energy is taken from different geometrical regions in the calorimeter. A list of definitions for the tracking variables implemented[4] at ATLAS can be found in Ref. [8].

4.4 τ reconstruction in the ATLAS Calorimeter

This section delineates the experimental context in which the study presented in this thesis has been done. The principal goal of the current study is to improve τ identification by using more of the calorimeter information. Even if the ATLAS calorimeter can never be as precise as the inner detector in reconstructing the direction and momentum of particles, it has a significant advantage: it can detect neutral particles such as photons and neutrons. Understanding how the ATLAS calorimeter can be used for τ/jet separation requires some background knowledge on the physical principles upon which it operates, which this section will provide. This section also describes the ATLAS experiment's implementation of calorimeter systems and the main software algorithms used to exploit the information they provide.

[3]The transverse impact parameter is defined as the point of closest approach of the interpolated track to the position of the collision point in the *x*-*y* plane.

[4]As the τ reconstruction software matures, new identification variables are implemented and existing definitions may be changed. The variable list found in Ref. [8] may not be up-to-date, but it still features most of the widely used variables.

4.4.1 Shower physics

Calorimeters are devices that measure the energy of impinging particles. Sufficiently energetic electrons and photons initiate a cascade of processes called electromagnetic showers, while charged pions, protons and neutrons initiate hadronic showers.

Electromagnetic showers

A highly energetic photon penetrating some material will most likely decay into an electron/positron pair. In the same situation, a highly energetic electron will most likely emit a bremsstrahlung photon. It is easy to see how a cascade of electrons and photons may form if the initiating particle has enough energy. Such a cascade is an electromagnetic shower.

A quantity called the radiation length X_0 is very convenient to describe electromagnetic showers. It is defined as the distance an electron must travel through matter in order for radiation losses to reduce its energy by a factor of $1/e$ [17], e being the base of natural logarithms. In pure materials (only one kind of atom), it can be approximated by [18]

$$X_0 \simeq 180 \frac{A}{Z^2}, \tag{4.1}$$

where A is the atomic weight and Z is the atomic number. The units of the radiation length are g/cm^2. It is customary to give the physical dimensions of electromagnetic showers in radiation lengths, since the number of radiation lengths is mostly independent of the material in which the showers occur.

Most features of an electromagnetic shower can be described qualitatively using a highly simplified model [17], which is illustrated in Fig. 4.2. An initial photon[5] has energy E_0. Once it enters the material, it undergoes pair creation after travelling one radiation length. The resulting two particles each have an energy of $E_0/2$. After another radiation length, the electron and the positron each emit a bremsstrahlung photon. There are now four particles with $E_0/4$ each. This process continues until the energy of the latest particles is low enough that other energy loss mechanisms dominate. It is mostly through these other mechanisms that the energy measurement is done in a calorimeter, which will be discussed in the context of the ATLAS calorimeter system.

[5] A shower initiated by an electron can also be described by the same model.

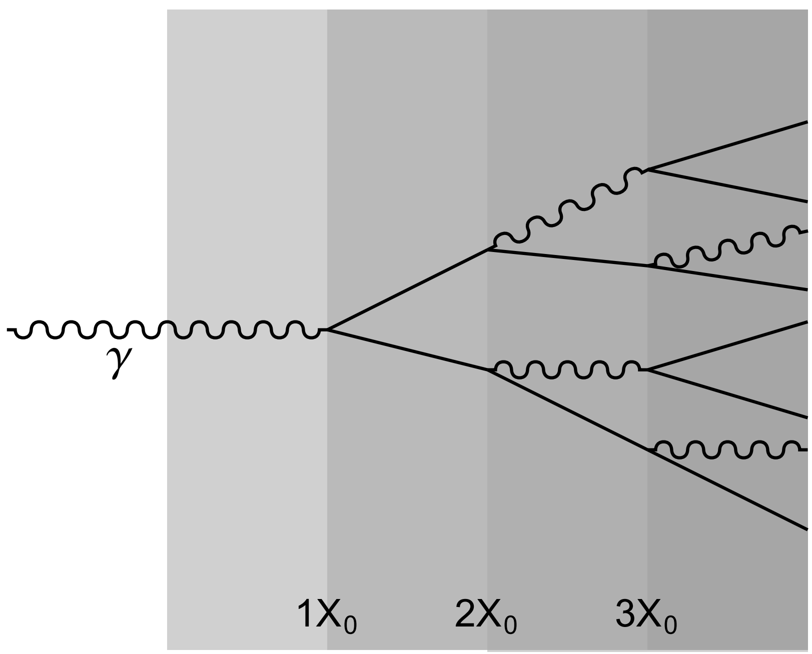


Figure 4.2: Development of an electromagnetic shower according to the simplified model. A wiggly line represents a photon, while a straight line represents an electron or a positron.

Since energy is deposited in the material wherever there are shower particles, an accurate energy measurement can only be achieved if the shower is contained. Knowledge of the longitudinal and transverse profiles of the expected showers is required for a good calorimeter design. The simplified model predicts that the depth of the electromagnetic shower is proportional to the energy of the incident particle. It is therefore sufficient for an electromagnetic calorimeter to be able to contain showers from the highest energy electrons or photons expected. To give a quantitative feel, $20X_0$ are necessary to completely contain the shower from a 30 GeV electron or photon [17].

The dependence of the transverse shower profile on the energy of the incident particle is actually very small. It is usually given in terms of the Moliere radius, defined as [17]

$$R_M = \frac{21.2\ \text{MeV}}{E_c} X_0, \tag{4.2}$$

where E_c is a critical energy below which the particles of the shower can no longer perpetuate the cascade. E_c is typically of the order of 100 MeV [18]. An electromagnetic shower will usually deposit 95% of its energy within two R_M of the shower axis.

An important consequence of the transverse shower profile is that it limits the maximum spatial resolution of an electromagnetic calorimeter. For the purposes of τ identification, the most relevant electromagnetic showers are the ones initiated by photons coming from neutral pion decays. It is very difficult to distinguish the two showers coming from a single π^0, but it can be done if the showers are sampled in the earliest stages of their development. However, reconstructing the energies of the individual photons is much more difficult than simply distinguishing them.

A real electromagnetic shower from a cosmic ray recorded using the new ICARUS T600 detector in Gran Sasso, Italy is shown in Fig. 4.3. The shower can be seen to be very large, due to the long radiation length of the medium in which it formed, which is liquid argon.

Hadronic showers

Hadrons also initiate showers, but the way they interact with the material is much more complicated. Since they are more massive, they are not disturbed much by the electrons of the medium. Hadronic showers are mostly caused by the strong interactions of the impinging particles with atomic nuclei. Since the cross-section for such interactions is lower than

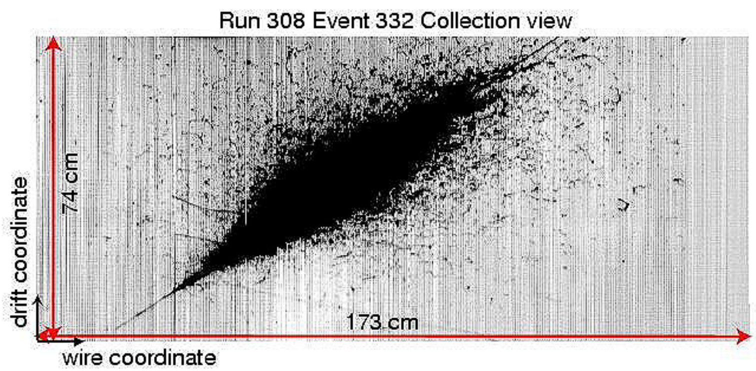


Figure 4.3: Real electromagnetic shower from a cosmic ray observed in the ICARUS T600 detector during a technical run in 2001. ICARUS Experiment Image: ©ETH Institute for Particle Physics.

for the simple interactions of the electromagnetic showers, hadronic showers take longer to develop. Containing hadronic showers requires a lot of material, which explains the need in some experiments to have an extra calorimeter surrounding the electromagnetic calorimeter.

While electromagnetic showers are governed by two simple processes and only two types of particles, hadronic showers involve the creation of many sorts of hadrons which all interact differently. For example, slow neutrons may wander for several milliseconds before they are absorbed by atomic nuclei and liberate their energy in a measurable form[19]. This implies that hadronic showers are not only difficult to contain in space, they are also difficult to contain in time. Another example of the variety of processes taking place in a hadronic shower is the generation of neutral pions, which result in secondary electromagnetic showers. In fact, about 20% of the energy deposited in a hadronic shower comes from neutral pions in a hadronic shower initiated by a 10 GeV proton [20]. A hadronic shower may be difficult to recognize as a single energy deposit.

While the radiation length X_0 is convenient to describe electromagnetic showers, hadronic showers are better described by the interaction length or absorption length λ. It is defined as [18]

$$\lambda = \frac{A}{N_A \sigma_{\text{abs}}}, \tag{4.3}$$

where A is the atomic weight, given in g/mol, N_A is Avogadro's number (~6.022×10^{23} atoms/mol) and σ_{abs} is an absorption cross-section defined as

$$\sigma_{\text{abs}} = \sigma_{\text{T}} - \sigma_{\text{el}} - \sigma_{\text{q}}, \tag{4.4}$$

where σ_{T} is the total interaction cross-section, σ_{el} is the cross-section for elastic scattering off a nucleus and σ_{q} is the cross-section for the quasi-elastic scattering off nucleons. For example, the longitudinal development of a hadronic shower from a 40 GeV hadron is mostly contained in 5λ. On the other hand, the lateral development of the shower does not relate to λ.

There is no simple model which can account for the basic characteristics of hadronic showers, because of the large number of possible processes involved. Even complex models based on Monte Carlo simulation still have limited domains of validity [21]. Test-beam experimentation is still necessary for successful hadronic calorimeter design. Hadronic showers can nevertheless be compared qualitatively with electromagnetic showers on many aspects. They are usually longer, wider, and they do not have a well-defined shape. An example of a real hadronic shower from the ICARUS experiment is shown in Fig. 4.4.

4.4.2 The ATLAS Calorimeter system

The ATLAS calorimeter system is represented in Fig.4.5. The electromagnetic calorimeter is divided into a cylinder (barrel) and two end-caps. Its active medium is liquid argon. The hadronic calorimeter is divided into one barrel, two extended barrel sections, and two end-caps. The active medium in the barrel and the extended barrels is a plastic scintillator, while the active medium in the hadronic end-caps is liquid argon. How these two different materials collect the energy of showers will be explained in this section. All ATLAS calorimeters are sampling calorimeters, which means that there is dead material of short radiation length interspaced in the active volume in order to precipitate the shower development. The calorimeters are also divided into cells, which are small volumes giving a single energy reading. Since the forward calorimeters are not used for τ identification, they will not be described here.

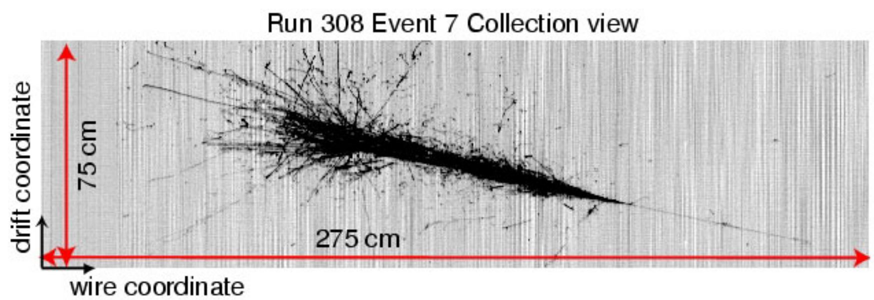


Figure 4.4: Real hadronic shower from a cosmic ray observed in the ICARUS T600 detector during a technical run in 2001. ICARUS Experiment Image: ©ETH Institute for Particle Physics.

Electromagnetic calorimeter

The barrel and the end-cap of the electromagnetic calorimeters are made of accordion shaped sheets of lead, stacked with copper electrodes with space in-between for the liquid argon. The biggest advantage of this accordion geometry is that the amplitude of the accordion wave can be varied in order for a stack of metallic sheets to fill volumes with circular symmetry, while keeping the space between the sheets constant (see Fig. 4.6). The result is that the electromagnetic calorimeter has no sensitivity gaps in ϕ.

The cell segmentation is implemented by delimiting conductive areas of copper on the electrode sheets, much like common electronic boards. The gaps between the electrode sheets and the lead sheets is filled with liquid argon, which is kept at a precise temperature by cryostats. The barrel and the end-caps are kept cold by different cryostats.

When a charged particle passes through the liquid argon, it ionizes the medium. The free charges liberated are collected by the copper electrodes placed under high voltage and an analog pulse is sent to the electronic read-outs. Ionization is another very important mechanism by which charged particles lose energy in a medium. The hundreds of electrons and positrons that are generated in electromagnetic showers all participate in ionization, but ionization only becomes a major source of energy loss at low particle energies. Even the

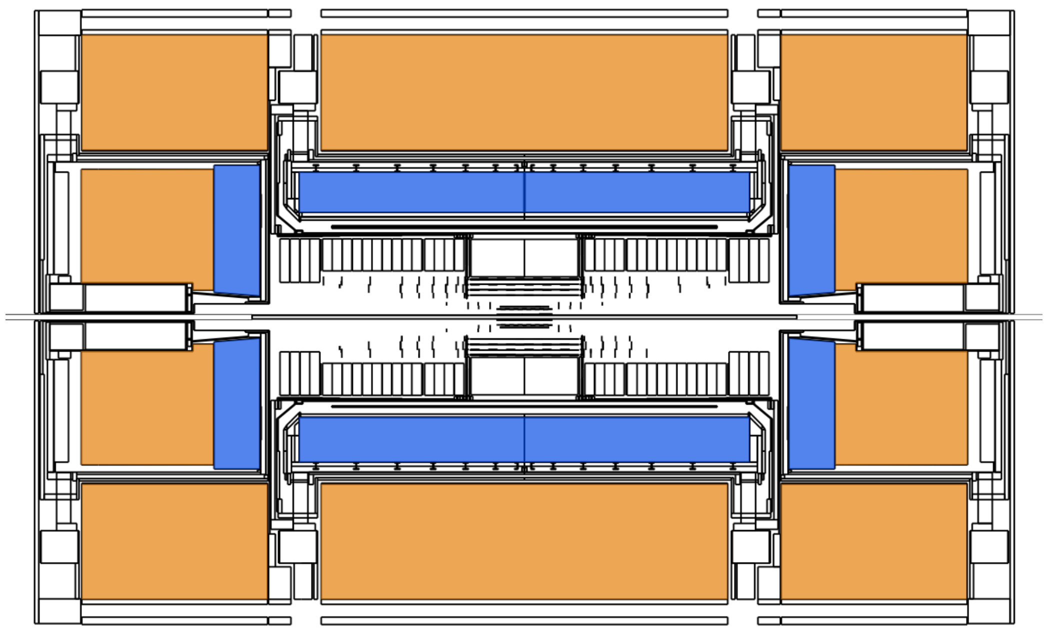


Figure 4.5: Longitudinal section of the inner detector and calorimeter system of the ATLAS detector. The sections highlighted in blue are the barrel and end-cap electromagnetic calorimeters. The sections highlighted in orange are the barrel, extended barrel and end-cap hadronic calorimeters. Not highlighted but also shown are the forward calorimeters, between the beam axis and the end-cap modules. ATLAS Experiment Image: ©CERN. [22]

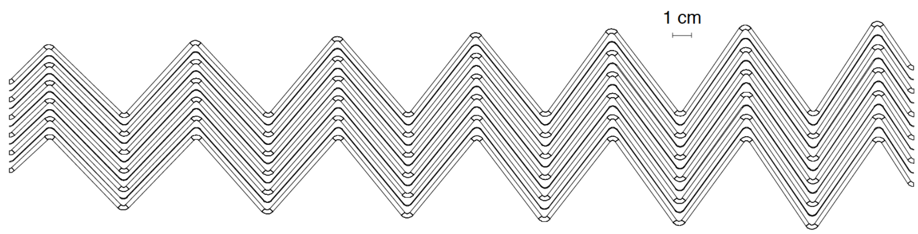


Figure 4.6: Cross-section view of the accordion geometry of the lead sheets and copper electrodes in the electromagnetic barrel. The total space taken by the sheets is seen to increase proportionally with the radius (from left to right). ATLAS Experiment Image: ©CERN. [22]

last particles created in the showering process will deposit their energy in the liquid argon.

The cell geometry of the electromagnetic calorimeter is based on pseudorapidity (see Fig. 4.7). Cell boundaries are aligned with the center of the detector, and the transverse width of the cells varies in θ to be constant in η. The granularity in ϕ is constant. The calorimeter is divided in three layers. The first is commonly called the η-strip layer, since it has an especially fine segmentation in η. In fact, this segmentation is fine enough to make γ/π^0 separation possible. The second layer is where most of the energy of typical electrons and photons from hard-scattering events is deposited. The third layer is more coarsely segmented, and not very uniform in thickness. Its purpose is to help getting a better energy resolution for showers by providing more active material. The granularity of the complete electromagnetic calorimeter is given in detail in Table 4.2.

The total depth of the electromagnetic calorimeter varies from 24 to $40X_0$. The width of the cells in the second layer of the barrel is designed to be roughly one Moliere radius [22]. This means that electromagnetic showers will not only span multiple cells in depth, but also in the lateral dimension. Of the ~190 000 read-out channels of the electromagnetic calorimeter, a few hundreds may return a signal caused by a τ hadronic decay. The fine granularity of the electromagnetic calorimeter makes it possible to study shower shapes and to resolve individual showers which are physically separated.

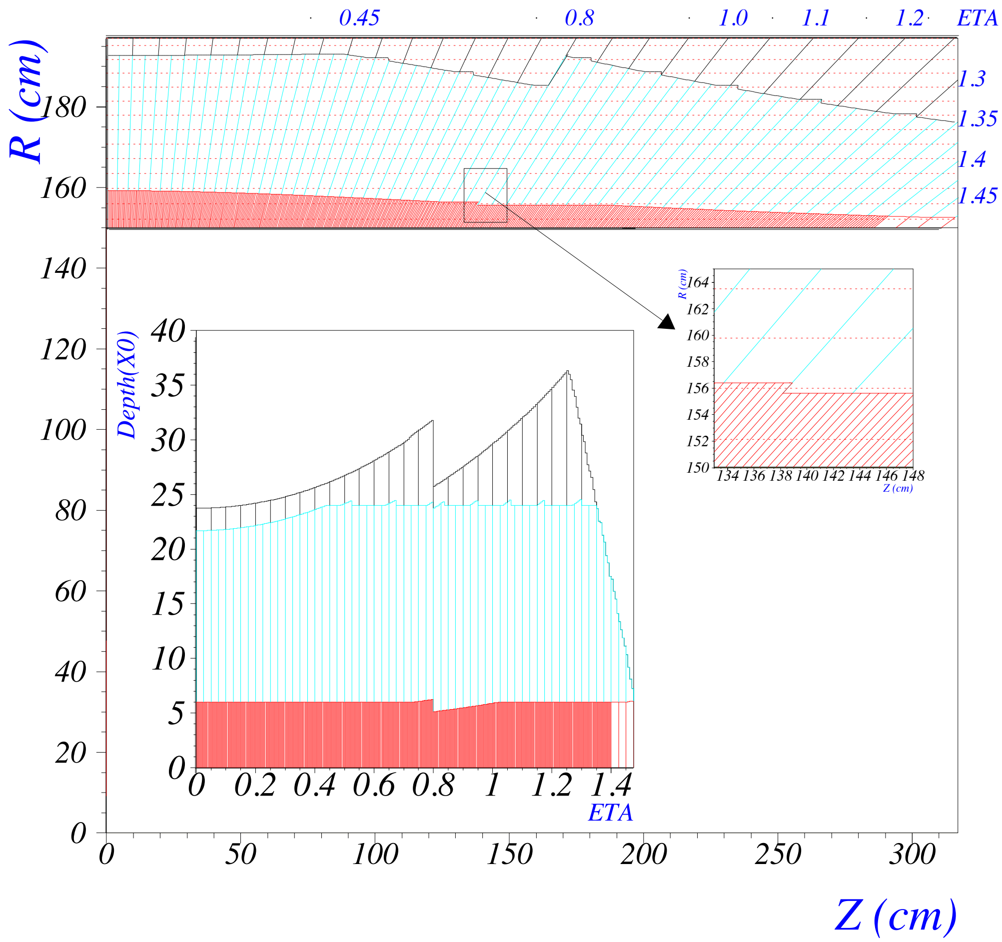


Figure 4.7: Granularity in η of the barrel of the electromagnetic calorimeter, and depth in radiation lengths. ATLAS Experiment Image: ©CERN. [22]

Table 4.2: Granularity and segmentation of the ATLAS electromagnetic and hadronic calorimeters [22]. The forward calorimeter is not shown, since it is not used in τ reconstruction.

Electromagnetic calorimeter	**Barrel**	**End-cap**	
Coverage	$\|\eta\| < 1.475$	$1.375 < \|\eta\| < 3.2$	
Longitudinal segmentation	3 layers	3 layers	$1.5 < \|\eta\| < 2.5$
		2 layers	$1.375 < \|\eta\| < 1.5$
			$2.5 < \|\eta\| < 3.2$
Granularity ($\Delta\eta \times \Delta\phi$)			
Layer 1	0.003×0.1	0.025×0.1	$1.375 < \|\eta\| < 1.5$
		0.003×0.1	$1.5 < \|\eta\| < 1.8$
		0.004×0.1	$1.8 < \|\eta\| < 2.0$
		0.006×0.1	$2.0 < \|\eta\| < 2.5$
		0.1×0.1	$2.5 < \|\eta\| < 3.2$
Layer 2	0.025×0.025	0.025×0.025	$1.375 < \|\eta\| < 2.5$
		0.1×0.1	$2.5 < \|\eta\| < 3.2$
Layer 3	0.05×0.025	0.05×0.025	$1.5 < \|\eta\| < 2.5$
Hadronic calorimeter	**Barrel**	**End-cap**	
Coverage	$\|\eta\| < 1.0$	$1.5 < \|\eta\| < 3.2$	
	$0.8 < \|\eta\| < 1.7^*$		
Longitudinal segmentation	3 layers	3 layers	
Granularity ($\Delta\eta \times \Delta\phi$)			
Layers 1 and 2	0.1×0.1	0.1×0.1	$1.5 < \|\eta\| < 2.5$
		0.2×0.2	$2.5 < \|\eta\| < 3.2$
Layers 3	0.2×0.1	0.1×0.1	$1.5 < \|\eta\| < 2.5$
		0.2×0.2	$2.5 < \|\eta\| < 3.2$

*Extended barrel.

Hadronic calorimeter

The barrel and the extended barrel of the hadronic calorimeter are often called the tile calorimeter as they rely on scintillator plastic tiles to provide an energy measurement. The cells are made of stacks of iron plates and scintillators. The iron plates are used to precipitate the shower formation.

There are a number of possible processes responsible for the absorption of the energy of particles entering a scintillator tile. All absorbed energy has to be liberated in some way. A scintillator has the very useful property that a significant fraction of that energy is released in the form of light to which the scintillator itself is transparent. This light is channelled out of the scintillator tiles by light-guides and collected by photo-multiplier tubes. The electrical signal provided by the tubes is then converted into the energy deposited in the scintillator medium.

The end-caps of the hadronic calorimeter use liquid argon as the sensitive medium. However, unlike the electromagnetic calorimeter, they do not use the accordion geometry. The liquid argon, the electrodes and the absorbing material are arranged in flat stacks oriented in the perpendicular direction to the beam pipe.

The hadronic calorimeter does not have to be as finely segmented as the electromagnetic calorimeter, since it will mostly be inhabited by wide hadronic showers. It does need to be very thick though, to capture all the energy of these showers and to act as a filter for the muon spectrometer. The total depth of the hadronic and electromagnetic calorimeters was required to be ~10λ to achieve this (see Fig. 4.9). The granularity and segmentation of the hadronic calorimeter is detailed in Table 4.2, and illustrated for the barrel section in Fig. 4.8.

