End-cap calorimeter performance

and

identification of the t-channel single top quark process with the ATLAS detector

by

Claudiu D. Cojocaru

A thesis submitted to the
Faculty of Graduate Studies and Research
in partial fulfillment of the requirements
for the degree of

Doctor of Philosophy

Department of Physics
Carleton University
Ottawa-Carleton Institute of Physics
Ottawa, Canada

March 2008

Your file Votre référence
ISBN: 978-0-494-40514-7
Our file Notre référence
ISBN: 978-0-494-40514-7

The undersigned recommend to
the Faculty of Graduate Studies and Research
acceptance of the thesis

"End-cap calorimeter performance and
identification of the t-channel single top quark process with the ATLAS detector"

submitted by Claudiu D. Cojocaru
in partial fulfillment of the requirements
for the degree of Doctor of Philosophy

Chair, Department of Physics

Thesis Supervisor

Carleton University

March 2008

Abstract

The LHC collider will provide proton-proton collisions with 14 TeV centre of mass energy and an expected peak luminosity of $10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$. ATLAS is one of the multipurpose detectors that will be used for particles detection and measurement of properties. The first part of this thesis focuses on the study of the response of the ATLAS electromagnetic and hadronic end-cap calorimeters (EMEC and HEC, respectively) in a beam test performed in the summer of 2002. For the EMEC, the dependence of the measured signal versus the beam energy was found to be linear and an electromagnetic conversion constant $\alpha_{em}^{EMEC} = (0.446 \pm 0.009)\,\mathrm{MeV/nA}$ was calculated. The energy resolution for the EMEC was $\frac{\sigma'(E_{reco})}{E_{reco}} = \frac{(11.8 \pm 0.2)\,\% \sqrt{\mathrm{GeV}}}{\sqrt{E_{reco}}} \oplus (0.4 \pm 0.1)\,\%$, while for the HEC it was $\frac{\sigma'(E_{reco})}{E_{reco}} = \frac{(87.7 \pm 0.6)\,\% \sqrt{\mathrm{GeV}}}{\sqrt{E_{reco}}} \oplus (3.0 \pm 0.2)\,\%$, where the reconstructed energy, E_{reco}, is in GeV. These results feed back into the tuning of the calorimeter Monte Carlo simulations. The second part of this thesis is on the analysis of the electroweak production of the top quark at the LHC, particularly the so-called single top t-channel, by investigating the possible isolation of signal events from the diverse background present at LHC energies. Kinematic characteristics of relevant physics objects were studied for their use as discriminant variables. After identifying these discriminants, a cut flow analysis was performed on both signal and backgrounds. It was found that the selection efficiency, $\epsilon\,(\%)$, for the t-channel was $\epsilon\,(\%) = 2.86 \pm 2.5\%_{Stat} \pm 2\%_{TopMass} \pm 5\%_{JES} \pm 5\%_{B-tagging} \pm 8\%_{ISR/FSR} \pm 2\%_{PDF}$, for an integrated luminosity of $1\,\mathrm{fb}^{-1}$. For the same luminosity, a signal to background ratio (S/B) of 0.28 and a signal significance ($\mathrm{S}/\sqrt{\mathrm{B}}$) of 21 were obtained. The single top production channels provide an unique opportunity to study directly properties of the top quark. These channels are also sensitive to new physics beyond the Standard Model of particle physics.

Acknowledgements

I would like to thank my supervisor, Dr. Manuella Vincter, for her continuous help, and for providing greatly appreciated guidance throughout the completion of this degree. For all her support, I am very grateful.

To all people in the ATLAS collaboration with which I have interacted over the years, thank you.

I would also like to thank my parents, Elena and Mihai, and my sister, Alina, for their love and support, and for always encouraging me to follow my dreams.

Foremost, I want to thank my wife, Simona, for everything; for keeping me focused, and for helping me all along the way. Our beautiful daughter, Mara, has added a new meaning to my life, and has made everything worthwhile.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 The Standard Model

The elementary particles are the components of matter not known to have substructure. Our current understanding of the elementary particles and their interactions is described by the Standard Model of particle physics. The basic constituents of the Standard Model are depicted in Figure 1.1. In the Standard Model, a widely tested theory, the elementary constituents of matter are grouped in two major classes: the leptons and the quarks (they are fermions - particles with a half integer spin). Each of these classes is composed of three generations of particles. The lepton first generation is represented by the electron, e, and its associated neutrino, ν_e, followed by the muon, μ, with its associated neutrino, ν_μ, and the third generation composed of the tau, τ, and its associated neutrino, ν_τ. In the Standard Model, these elementary particles interact amongst themselves through the exchange of another class of particles, the force carriers (they are bosons - particles with an integer spin). The electromagnetic force existing due to the electrical charge of particles is mediated by

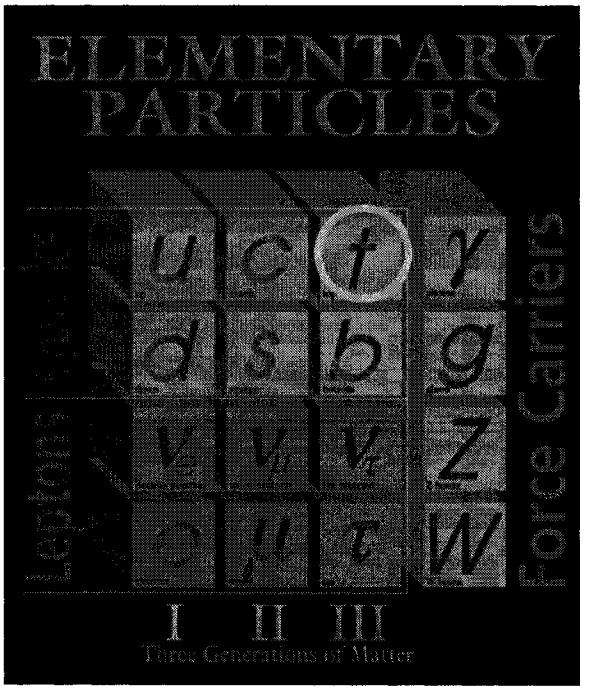


Figure 1.1: The elementary constituents of the Standard Model of particle physics [1].

the photon, γ, the weak force occurring due to the weak isospin is mediated by the W and Z bosons, and the strong force made possible by the color of particles is mediated by the gluons, g. There is also the force of gravity, mediated by the graviton (a hypothetical elementary particle), but it is not part of the Standard Model.