A total of ~15 000 read-out channels return the hadronic calorimeter information. When the electromagnetic and hadronic calorimeters are combined, it is possible to reconstruct complete showers and measure their energy, longitudinal and lateral profiles, energy density, etc. Since showers originate from energetic particles, the calorimeter can thus be used to reconstruct these particles. Of course, there are complications when there are many particles showering in a small calorimeter region, but the goal is not usually to reconstruct accurately individual showers. The amount of information provided by the calorimeter system to discriminate τ hadronic decays from hadronic jets is potentially greater than the

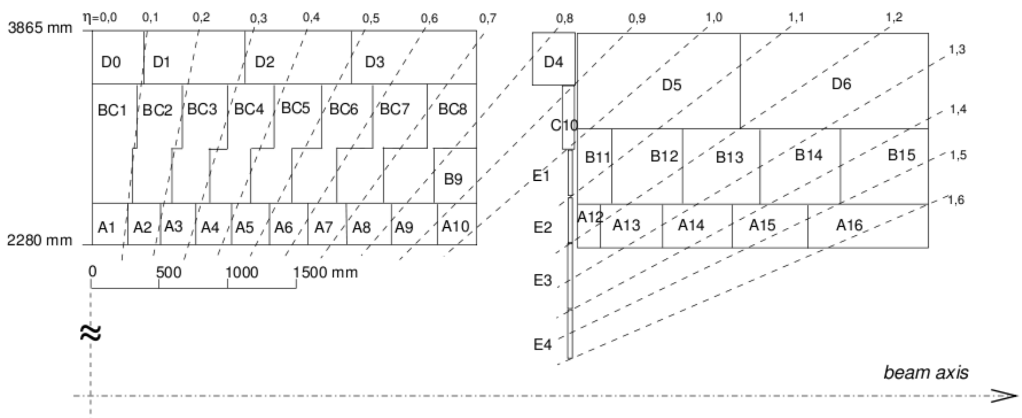


Figure 4.8: Cell division of the tile calorimeter in η. ATLAS Experiment Image: ©CERN. [22]

information provided by the inner detector (although not as precise) since it can reconstruct neutral pions.

4.4.3 Shower reconstruction

The topological clustering algorithm

The topological clustering algorithm [23] scans the whole calorimeter system in a search for spatially connected energy deposits. It reconstructs the energy and direction of particles initiating showers. It runs in two steps: making and splitting cell clusters.

Making the clusters works by first identifying cells which have a signal significance over the expected noise of $|E|/\sigma \geq 4$. Each of those cells becomes a seed for a new cluster. The algorithm then looks for neighbouring cells and it adds them to the seed cell. Neighbouring cells are (for most of the calorimeter) the eight surrounding cells in the same layer, as well as the cells of adjacent layers[6] which are at least partially overlapping in η and

[6]The topological clustering algorithm even looks for neighbours across the boundary between the electromagnetic and hadronic calorimeters.

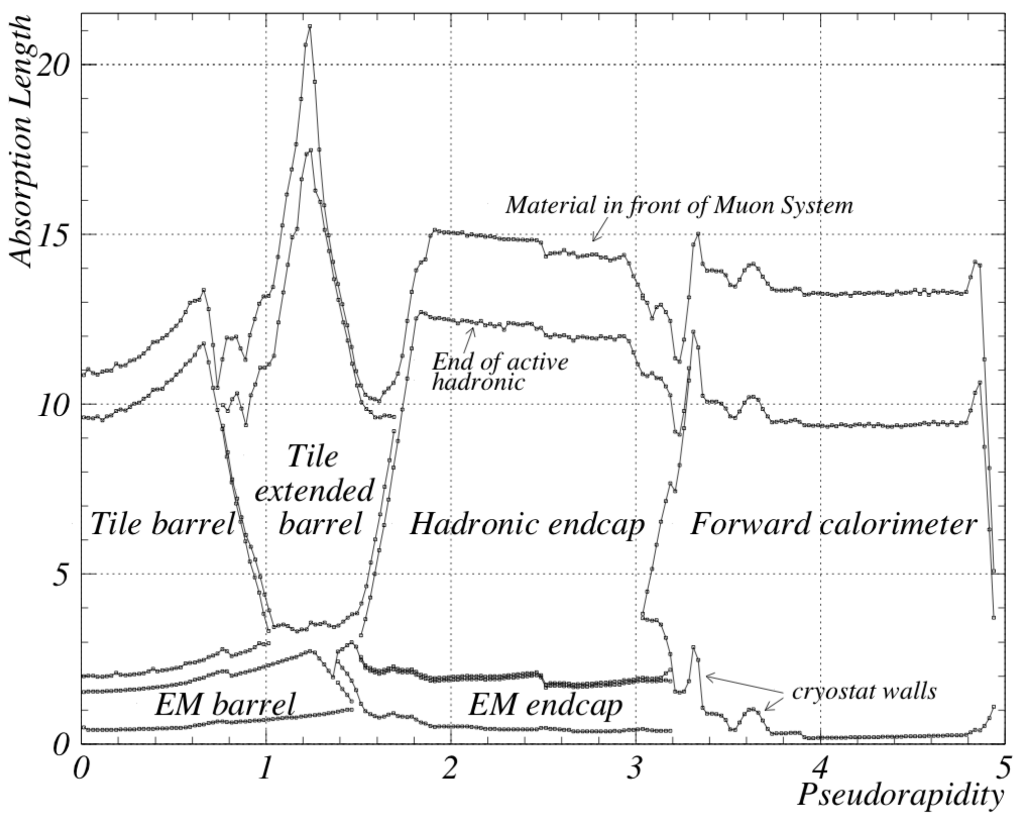


Figure 4.9: Amount of material composing the ATLAS calorimeters as a function of η, given in absorption lengths. ATLAS Experiment Image: ©CERN. [22]

ϕ with the seed cell. If a newly added neighbouring cell has $|E|/\sigma \geq 2$, it then adds the neighbours of this cell to the cluster. It will keep adding new neighbours to the cluster until no new cell with $|E|/\sigma \geq 2$ is encountered. If this procedure ends up connecting two seed cells, a single cluster containing both seeds will be formed. This first step of the topological clustering algorithm has been optimized to give the best performance with single beam tests [24].

Splitting the clusters involves searching for local maxima in clusters built during the cluster making. A local maximum is defined as a cell which has an energy greater than 500 MeV and greater than the energy of any neighbouring cells. A minimum energy value is also required for the neighbouring cells. Each such maximum will then become its own cluster, with no possible merging. Original clusters possessing many of these maxima will be split. New clusters are formed around the maxima by the same procedure used in the cluster making. This time, no threshold on $|E|/\sigma$ is applied since only the cells found during the cluster making are used. Some cells will end up being shared between two clusters. In this case, weights based on the probability of the cells being part of a given cluster are used to divide the energy of these cells among the clusters.

One of the biggest advantages of the topological clustering algorithm is that it rejects cells containing only noise, improving the energy resolution of more complicated objects made of topological clusters. In spite of the fact that its default implementation at ATLAS has been optimized to gather the energy depositions of hadronic showers from charged pions, it can also reconstruct electromagnetic showers.

Trying to resolve the individual showers coming from a single collimated object like a τ decay or a hadronic jet is not the intended purpose of the topological clustering algorithm. Its main goal is to collect the energy. In a busy environment, it may happen that multiple close-by showers are reconstructed as a single cluster, or that hadronic showers showing strong irregularities are reconstructed as several clusters. Nevertheless, the information provided by the topological clustering algorithm is still very useful for τ identification, as will be shown in the next chapters.

4.4.4 Jet reconstruction

The anti-k_t jet algorithm

In order to reconstruct objects with many decay products, such as hadronic jets or τ decays, a second level of clustering must be performed. Each topological cluster reconstructed has three basic values: an energy E and two angular coordinates: η and ϕ. In the LHC collisions, the outgoing particles often satisfy $E \gg m$. Assuming this, full 4-vectors can be calculated. The list of 4-vectors from all topological clusters is the input to the jet clustering algorithm.

There are many jet algorithms which all have different behaviours and properties. In this study, the anti-k_t algorithm [25] is used. The primary purpose of a jet algorithm is to reconstruct jets with large transverse momentum. For this purpose, the outcome of the jet algorithm must be independent of low energy radiation (infrared safety). It should also be independent of collinear splitting of hadronic jets (collinear safety). The anti-k_t algorithm satisfies both of these requirements, and it has a few additional benefits.

The execution of the anti-k_t algorithm is based on the calculation of two distances:

$$d_{ij} = \min\left(\frac{1}{k_{t,i}^2}, \frac{1}{k_{t,j}^2}\right)\frac{\Delta R_{ij}^2}{R^2}, \tag{4.5}$$

$$d_{iB} = \frac{1}{k_{t,i}^2}. \tag{4.6}$$

Here, $k_{t,i}$ is the transverse momentum of cluster i, R is the jet reach parameter and ΔR_{ij} is the cartesian distance in the η,ϕ-space between cluster i and cluster j given by

$$\Delta R_{ij} = \sqrt{(\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2}. \tag{4.7}$$

In the first iteration of the anti-k_t algorithm, d_{ij} is calculated for each possible pair of clusters ($i \neq j$), and d_{iB} is calculated for each cluster. Then, the smallest of all d_{ij} and d_{iB} is found. If the smallest distance is a d_{ij}, cluster i and cluster j are combined by summing their 4-momenta together. Then, clusters i and j are removed from the input cluster list and they are replaced by the new "protojet" composed of these two clusters. If the smallest distance turns out to be a d_{iB} instead, cluster i is removed from the input cluster list and placed into the final jet list. After following these rules, the distances are recalculated on the updated input list and another iteration begins. The algorithm keeps iterating until no cluster is left in the input list.

A most interesting property of the anti-k_t jet algorithm is how it behaves with low energy clusters. Consider the following scenario. Cluster i is very energetic while cluster j has negligible energy. In this case, we have

$$d_{ij} = \frac{1}{k_{t,i}^2} \frac{\Delta R_{ij}^2}{R^2}, \tag{4.8}$$

$$d_{iB} = \frac{1}{k_{t,i}^2}. \tag{4.9}$$

The only factor determining which of d_{ij} or d_{iB} is the smallest distance is ΔR_{ij}. Cluster j will be clustered with cluster i only if $\Delta R_{ij} < R$. Therefore, if energetic jets are narrower than R and unwanted low energy clusters are disseminated all over the η, ϕ-space, anti-k_t will still have a well-defined circular catchment area of radius R. It is much easier to devise schemes to minimize the influence of unwanted low energy clusters when they are contained in a well-defined region.

On the other hand, anti-k_t is not prevented from clustering highly energetic clusters which are further apart than R. A correct definition for the jet reach R is then the minimal distance between two final jets.

At ATLAS, anti-k_t is the default jet algorithm. It is run over the whole calorimeter for each recorded event. Even if it is designed for hadronic jets, it will also reconstruct more or less accurately any other object leaving significant energy deposits, such as τ hadronic decays.

4.4.5 τ calorimeter seed

The definition of a τ calorimeter seed can now be stated. Any anti-k_t jet with $R = 0.4$ calculated from the topological clusters having a transverse energy[7] $E_T > 10$ GeV and within $|\eta| < 2.5$ is considered a τ candidate.

For each calorimeter seed, a number of identification variables based on the calorimeter information are calculated. The six baseline variables soon to be introduced have been the

[7] The transverse energy is jargon for the projection of the calorimeter energy in the plane transverse to the beam axis. It is related to the energy E by the relation $E = E_T \cosh\eta$.

main quantities used to exploit the calorimeter information before the work presented in this thesis took place. They are calculated by iterating over the cells from topological clusters in pre-defined geometrical regions. For this reason, they are referred to as the cell-based variables. The information contained in these six variables defines the baseline from which any improvement brought by new calorimeter-based variables must be measured.

R_{EM} and R_{Had}

The electromagnetic radius R_{EM} and hadronic radius R_{Had} are defined as

$$R_{\text{EM,Had}} = \frac{\sum_i E_{T,i} \Delta R_i}{\sum_i E_{T,i}}, \tag{4.10}$$

where ΔR_i is taken with respect to the calorimeter seed centroid and $E_{T,i}$ is the transverse energy of cell i, which is any cell in the first two layers of the electromagnetic calorimeter, with $\Delta R_i < 0.4$. The hadronic radius R_{Had} is defined in an identical way, except that the cells entering its calculation are from the remaining layers of the calorimeter system. These variables take advantage of the small width of hadronic tau decays. Fig. 4.10(a) shows the signal and background distributions for R_{EM} (signal and background will be properly defined in section 5.1), and Fig. 4.10(b) shows these same distributions for R_{Had}.

f_{core} and f_{ring}

The centrality fraction f_{core} is defined as

$$f_{\text{core}} = \frac{\sum_i E_{T,i}}{\sum_j E_{T,j}} \tag{4.11}$$

where i designates all cells with $\Delta R_i < 0.1$ and j designates all cells with $\Delta R_j < 0.4$. The isolation fraction f_{ring} is also defined by Eq. 4.11 but i now designates all cells satisfying $0.1 < \Delta R_i < 0.2$. These variables also take advantage of the small width of the τ decays, but they use all layers of the calorimeter to do so. The distributions for f_{core} and f_{ring} are shown in Fig. 4.10(c) and Fig. 4.10(d) respectively.

W_{strip} and N_{strip}

The width in the strip layer W_{strip} is defined as

$$W_{\text{strip}} = \sqrt{\frac{\sum_i E_{T,i}(\eta_i - \eta_0)^2}{\sum_i E_{T,i}}} \tag{4.12}$$

where i runs over all cells with $\Delta R_i < 0.4$ within the η strip layer of the electromagnetic calorimeter. This variable is intended to take advantage of the finer granularity of the strip layer in η to measure the width of the τ candidate.

The number of hits in the strip layer N_{strip} is defined as the number of cells with $\Delta R_i < 0.4$ having $E_T > 200$ MeV. This variable also uses the finer granularity of the strip layer. It is useful in measuring neutral pions, which decay to two photons 99% of the time. Charged pions are less likely to leave a significant energy deposit in the first layer of the electromagnetic calorimeter.

The distributions for W_{strip} and N_{strip} are shown in Fig. 4.10(e) and Fig. 4.10(f) respectively.

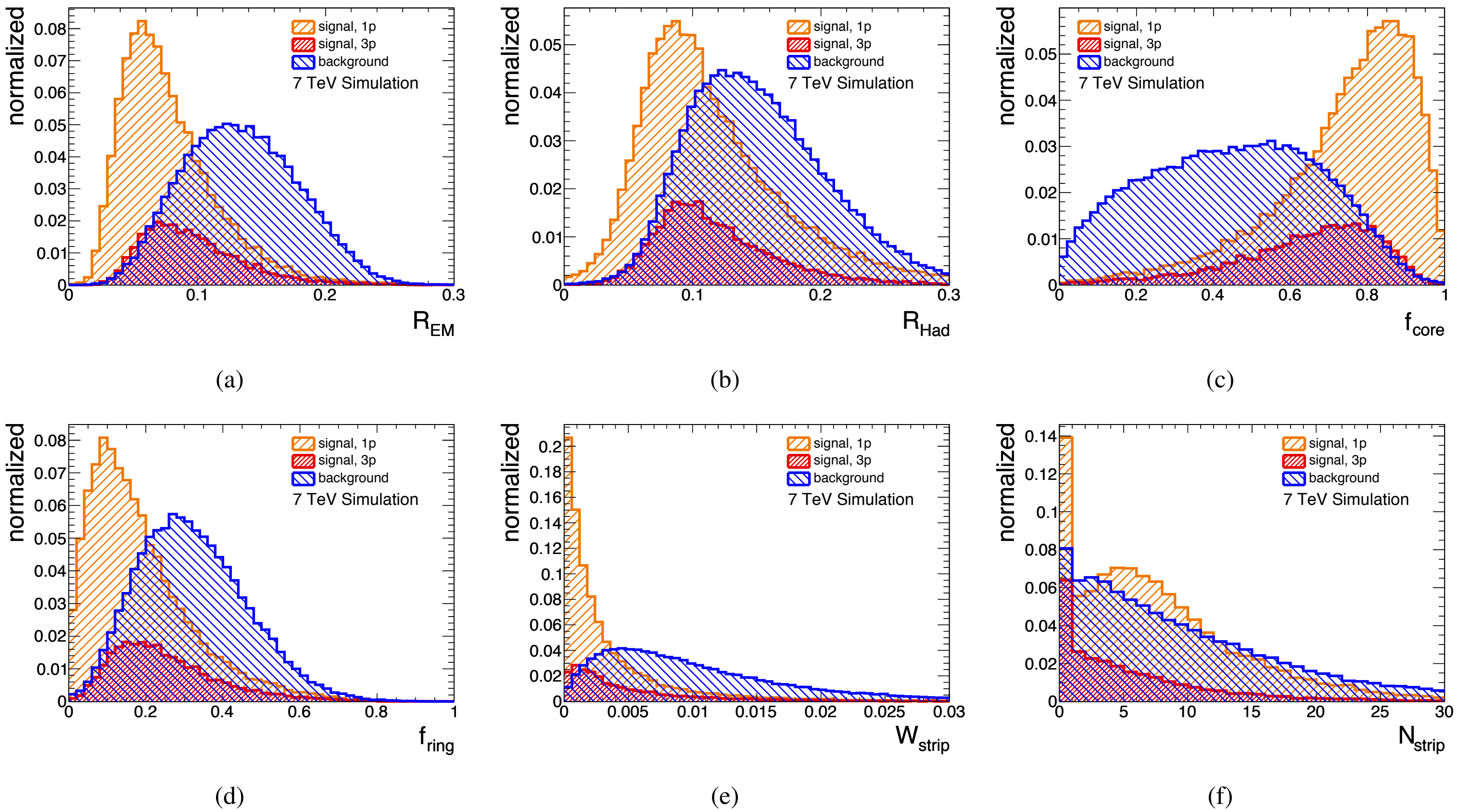


Figure 4.10: Calorimeter cell-based τ identification variables: (a) Electromagnetic radius , (b) hadronic radius, (c) centrality fraction, (d) isolation fraction, (e) η width in the strip layer, (f) number of hits in the strip layer.

Chapter 5

cluster-based τ identification variables

The τ identification variables based on topological clusters, or cluster-based variables, are calculated by iterating over the clusters composing an anti-k_T jet. They are divided in two classes. They can either be based on the 4-momenta of the clusters or on a list of parameters called cluster moments. The variables presented in this chapter have been selected from a larger initial set of variables based on their relative discrimination power. How they were selected is explained in Chapter 6. In this chapter, the parameters defining the scope of this study will be explained first. Then, each new cluster-based variable will be defined and discussed.

5.1 Parameters of the study

In order to maximize the relevance of this study for the near future, the events used have been simulated assuming 7 TeV collisions[1]. The signal τ candidates come from ~35 000 $W \to \tau\nu$ events. QCD dijets are used as a sample of fake τ candidates. The initial 2-2 hard-scatter process is constrained such that parton energy is roughly in the range of 70 to 140 GeV. However, effects from initial and final-state radiation mean that the p_T spectrum for all jets in the event ranges from about 10 GeV to 120 GeV, with a local maximum at about 60 GeV.

[1] At the time of writing, the LHC is colliding protons at 7 TeV. It will keep doing so until a given quantity of data have been recorded by the experiments. It will then undergo a year-long shut-down and prepare to deliver 14 TeV collisions.

In the LHC collisions, the probability of a hard scatter event decreases very fast with the transverse momentum of the final state. Therefore, using a realistic sample of hadronic jets forces the focus of the study only on low p_T τ candidates. The limited hadronic jet sample described in the previous paragraph has similar numbers of jets in the range 10 GeV $< p_T <$ 35 GeV and in the range 35 GeV $< p_T <$ 70 GeV. Since the jets in the 10 GeV $< p_T <$ 35 GeV region are mostly from initial and final state radiation, they contain an unrealistically large proportion of gluon jets, which are wider and contain more particles than quark jets. However, even the hard-scatter jets are mostly gluon jets at the LHC [14].

Both the signal and the background samples feature underlying event. Underlying event can be described as scattering between the partons not taking part in the hard-scatter in a proton-proton collision. It does not provide useful information about the hard-scatter, but it still results in measurable energy deposits in the calorimeter. The cross-section of the underlying event is one of the most uncertain predictions for hadron collisions beyond the Tevatron[2] energy. Therefore, any quantity which is independent of the amount of underlying event will be free of a major source of theoretical uncertainty.

The phase space of τ candidates at ATLAS is limited in η to the inner detector and the regions of the calorimeter with finer granularity ($|\eta| < 2.5$). Also, the τ candidates included in this study are restricted to have 10 GeV $< E_T <$ 70 GeV.

5.2 Variables constructed from the 4-momenta of topological clusters

The design of the cell-based variables follows the geometry of the ATLAS calorimeter. The cluster-based 4-momentum variables do not have this restriction. They inherit from the flexibility of the topological clustering and anti-k_t algorithms. The primary concern in their design is the physics of τ hadronic decays and hadronic jets.

The cluster-based 4-momentum variables are divided into "all-cluster" and "effective-

[2] The Tevatron is a hadron collider based at Fermilab near Chicago. It collides protons with anti-protons at a center of mass energy of 1.96 TeV and it was at the energy frontier of experimental particle physics before the LHC started.

cluster" variables. The all-cluster variables are calculated from the complete set of clusters $\{\text{all}\}$ provided by the jet algorithm, while the effective-cluster variables are calculated from a restricted set $\{\text{eff}\}$. How this restricted set is built is explained in section 5.2.2.

Figs. 5.1 to 5.6 require an early description. The purpose of these plots is first to show the effect of the underlying event on the variables, and then to show how well the distributions are reproduced after full detector simulation and reconstruction. However, looking at the underlying event cannot be done in any simple way for background since hadronic jets are defined by the jet algorithm used to construct them. The effect of underlying event on the variables is shown for signal only.

In Figs. 5.1 to 5.6, plots (a), (b) and (c) show signal and background distributions where the signal distribution is a stack of 1 prong and 3 prong τ decays. Plots (d), (e) and (f) show the dependence of the mean of these distributions on E_T.

For signal in plots (a) and (d), the constituents entering the calculation of the variables are the visible final state particles from the monte carlo event generator, or "truth" particles. For signal, the truth particles are the hadrons of the τ decay making potentially distinct showers in the calorimeter[3].

For background in these same plots, the constituents are obtained by running the anti-k_T jet algorithm with $R = 0.4$ on all truth particles in the event, making so-called truth jets. The particles in the closest truth jet in η, ϕ space to the τ candidate are retrieved. If the truth jet has $\Delta R > 0.2$ with respect to the τ candidate, the candidate is discarded from the study.

When 4-momentum variables are calculated from truth particles, these particles are taken to be massless to better reflect calorimeter clusters, which are always assumed to come from the showering of massless particles.

In plots (b) and (e), both signal and background are calculated from truth jets, meaning that underlying event particles are now included in the signal (the background stays the same as in plots (a) and (d)).

Finally, in plots (c) and (f), the variables are calculated from the topological clusters reconstructed in the calorimeter (detector-level). These plots show the result of the complete simulation of what these variables should look like in real data.

Showing the cluster-based 4-momentum variables at these different levels makes it pos-

[3]For neutral pions, the two photons resulting from their decay are combined, so that a neutral pion counts as a single particle.

sible to test how well the topological clustering and the jet algorithm can resolve the substructure of τ decays and jets. It also reveals the strengths and weaknesses of the variables, indicating the way towards potential future improvements.

5.2.1 Number of clusters

The number of clusters N_{all} included in a τ candidate is the simplest quantity that can be obtained from the jet algorithm. This variable should ideally count the number of visible final state particles of the physics object. This variable is plotted in Figs. 5.1(a) and 5.1(d). Fig. 5.1(d) shows clearly that the number of decay products is independent of E_T for signal, while it grows with E_T for background due to the hadronization process. This can still be seen at detector-level (Fig 5.1(f)).

In Figs. 5.1(b) and 5.1(e), N_{all} for signal becomes larger, which is an effect of the underlying event. It affects the discrimination power of N_{all} by making the signal distributions move towards the background distribution.

At detector-level (Figs. 5.1(c) and 5.1(f)), the signal distributions spreads even more. The overestimation of the number of clusters can be caused by a number of detector, showering and clustering algorithm effects. Most notably, showers from single particles are not always reconstructed as a single topological cluster, especially hadronic showers. Calorimeter noise can also be responsible for the creation of additional clusters.

In Fig. 5.1(f), it can be seen that N_{all} for background is over-estimated at low E_T, and under-estimated at large E_T relative to generator-level (Fig. 5.1(e)). When the density of showers is too high in the calorimeter, the topological clustering algorithm has trouble distinguishing individual showers. At large E_T, this effect counterbalances the shower splitting and spontaneous noise cluster creation. For signal, N_{all} is overestimated over the whole E_T range. Since τ decays are narrower than jets, there is more room around the τ decay products within the jet reach $R = 0.4$ for noise and underlying event particles to be resolved and counted as separate clusters.