The physical properties of the leptons and quarks in the Standard Model are listed in Figure 1.2, while those of the force mediating bosons are in Figure 1.3. The leptons are either neutral or have a unit of elementary charge, while the quarks have either $+2/3$ or $-1/3$ units of elementary charge. The masses of these particles cover a large range. For leptons, the neutrinos have a mass smaller than 10^{-10} GeV while the heaviest of them, the τ lepton, reaches 1.777 GeV. For quarks, the lightest one, the up quark, has a mass of 0.002 GeV, while the heaviest one, the top quark, reaches about 171 GeV. The bosons are either massless (γ and g - which are also neutral) or

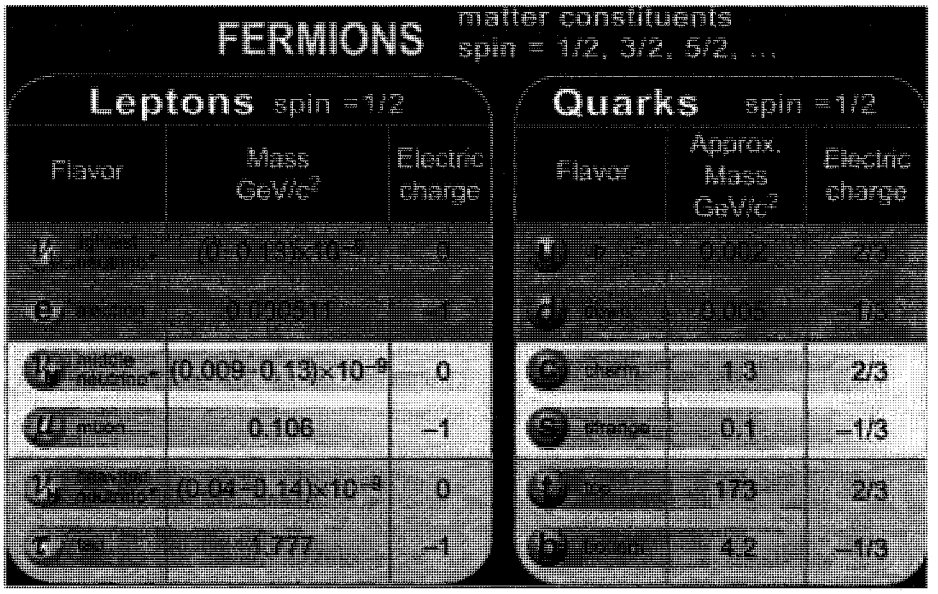


Figure 1.2: The physical properties of the leptons and quarks in the Standard Model of particle physics [1]. * currently allowed mass ranges.

they have masses close to 100 GeV (80.39 GeV for W's - charged; or 91.188 GeV for Z - neutral).

The top quark is the heaviest of all known elementary particles, about 40 times heavier than the next heaviest quark. This thesis will focus on the study of this special quark, and in particular will look at the production of the top quark as a single particle (unaccompanied by an anti-top quark) using the ATLAS detector at the Large Hadron Collider (LHC) of the European Centre for Nuclear Research (CERN), near Geneva, at the Swiss-French border.

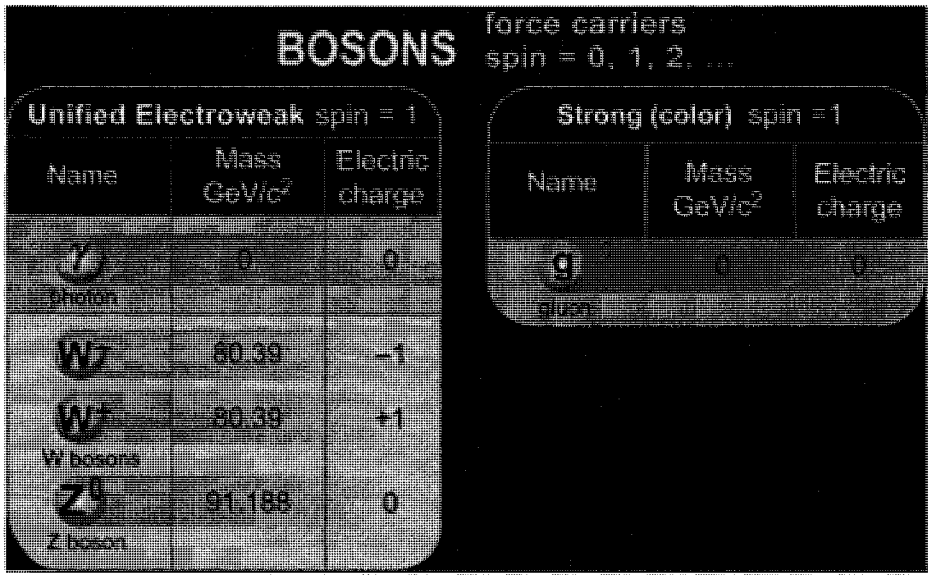


Figure 1.3: The physical properties of the force mediating bosons in the Standard Model of particle physics [1].

1.2 The top quark

1.2.1 Introduction

The top quark is one of the elementary particles of the Standard Model of particle physics. It was simultaneously discovered in 1995 by two groups, CDF [2] and D0 [3], operating at the Tevatron (a particle accelerator at the Fermi National Accelerator Laboratory in Batavia, Illinois, USA). The first measurements of the mass of the top were found to be $176 \pm 8\,(stat) \pm 10\,(syst)$ GeV (by CDF) and $199^{+19}_{-21}\,(stat) \pm 22\,(syst)$ GeV (by D0). With new data being recorded and more precise analyses being performed, the value of the top mass has slightly fluctuated in the last 12 years. The current world average value for the top mass is $170.9 \pm 1.1\,(stat) \pm 1.5\,(syst)$ GeV [4]. This value has been obtained by combining the published results

from Run-I (1992-1996) with the most recent preliminary and published Run-II (2001-present) measurements using up to $1\,\mathrm{fb}^{-1}$ of data (see Figure 1.4). Due to the limited number of top quarks produced in the events (statistics of a few hundred events containing the top quark - an event represents the collision of two particles), the properties of the top quark (the charge conjugates of particles are implicitly included everywhere in this thesis) are not yet very precisely determined. For example, the upper limit on the top width is estimated at 12.7 GeV at 95 % confidence level [7] with $1.0\,\mathrm{fb}^{-1}$ of Tevatron Run-II data, for a top mass of 175.0 GeV.

The precise determination of the top quark mass is very important for the restrictions it can put on the mass of the Higgs boson (introduced by the Higgs mechanism which is capable of giving mass to particles) through its role in the electroweak radiative corrections. Due to its short life time, the top quark decays before hadronizing (forming hadrons due to the color confinement) [6], providing a unique opportunity to measure the properties of a quark without interference from hadronization effects. Thus its spin properties are transmitted unchanged to its decay products. In the Standard Model, the top quark should have a fractional electric charge of $+2/3\,e$, where e is the absolute value of the electron charge. Previous results from Tevatron have confirmed this value to a level of 92 % confidence level (the hypothesis that only a $\pm 4/3\,e$ is produced has been excluded at the 92 % confidence level) [8]. The top quark constitutes a major background for other physics channels such as the Higgs and supersymmetric (SUSY) particle production. Also, the top quark appears in the decay products of processes beyond the Standard Model.

The properties of the top quark are studied by observing its decay products. In the Standard Model, the top decays about 100 % of the time into a W boson and a b-quark. The b-quark will hadronize and produce a spray of particles in a narrow cone

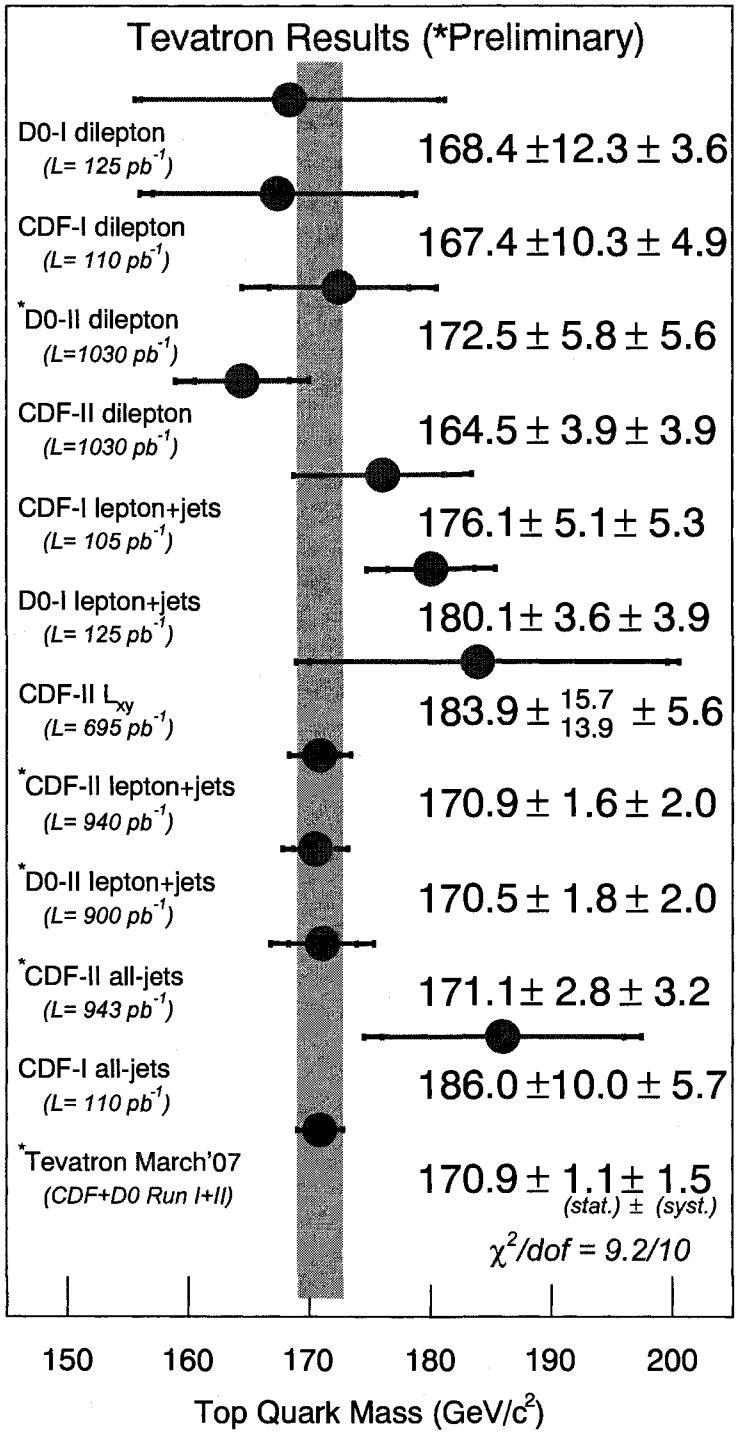


Figure 1.4: The current world average for the top mass using published and preliminary results from CDF and D0 collaborations at the Tevatron as of March 2007 [5].

(called a jet) in the direction of the initial b-quark. The W boson decays leptonically into a lepton (e, μ, τ) and its corresponding neutrino (ν_e, ν_μ, ν_τ) or hadronically into two quarks, which hadronize into jets. The branching ratios of the W boson decays are given in Table 1.1 (Particle Data Group values [25]).

W $\rightarrow$	$e\nu_e$	$\mu\nu_\mu$	$\tau\nu_\tau$	hadrons
BR (%)	10.75 $\pm$ 0.13	10.57 $\pm$ 0.15	11.25 $\pm$ 0.20	67.60 $\pm$ 0.27

Table 1.1: Branching ratios for the decays of the W boson (Particle Data Group values [25]).

At the Tevatron, protons collide with anti-protons at a centre of mass energy of 1.98 TeV, and so enough energy is available to produce the top quark. Given the available luminosity, a few hundred top quark events have so far been observed. At the LHC (a proton-proton collider), centre of mass energies of 14 TeV will be available, with an expected peak design luminosity of $10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$. Under these conditions, about 11 million events containing top quarks are expected for $10\,\mathrm{fb}^{-1}$ integrated luminosity (about one year at low luminosity of $10^{33}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$). This large available statistic will allow for the precise determination of the Standard Model properties of the top quark. For example, the expected precision of the top quark mass with ATLAS is at the 1 GeV level [9].

1.2.2 Top-anti-top quark pair production

Top quarks can be produced through two types of interactions: strong and electroweak. When two gluons or sea quarks belonging to the colliding protons interact

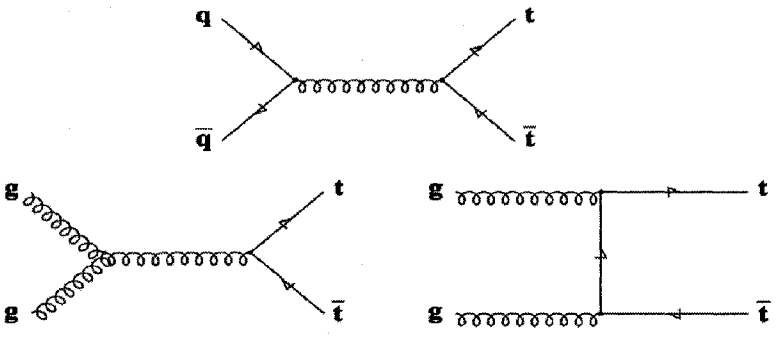


Figure 1.5: Leading order Feynman diagrams for the production of top anti-top quark pairs at proton colliders.

through the strong interaction, a top-anti-top quark ($t\bar{t}$) pair can be created. The leading order (LO) Feynman diagram of the process is shown in Figure 1.5. The $t\bar{t}$ pair production is the dominant mechanism to produce top quarks at both the Tevatron and the LHC.

At the Tevatron, the quark-anti-quark ($q\bar{q}$) production mechanism accounts for about 85 % of the $t\bar{t}$ production, while the gluon-gluon (gg) diagrams contribute the remaining 15 %. The Standard Model theoretical prediction for the top-anti-top quark pair production at the Tevatron is $6.8 \pm 0.6\,\text{pb}$ [17]. The first measurement of the cross section was $6.8^{+3.6}_{-2.4}\,\text{pb}$ by CDF [2] and $6.4 \pm 2.2\,\text{pb}$ by D0 [3]. The most recent studies show a cross section around 7 pb (averaged over various analysis channels) for both CDF [10] and D0 [11].

The predicted cross sections at hadron colliders are shown in Figure 1.6. As can be seen, the $t\bar{t}$ production cross section at the LHC is about 100 times higher than at the Tevatron. At the LHC, the gluon-gluon production mechanism becomes dominant contributing about 90 % to the total $t\bar{t}$ cross section, while the quark-anti-

quark production mechanism accounts for about 10 %. The predicted cross section for the $t\bar{t}$ process is about 850 pb (NNLL - next to next to leading logarithm - calculations predict 872.8 $\pm$ 15.0 pb [17], where the error is based on theoretical calculations). For an integrated luminosity of 10 fb^{-1}, there will be about 8.5 million $t\bar{t}$ events produced.

1.2.3 Single top quark production

Although the $t\bar{t}$ production channel is dominant at both Tevatron and LHC, there is also the so-called "single top" production mechanism. This process is mediated by the electroweak interaction. Based on the exchanged particle involved in the interaction and on the final products, one distinguishes three single top channels: the t-channel (or the W-g-channel), the s-channel (or the W*-channel), and the W+t-channel (or the associated channel). The leading order Feynman diagrams for these channels are presented in Figure 1.7. The single top quark channels offer the opportunity to directly measure the Cabibbo-Kobayashi-Maskawa (CKM) matrix element V_{tb} of the $W - t - b$ vertex. These channels are also sensitive to new physics beyond the Standard Model [22]. For example, an additional heavy W' boson will increase the cross section of the the s-channel, while anomalous flavor changing neutral current couplings or an anomalous chromo-magnetic moment in the top-gluon vertex can modify the t-channel cross section.