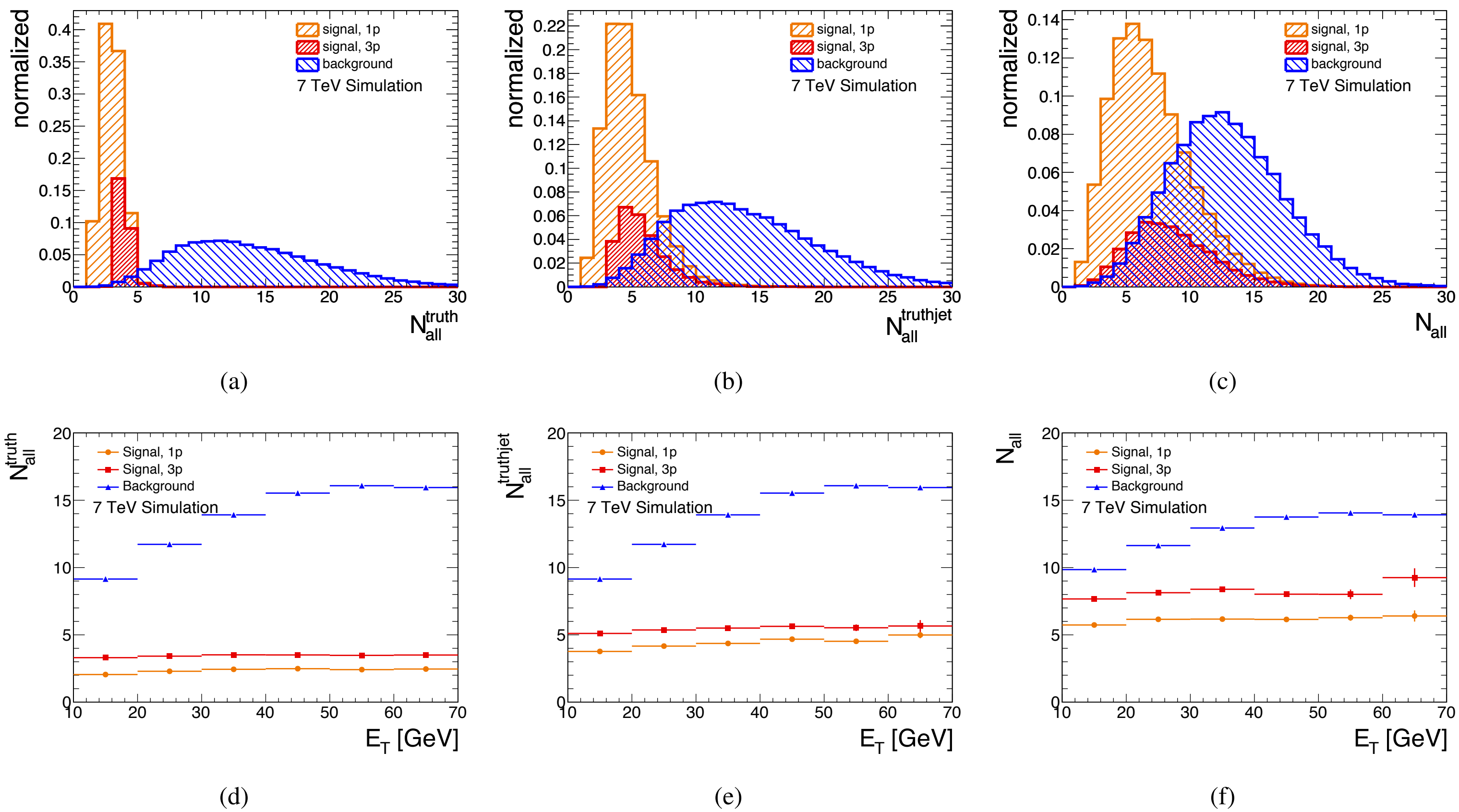


Figure 5.1: Number of clusters/particles. (a), (b) and (c) show the integrated distributions for 10 GeV $< E_T <$ 70 GeV and $-2.5 < \eta < 2.5$ for visible final state τ decay products only (signal) and final state particles including underlying event (background), final state particles including underlying event (signal and background) and reconstructed clusters respectively. (d), (e) and (f) show the E_T profiles for the same selections as for (a), (b) and (c) respectively.

5.2.2 Effective number of clusters

The effective number of constituents N_{eff} is defined as:

$$N_{\text{eff}} = \frac{\left(\sum_j E_j\right)^2}{\sum_j E_j^2}, \qquad j \in \{\text{all}\}, \tag{5.1}$$

where E_j is the energy of cluster j. This variable characterizes how the energy is distributed among the clusters. It can be interpreted as the number of clusters carrying most of the energy. If all clusters share the same energy, $N_{\text{eff}} = N_{\text{all}}$. If all but one cluster have negligible energy, $N_{\text{eff}} \simeq 1$.

The merit of this variable becomes apparent when comparing Fig. 5.2(a) to Fig. 5.2(b) and Fig. 5.2(d) to Fig. 5.2(e). The change in the signal distribution caused by the underlying event is significantly smaller than it is for N_{all}.

In Fig. 5.2(c), the signal distributions are seen to be very similar to their representation at generator-level in Fig. 5.2(b). The features of the signal distributions, which are attributed to the different modes of hadronic τ decays, are still visible at detector-level. However, the background gets more distorted than for N_{all}. This becomes especially obvious when comparing Fig. 5.2(e) to Fig. 5.2(f).

In Fig. 5.2(f), N_{eff} for signal can be seen to get smaller as E_T gets larger. When τ decays become more energetic, the ratio of the energy of the decay products to the energy of the underlying event particles (or noise) grows larger, and N_{eff} becomes better at dismissing the undesired clusters.

We can obtain a list of clusters deemed most likely to come from the τ decay by rounding N_{eff} to its nearest larger integer $N_{\text{eff}}^{\text{ceil}}$, ordering the list of clusters by energy, and keeping the first $N_{\text{eff}}^{\text{ceil}}$ clusters[4]. The clusters making this new list are referred to as the "effective clusters". They are denoted by {eff}.

[4]Using the ceiling as opposed to ordinary rounding is a conservative attempt at not eliminating clusters which may really be from the τ decay. This has the effect that underlying event and noise clusters are accepted as effective clusters more often. The rounding may be modified to obtain a better suppression of unwanted clusters but separation power may suffer.

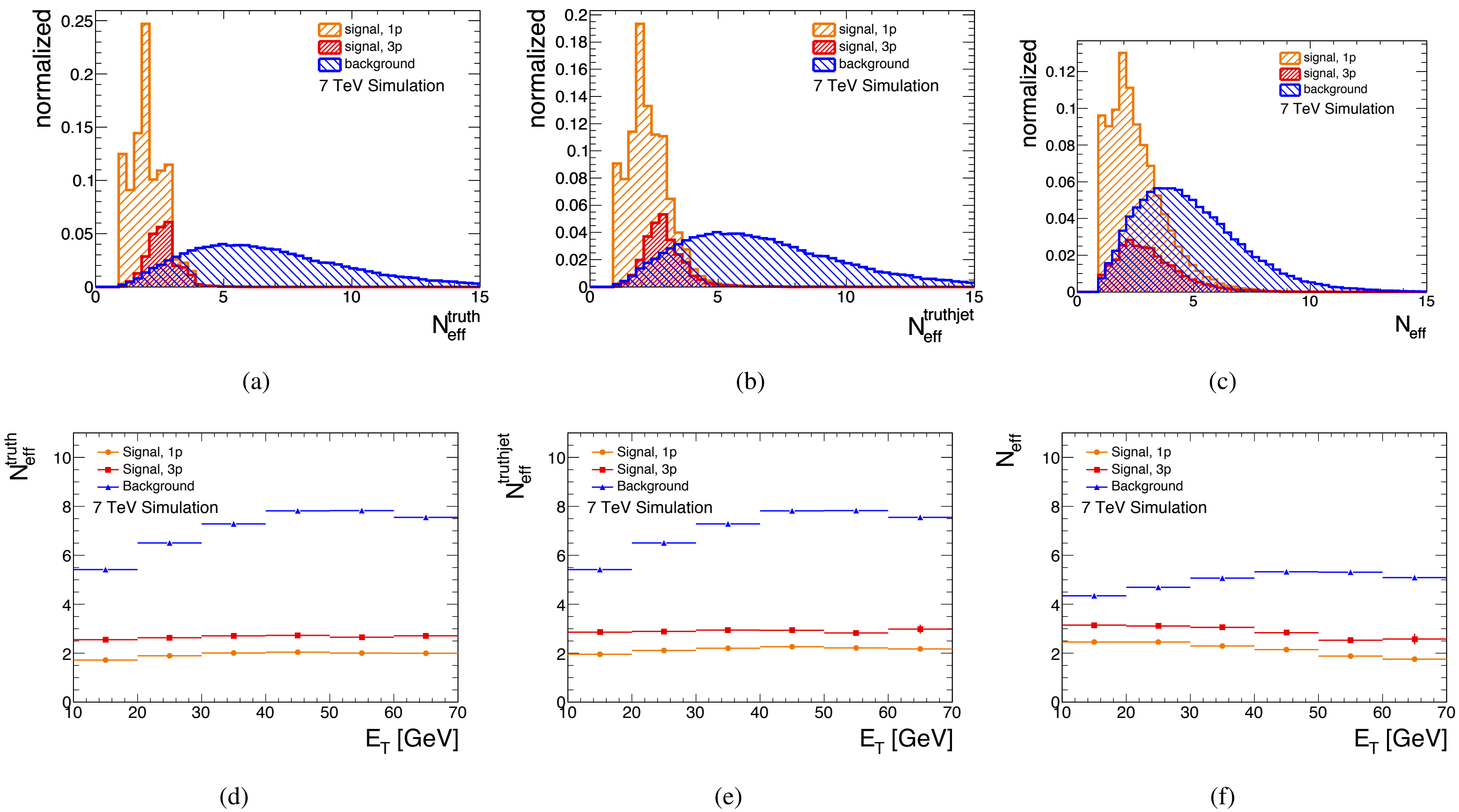


Figure 5.2: Effective number of clusters/particles. (a), (b) and (c) show the integrated distributions for 10 GeV $< E_T <$ 70 GeV and $-2.5 < \eta < 2.5$ for visible final state τ decay products only (signal) and final state particles including underlying event (background), final state particles including underlying event (signal and background) and reconstructed clusters respectively. (d), (e) and (f) show the E_T profiles for the same selections as for (a), (b) and (c) respectively.

5.2.3 Invariant mass of all clusters

The invariant mass of all clusters m_{all} is defined as

$$m_{\text{all}} = \sqrt{\Big(\sum_j E_j\Big)^2 - \Big(\sum_j \vec{p}_j\Big)^2}, \qquad j \in \{\text{all}\}. \tag{5.2}$$

τ candidates with a single cluster automatically get $m_{\text{all}} = 0$. These candidates are not shown in the distributions in Fig. 5.3. Since every τ decay involves neutrinos, the reconstructed mass for signal is expected to be somewhere below 1.777 GeV. The background mass should be made arbitrarily large by the hadronization process. This is exactly what is seen in Fig. 5.3(a). Fig. 5.3(d) also shows how the signal mass is independent of E_T while the jet mass strongly depends on it.

In Fig. 5.3(b), a tail appears in the signal distributions which is caused by the underlying event. The rest of the distributions don't change much, showing that only a fraction of τ decays are affected. For the affected decays, their mass is seen in Fig. 5.3(e) to acquire a dependence on E_T. This can be explained by the following. In order for the invariant mass of an object to remain independent of energy, all its decay products should collimate in concert as the object becomes more boosted. The τ decay products become more collimated with energy but the underlying event particles do not.

In Fig. 5.3(c), the signal distributions are seen to suffer a similar loss in resolution as the one observed for N_{all}. The background distribution shows some structure, which is just an artefact of the energy distribution of the background. Also, m_{all} becomes more correlated with energy due the calorimeter's limited resolution. This can be seen more clearly in 5.3(f).

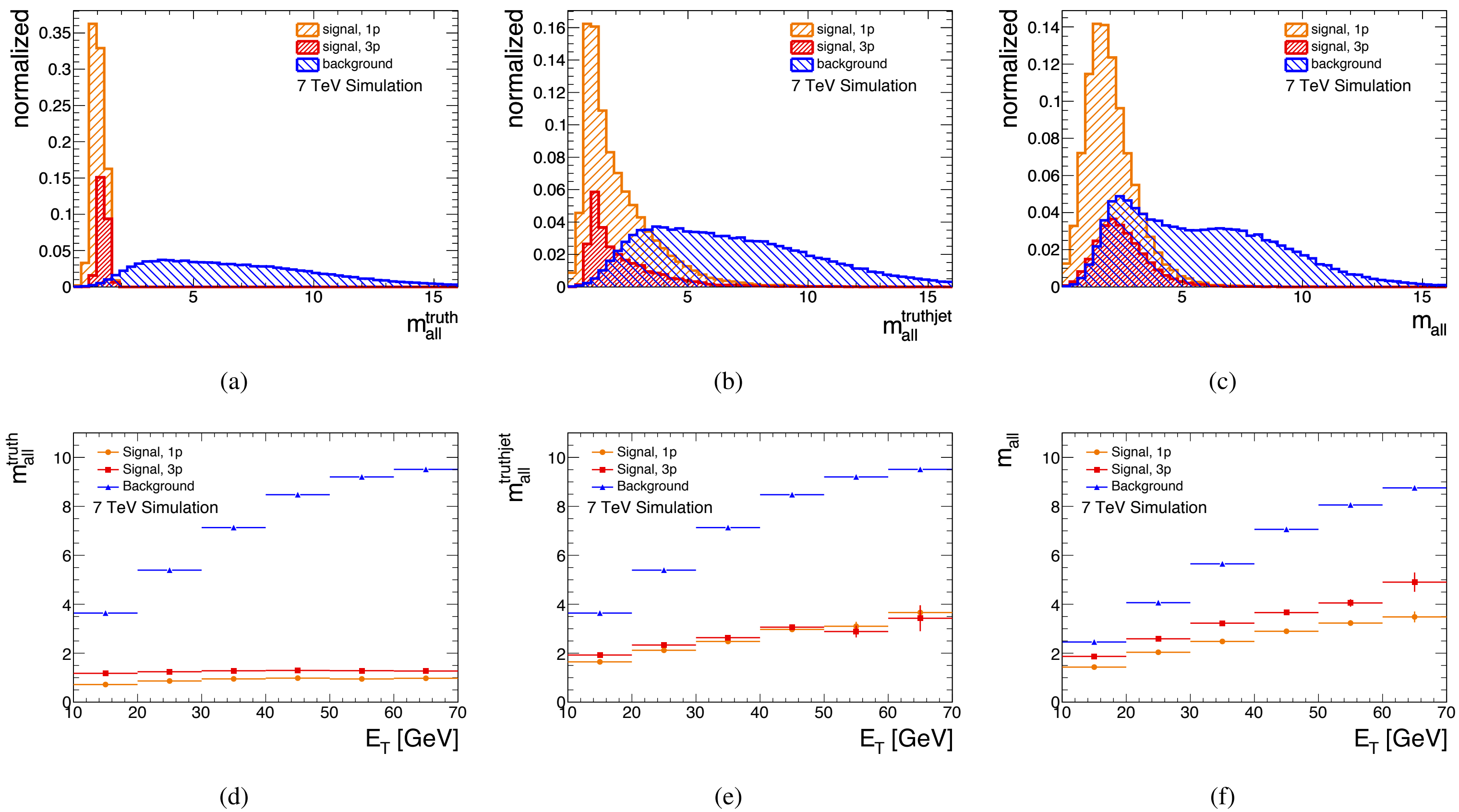


Figure 5.3: Invariant mass of all clusters/particles. (a), (b) and (c) show the integrated distributions for 10 GeV $< E_T <$ 70 GeV and $-2.5 < \eta < 2.5$ for visible final state τ decay products only (signal) and final state particles including underlying event (background), final state particles including underlying event (signal and background) and reconstructed clusters respectively. (d), (e) and (f) show the E_T profiles for the same selections as for (a), (b) and (c) respectively.

5.2.4 Invariant mass of the effective clusters

The invariant mass of the effective clusters m_{eff} is defined as

$$m_{\text{eff}} = \sqrt{\left(\sum_j E_j\right)^2 - \left(\sum_j \vec{p}_j\right)^2}, \qquad j \in \{\text{eff}\}. \tag{5.3}$$

τ candidates with a single cluster automatically get $m_{\text{eff}} = 0$, and they are not shown in the distributions in Fig. 5.4. Since fewer constituents than for m_{all} enter the calculation of this variable, $m_{\text{eff}} \leq m_{\text{all}}$ for both signal and background. However, m_{eff} retains most of the desirable features of m_{all}.

Signal distributions in Fig. 5.4(a) and Fig. 5.4(b) are very similar. By comparing Fig. 5.4(b) to Fig. 5.3(b), it can be seen that the underlying event tail is suppressed by selecting effective constituents.

At detector-level, the signal mass still rises with E_T, but not as quickly as for m_{all}. The signal distributions suffer the usual resolution loss while the background distribution gets slightly deformed. The structure of the background no longer mirrors the energy distribution of the background, indicating that the E_T dependence of this variable is not affected as badly as the E_T dependence of m_{all} by the calorimeter imperfections.

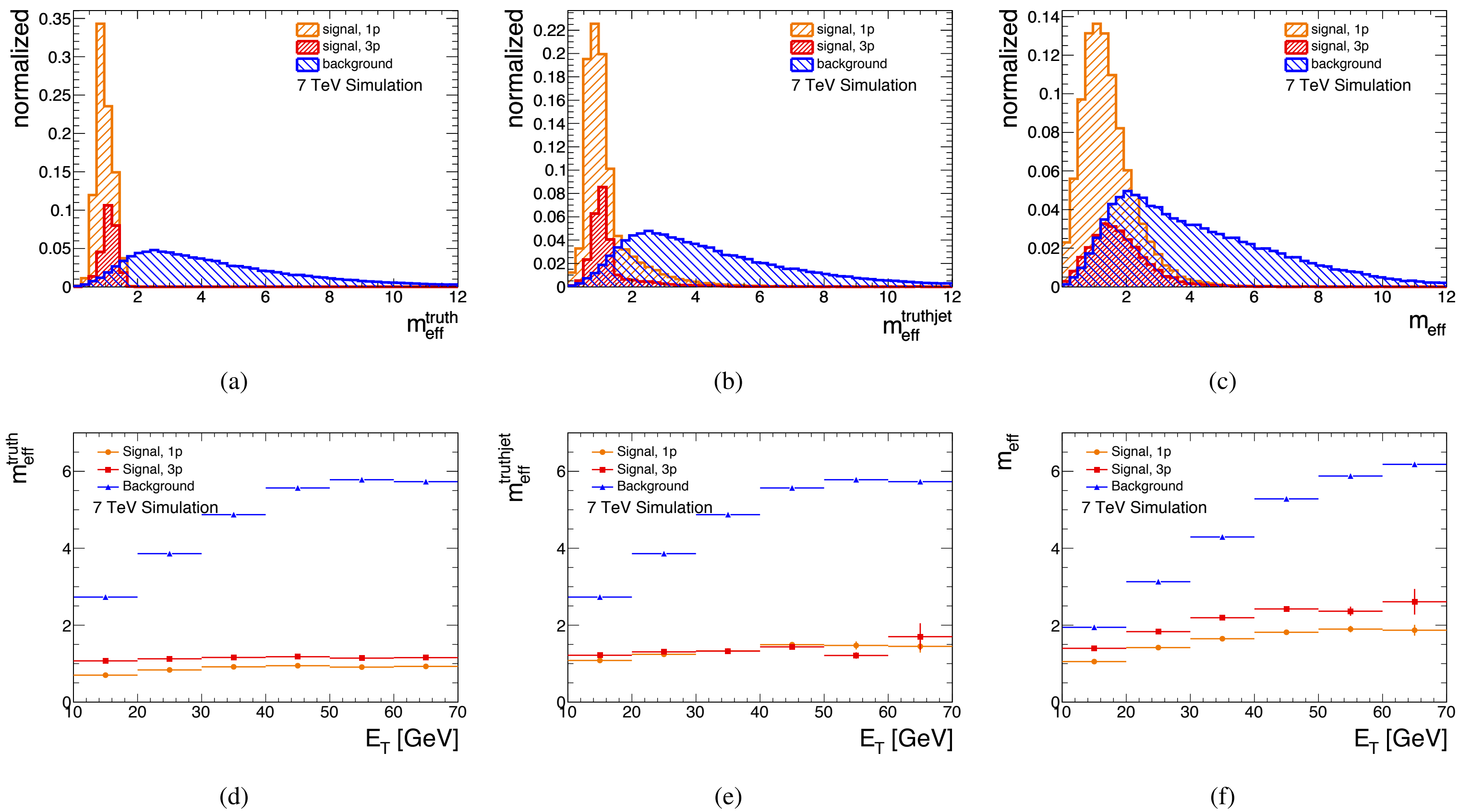


Figure 5.4: Invariant mass of effective clusters/particles only. (a), (b) and (c) show the integrated distributions for 10 GeV $< E_T <$ 70 GeV and $-2.5 < \eta < 2.5$ for visible final state τ decay products only (signal) and final state particles including underlying event (background), final state particles including underlying event (signal and background) and reconstructed clusters respectively. (d), (e) and (f) show the E_T profiles for the same selections as for (a), (b) and (c) respectively.

5.2.5 Average ΔR to the geometrical center of all clusters

The average ΔR to the geometrical center of all clusters $\langle R\rangle_{\text{all}}$ is defined as

$$\langle R\rangle_{\text{all}} = \frac{\sum_j \sqrt{\left(\eta_j - \eta_{\text{all}}^{\text{geo}}\right)^2 + \left(\phi_j - \phi_{\text{all}}^{\text{geo}}\right)^2}}{N_{\text{all}}}, \qquad j \in \{\text{all}\}, \tag{5.4}$$

where $\left(\eta_{\text{all}}^{\text{geo}}, \phi_{\text{all}}^{\text{geo}}\right)$ are obtained from

$$\vec{p}_{\text{all}}^{\,\text{geo}} = \sum_j \frac{\vec{p}_j}{|\vec{p}_j|}, \qquad j \in \{\text{all}\}. \tag{5.5}$$

This variable is intended to be a purely geometrical measure of the width of the physics object (no energy value enters its calculation). It is vulnerable to a number of effects, which makes it a good quantity to test the simulation against real data. It also means that this variable is subject to larger systematic and theoretical uncertainties.

As can be seen in Fig. 5.5(d), $\langle R\rangle_{\text{all}}$ gets smaller with E_T. This is a consequence of the tendency of decay products to get more collimated as the decaying object becomes more boosted.

This variable is especially vulnerable to underlying event, as can be seen in Fig. 5.5(b). In the signal distributions, a wide bump is seen to grow where the background distribution is, badly affecting the discrimination power. The collimation effect is almost completely hidden by underlying event, as can be seen in Fig. 5.5(e).

At detector-level, detector effects cause jets and τ decays to appear even wider. It looks like the bump from underlying event in the signal distributions of Fig. 5.5(b) now dominates. In Fig. 5.5(f), the collimation effect is no longer visible.