At the Tevatron, the Standard Model next-to-leading-order (NLO) predicted cross sections are 1.98 $\pm$ 0.25 pb for the t-channel, 0.88 $\pm$ 0.11 pb for the s-channel, and 0.12 $\pm$ 0.05 pb for the W+t-channel [14]. First evidence for this electroweak production was presented last year by the D0 collaboration [14]. After the analysis of 0.9 fb of collected data, a combined t-channel and s-channel single top cross section

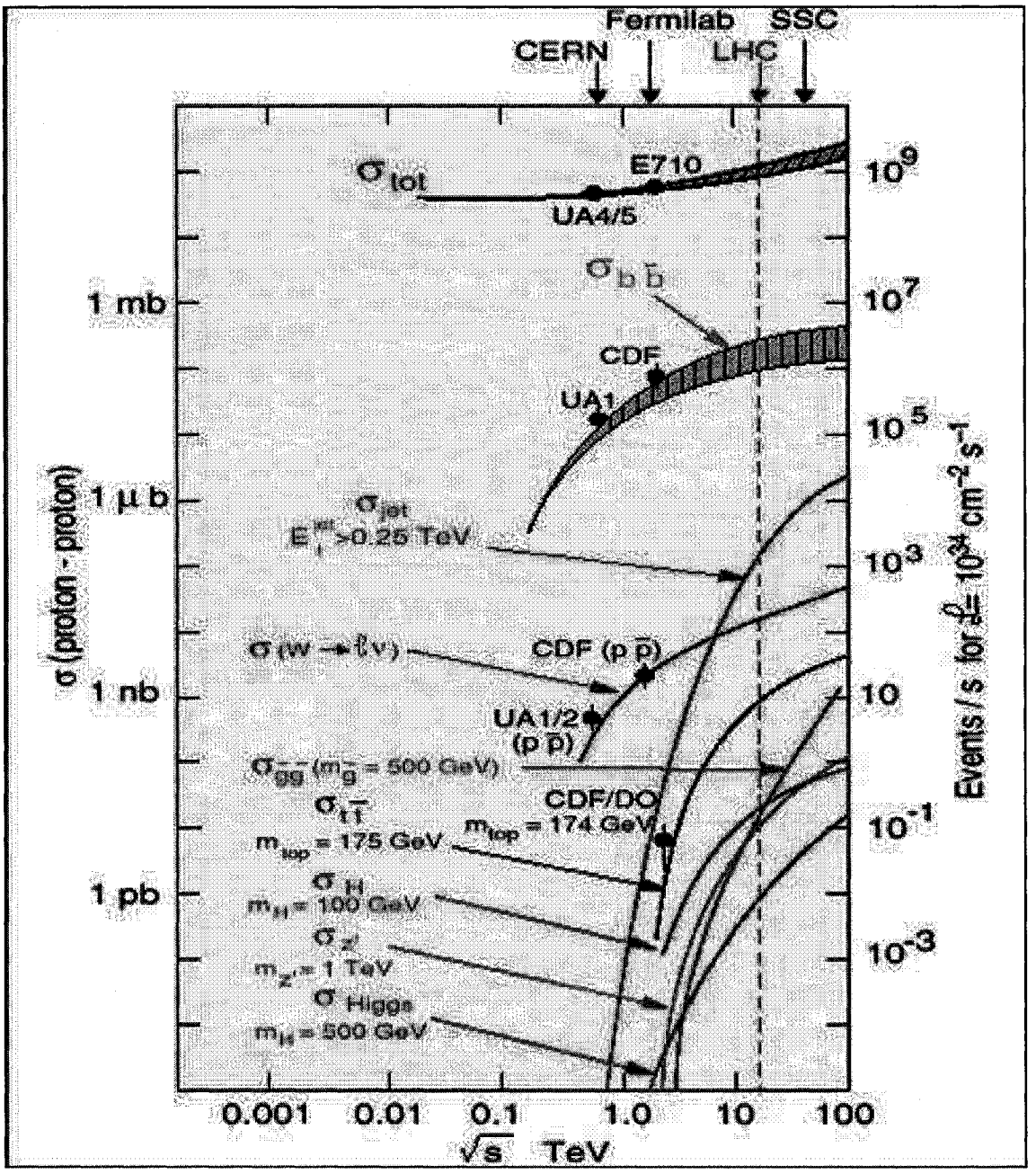


Figure 1.6: Predicted Standard Model cross sections and event rates at hadron colliders [12].

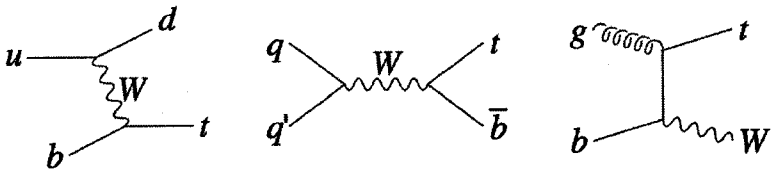


Figure 1.7: Leading order Feynman diagrams for the production of single top quark at proton colliders [15] (left: t-channel; middle: s-channel; right: W+t-channel).

	t-channel	W+t	s-channel
NLO cross section (pb)	$246.6^{+9.3}_{-10.2}$ [15]	66 ± 2 [16]	$10.65^{+0.65}_{-0.64}$ [15]
Total events ($10\,\text{fb}^{-1}$)	$\sim 2.466 \times 10^6$	$\sim 0.66 \times 10^6$	$\sim 0.1065 \times 10^6$

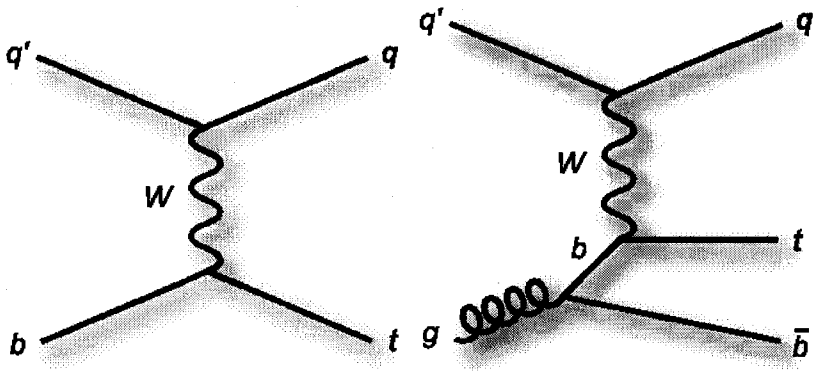
Table 1.2: Production cross sections for the single top quark channels at the LHC and the estimated number of events for $10\,\text{fb}^{-1}$ of integrated luminosity.

of $4.9 \pm 1.4\,\text{pb}$ was reported, using about 70 events. This cross section measurement allowed the first direct calculation of the V_{tb}: $0.68 < |V_{tb}| < 1$ at 95 % confidence level. The CDF group also measured a combined cross section of $2.7^{+1.5}_{-1.3}\,\text{pb}$ [13].

The single top quark cross sections at LHC energies and the number of events expected for $10\,\text{fb}^{-1}$ are summarized in Table 1.2. A total number of about 3 million events are expected to be produced with this integrated luminosity.