5.2.6 Average ΔR to the geometrical center of the effective clusters

The Average ΔR to the geometrical center of the effective clusters $\langle R\rangle_{\text{eff}}$ is defined as:

$$\langle R\rangle_{\text{eff}} = \frac{\sum_j \sqrt{\left(\eta_j - \eta_{\text{eff}}^{\text{geo}}\right)^2 + \left(\phi_j - \phi_{\text{eff}}^{\text{geo}}\right)^2}}{N_{\text{eff}}^{\text{ceil}}}, \qquad j \in \{\text{eff}\}, \tag{5.6}$$

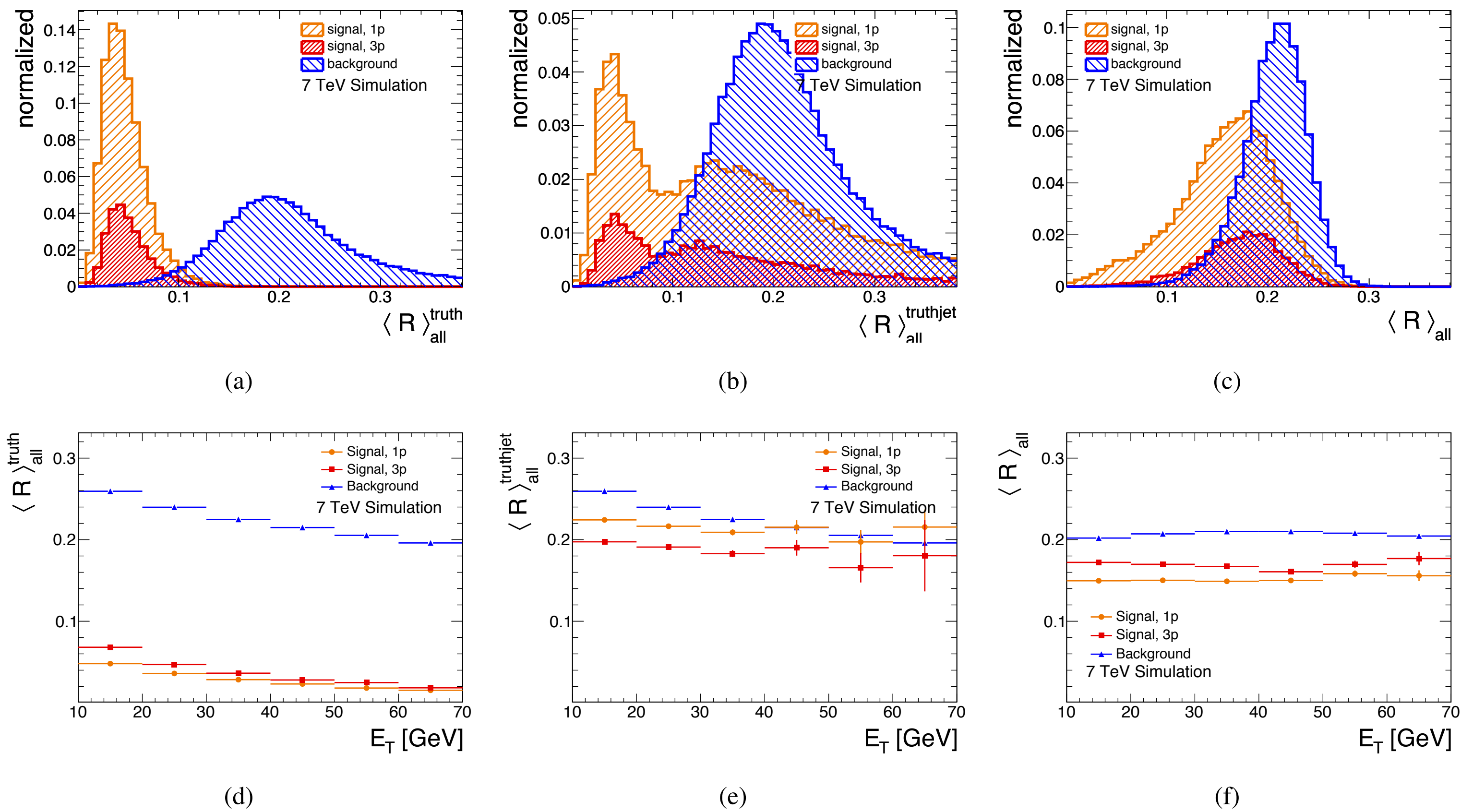


Figure 5.5: Average ΔR to the geometrical center for all clusters/particles. (a), (b) and (c) show the integrated distributions for 10 GeV $< E_T <$ 70 GeV and $-2.5 < \eta < 2.5$ for visible final state τ decay products only (signal) and final state particles including underlying event (background), final state particles including underlying event (signal and background) and reconstructed clusters respectively. (d), (e) and (f) show the E_T profiles for the same selections as for (a), (b) and (c) respectively.

where $\left(\eta_{\text{eff}}^{\text{geo}}, \phi_{\text{eff}}^{\text{geo}}\right)$ are obtained from

$$\vec{p}_{\text{eff}}^{\text{geo}} = \sum_j \frac{\vec{p}_j}{|\vec{p}_j|}, \qquad j \in \{\text{eff}\}. \tag{5.7}$$

In Fig. 5.6(a), it can be seen that this variable is potentially weaker than for $\langle R \rangle_{\text{all}}$by comparing Fig. 5.6(a) to Fig. 5.5(a). However, $\langle R \rangle_{\text{eff}}$ is much more resistant to the effect of underlying event, which can be seen by comparing Fig. 5.6(a) and Fig. 5.6(b) to Fig. 5.5(a) and Fig. 5.5(b).

At detector-level (Fig. 5.6(c)), we see that the effective clusters also have a tendency to get wider than at generator-level. However, the growth in size is controlled in such a way that the signal and background distributions are now peaked further away from each other. This makes $\langle R \rangle_{\text{eff}}$ the only effective-cluster variable having a better discrimination power than its all-cluster counterpart. In Fig. 5.6(f), the collimation effect is seen to have survived detector simulation and reconstruction.

The signal distribution in Fig. 5.6(c) is composed of two bumps. The smaller bump at $\langle R \rangle_{\text{eff}} \simeq 0.15$ arises because of a few different effects. First, when N_{eff} is only slightly greater than its floor integer value, it is likely that the last particle included in the effective constituents is not part of the τ decay. If this extra constituent is an underlying event particle, it will probably be far from the τ decay. Since underlying event particles are uniformly spread in η, ϕ-space, the probability of finding these particles grows linearly with the ΔR distance to the τ decay. This extra particle is thus likely to make $\langle R \rangle_{\text{eff}}$ bigger than it should be, which explains the tail in Fig. 5.6(b).

The other effects are more difficult to investigate unless one goes back to the simulation of the showering in the detector. As mentioned in section 5.2.1, it happens rather frequently that showers are reconstructed as multiple topological clusters. The several clusters coming from the same shower may all have significant energy and they may be included in the list of effective clusters. However, they will not reflect the position of the decay products anymore, leading to a value for $\langle R \rangle_{\text{eff}}$ which does not match the generator-level value.

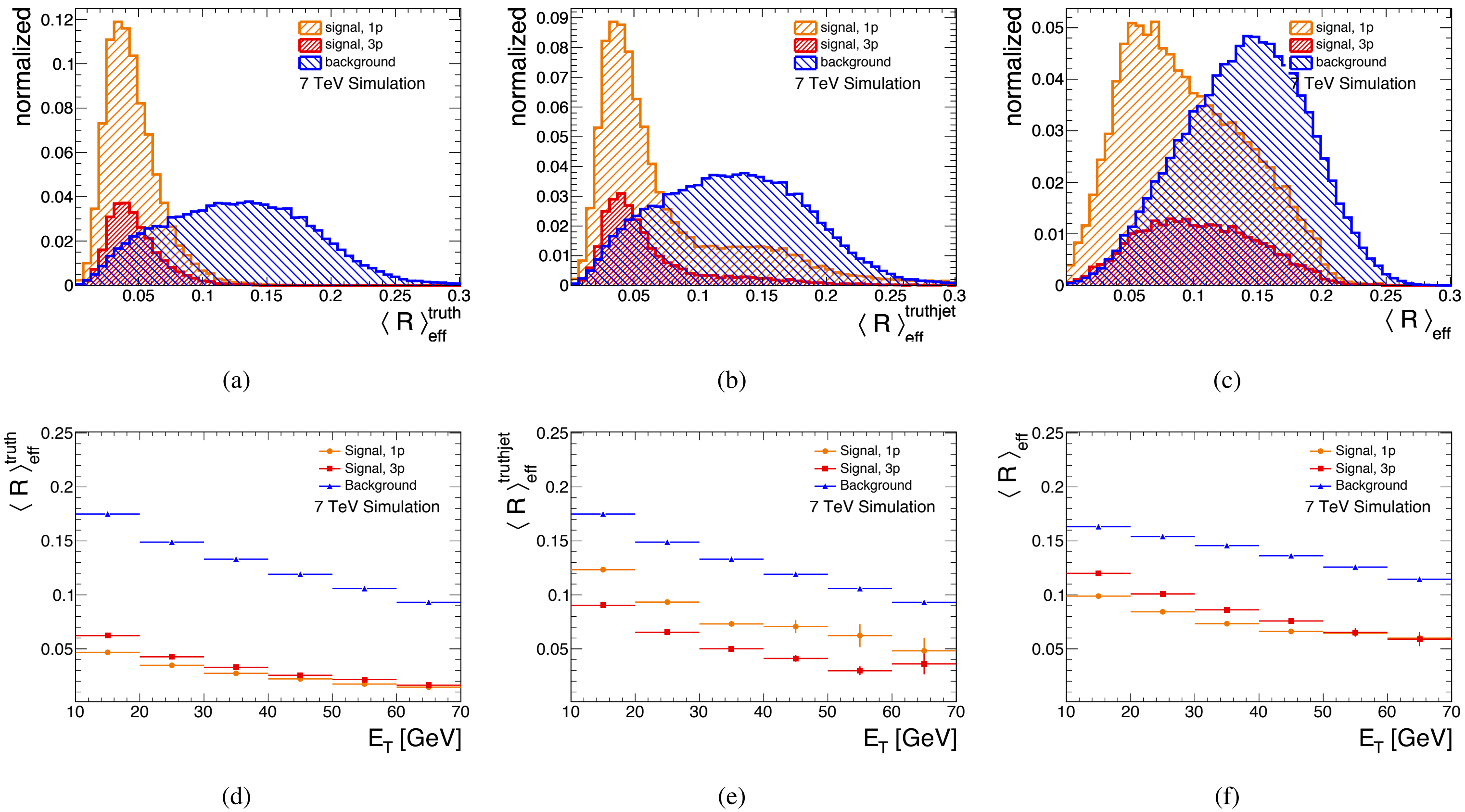


Figure 5.6: Average ΔR to the geometrical center for effective clusters/particles only. (a), (b) and (c) show the integrated distributions for 10 GeV $< E_T <$ 70 GeV and $-2.5 < \eta < 2.5$ for visible final state τ decay products only (signal) and final state particles including underlying event (background), final state particles including underlying event (signal and background) and reconstructed clusters respectively. (d), (e) and (f) show the E_T profiles for the same selections as for (a), (b) and (c) respectively.

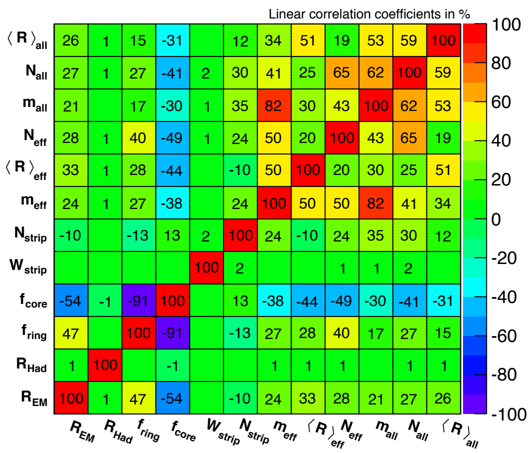
5.2.7 Linear correlations

The linear correlation coefficients between all calorimeter-based τ identification variables presented so far are shown in Fig. 5.7.

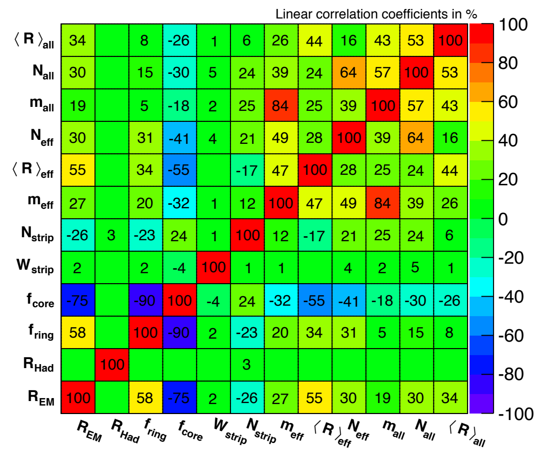
It can be seen by looking at the correlations between cell-based variables that R_{Had} is very weakly correlated with the other variables, even R_{EM}. However, the energy deposits in the cells considered in the calculation of R_{Had} are most likely coming from hadronic showers. The large statistical fluctuations in the geometry of hadronic showers can provide a partial explanation for this lack of correlation.

It can also be seen that f_{core} and f_{ring} are strongly anti-correlated. When the energy in an anti-k_t jet is concentrated in the core ($f_{\mathrm{core}} \simeq 1$), there is not a lot of energy left in the ring just outside this core ($f_{\mathrm{ring}} \simeq 0$). The variables based on the η strip layer show little correlation with the other variables.

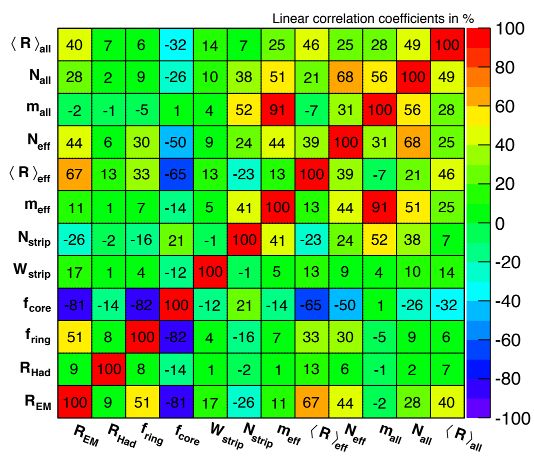
It can be seen that generally, the 4-momenta variables show more correlation among themselves than with cell-based variables. f_{core} is an exception, since it shows strong anti-correlation with most of the 4-momenta variables. Effective-cluster variables tend to have a strong correlation with their all-cluster counterpart, with the exception of $\langle R \rangle_{\mathrm{all}}$ and $\langle R \rangle_{\mathrm{eff}}$.



(a)



(b)



(c)

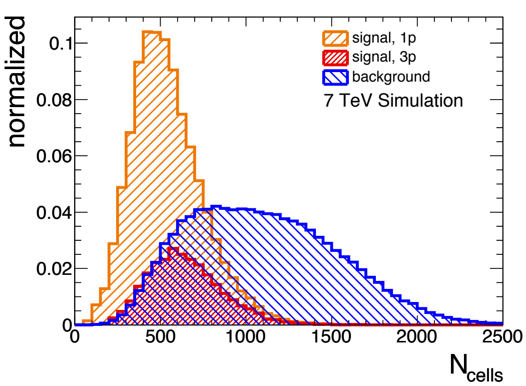
Figure 5.7: Linear correlation coefficients between cell-based and 4-momentum variables for (a) 1prong signal, (b) 3 prongs signal, (c) background. An empty square represents a correlation factor of 0.

5.3 Variables constructed from cluster moments

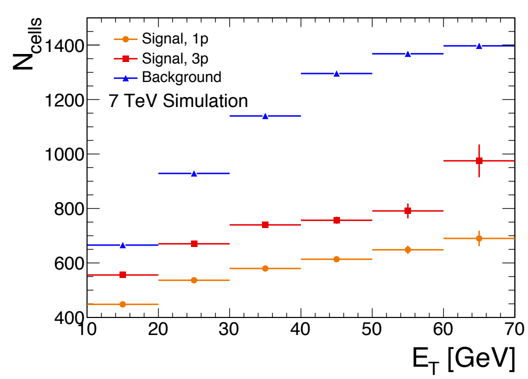
For each cluster found in the calorimeter, properties are calculated to characterize its geometry and spatial energy distribution. These properties are referred to as "cluster moments" [19]. Unlike the 4-momentum variables, variables based on cluster moments cannot be calculated from the truth particles since they depend on the showering process and the distribution of the material making up the calorimeter.

5.3.1 Number of cells

The number of cells N_{cells} is simply the sum of the number of cells in all the clusters of the τ candidate. This variable is arguably the most difficult to model, since it depends on almost every detail of the simulation. It is nevertheless a very powerful identification variable which reflects the smaller width of the τ decay. Its distributions and E_T profiles are shown in Fig. 5.8.



(a)



(b)

Figure 5.8: Number of cells. (a) shows the integrated distribution for 10 GeV $< E_T <$ 70 GeV and $-2.5 < \eta < 2.5$ for reconstructed clusters and (b) shows the E_T profile for the same τ candidates.

5.3.2 Average over effective clusters of the second moment in radial distance to the shower axis

This variable is an average of the transverse width of the most energetic clusters of a τ candidate. Many steps are involved in its calculation. First, a shower axis must be determined for each cluster, which is done by calculating the eigenvector $\vec{s}$ of the matrix

$$C = \begin{pmatrix} C_{xx} & C_{xy} & C_{xz} \\ C_{yx} & C_{yy} & C_{yz} \\ C_{zx} & C_{zy} & C_{zz} \end{pmatrix} \tag{5.8}$$

which is the closest in direction to the line joining the interaction point to the shower center $\vec{c}$. The coefficients of matrix C are given by

$$C_{ab} = \frac{1}{w}\sum_i E_i^2 \left(a_i - \langle a \rangle\right)\left(b_i - \langle b \rangle\right), \tag{5.9}$$

where a_i and b_i stand for x_i, y_i and z_i, the spatial coordinates of cell i and

$$w = \sum_i E_i^2. \tag{5.10}$$

The shower center is defined as

$$\vec{c} = \frac{\sum_i E_i \vec{x}_i}{\sum_i E_i}, \qquad \vec{x}_i = (x_i, y_i, z_i). \tag{5.11}$$

In Eq. 5.9 to 5.11, cell i must have $E_i > 0$. The distance of cell i to the shower axis is then given by

$$r_i = |(\vec{x}_i - \vec{c}) \times \vec{s}|, \tag{5.12}$$

the second moment in radial distance to the shower axis $\left\langle r^2 \right\rangle_j$ of cluster j is defined as

$$\left\langle r^2 \right\rangle_j = \frac{\sum_i E_i r_i^2}{\sum_i E_i}. \tag{5.13}$$

Finally, the average over effective clusters of this moment $\langle r^2 \rangle$ is given by

$$\langle r^2 \rangle = \frac{\sum_j \langle r^2 \rangle_j}{N_{\text{eff}}^{\text{ceil}}}, \qquad j \in \{\text{eff}\}. \tag{5.14}$$

$\langle r^2 \rangle$ shows that showers tend to get wider with energy. This variable offers better separation for 1 prong τ decays as can be seen from Fig. 5.9. Since the majority of 1 prong τ decays have neutral pions, they get a larger contribution from narrow electromagnetic showers than 3 prong τ decays do.

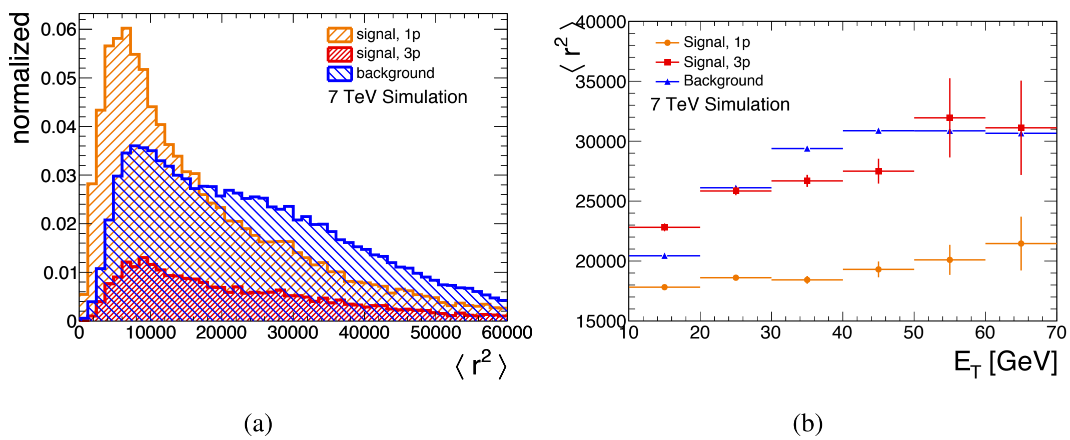


Figure 5.9: Average over effective clusters of the second moment in radial distance to the shower axis. (a) shows the integrated distribution for 10 GeV $< E_T <$ 70 GeV and $-2.5 < \eta < 2.5$ for reconstructed clusters and (b) shows the E_T profile for the same τ candidates.

5.3.3 Average of the electromagnetic energy fraction of the effective clusters

The average of the electromagnetic energy fraction of the effective clusters is defined as

$$f_{\text{EM}} = \frac{\sum_j x_j^{\text{EM}}}{N_{\text{eff}}^{\text{ceil}}}, \qquad j \in \{\text{eff}\}, \tag{5.15}$$

where x_j^{EM} is the fraction of the energy of cluster j in the EM calorimeter, which is calculated from positive energy cells only.

This variable may not be a very strong discriminant, but the structure of the signal distributions contains interesting information. It can be seen in Fig. 5.10(a) that the decay modes of the τ leptons are partly revealed. This is possible since equal weight is given to all clusters entering its calculation. Generally, neutral pions will produce clusters having $x_{\pi^0}^{\text{EM}} \simeq 1$, while charged pions will produce clusters with $x_{\pi^\pm}^{\text{EM}} < x_{\pi^0}^{\text{EM}}$. The number of decay products in the most important τ decay modes is small. Given that the difference between $x_{\pi^0}^{\text{EM}}$ and $x_{\pi^\pm}^{\text{EM}}$ is large enough, separated peaks are expected in the signal distribution.

It should be mentioned that charged pions can deposit a large fraction of their energy in the electromagnetic calorimeter. Even if the decay modes not involving neutral pions have a total branching fraction that is quite large (20.2%), only a small fraction of τ candidates have $f_{\text{EM}} < 0.5$.

It looks from Fig. 5.10(b) that 3 prong τ decays are better served by this variable. However, since the 1 prong distribution is not gaussian, its mean values should not be taken too seriously.

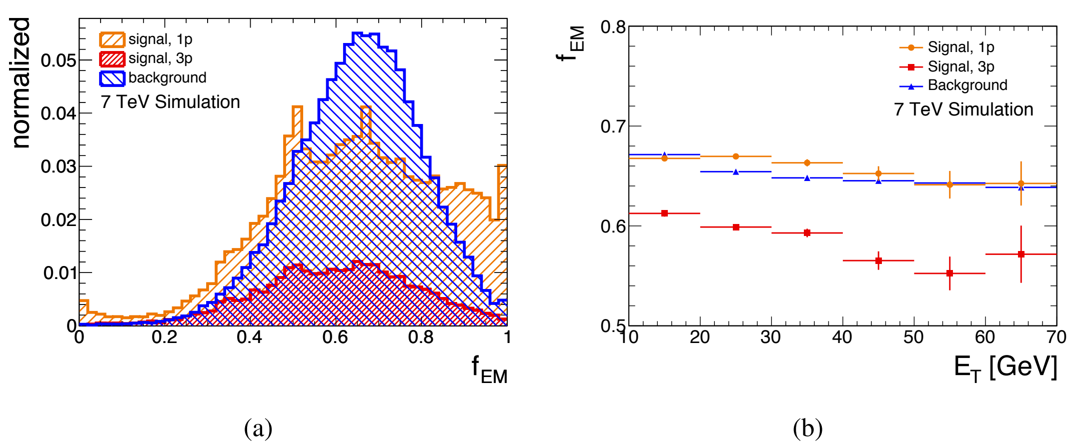


Figure 5.10: Average over effective clusters of the fraction of energy in the EM calorimeter. (a) shows the integrated distribution for 10 GeV $< E_T <$ 70 GeV and $-2.5 < \eta < 2.5$ for reconstructed clusters and (b) shows the E_T profile for the same τ candidates.

5.3.4 Isolation of the effective clusters

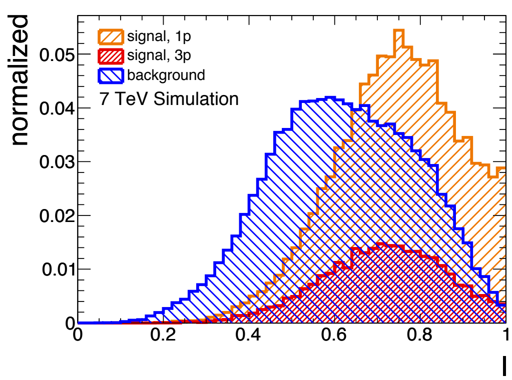
The isolation of cluster j is defined as

$$I_j = \frac{\sum_k E_k \left(N_k^{\text{empty}} / N_k^{\text{all}}\right)}{E_j} \tag{5.16}$$

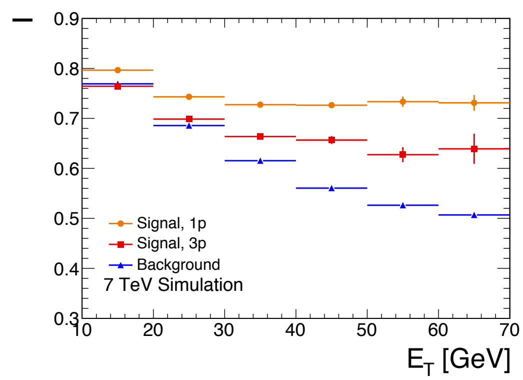
where k refers to the different calorimeter layers that cluster j occupies, E_k is the energy of cluster j in the k layer, E_j is the energy of the whole cluster, N_k^{all} is the total number of neighbouring cells in layer k and N_k^{empty} is the number of neighbouring cells in layer k that don't belong to any clusters. The isolation of the effective clusters I is defined as

$$I = \frac{\sum_j I_j}{N_{\text{eff}}^{\text{ceil}}}, \qquad j \in \{\text{eff}\}. \tag{5.17}$$

From Fig. 5.11(b), one can see that there is a tendency for objects with fewer constituents to be more isolated, which explains where this variable gets its discrimination power.