1.2.4 Single top t-channel

At the LHC, the highest production rate among the single top channels is given by the W-g or t-channel [22]. Two main diagrams (Figure 1.8) contribute to the total cross section for this channel: the leading order, involving two vertices ($2 \rightarrow 2$ process), and the next-to-leading order, involving three vertices ($2 \rightarrow 3$ process). In the LO



(a) (b)

Figure 1.8: Leading order (a) and next-to-leading order (b) Feynman diagrams for the t-channel single top quark production at proton colliders.

diagram ($q'b \rightarrow qt$), a light quark from the sea inside one proton interacts with a b-quark from the other proton. A W boson is exchanged and a light quark and a top quark are produced. For the NLO process ($q'g \rightarrow qt\bar{b}$), a light quark from one proton interacts with a gluon from the other proton through the exchange of a W boson and a b-quark (from the gluon that splits into a $b\bar{b}$ pair). A light quark, a top quark and an anti-b quark are produced.

In both diagrams, the top quark decays into a W boson and a b-quark, with the W boson decaying into a lepton and the corresponding neutrino or two jets (see Table 1.1), and the b-quark hadronizing into a b-jet. Due to the large number of background jets (particularly quantum chromodynamics, QCD, multi-jets) and the difficulties of precisely calibrating them, the hadronic W boson decay channel is not considered in this study. Thus, in the final LO process configuration with the W boson decaying leptonically ($q'b \rightarrow qt \rightarrow q(Wb) \rightarrow q(l\nu_l b)$), there are: a lepton, a neutrino, and at least two jets (one light jet and one b-jet). For the NLO process ($q'g \rightarrow qt\bar{b} \rightarrow q(Wb)\bar{b} \rightarrow q(l\nu_l b)\bar{b}$), there is an additional b-quark present in the

final state, increasing the number of jets in the final state to at least three (one light jet and two b-jets).

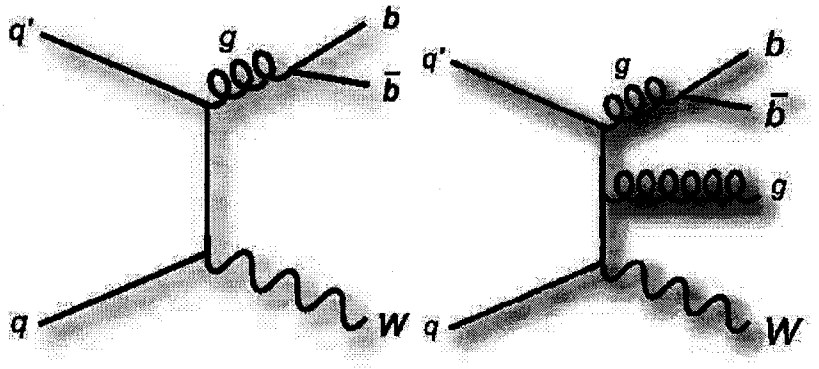
The decay products for this channel (leptons, neutrinos - missing energy, jets and b-jets) are common to other physics channels. One needs to use specific topological and kinematic features of the t-channel in order to isolate it from the background events present in the LHC environment.

1.2.5 Backgrounds for the t-channel

There are various physics processes for which the final states mimic that of the single top t-channel. The first category of backgrounds is represented by the other two single top channels: the W+t-channel and the s-channel. The second category is top-anti-top pair production and the third one is non-top background [22].

The final state for the W+t-channel includes a top quark and a W boson. The top quark decays into a W boson and a b-quark. When one of the W bosons decays leptonically and the other one hadronically ($gb \rightarrow tW \rightarrow (Wb)W \rightarrow (l\nu_l b)(qq')$), in the final state a lepton, a neutrino and at least three jets (two light jets and a b-jet) are present. For the single top s-channel, the two final particles are a top quark and a b-quark. The top decays into a W boson and a b-quark, with the W boson decaying into a lepton and the corresponding neutrino and the b-quark hadronizing into a b-jet ($qq' \rightarrow t\bar{b} \rightarrow (Wb)\bar{b} \rightarrow (l\nu_l b)\bar{b}$). Thus, one looks for the presence of a lepton, a neutrino, and at least two jets (both being b-jets). The various single top channels are therefore difficult to distinguish from each other as they differ mostly through jet multiplicity and light-jet/b-jet identification.

The number of $t\bar{t}$ events produced at the LHC will be about 3.4 times bigger



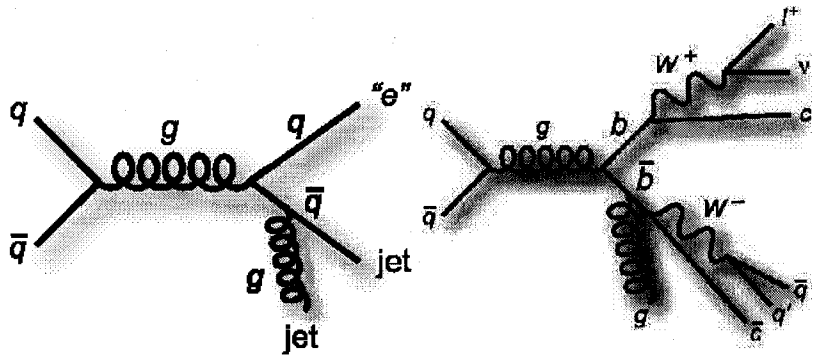
(a) (b)

Figure 1.9: Examples of two Feynman diagrams for the production of the W+jets background: (a) $W + b\bar{b}$ (b) $W + b\bar{b}g$.

than the number of t-channel events. With an event topology very similar to the t-channel, the $t\bar{t}$ background will be very challenging to eliminate. With one of the W boson decaying leptonically and the other one hadronically (gg or $q\bar{q} \rightarrow t\bar{t} \rightarrow (Wb)(W\bar{b}) \rightarrow (l\nu_l b)(qq'\bar{b})$), there are one lepton, one neutrino and at least four jets in the final state (two b-jets and two light jets).

The electroweak W boson production in association with jets (W+jets) has at LHC energies a Monte Carlo (MC) estimated cross section of about 36×10^3 pb. Due to the presence of the W boson and jets in the final state, this background is very difficult to isolate from the signal. Two examples of Feynman diagrams for the production of W+jets are shown in Figure 1.9. The associated jets are usually divided in two categories: light jets (u, d, s and gluon) and heavy jets (c and b).

The process in which a $b\bar{b}$ pair is produced at the LHC has a very high cross section of about 0.5 mb. Under these conditions, events with fake reconstructed leptons or wrongly reconstructed missing transverse energy, E_T^{miss}, may pass the selection criteria and contaminate the signal. Two examples of Feynman diagrams for the



(a) (b)

Figure 1.10: Examples of two Feynman diagrams for the production of QCD multi-jet background: (a) 2 jets and a fake electron and (b) QCD $b\bar{b}g$.

production of QCD multi-jets are shown in Figure 1.10.

A summary of the cross sections, expected events for $10\,\mathrm{fb}^{-1}$ integrated luminosity and the final state for all channels are presented in Table 1.3.

1.3 Outline of the thesis

The remainder of the thesis is structured as follows:

Chapter 2

In this chapter, the Large Hadron Collider (LHC) and the ATLAS detector are described. The general characteristics of the CERN proton-proton collider are presented followed by the description of the main component parts of the ATLAS detector and their role in particle detection. The physics principles involved in the interactions of particles with specific components are explained with an emphasis on the liquid

Process	σ_{total} (pb)	Expected events ($10\,\text{fb}^{-1}$)	Final state	$\sigma_{e,\mu,\tau}$(pb)
t-channel	246.6	2,466,000	$(l\nu_l b)q$	81.3
s-channel	10.65	106,500	$(l\nu_l b)\bar{b}$	3.3
W+t	66	660,000	$(l\nu_l b)(qq')$	26.7
$t\bar{t}$	833	8,330,000	$(l\nu_l b)(qq'\bar{b})$	461
W+jets	30,500	30.5×10^7	$(l\nu_l) + jets$	9,820
QCD $b\bar{b}$	$\sim 10^8$	$\sim 10^{12}$	jets	

Table 1.3: NLO cross sections, expected events for $10\,\text{fb}^{-1}$ integrated luminosity and final state for the signal (t-channel) and the major backgrounds; $\sigma_{e,\mu,\tau} = BR_{e,\mu,\tau} \cdot \sigma_{total}$, where $BR_{e,\mu,\tau}$ is the branching ratio for the leptonic final state (a lepton: e, μ or τ, present in the final state). The W+jets and the QCD $b\bar{b}$ cross sections are estimated with MC generators.

argon calorimetry in the end-cap region.

Chapter 3

The testing of components of the ATLAS detector was performed over a period of a few years. This chapter presents some results of one of these beam tests that took place in the summer of 2002. This study analyzes the interaction of electrons and pions with modules of the end-cap calorimeters by investigating the modules' response to electrons and pions. Calibration constants are derived and the detector energy resolution dependence on the detected energy is obtained for both the electromagnetic and hadronic components.

Chapters 4, 5, 6

The physics potential of the study of the single top production mechanism is investigated in these chapters. In Chapter 4, the Monte Carlo generation of physics events is described, as well as the methods used for the reconstruction of the physics objects necessary for the analysis, such as electrons, muons and jets. In Chapter 5, the reconstruction performance of various physics objects is analyzed for various Monte Carlo data sets. A comparison between two different types of detector response simulations is presented in order to accommodate the coherent use of samples created using the different detector simulations for physics studies. The physics study performed in this thesis is on the single top quark production, particularly on the investigation of the detection of events created through the t-channel. In Chapter 6, a selection strategy for the signal over the backgrounds is developed that allows the measurement of the t-channel cross section. Various kinematic cuts are investigated in order to maximize the signal selection efficiency and minimize the selection of background events. The statistical and systematic errors are estimated for the signal selection efficiency and then propagated into the t-channel cross section uncertainty.

Chapter 7

The conclusions of these studies are presented. The importance of the beam test results for the calibration of the calorimeter response to various types of particles is presented, as well as the analysis potential of the ATLAS detector in the study of the single top quark production, particularly the so-called t-channel.

Appendix A

My participation in the ATLAS experiment is summarized at the end of this thesis.

Appendix B

A list of abbreviations.

Chapter 2

The ATLAS experiment

2.1 LHC

The new generation of particle colliders are accelerating particles to centre of mass energies of more than 1 TeV. At the Large Hadron Collider (LHC) at CERN (the European Centre for Nuclear Research), located near Geneva, Switzerland, proton-proton collisions will take place at a centre of mass energy of 14 TeV, representing the highest man made energy in a particle accelerator. The proton beams will each contain about 10^{14} protons, with a peak design luminosity of $10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ [22].

The various accelerator components of the LHC are shown in Figure 2.1. First, protons with an energy of 92 keV are produced from hydrogen gas in a duoplasmatron. A radio frequency quadrupole is then used for shaping the beam and increasing the energy to 750 keV. Next, the Linac-2 increases the beam energy to 50 MeV. From here the beam is injected into the Proton Synchrotron Booster (PSB) where its energy reaches 1.4 GeV. The Proton Synchrotron (PS) increases this beam's energy to

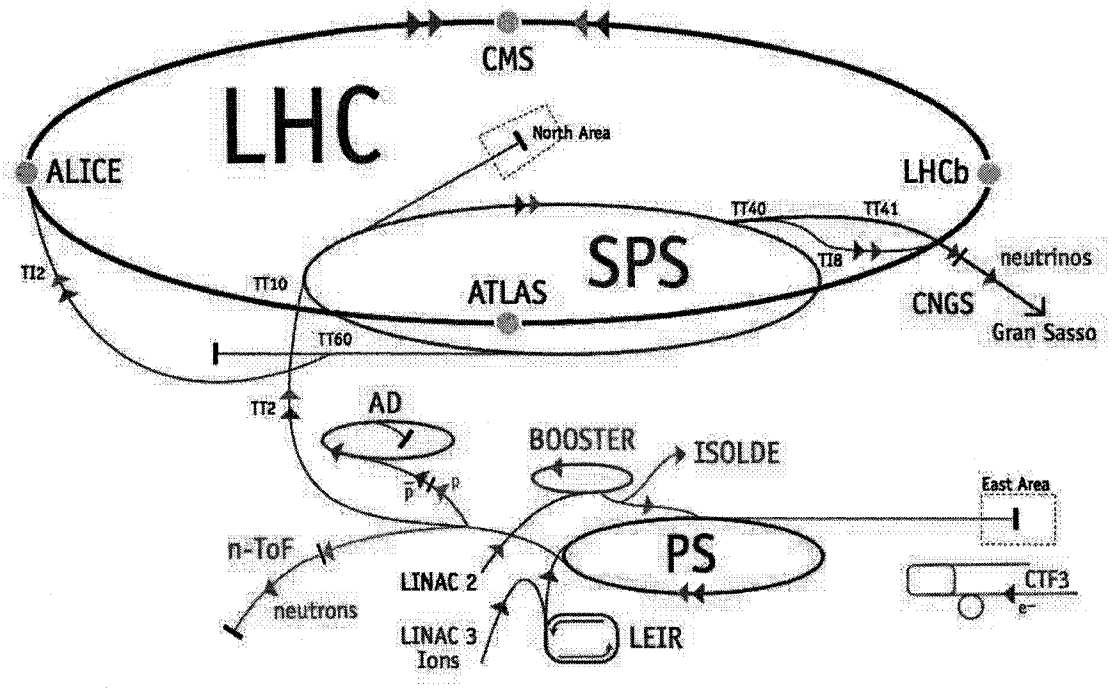


Figure 2.1: The components of the CERN accelerator used in beam production [18]. The beam test experiments took place in the North Area in the H6 beam line of the SPS (see Section 3.1 for details).

28 GeV. The last step is performed by the Super Proton Synchrotron (SPS) that boosts protons up to 450 GeV for LHC injection in general and for the beam tests in particular. The LHC will be able to further increase the protons' energy to 7 TeV.

The two counter rotating proton beams have a bunch structure. These bunches contain about 10^{11} particles/bunch and are separated by 25 ns. The luminosity of the collider is given by:

$$L = \frac{1.07 \times 10^{-4}}{\Delta t} \frac{N^2 E \epsilon}{\beta} \tag{2.1}$$

where Δt is the bunch separation in time, N is the number of protons per bunch, E is the beam energy, ϵ is the emittance (given by the dimension of the region in

space and momentum phase space of the beam particles) and β is the beta function (a property of the external focusing arrangement - proportional to the beam size at the interaction point). A simplified version of expression for the luminosity is:

$$L = f_n \frac{N_1 N_2}{A} \tag{2.2}$$

where f_n is the revolution frequency, $N_{1,2}$ are the number of protons per bunch and A is the cross section of the beam.