(a)



(b)

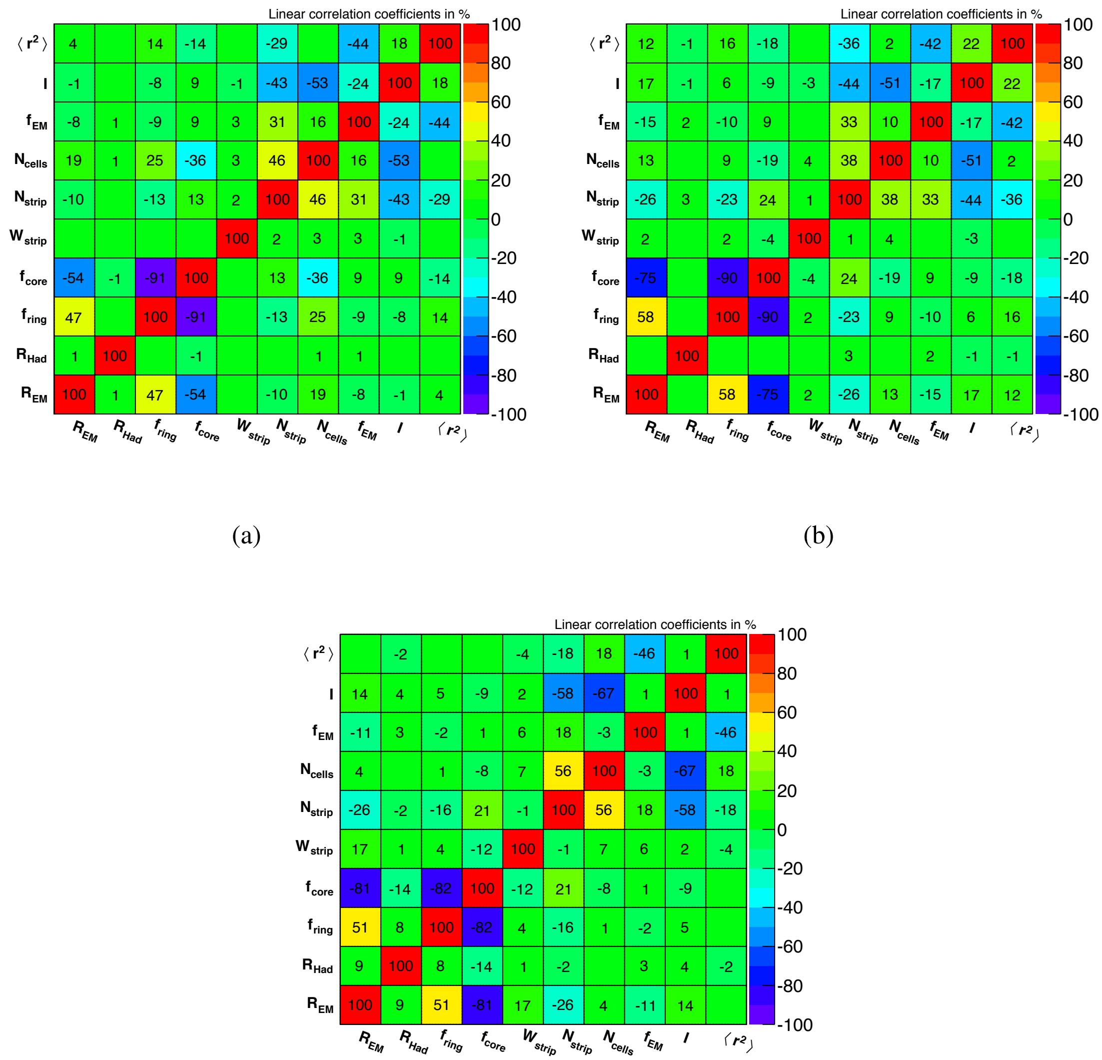
Figure 5.11: Average over effective clusters of the isolation of the clusters. (a) shows the integrated distribution for 10 GeV $< E_T <$ 70 GeV and $-2.5 < \eta < 2.5$ for reconstructed clusters and (b) shows the E_T profile for the same τ candidates.

5.3.5 Linear correlations

Cluster moment variables show very little correlation with the cell-based variables (Fig. 5.12) which indicates that they may hold distinct information.

On the other hand, Fig. 5.13 reveals that they correlate more strongly with the 4-momenta variables, especially N_{cells} and I. They may have little to offer on top of the 4-momenta variables.

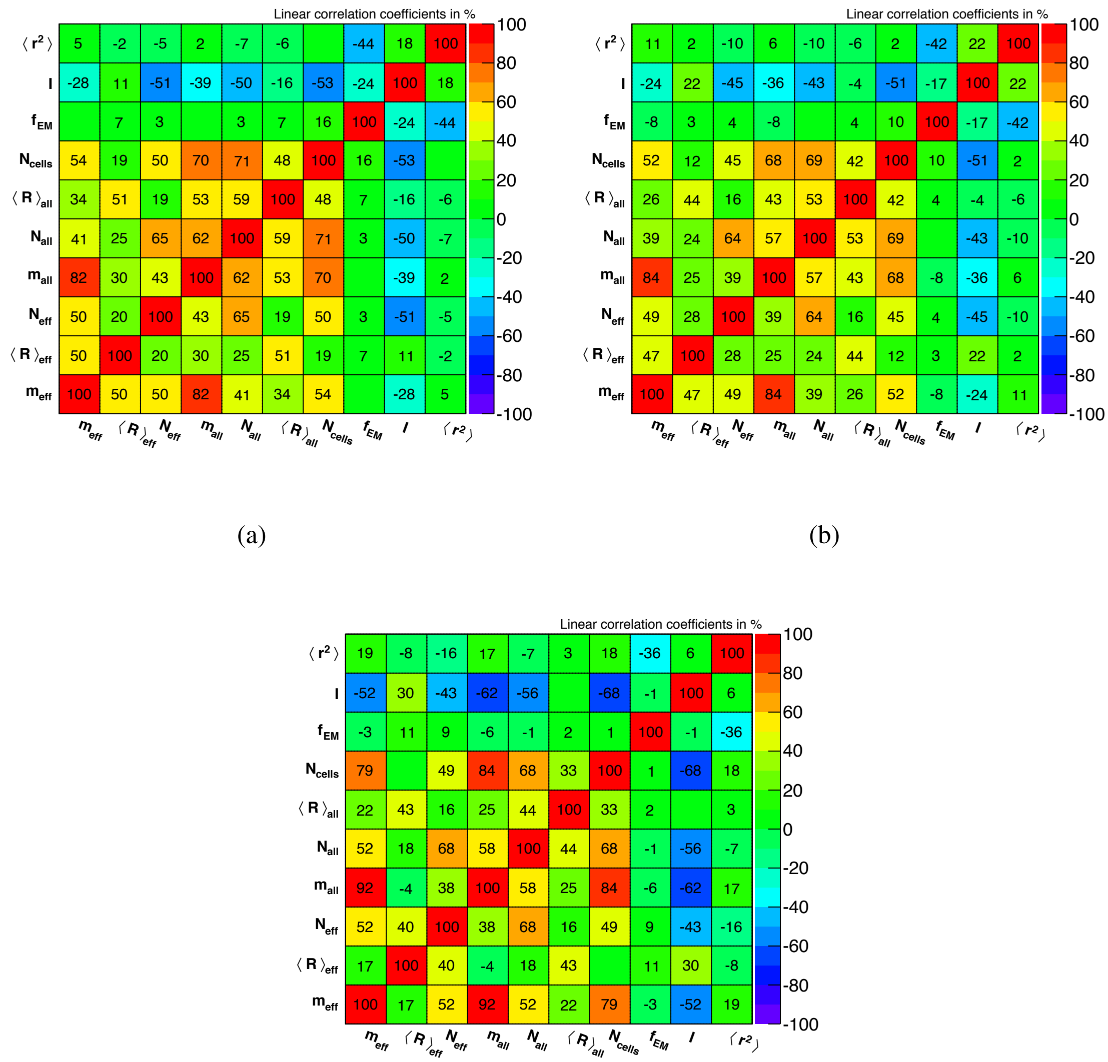
Among themselves, N_{cells} and I are significantly correlated, while f_{EM} and $\langle r^2 \rangle$ also show some correlation. Both correlations are easily explained by the facts that a larger number of cells in a jet of constant size will lead to less isolated clusters, and that electromagnetic showers are narrower.



(a) (b)

(c)

Figure 5.12: Linear correlation coefficients between cell-based and cluster moment variables for (a) 1prong signal, (b) 3 prongs signal, (c) background. An empty square represents a correlation factor of 0.



(a) (b)

(c)

Figure 5.13: Linear correlation coefficients between 4-momentum and cluster moment variables for (a) 1prong signal, (b) 3 prongs signal, (c) background. An empty square represents a correlation factor of 0.

Chapter 6

Performance

The difficult task of τ identification in the LHC collisions requires dedicated techniques to combine the discrimination power from many sources of information. At ATLAS, the information takes the form of identification variables. The discrimination power of these variables is consolidated using multi-variate techniques. The variables presented in Chapter 5 have been selected among larger sets of 4-momentum and cluster-moment variables based on how much distinct information they hold with respect to the cell-based variables. Measuring the distinct information held in a variable is a complex issue in itself, but it can be done using the same multi-variate techniques used for identification.

In Section 6.1, a multi-variate technique called boosted decision trees is introduced as a tool to measure the distinct information content of identification variables. The elimination procedures leading to the variables introduced in Chapter 5 are explained in Section 6.2. Finally, Section 6.3 shows a quantitative estimate of how much discrimination power can be gained by combining the cell-based variables with different sets of cluster-based variables.

6.1 Boosted decision trees

Boosted decision trees (BDT) [26] is one of the standard methods used for τ identification at ATLAS. How it works is best explained by first having a look at a single decision tree. A decision tree is a classifier with a binary tree-like structure. Each node in the decision tree corresponds to a selection cut on a single variable. The decision tree splits the multi-dimensional space defined by the input variables into many regions. A "purity" index is calculated for each of these regions, which quantifies the proportions of signal and background in the region. The BDT is simply a linear combination of weighted decision trees giving a continuous score.

The construction or "training" of a specific BDT is done by a specialized algorithm that uses provided samples of signal and background. The parameters used for BDT training in this study have been carefully optimized. They are the same parameters used for the standard BDT dedicated to τ identification at ATLAS.

A particular cut on the BDT score corresponds to choosing a signal efficiency, and the performance of the BDT is judged by the background rejection achieved at this signal efficiency. The signal (background) efficiency is defined as the fraction of all signal (background) τ candidates kept when cutting on the BDT score. The background rejection corresponds to how many background τ candidates are rejected for each one accepted. It is calculated from the background efficiency with

$$\text{background rejection} = \frac{1}{\text{background efficiency}} - 1. \tag{6.1}$$

The background rejection obtained from a BDT gives a good estimate of the discrimination power of any given set of identification variables.

The TMVA framework [27] has been used in this study to train and test boosted decision trees. In addition to calculating the background rejection as a function of signal efficiency, TMVA also ranks the variables according to the number of cuts made on them in the trees, weighted by the purity index of each cut. Removing a highly ranked variable causes a larger drop in performance than removing a variable that is poorly ranked, provided that the variables are not correlated. However, variables generally have some degree of correlation. The ranking may change drastically depending on the presence of strongly correlated variables in the set. Correlated variables will tend to lower each other's ranking.

On the other hand, the background rejection of a well-tuned BDT is independent of correlations, so that one of two highly correlated variables may be removed without much impact. It follows that if the background rejection increases when a few variables are added, it can be concluded that these new variables bring new information.

6.2 Selection of the cluster-based variables

6.2.1 Selection of the 4-momentum variables

The sixteen 4-momentum variables which have been considered in the early stages of the current study have strong correlations. To see if these 16 variables can be reduced to only a few, all 16 variables along with the 6 cell-based variables are provided as input to the BDT, and a background rejection versus signal efficiency curve is produced. This curve provides a reference from which any drop would indicate a loss of discrimination power.

From this point, the last variable in the BDT ranking is removed, and the performance is evaluated again. This step is repeated as long as the performance does not deviate visibly from the reference performance[1]. All 4-momentum variables except the ones presented in section 5.2 can be removed from the complete set without causing a visible drop in performance.

6.2.2 Selection of the cluster moment variables

There are 13 cluster moments made available for shower characterization and of each of these 13 moments, 6 different ways of combining the moment values of all clusters in a τ candidate have been explored. No attempt was made at combining different moments together.

The information held in the cluster moment variables which is distinct from that of the cell-based variables is spread thin among the variables. Adding all cluster moment variables to the cell-based variables causes a visible increase in performance, but the individual contributions are very small (with the exception of N_{cells}). In such a case, an elimination procedure like the one followed for the 4-momentum variables yields ambiguous results.

Instead, the cluster moment variables were added one by one to the cell-based variables and the area under the background rejection versus signal efficiency curve was monitored. Any variable that didn't cause a visible increase was dismissed. This procedure was done separately on the 6 different classes of cluster moment variables. It turned out that the same moments were responsible for the increase in performance in 5 of the 6 classes. Since the

[1]This procedure does not guarantee that the final subset will be the only subset with the maximum discrimination power available for any subset.

different versions of the same cluster moment variables were highly correlated, only the most powerful version was kept, resulting in the list of variables presented in section 5.3.

6.3 Performance of the cluster-based variables

In this section, background rejection versus signal efficiency curves calculated from the BDT described in Section 6.1 are used to show how much the new variables described in Chapter 5 can affect the performance of a calorimeter-based τ identification. In Figs. 6.1 to 6.6, plots (a), (b) and (c) ((d), (e) and (f)) show τ candidates for signal and background having $E_T < 35$ GeV ($E_T > 35$ GeV). Plots (a) and (d) show the performance for τ candidates associated with a single reconstructed track while plots (b) and (e) show the performance for τ candidates with many reconstructed tracks. While τ decays producing a single track constitute the majority of hadronic decays, most of the background will be categorized as having many tracks. Plot (c) and (f) shows the performance for τ candidates with any number of reconstructed tracks.

6.3.1 Performance of the 4-momentum variables

Performance at reconstruction level

Fig. 6.1 shows by how much the 4-momentum variables can improve the background rejection of the cell-based variables at a given signal efficiency. Below 35 GeV, the 4-momentum variables do not bring much of an improvement, but above 35 GeV, they significantly improve the background rejection. The 4-momentum variables make a larger improvement for many-track candidates than for single-track candidates.

A typical value of signal efficiency at which background rejection should be optimized for different analyses is 30%. Table 6.1 shows the gain in background rejection at 30% signal efficiency in each plot of Fig. 6.1.

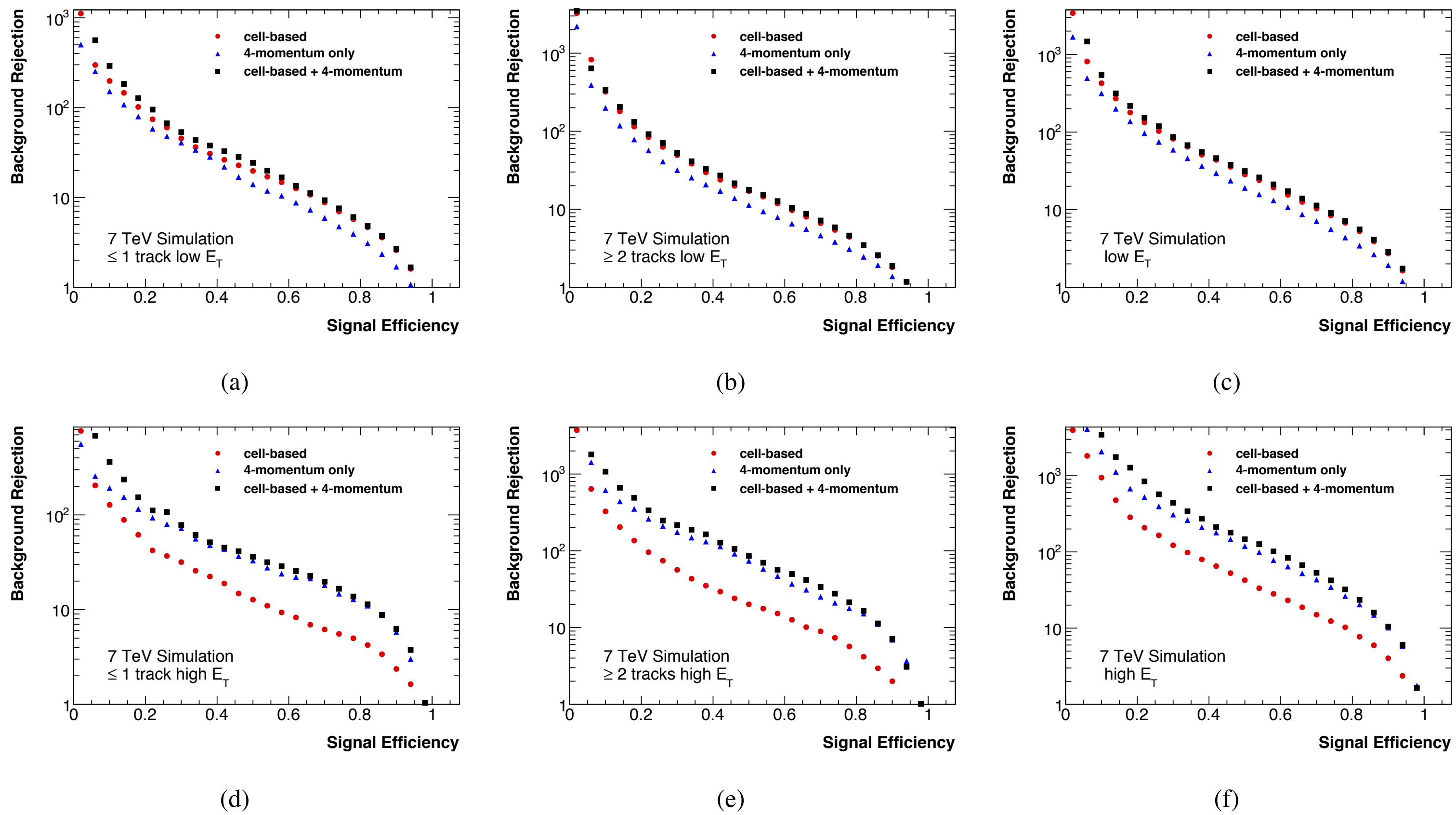


Figure 6.1: Comparison of the performance of the 4-momentum variables and the cell-based variables. (a) shows 1 prong candidates only, (b) shows 3 prong candidates only and (c) shows all candidates, each for $E_T < 35$ GeV. (d), (e) and (f) show the same but for $E_T > 35$ GeV.

Table 6.1: Gain in background rejection at 30% signal efficiency when adding the 4-momentum variables to the cell-based variables.

track selection	$E_T < 35$ GeV	$E_T > 35$ GeV
≤ 1 track	1.2	2.5
≥ 2 tracks	1.1	3.8
all	1.1	3.6

Table 6.2 shows the variable rankings associated with the six plots of Fig. 6.1. In these rankings, the top variable is usually used much more often than the other variables. Therefore, the order in which the subsequent variables appear is strongly dependent on correlations between variables and is not necessarily a good indicator of how much discrimination power they contribute.

At low E_T, f_{core} and R_{EM} are on top of the rankings, but the top five in each track selection is dominated by 4-momentum variables. At high E_T, m_{all} is deemed the most useful variable by the BDT, meaning that a large fraction of the discrimination power that the 4-momentum variables contribute comes from this variable.

In Fig. 6.2, adding the all-cluster variables to the cell-based variables is compared to adding the effective-cluster variables to the cell-based variables. By looking at the corresponding variable rankings in Tables 6.3 and 6.4, one can see that m_{eff} is also a very powerful variable on its own as it replaces m_{all} in the role of the most useful variable at high E_T.

Another important observation is that the subset of effective-cluster variables is better than the subset of all-cluster variables at improving over the cell-based variables. This is despite the fact that effective-cluster variables appear more correlated with cell-based variables, as can be seen in Fig. 5.7. According to the rankings, the geometrical radius variables are responsible for most of this difference.

Table 6.2: Variable rankings for 4-momentum variables and cell-based variables

Rank	$E_T < 35$ GeV			$E_T > 35$ GeV		
	≤ 1 track	≥ 2 tracks	all	≤ 1 track	≥ 2 tracks	all
1	f_{core}	R_{EM}	f_{core}	m_{all}	m_{all}	m_{all}
2	N_{eff}	m_{all}	N_{eff}	f_{core}	f_{ring}	f_{ring}
3	R_{EM}	N_{all}	R_{EM}	m_{eff}	f_{core}	m_{eff}
4	$\langle R\rangle_{\text{eff}}$	N_{eff}	$\langle R\rangle_{\text{eff}}$	N_{strip}	$\langle R\rangle_{\text{all}}$	f_{core}
5	m_{all}	R_{Had}	m_{all}	R_{EM}	$\langle R\rangle_{\text{eff}}$	N_{eff}
6	N_{strip}	f_{core}	N_{strip}	N_{eff}	m_{eff}	R_{EM}
7	N_{all}	N_{strip}	N_{all}	W_{strip}	R_{EM}	N_{all}
8	W_{strip}	m_{eff}	W_{strip}	$\langle R\rangle_{\text{eff}}$	W_{strip}	N_{strip}
9	f_{ring}	f_{ring}	f_{ring}	$\langle R\rangle_{\text{all}}$	R_{Had}	R_{Had}
10	$\langle R\rangle_{\text{all}}$	$\langle R\rangle_{\text{all}}$	$\langle R\rangle_{\text{all}}$	f_{ring}	N_{eff}	W_{strip}
11	m_{eff}	$\langle R\rangle_{\text{eff}}$	m_{eff}	R_{Had}	N_{all}	$\langle R\rangle_{\text{eff}}$
12	R_{Had}	W_{strip}	R_{Had}	N_{all}	N_{strip}	$\langle R\rangle_{\text{all}}$

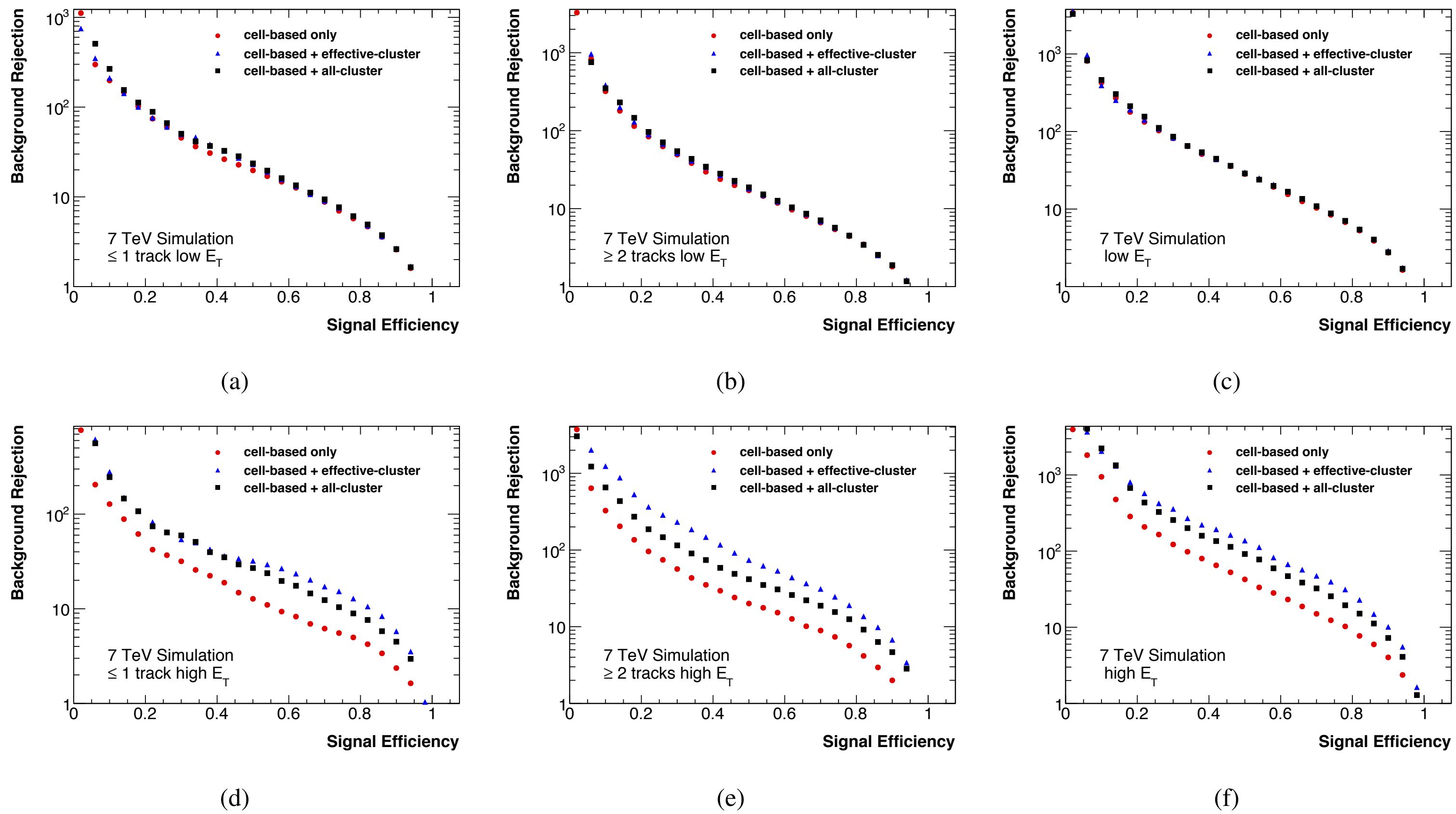


Figure 6.2: Comparison of the performance of the effective-cluster variables and the all-cluster variables when added to the cell-based variables. (a) shows 1 prong candidates only, (b) shows 3 prong candidates only and (c) shows all candidates, each for $E_T < 35$ GeV. (d), (e) and (f) show the same but for $E_T > 35$ GeV.