Protons are composite particles. When they collide, there is an elastic and an inelastic part to the collision. What follows, refers to the inelastic part. The quarks and the gluons are the constituents of the protons and they each carry a certain fraction of the proton's momentum. These momentum fractions are described by the parton density functions (PDFs). The processes of interest are produced when proton constituents interact through a hard scattering, in which a heavy particle is produced or the resulting partons have a large transverse momentum. The other type of parton interactions are called soft and they are not of interest for these physics studies. The average number of inelastic collisions per bunch crossing, $< n >$, can be estimated as:

$$< n > = \sigma_{inelastic} \times L \times \frac{\Delta t}{\epsilon_{bunch}} \approx 75\,\text{mb} \times 10^{34}\,\text{cm}^{-2}\,\text{s}^{-1} \times 25\,\text{ns}\,/\,0.8 = 23 \tag{2.3}$$

where $\sigma_{inelastic}$ is the inelastic cross section and ϵ_{bunch} is the efficiency of filling the beams with bunches (note that for a fixed collision rate of 750 million particle collisions per second, a reduced "bunch filling efficiency" results in more collisions per bunch). There will be about 23 minimum bias events per bunch crossing. Also, the rest of the proton's constituents (proton remnants, remaining beam fragments) also interact producing a multitude of particles that will mask the processes of interest in the same interaction as the hard scattering, creating the so-called underlying event.

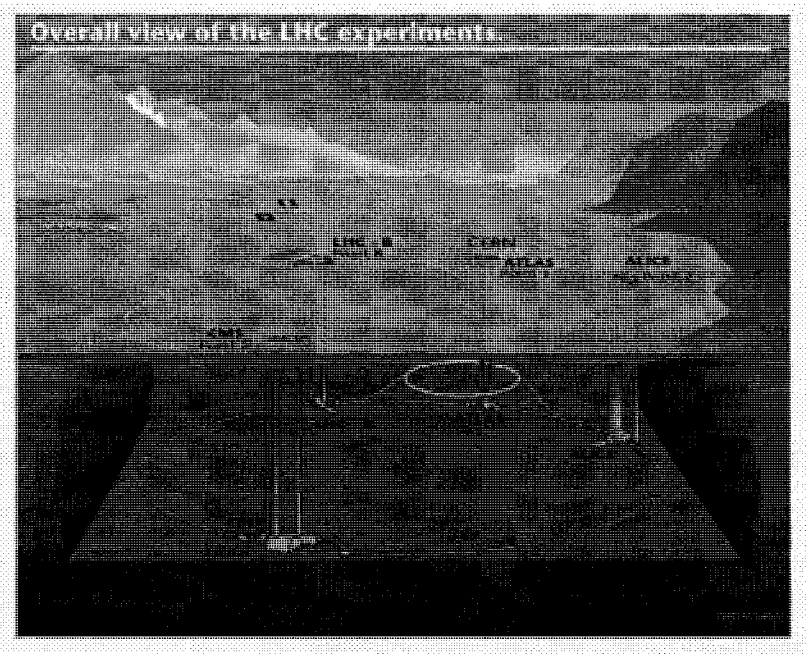


Figure 2.2: The underground LHC tunnel and the locations of the four experiments: ATLAS, CMS, LHC-B and ALICE.

Furthermore, due to the high luminosity, the detectors sample several bunch crossings simultaneously. These effects contribute to the pileup of events in the detector.

The accelerator is situated at the border between France and Switzerland, about 100 m underground, in the 27 km circumference tunnel of the former LEP accelerator. Four particle detectors (ATLAS, CMS, LHC-B and ALICE) are placed at four locations where the two protons beams will be made to collide. The locations of the four detectors are shown in Figure 2.2. The first collisions are expected to take place in the fall of 2008.

2.2 ATLAS detector

One of the four detectors for the LHC is the ATLAS detector (A Toroidal LHC ApparatuS) [21]. The detector has cylindrical symmetry and a weight of 7,000 tons. It has a diameter of 22 m and a length of 46 m, see Figure 2.3. The high luminosity environment at the LHC, the 25 ns bunch crossing and the event pileup demand fast electronics and triggering system, as well as high granularity in order to minimize the detector pileup. The high radiation environment around the interaction point requires the use of radiation hard hardware and electronics, as well as liquid argon as the calorimeter active medium.

The ATLAS coordinates system is defined in the x, y and z coordinates. The x axis points towards the centre of the accelerator ring, the y axis is perpendicular to the ground and pointing up, while the z axis is defined along the beam direction, using a normal right-handed coordinate system. In addition, the polar angle θ (the angle from the beam axis) and the azimuthal angle ϕ (measured around the beam axis) are defined. A special variable is introduced, the pseudorapidity η, defined as $-ln[tan(\theta/2)]$, that is Lorentz invariant. This implies an almost constant physics versus η at a fix p_T for the η range of the experiment. For example, the multiplicities $\frac{\delta n_{charged}}{\delta \eta}$ and $\frac{\delta n_{neutral}}{\delta \eta}$ are both about 7.5 per unit of pseudorapidity $\Delta\eta = 1$. The η coordinate varies from 0, for $\theta = 90\,°$, to ∞ for $\theta = 0\,°$. However, ATLAS provides coverage in the range $|\eta| < 5$ or θ down to $0.8\,°$.

The ATLAS detector, shown in Figure 2.3, has an almost full spatial coverage around the interaction point. Radially, closest to the beam pipe is the inner detector (ID). The next layer is represented by the calorimeters. The outer most layer is the muon spectrometer.

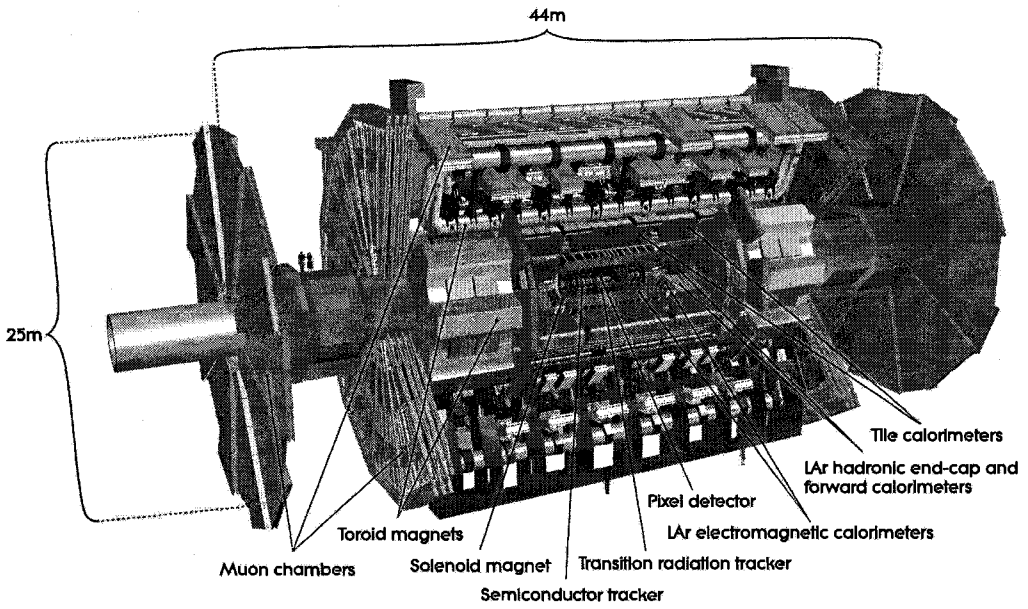


Figure 2.3: The ATLAS detector. Shown are the various components that make up the detector [24].