Table 6.3: Variable rankings for effective-cluster and cell-based variables.

Rank	$E_T < 35$ GeV			$E_T > 35$ GeV		
	≤ 1 track	≥ 2 tracks	all	≤ 1 track	≥ 2 tracks	all
1	f_{core}	R_{EM}	f_{core}	m_{eff}	m_{eff}	m_{eff}
2	R_{EM}	f_{core}	R_{EM}	f_{core}	f_{core}	f_{core}
3	m_{eff}	$\langle R \rangle_{\text{eff}}$	m_{eff}	W_{strip}	$\langle R \rangle_{\text{eff}}$	N_{strip}
4	N_{eff}	m_{eff}	N_{eff}	$\langle R \rangle_{\text{eff}}$	W_{strip}	$\langle R \rangle_{\text{eff}}$
5	$\langle R \rangle_{\text{eff}}$	N_{eff}	$\langle R \rangle_{\text{eff}}$	N_{strip}	N_{eff}	W_{strip}
6	W_{strip}	W_{strip}	W_{strip}	R_{EM}	N_{strip}	f_{ring}
7	f_{ring}	R_{Had}	f_{ring}	N_{eff}	f_{ring}	R_{EM}
8	N_{strip}	f_{ring}	N_{strip}	f_{ring}	R_{Had}	N_{eff}
9	R_{Had}	N_{strip}	R_{Had}	R_{Had}	R_{EM}	R_{Had}

Performance at generator-level

Since 4-momentum variables can be calculated from the truth particles, it is possible to use them to estimate how a calorimeter-based τ identification would do given a perfect visible decay product reconstruction[2]. Fig. 6.3 shows the performance of the BDT using only the six 4-momentum variables as input, calculated from truth particles (including underlying event particles) and from reconstructed clusters. These plots provide an estimate of how much room there is left for improvement. However, one should keep in mind that the blue curve can never be attained, since some detector-related effects such as showering impose practical limits on position and energy resolution. Table 6.5 show the ratios of the background rejections at 30% signal efficiency in Fig. 6.3.

[2]Although it is assumed that the mass of the decay products cannot be reconstructed, so that the decay products have massless 4-momenta.

Table 6.4: Variable rankings for all-cluster and cell-based variables.

Rank	$E_T < 35$ GeV			$E_T > 35$ GeV		
	≤ 1 track	≥ 2 tracks	all	≤ 1 track	≥ 2 tracks	all
1	f_{core}	R_{EM}	f_{core}	m_{all}	m_{all}	m_{all}
2	R_{EM}	f_{core}	R_{EM}	f_{core}	f_{ring}	f_{ring}
3	f_{ring}	N_{all}	m_{all}	$\langle R \rangle_{\text{all}}$	f_{core}	f_{core}
4	N_{strip}	m_{all}	N_{all}	f_{ring}	$\langle R \rangle_{\text{all}}$	N_{strip}
5	$\langle R \rangle_{\text{all}}$	R_{Had}	W_{strip}	N_{strip}	R_{Had}	R_{EM}
6	m_{all}	f_{ring}	f_{ring}	R_{EM}	R_{EM}	N_{all}
7	N_{all}	N_{strip}	$\langle R \rangle_{\text{all}}$	N_{all}	W_{strip}	W_{strip}
8	W_{strip}	$\langle R \rangle_{\text{all}}$	N_{strip}	R_{Had}	N_{strip}	R_{Had}
9	R_{Had}	W_{strip}	R_{Had}	W_{strip}	N_{all}	$\langle R \rangle_{\text{all}}$

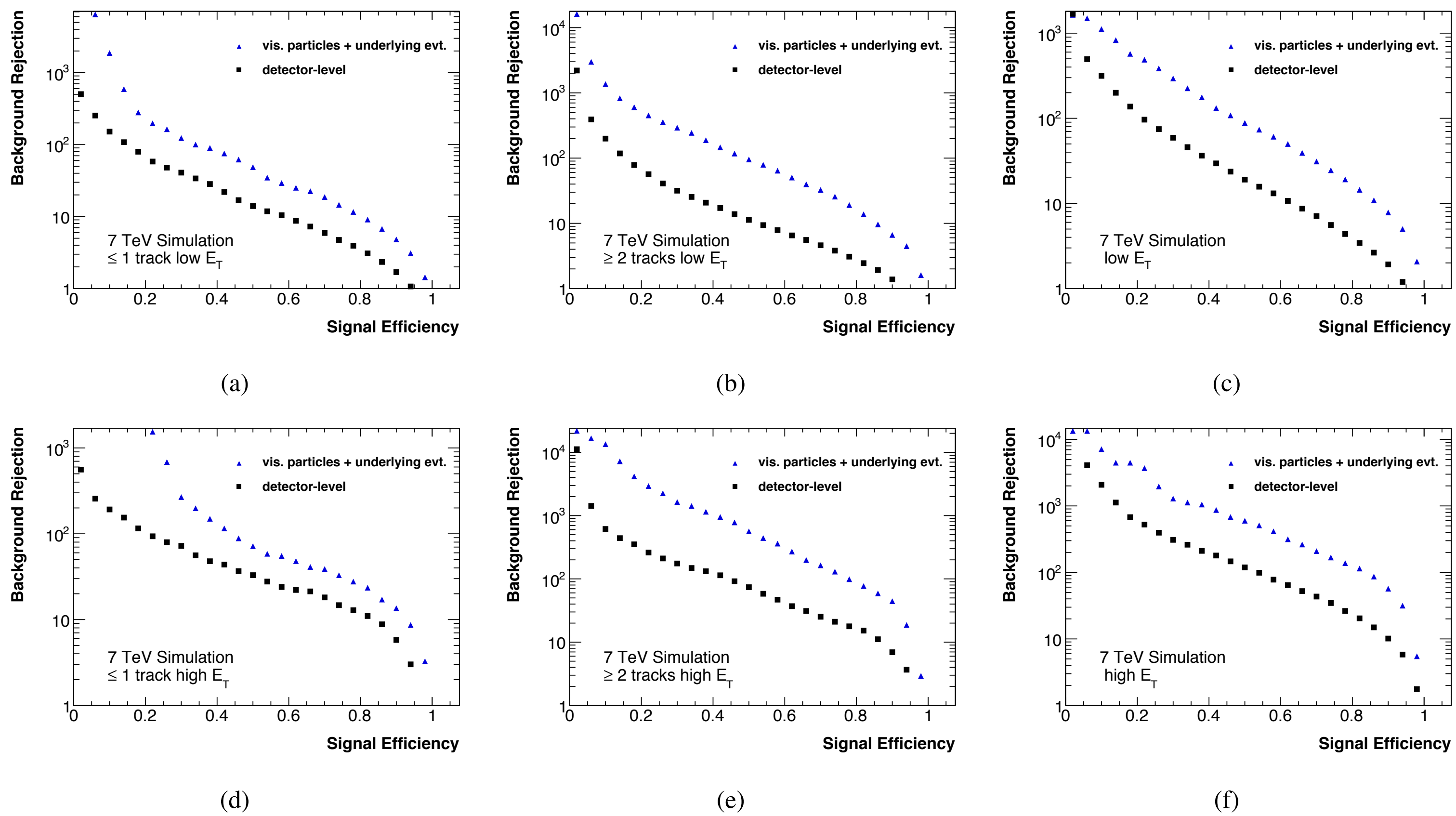


Figure 6.3: Comparison of the performance of the 4-momentum variables when calculated from truth particles and from reconstructed clusters. (a) shows 1 prong candidates only, (b) shows 3 prong candidates only and (c) shows all candidates, each for $E_T < 35$ GeV. (d), (e) and (f) show the same but for $E_T > 35$ GeV.

Table 6.5: Ratio of the in background rejection at 30% signal efficiency based on the 4-momentum variables between an ideal calorimeter and the ATLAS calorimeter.

track selection	$E_T < 35$ GeV	$E_T > 35$ GeV
≤ 1 track	3.0	3.7
≥ 2 tracks	9.2	9.4
all	5.0	4.2

6.3.2 Performance of the cluster moment variables

Performance at reconstruction level

Fig. 6.4 and Table 6.6 show the gain in performance obtained by combining cluster moment variables with cell-based variables. At low E_T, cluster moment variables are rather weak and do not contribute any significant amount of discrimination power. At high E_T, they are seen to outperform the cell-based variables. Table 6.7 reveals that N_{cells} is responsible for this. This does not come as a surprise since N_{cells} brings much of the same information provided by the invariant mass variables (see Fig. 5.13), which are responsible for a similar behavior when 4-momentum variables are considered instead.

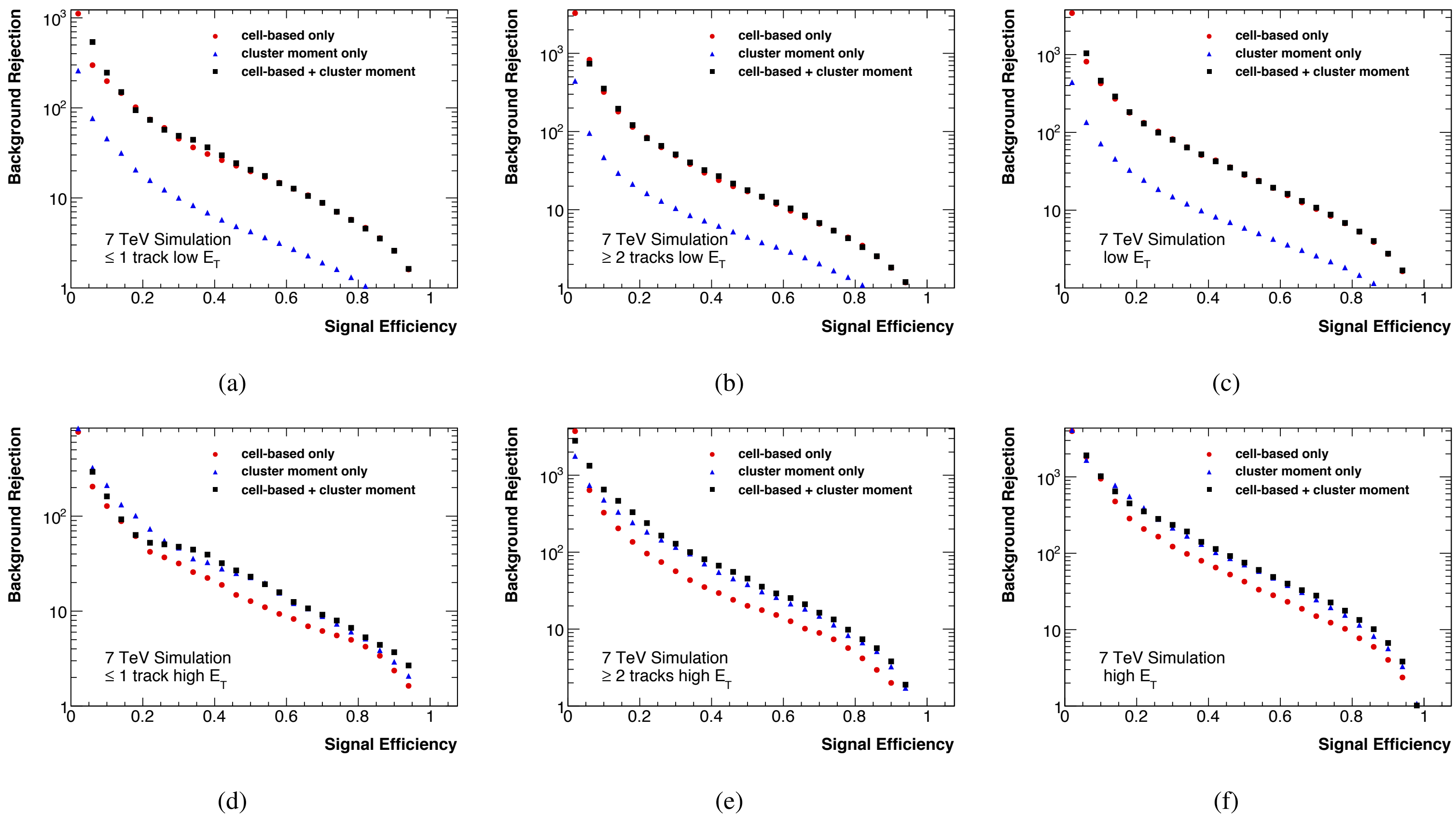


Figure 6.4: Comparison of the performance of cluster moment variables and the cell-based variables. (a) shows 1 prong candidates only, (b) shows 3 prong candidates only and (c) shows all candidates, each for $E_T < 35$ GeV. (d), (e) and (f) show the same but for $E_T > 35$ GeV.

Table 6.6: Gain in background rejection at 30% signal efficiency when adding the cluster moment variables to the cell-based variables.

track selection	$E_T < 35$ GeV	$E_T > 35$ GeV
≤ 1 track	1.1	1.5
≥ 2 tracks	1.0	2.3
all	1.0	1.9

Table 6.7: Variable rankings for cluster moment variables and cell-based variables

Rank	$E_T < 35$ GeV			$E_T > 35$ GeV		
	≤ 1 track	≥ 2 tracks	all	≤ 1 track	≥ 2 tracks	all
1	f_{core}	R_{EM}	f_{core}	f_{core}	N_{cells}	N_{cells}
2	R_{EM}	f_{core}	N_{cells}	N_{cells}	f_{core}	R_{EM}
3	f_{EM}	N_{cells}	R_{EM}	R_{EM}	f_{ring}	f_{core}
4	f_{ring}	R_{Had}	f_{EM}	N_{strip}	$\langle r^2 \rangle$	f_{ring}
5	N_{strip}	W_{strip}	N_{strip}	I	I	N_{strip}
6	W_{strip}	f_{EM}	W_{strip}	f_{EM}	W_{strip}	$\langle r^2 \rangle$
7	N_{cells}	N_{strip}	f_{ring}	W_{strip}	R_{Had}	R_{Had}
8	I	f_{ring}	I	f_{ring}	f_{EM}	I
9	$\langle r^2 \rangle$	I	$\langle r^2 \rangle$	R_{Had}	R_{EM}	f_{EM}
10	R_{Had}	$\langle r^2 \rangle$	R_{Had}	$\langle r^2 \rangle$	N_{strip}	W_{strip}

Fig. 6.5 and Table 6.8 show the impact of adding the cluster moment variables to the 4-momentum variables. They do not increase the performance by much. Table 6.9 shows that 4-momentum variables are dominating the ranking. The cluster moment variable achieving the highest ranking is N_{cells}.

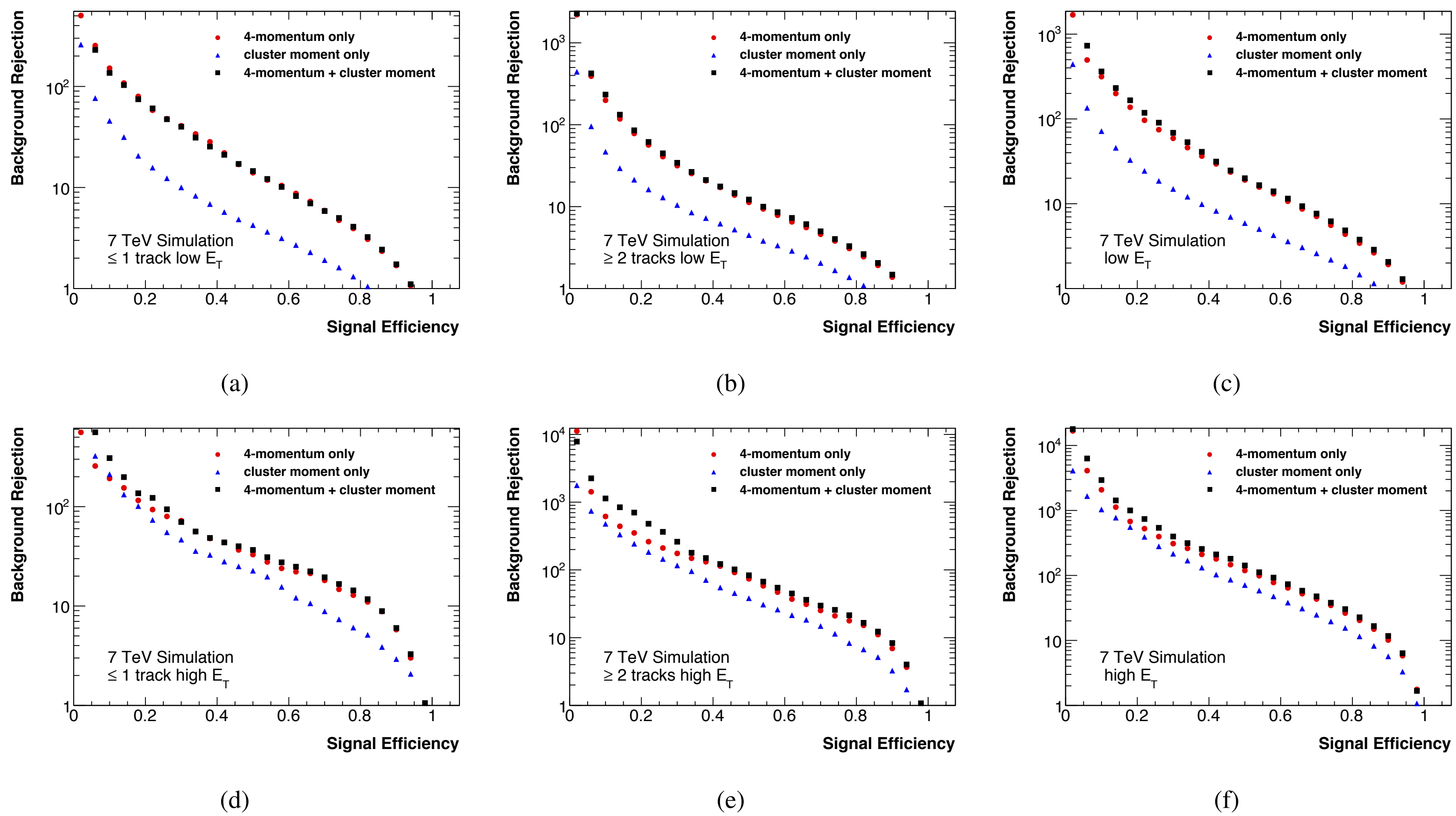


Figure 6.5: Comparison of the performance of the cluster moment variables and the 4-momentum variables. (a) shows 1 prong candidates only, (b) shows 3 prong candidates only and (c) shows all candidates, each for $E_T < 35$ GeV. (d), (e) and (f) show the same but for $E_T > 35$ GeV.

Table 6.8: Gain in background rejection at 30% signal efficiency when adding the cluster moment variables to the 4-momentum variables.

track selection	$E_T < 35$ GeV	$E_T > 35$ GeV
≤ 1 track	1.0	1.0
≥ 2 tracks	1.1	1.5
all	1.2	1.3

Table 6.9: Variable rankings for cluster moment variables with 4-momentum variables

Rank	$E_T < 35$ GeV			$E_T > 35$ GeV		
	≤ 1 track	≥ 2 tracks	all	≤ 1 track	≥ 2 tracks	all
1	$\langle R\rangle_{\text{eff}}$	$\langle R\rangle_{\text{eff}}$	$\langle R\rangle_{\text{eff}}$	m_{all}	m_{all}	m_{all}
2	$\langle R\rangle_{\text{all}}$	m_{all}	N_{cells}	N_{cells}	f_{EM}	m_{eff}
3	N_{eff}	N_{cells}	N_{eff}	$\langle r^2\rangle$	N_{cells}	N_{cells}
4	N_{all}	N_{eff}	$\langle R\rangle_{\text{all}}$	m_{eff}	m_{eff}	N_{eff}
5	m_{all}	$\langle R\rangle_{\text{all}}$	f_{EM}	$\langle R\rangle_{\text{eff}}$	$\langle r^2\rangle$	N_{all}
6	m_{eff}	f_{EM}	m_{all}	N_{eff}	N_{eff}	I
7	$\langle r^2\rangle$	$\langle r^2\rangle$	$\langle r^2\rangle$	$\langle R\rangle_{\text{all}}$	I	$\langle R\rangle_{\text{eff}}$
8	N_{cells}	I	m_{eff}	I	N_{all}	$\langle r^2\rangle$
9	I	N_{all}	I	f_{EM}	$\langle R\rangle_{\text{eff}}$	f_{EM}
10	f_{EM}	m_{eff}	N_{all}	N_{all}	$\langle R\rangle_{\text{all}}$	$\langle R\rangle_{\text{all}}$

6.3.3 Combining all Calorimeter-based variables

Fig. 6.6 shows the result of successively adding the 4-momentum variables and the cluster moment variables to the cell-based variables. Once the 4-momentum variables are added, little remains to be gained from the cluster moment variables. Table 6.11 shows the corresponding rankings and Table 6.10 shows the gain in rejection at 30% signal efficiency when adding the cluster moment variables to the cell-based and 4-momentum variables.

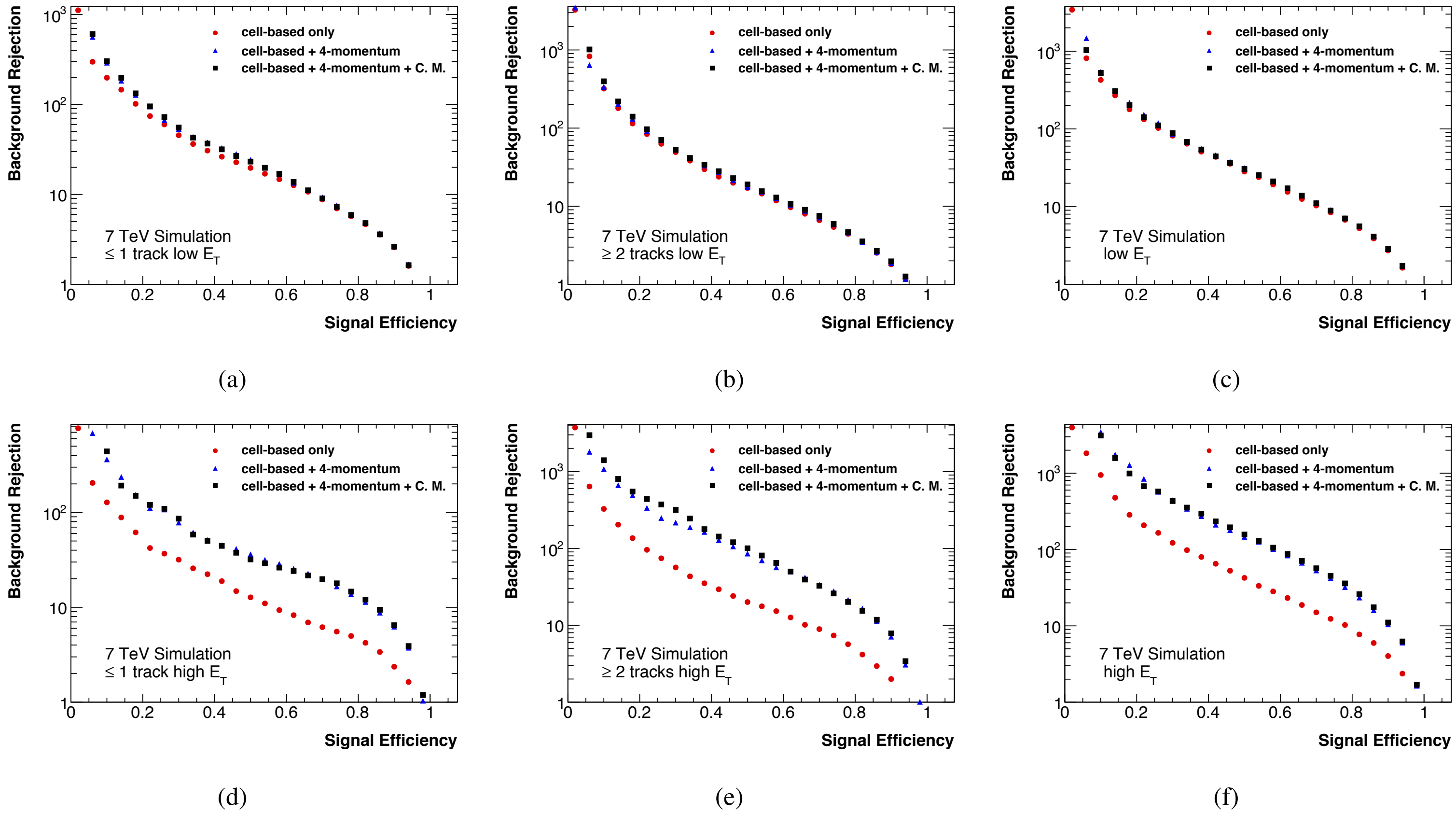


Figure 6.6: Comparison of the performance of the cell-based variables, the 4-momentum variables and the cluster moment variables when put together. (a) shows 1 prong candidates only, (b) shows 3 prong candidates only and (c) shows all candidates, each for $E_T < 35$ GeV. (d), (e) and (f) show the same but for $E_T > 35$ GeV.

Table 6.10: Gain in background rejection at 30% signal efficiency when adding the cluster moment variables to the cell-based and 4-momentum variables.

track selection	$E_T < 35$ GeV	$E_T > 35$ GeV
≤ 1 track	1.0	1.1
≥ 2 tracks	1.0	1.5
all	1.0	1.0

Table 6.11: Variable rankings for all variables

Rank	$E_T < 35$ GeV			$E_T > 35$ GeV		
	≤ 1 track	≥ 2 tracks	all	≤ 1 track	≥ 2 tracks	all
1	f_{core}	R_{EM}	f_{core}	m_{all}	m_{all}	m_{all}
2	R_{EM}	f_{core}	N_{cells}	N_{cells}	f_{ring}	m_{eff}
3	$\langle R \rangle_{eff}$	N_{cells}	R_{EM}	f_{core}	f_{core}	N_{cells}
4	f_{EM}	N_{all}	$\langle R \rangle_{eff}$	$\langle r^2 \rangle$	f_{EM}	f_{ring}
5	$\langle R \rangle_{all}$	m_{all}	N_{eff}	N_{strip}	I	f_{core}
6	m_{eff}	N_{eff}	f_{EM}	$\langle R \rangle_{all}$	$\langle R \rangle_{eff}$	R_{Had}
7	N_{strip}	f_{EM}	N_{strip}	W_{strip}	W_{strip}	I
8	f_{ring}	$\langle R \rangle_{eff}$	W_{strip}	N_{eff}	N_{cells}	N_{strip}
9	N_{eff}	f_{ring}	m_{all}	I	$\langle R \rangle_{all}$	$\langle r^2 \rangle$
10	m_{all}	$\langle R \rangle_{all}$	N_{all}	m_{eff}	N_{all}	W_{strip}
11	W_{strip}	I	$\langle R \rangle_{all}$	$\langle R \rangle_{eff}$	$\langle r^2 \rangle$	N_{eff}
12	N_{cells}	W_{strip}	f_{ring}	R_{EM}	R_{EM}	$\langle R \rangle_{eff}$
13	N_{all}	N_{strip}	$\langle r^2 \rangle$	R_{Had}	N_{eff}	f_{EM}
14	I	$\langle r^2 \rangle$	m_{eff}	f_{ring}	m_{eff}	N_{all}
15	$\langle r^2 \rangle$	m_{eff}	I	f_{EM}	R_{Had}	$\langle R \rangle_{all}$
16	R_{Had}	R_{Had}	R_{Had}	N_{all}	N_{strip}	R_{EM}

6.4 Effect of pileup

As the understanding of the LHC improves, the running conditions are changed to achieve better performance. As the instantaneous luminosity increases, the probability of multiple simultaneous interactions (pileup) also increases. Pileup can be divided in two categories. Multiple collisions happening every bunch crossing are usually referred to as in-time pileup. Out-of-time pileup refers to successive bunch crossings that are very close in time, causing the response of the detector to one bunch crossing to overlap with the response of the next. In-time pileup is present in the 7 TeV collisions studied to date, but not out-of-time pileup. The pileup conditions considered in this section are 5 collisions on average per bunch crossing every 900ns, which means essentially no out-of-time pileup. In the rest of this thesis, pileup will refer to these specific conditions.

Pileup is known to have a negative impact on τ identification. It tends to make reconstructed objects wider and impairs the energy resolution. The additional pileup particles ending up with the objects of interest tend to have a low energy, since they come mostly from minimum bias collisions. Cell-based variables already have protection mechanisms against these extra particles. First, since they are calculated from cells within topological clusters only, they benefit from the noise suppression[3] which also suppress the influence of the lowest energy pileup particles. Second, R_{EM}, R_{Had} and W_{strip} are calculated by weighting the contribution of each cell with E_T so that small weights are given to additional low energy cells from pileup. f_{core} and f_{ring} have the most effective protection mechanism, which relies on the uniform spreading of the pileup particles in η,ϕ-space. The additional energy from pileup is proportional to any area in η,ϕ-space. By taking the ratio of the E_T in two different areas, most of the pileup contribution partially cancels out over many events.

These tricks used for the cell-based variables do not necessarily apply to cluster-based variables. From a topological cluster perspective, pileup can be separated in two effects: more energy per cluster (pileup particles ending up in the clusters of the object of interest) and more clusters per jet (pileup particles creating their own clusters). In this section, it will be demonstrated that the effective-cluster variables have a built-in resistance against

[3]The expected noise σ used in the topological clustering algorithm can be adjusted to the changing conditions of the LHC. However, no such adjustment has been made in the study presented in this chapter. Only electronic noise is assumed.

the latter of these effects by comparing them to cell-based and all-cluster variables.

Generally, for the cluster-based variables, the E_T-dependence is not changed much by pileup. For this reason, the effect of pileup has been evaluated on the complete set of candidates, without breaking it down in two E_T bins like in the previous chapter.

There are two definitions of pileup resistance that should be considered. First, an identification variable is deemed resistant to pileup if it doesn't lose discrimination power in pileup conditions. This statement is independent of what happens to the shape of the signal and background distributions. The second definition is the requirement that distributions themselves remain unaffected. Which definition is most relevant depends on how quickly it is possible to adapt the τ identification techniques used by the ATLAS collaboration to pileup conditions. Section 6.4.1 quantifies the resistance to pileup of the 4-momentum variables following the second definition, while section 6.4.2 does the same but following the first definition.

6.4.1 Effect on the distributions

Fig. 6.7 shows the effect of pileup on N_{all}. By looking at the effect of pileup on signal (Fig. 6.7(a)) and on background (Fig. 6.7(b)), it can be seen that the whole distributions are shifted. This means that pileup affects most candidates. Also, background is less affected than signal, which is made more explicit by looking at the relative mean shift defined as

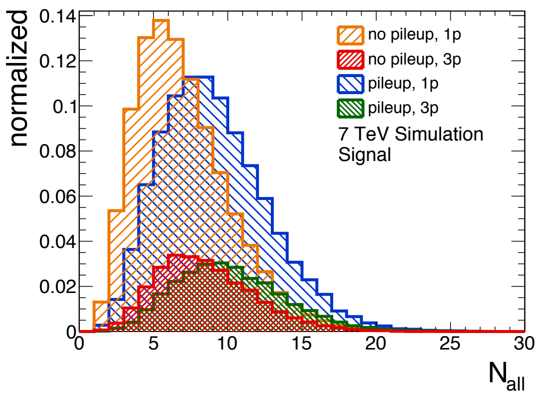
$$\text{rel. mean shift} = \frac{\langle x \rangle_{\text{pileup}} - \langle x \rangle_{\text{no pileup}}}{\langle x \rangle_{\text{no pileup}}} \tag{6.2}$$

and the relative RMS shift defined as

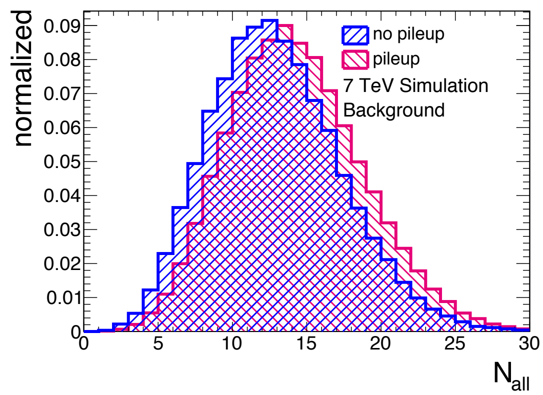
$$\text{rel. RMS shift} = \frac{\text{RMS}_{\text{pileup}} - \text{RMS}_{\text{no pileup}}}{\text{RMS}_{\text{no pileup}}}. \tag{6.3}$$

The relative mean and RMS shift values for cell-based and 4-momentum variables are listed in Table 6.12. These two numbers quantify the change in shape of the distributions with pileup. However, they should not be taken too seriously since the distributions from which they are calculated are non-gaussian.

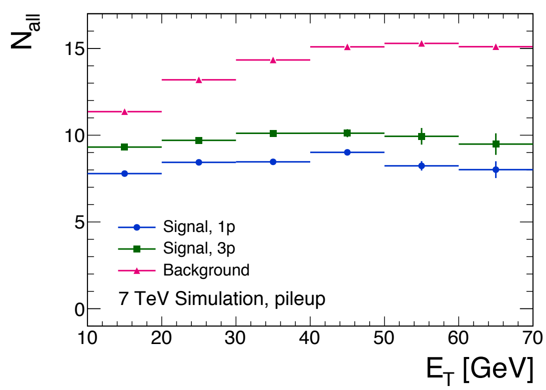
The fact that narrower objects are more affected by pileup is an effect of the way anti-k_T gathers everything within the limit of its reach parameter R. Since τ decays are narrower than jets, they have more space within R which is unoccupied, and potentially filled by pileup clusters. Hadronic jets don't have much room for additional clusters, so the pileup contribution to these jets cannot change their number of clusters as much.



(a)



(b)

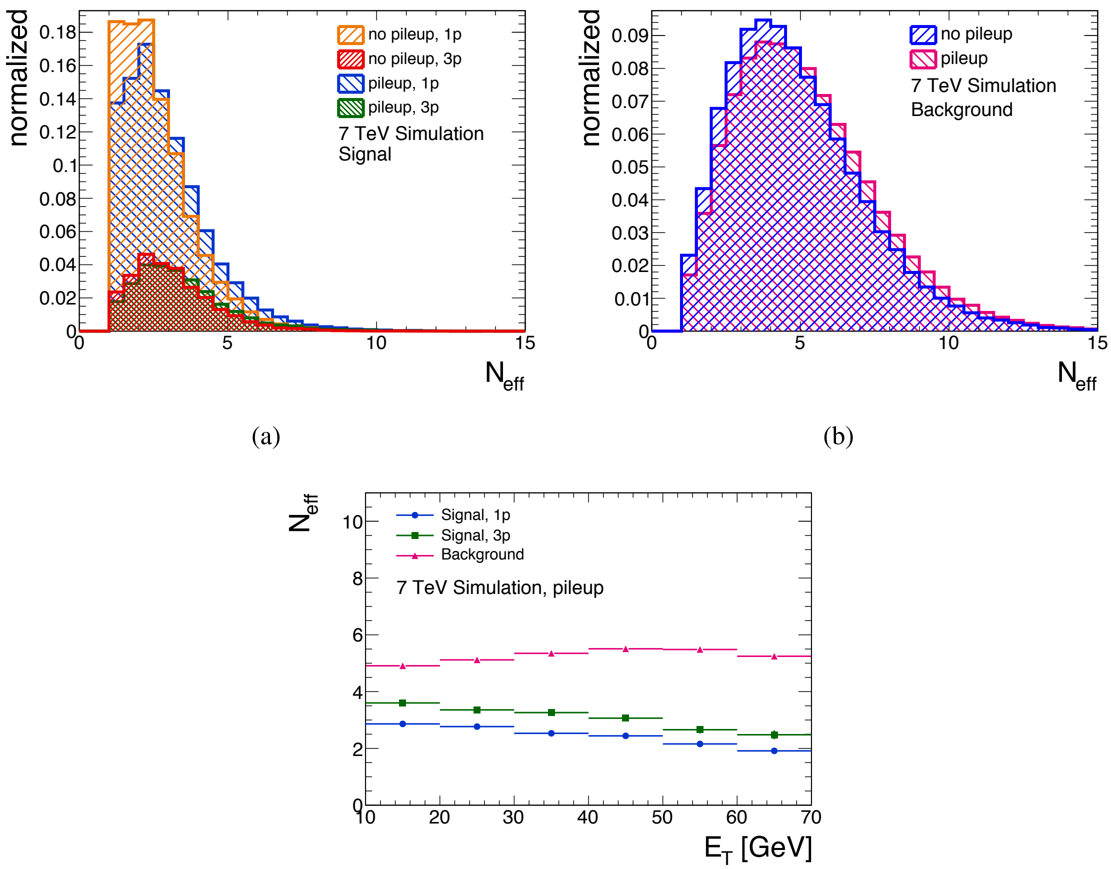


(c)

Figure 6.7: Number of clusters. (a) shows effect of pileup on signal, (b) shows effect of pileup on background, (c) shows E_T profiles when pileup is present.

Since particles coming from the pileup collisions are expected to have a significantly lower energy than the object of interest, they are not supposed to affect N_{eff} much, which is what is seen in Fig. 6.8. However, the structure of the signal distributions gets smeared by the pileup particles. When compared to Fig. 5.2(f), Fig. 6.8(c) shows that pileup negatively affects N_{eff} mostly at low E_T.

Even if N_{eff} has a good resistance to pileup, it does not guarantee that the selection



(a) (b)

(c)

Figure 6.8: Effective number of clusters. (a) shows effect of pileup on signal, (b) shows effect of pileup on background, (c) shows E_T profiles when pileup is present.

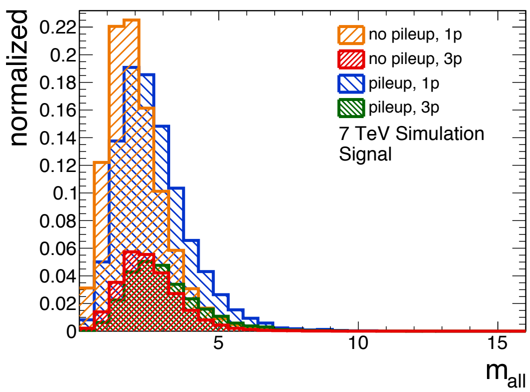
Table 6.12: Relative mean and RMS shifts of the cell-based and 4-momenta variables caused by the presence of pileup. Smaller values are consistent with pileup resistance.

Variable	Signal				Background	
	1 prong		3 prong			
	mean	RMS	mean	RMS	mean	RMS
R_{EM}	0.19	0.20	0.12	0.12	0.03	0.01
R_{Had}	0.15	0.10	0.07	0.08	0.03	0.02
f_{core}	-0.06	0.13	-0.06	0.07	-0.02	0.00
f_{ring}	0.08	0.08	0.02	0.02	0.00	-0.01
W_{strip}	0.71	0.34	0.40	0.21	0.08	0.02
N_{strip}	0.03	0.03	0.04	0.02	0.01	0.01
N_{all}	0.37	0.19	0.21	0.12	0.11	0.02
m_{all}	0.35	0.25	0.20	0.22	0.05	0.01
$\langle R\rangle_{all}$	0.22	-0.12	0.12	-0.06	0.03	-0.07
N_{eff}	0.14	0.18	0.10	0.14	0.06	0.04
m_{eff}	0.24	0.26	0.16	0.21	0.03	0.01
$\langle R\rangle_{eff}$	0.20	0.09	0.13	0.07	0.04	0.01

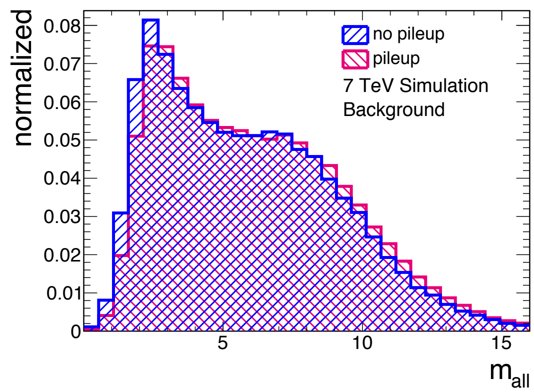
of effective constituents used in the calculation of m_{eff} and $\langle R\rangle_{eff}$ is equally resistant. By comparing Fig. 6.9 and 6.10 and by looking at Table 6.12, it can be seen that m_{eff} seems to have a better behaviour against pileup than m_{all}. This indicates that the selection of the effective constituents works to a certain degree as a protection mechanism against pileup.

Fig. 6.11 shows the effect of pileup on $\langle R\rangle_{all}$. This variable loses a lot of its discrimination power with pileup. This is expected since it has no protection at all against the addition of low energy clusters.

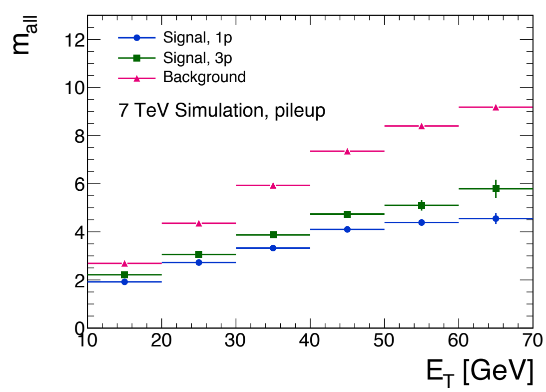
$\langle R\rangle_{eff}$ is the weakest effective-cluster variable against pileup. Some τ decays previously unaffected by underlying event particles now harbour pileup particles in their vicinity, making the second bump in the signal distribution more important (see Section 5.2.6).



(a)

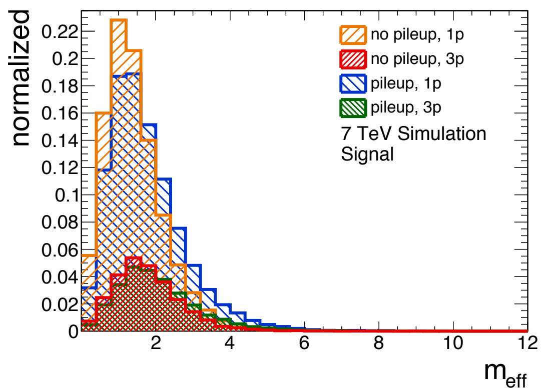


(b)

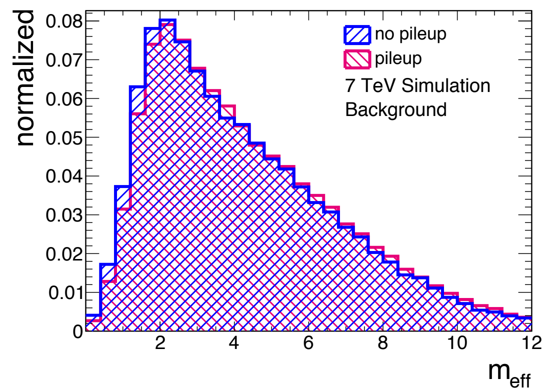


(c)

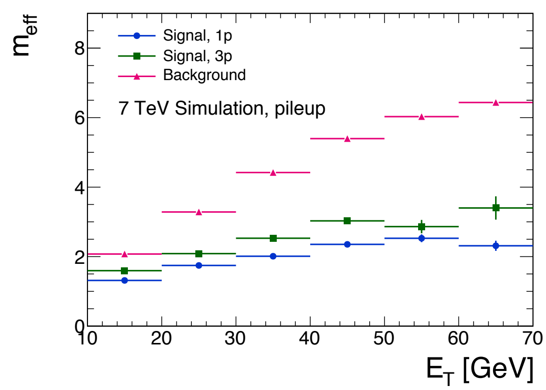
Figure 6.9: Invariant mass of all clusters. (a) shows effect of pileup on signal, (b) shows effect of pileup on background, (c) shows E_T profiles when pileup is present.



(a)

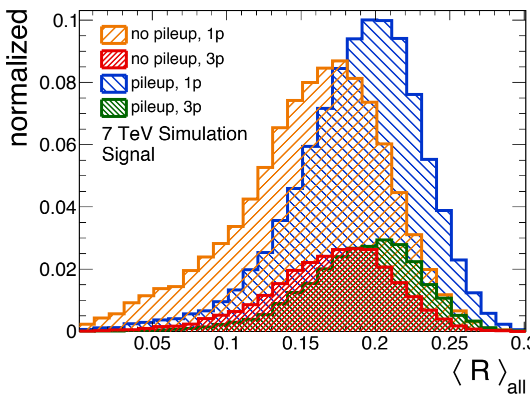


(b)

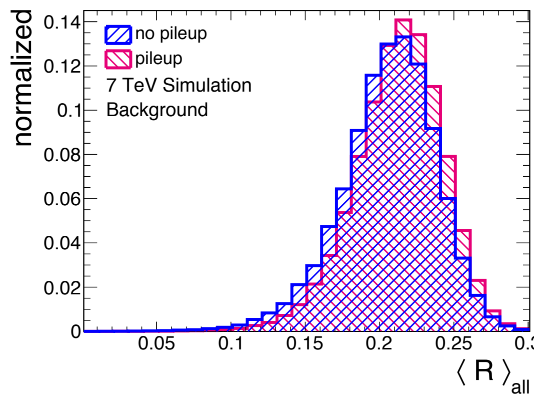


(c)

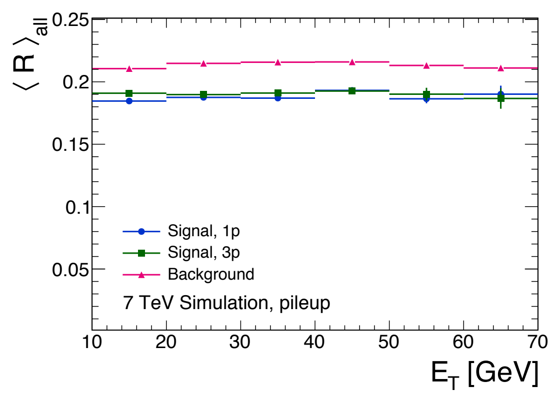
Figure 6.10: Invariant of the effective clusters. (a) shows effect of pileup on signal, (b) shows effect of pileup on background, (c) shows E_T profiles when pileup is present.



(a)



(b)



(c)

Figure 6.11: Average ΔR to the geometrical center of all clusters. (a) shows effect of pileup on signal, (b) shows effect of pileup on background, (c) shows E_T profiles when pileup is present.

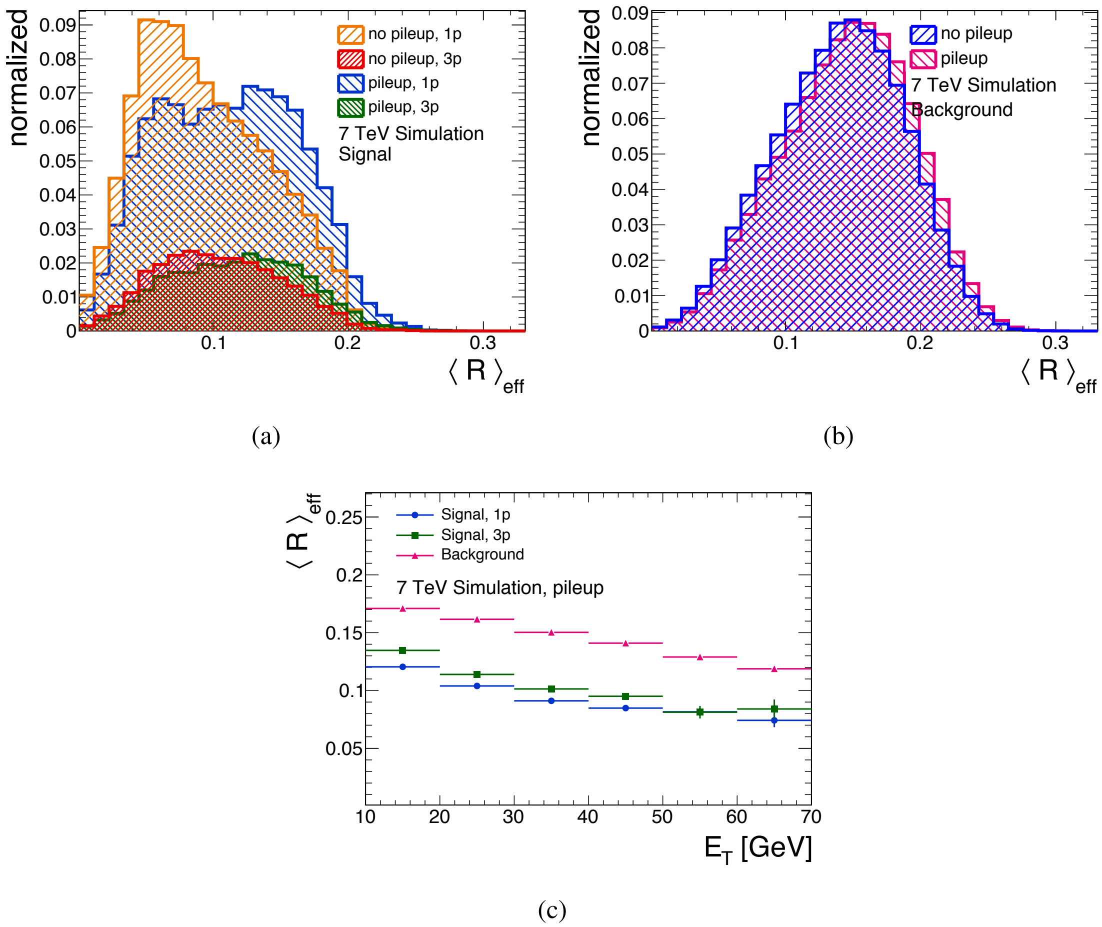


Figure 6.12: Average ΔR to the geometrical center of the effective clusters. (a) shows effect of pileup on signal, (b) shows effect of pileup on background, (c) shows E_T profiles when pileup is present.

6.4.2 Effect on the performance

Fig. 6.13 shows the ratio of the background rejection with pileup to the background rejection without pileup for different sets of variables. It should be noted that the boosted decision trees used to calculate the background rejection with pileup have been retrained on the pileup samples in order to measure the true loss of discrimination power due to pileup. In general, the effective-cluster variables are seen to have the smallest loss of performance.

Figs. 6.14 and 6.15 show the ratio of the background rejection of a given subset of variable to the background rejection of the whole set (cell-based and 4-momentum variables). Fig. 6.14 shows this ratio when there is no pileup, and Fig. 6.15 shows what happens to it when pileup is present. The effective-cluster variables are seen to stand higher above the cell-based and all-cluster variables when pileup is present.

Table 6.13 shows the variable ranking with pileup. It is seen that most of the resistance to pileup of the effective-cluster variables comes from N_{eff} since it rises in the ranking consistently and significantly. It is hard to tell how much m_{eff} contributes to the pileup resistance because of its correlation to m_{all}. Also, when m_{all} is taken out, m_{eff} is always at the top of the rankings. It can be seen that $\langle R \rangle_{\text{eff}}$ rises further above $\langle R \rangle_{\text{all}}$ consistently with pileup, even if its absolute ranking goes down.

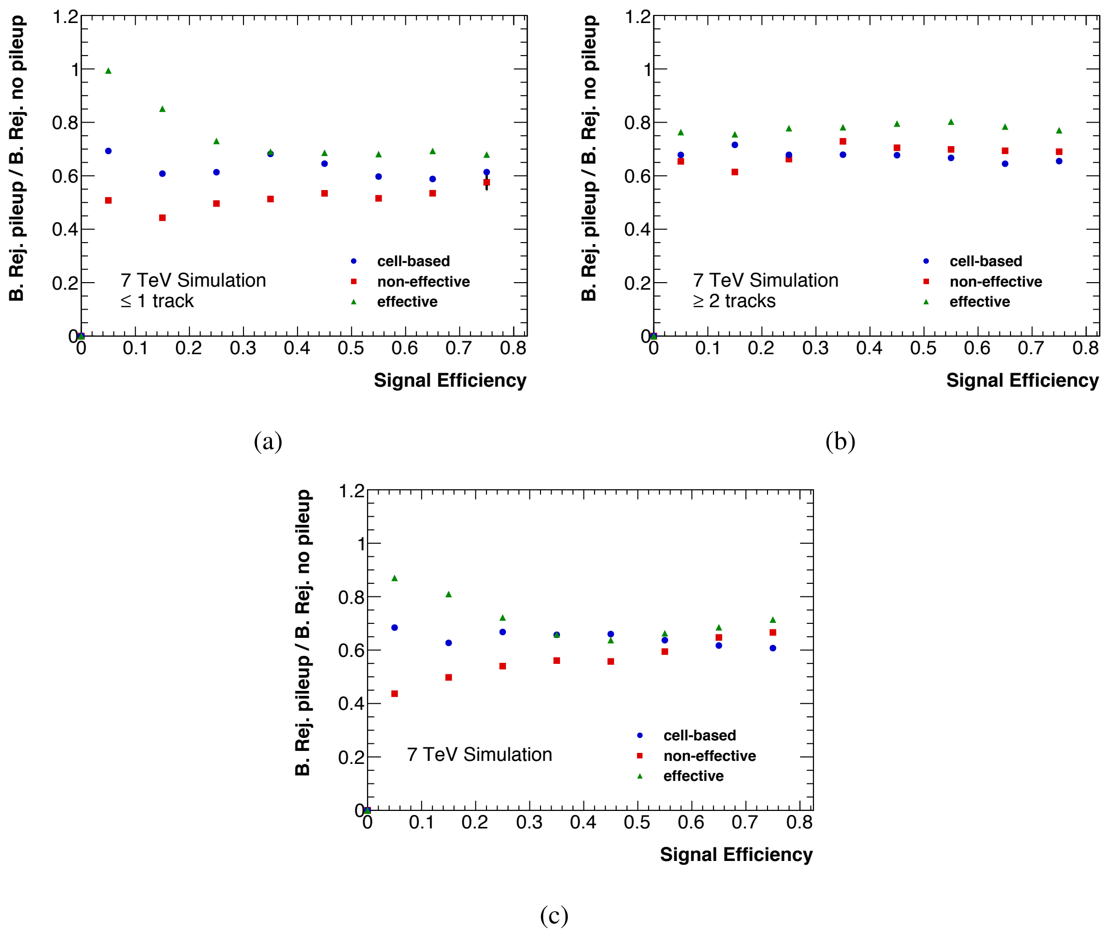


Figure 6.13: Fraction of background rejection retained when pileup is taken into account with respect to signal efficiency for different sets of variables. (a) 1 prong candidates only, (b) 3 prong candidates only, (c) all candidates. Statistical error bars on the background rejection are plotted. Above a signal efficiency of 0.8, they become larger than the difference between the points.

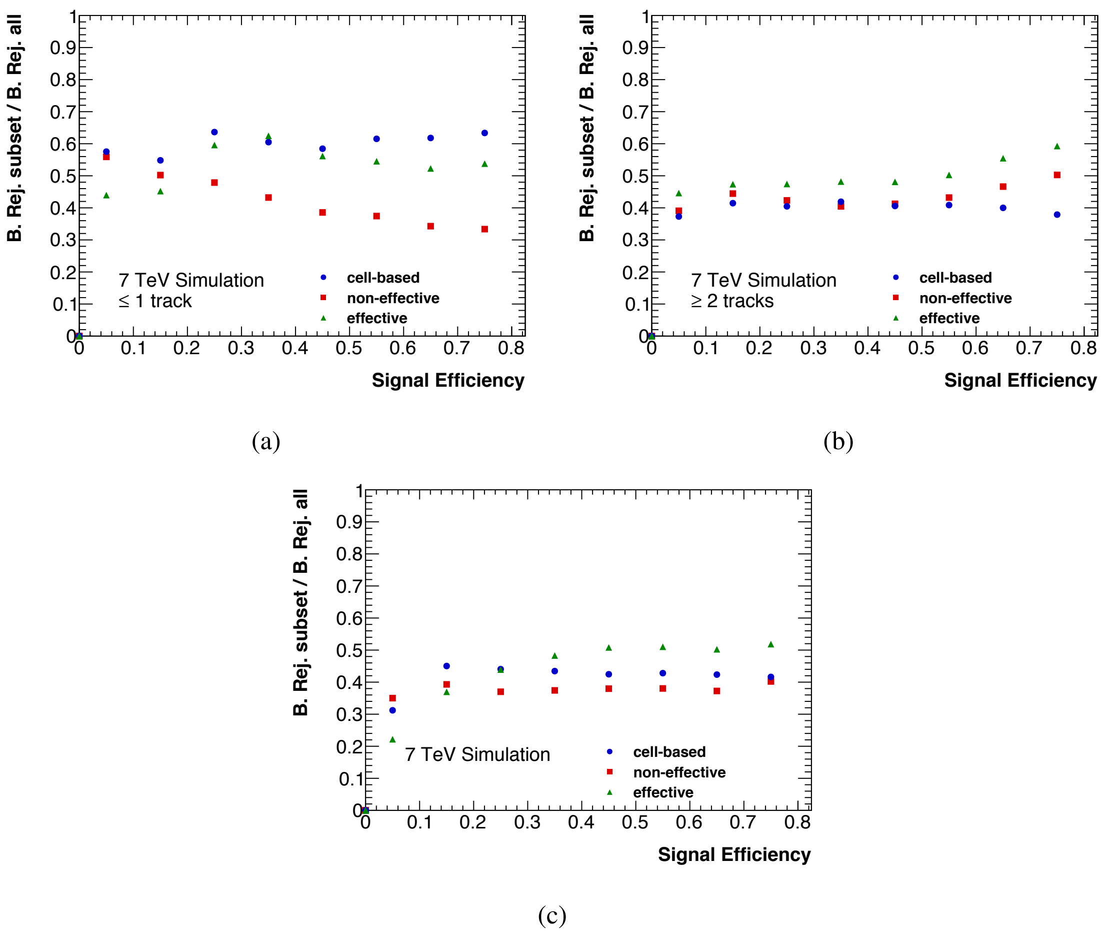


Figure 6.14: Ratio of background rejection for a subset of variable to the complete set of variables (cell-based and 4-momenta based) with no pileup conditions. (a) 1 prong candidates only, (b) 3 prong candidates only, (c) all candidates. Statistical error bars on the background rejection are plotted. Above a signal efficiency of 0.8, they become larger than the difference between the points.

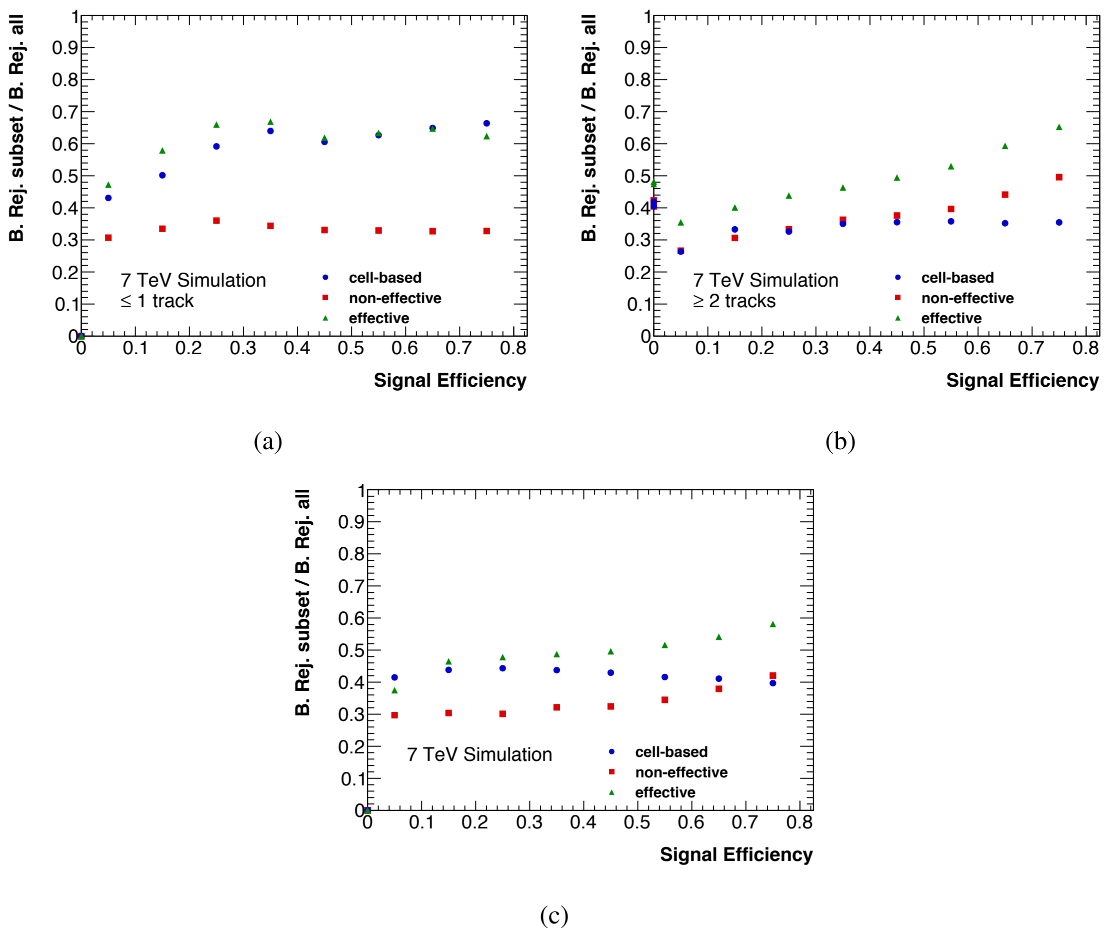


Figure 6.15: Ratio of background rejection for a subset of variable to the complete set of variables (cell-based and 4-momenta based) with pileup conditions. (a) 1 prong candidates only, (b) 3 prong candidates only, (c) all candidates. Statistical error bars on the background rejection are plotted. Above a signal efficiency of 0.8, they become larger than the difference between the points.

Table 6.13: Variable rankings for 4-momentum variables and original variables, pileup vs. no pileup

Rank	≤ 1 track		≥ 2 tracks		all	
	no pileup	pileup	no pileup	pileup	no pileup	pileup
1	R_{EM}	m_{all}	m_{all}	m_{all}	m_{all}	m_{all}
2	m_{all}	f_{core}	R_{EM}	f_{core}	f_{core}	f_{core}
3	f_{core}	N_{eff}	f_{core}	R_{EM}	R_{EM}	N_{eff}
4	m_{eff}	R_{EM}	N_{eff}	N_{eff}	$\langle R\rangle_{\mathrm{eff}}$	R_{EM}
5	f_{ring}	m_{eff}	$\langle R\rangle_{\mathrm{eff}}$	$\langle R\rangle_{\mathrm{eff}}$	m_{eff}	$\langle R\rangle_{\mathrm{eff}}$
6	$\langle R\rangle_{\mathrm{all}}$	R_{Had}	f_{ring}	m_{eff}	N_{eff}	f_{ring}
7	$\langle R\rangle_{\mathrm{eff}}$	$\langle R\rangle_{\mathrm{eff}}$	$\langle R\rangle_{\mathrm{all}}$	N_{strip}	f_{ring}	m_{eff}
8	N_{eff}	f_{ring}	N_{strip}	f_{ring}	$\langle R\rangle_{\mathrm{all}}$	N_{strip}
9	N_{all}	$\langle R\rangle_{\mathrm{all}}$	W_{strip}	N_{all}	N_{all}	N_{all}
10	N_{strip}	N_{all}	N_{all}	W_{strip}	R_{Had}	$\langle R\rangle_{\mathrm{all}}$
11	W_{strip}	N_{strip}	m_{eff}	$\langle R\rangle_{\mathrm{all}}$	N_{strip}	W_{strip}
12	R_{Had}	W_{strip}	R_{Had}	R_{Had}	W_{strip}	R_{Had}

Chapter 7

Summary, conclusions and future directions

7.1 Summary and conclusions

A robust τ identification at the ATLAS experiment can help significantly in the search for the production of Standard Model and supersymmetric Higgs bosons in the LHC collisions. Since hadronic jets are the main source of background for τ leptons, their rejection must be optimized by considering all sources of discriminating information in the ATLAS detector.

In this thesis, new τ identification variables based on topological clusters in the ATLAS calorimeter are studied. The variables based on the 4-momenta of the topological clusters are shown to possess discriminating power against hadronic jets that is distinct from the information already provided by the cell-based τ identification variables. When they are combined with the cell-based variables in a BDT, they can increase the background rejection by a factor of at least 3 for candidates with E_T above 35GeV. The improvement below 35 GeV is much smaller. The cluster-based variables are therefore especially relevant for the search of massive new particles producing very energetic τ leptons when they decay, such as Higgs bosons.

The cluster moment variables are also capable of adding to the discrimination power of the cell-based variables, but their information is not distinct from that of the 4-momentum variables. Since 4-momentum variables are easier to understand and more powerful, their use is recommended over the use of the cluster moment variables.

The 4-momentum variables are also used to give an estimate of the current status of the performance of τ identification using the calorimeter. They show that a τ identification based on the 4-momentum variables only is a factor between 3 and 9 (depending on the number of associated tracks) away from the background rejection achievable with an ideal calorimeter.

The effective-cluster 4-momentum variables have good behavior against pileup and underlying event, since additional low energy clusters have a limited influence on their calculation. On the other hand, the way that the low energy extra clusters are discarded from the calculation of these variables leaves room for optimization.

The six cluster-based variables have been integrated into the official ATLAS reconstruction software. They will be available in all data formats as of September 2010. m_{all} is already part of the default ATLAS identification.

7.2 Future directions

There are a number of directions that can be investigated to complement and improve the study presented in this thesis. In this section, three of the more promising directions are briefly introduced.

Selection of the effective clusters

As previously discussed, the selection of effective clusters is currently made by simply rounding N_{eff} to the next integer $N_{\text{eff}}^{\text{ceil}}$. This has been shown to be responsible for a threshold effect, which mostly affects $\langle R \rangle_{\text{eff}}$. It seems that this specific selection accepts clusters that don't belong to the τ decay too often. One possible solution is to define $N_{\text{eff}}^{\text{ceil}}$ the following way:

$$N_{\text{ceil}}^{\text{int}} = \text{ceil}\left(N_{\text{eff}} - x\right). \tag{7.1}$$

The parameter x can then be adjusted to set the strength of the threshold on the low energy constituents. However, this will not remove the threshold effect. Another possible approach is to modify the definition of $\langle R \rangle_{\text{eff}}$ and m_{eff} to include all clusters up to $\text{floor}(N_{\text{eff}})$, and then include the next cluster in the energy ordered list using a weight. A possible weight that can be used is $N_{\text{eff}} - \text{floor}(N_{\text{eff}})$.

One should be careful about the purpose of the optimization of the selection of effective clusters. It is possible that the behavior against pileup and underlying event is improved at the cost of some discrimination power. For example, m_{all} is already more powerful than m_{eff} because it includes more clusters. Tightening the selection of effective clusters may reduce the discrimination power of m_{eff} but it may reduce its correlation with m_{all} at the same time. All of this may depend on the context in which the variables are used.

Alternative clustering schemes

The topological clustering was designed primarily to collect the energy of showers. However, in test-beam studies where the energy reconstruction of single charged pions in the calorimeter is tested, it was sometimes necessary to sum the energy of many topological clusters [24]. There is also a number of hints in Chapter 5 of this thesis that the number of showers is overestimated by the topological clustering algorithm. It is quite possible that there is a better alternative when the goal is to reconstruct accurately the individual showers.

One possible alternative considered is to study the clustering history of the anti-k_t algorithm. By looking at intermediate steps in this history, one may find that the many clusters from a single shower are put back together. Other alternatives include adjusting the parameters of the topological clustering algorithm and trying different clustering procedures.

In the course of this study, some of these alternatives have quickly been tried without seeing any significant improvement of the discrimination power of the cluster-based variables. The effect of changing the thresholds in the cluster making of topological clustering only has a very small impact on cluster configuration. No improvement could be achieved either by tightening or loosening the definition of local maxima in the cluster splitting.

The main difficulty in this kind of study comes from the fact that the individual showers considered have a very busy neighbourhood, which is different for τ decays and hadronic jets. It is possible that one alternative clustering procedure reconstructs accurately the showers from a τ decay, while it fails at reconstructing the showers of a jet. In such a case, it is unclear if the discrimination power improves.

Using tracks

The 4-momenta variables can benefit from a more accurate reconstruction of the individual decay products. One of the things the inner detector does very well is to reconstruct charged particles. If an accurate way of recognizing neutral pion clusters in the calorimeter is devised, one can use these clusters in conjunction with the tracks from the inner detector to calculate the 4-momenta variables. Such variables are potentially very powerful and they may even bring some new discriminating power.

Bibliography

[1] David Griffiths. *Introduction to Elementary Particles*. Wiley-VCH, 2008.

[2] Claude Amsler et al. Review of particle physics. *Phys. Lett.*, B667:1, 2008.

[3] Louis Marchildon. *Quantum Mechanics: From Basic Principles to Numerical Methods and Applications*. Springer, 2000.

[4] F. Mandl and Shaw. G. *Quantum Field Theory*. John Wiley & Sons, 1984.

[5] M. Roos. Dark Matter: The evidence from astronomy, astrophysics and cosmology. *ArXiv e-prints*, January 2010.

[6] S. P. Martin. A Supersymmetry Primer. In G. L. Kane, editor, *Perspectives on Supersymmetry*, 1998.

[7] A. Kaczmarska. Tau leptons as a probe for new physics at lhc. Technical Report CERN-ATL-PHYS-CONF-2006-015. ATL-COM-PHYS-2006-078, CERN, Geneva, Nov 2006.

[8] G. Aad et al. *Expected performance of the ATLAS experiment: detector, trigger and physics*. CERN, Geneva, 2009.

[9] ALEPH Collaboration, DELPHI Collaboration, L3 Collaboration, OPAL Collaboration, and The LEP Working Group For Higgs Boson Searches. Search for the Standard Model Higgs boson at LEP. *Physics Letters B*, 565:61–75, July 2003.

[10] Elias Coniavitis. Search for the higgs boson at the atlas experiment. Technical Report ATL-PHYS-PROC-2010-037. ATL-COM-PHYS-2010-388, CERN, Geneva, Jun 2010.

[11] M. Benedikt. *LHC Design Report - The Injector Chain - Introduction*. CERN, November 2004.

[12] CERN. Atlas detector and physics performance: Technical design report, 1. Technical report, CERN, Geneva, 1999. CERN-LHCC-99-14.

[13] D. Adams et al. The atlas computing model. oai:cds.cern.ch:811058. Technical Report CERN-ATL-COM-SOFT-2004-009, CERN, Geneva, Dec 2004.

[14] J. M. Campbell, J. W. Huston, and W. J. Stirling. Hard interactions of quarks and gluons: a primer for lhc physics. *Reports on Progress in Physics*, 70(1), 2007.

[15] B. Andersson, G. Gustafson, G. Ingleman, and T. Sjorstrand. Parton fragmentation and string dynamics. *Physics Reports*, 97(2-3), 1983.

[16] Torbjorn Sjostrand, Stephen Mrenna, and Peter Z. Skands. PYTHIA 6.4 Physics and Manual. *JHEP*, 05:026, 2006.

[17] W. R. Leo. *Techniques for Nuclear and Particle Physics Experiments*. Springer-Verlag, 1994.

[18] R. C. Fernow. *Introduction to experimental particle physics*. Cambridge University Press, 1986.

[19] E. Abat et al. Response and shower topology of 2 to 180 gev pions measured with the atlas barrel calorimeter at the cern test-beam and comparison to monte carlo simulations. Technical Report ATL-CAL-PUB-2010-001. ATL-COM-CAL-2010-004, CERN, Geneva, May 2010.

[20] T. A. Gabriel and W. Schmidt. Calculated Performance of Iron-Argon and Iron-Plastic Calorimeters for Incident Hadrons with Energies of 5-GeV to 75-GeV. *Nucl. Instr. Meth.*, 134:271, 1976.

[21] G. Mavromanolakis and D. Ward. Comparisons of hadronic shower packages. *ArXiv Physics e-prints*, September 2004.

[22] CERN. *ATLAS calorimeter performance: Technical Design Report*. Technical Design Report ATLAS. CERN, Geneva, 1996.

[23] W. Lampl, S. Laplace, D. Lelas, P. Loch, H. Ma, S. Menke, S. Rajagopalan, D. Rousseau, S. Snyder, and G. Unal. Calorimeter clustering algorithms: Description and performance. Technical Report ATL-LARG-PUB-2008-002. ATL-COM-LARG-2008-003, CERN, Geneva, Apr 2008.

[24] T. Carli and P. Speckmayer. Study of energy reconstruction algorithms for pions from 2 to 180 gev measured with the atlas barrel calorimeter at the cern sps test-beam. Technical Report ATL-COM-CAL-2010-001, CERN, Geneva, Jan 2010. This note contains material from ATL-COM-CAL-2009-004. The referee C. Roda and M. Aleksa had requested to separate the material of ATL-COM-CAL-2009-004 in two note.

[25] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. The anti-kt jet clustering algorithm. *JHEP*, 04:063, 2008.

[26] Trevor Hastie, Robert Tibshirani, and Jerome Friedman. *The Elements of Statistical Learning: Data Mining, Inference and Prediction*. Springer, 2009.

[27] A. Hoecker et al. TMVA - Toolkit for Multivariate Data Analysis. *ArXiv Physics e-prints*, 2007.