2.2.1 Inner detector

The inner detector [21] measures charged tracks with very high precision. It has very good momentum resolution (the required resolution is given as: $5 \times 10^{-4} p_T \oplus 1\,\%$, where p_T is in GeV [21]) and pattern recognition capabilities, and can measure accurately the positions of primary and secondary vertices of charged tracks (e, μ, τ, π, K), and it is also used for b-tagging. The inner detector covers the η range, $|\eta| < 2.5$, matching the precision calorimeter coverage, and can measure the momentum of charged tracks with transverse momentum p_T as low as 0.1 GeV. Electrons with energies between 0.5 GeV and 150 GeV can be identified within the η range, $|\eta| < 2.0$.

The inner detector is submerged in a magnetic field, generated by a 2 T solenoid. There are three sub-components making up the inner detector: a pixel detector, a semiconductor tracker and a transition radiation tracker (TRT). An ID

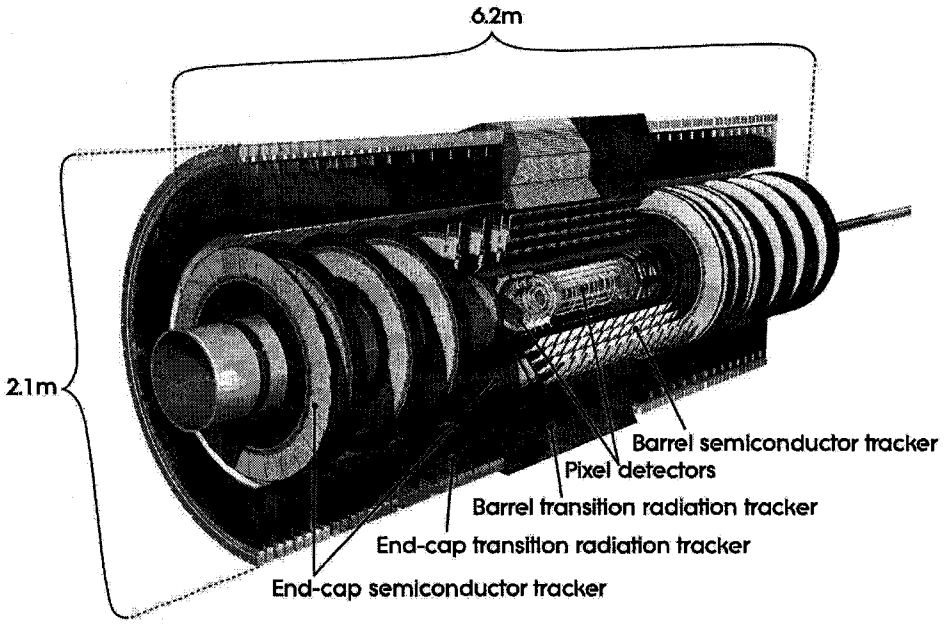


Figure 2.4: The inner detector components: the pixel detector, the semiconductor tracker and the transition radiation tracker [24].

diagram with the components is shown in Figure 2.4.

The pixel detector and the semiconductor tracker are fine-granularity detectors using discrete space-points and pixel and silicon microstrip technologies. The transition radiation tracker enhances the pattern recognition and improves the momentum resolution. It has 73 straw planes in the barrel and 160 straw planes in the end-cap of gaseous straw tube elements, interleaved with transition radiation material and provides continuous tracking (about 35 hits per track). In the range $|\eta| < 2.0$, the electron identification capabilities are enhanced by the detection of transition-radiation photons in the xenon-based gas mixture of the straw tubes.

2.2.2 Calorimeters

Calorimeters [19], [20] are used as part of the ATLAS detector for fast readout for triggering and for measuring the energy of electromagnetic (e, γ) and hadronic (π, jets) objects produced in the LHC, as well as calculating the missing transverse energy (E_T^{miss}). When a particle travels through these detectors it creates a shower of particles. This effect is used to measure the energy of the initial particle that instigated the shower, by sampling it in consecutive layers in the calorimeter. Alternating layers of high Z materials for the electromagnetic calorimeters or high density material for the hadronic calorimeters act as the absorbers used to produce the showers while active layers such as liquid argon or scintillators are the sensitive materials used to measure the deposited energy. A calorimeter that samples the energy of the shower at different points in its development is called a sampling calorimeter. It is characterized by the sampling fraction defined as the ratio between the energy deposited in the active layer and the total energy. All the ATLAS calorimeters are sampling calorimeters.

Due to the different physics involved in the development of the two types of showers, the response of the calorimeters to electromagnetic and hadronic showers is different. For example, for the same initial energy, the response to electrons is higher that the response to pions. These calorimeters are called non-compensating and are characterized by the ratio of the calorimeter response to electrons over the response to pions, e/π. Typical values for the ATLAS calorimeters are $e/\pi = 1.1 - 1.3$ [21].

In order to assess how well the energy is measured in a calorimeter, the energy resolution is introduced as the ratio between the width over the mean of the measured energy distribution, E. The general form for the energy dependence of the energy

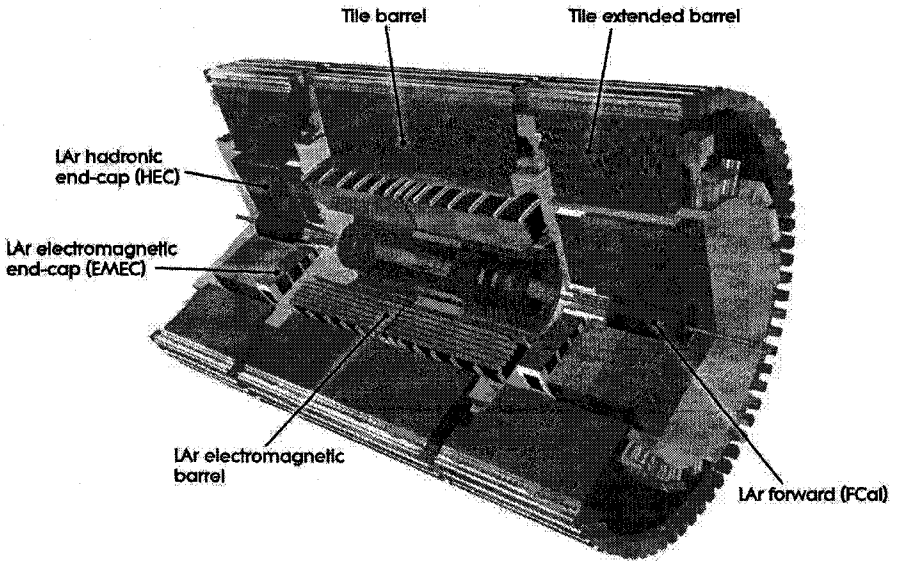


Figure 2.5: The layout of the calorimeter system including the electromagnetic, hadronic (HEC and Tile) and forward calorimeters [24].

resolution is given by the Equation 2.4:

$$\frac{\sigma(E)}{E} = \frac{a}{\sqrt{E}} \oplus b \oplus \frac{c}{E} \tag{2.4}$$

where $\sigma(E)$ is the width of the energy distribution (approximately Gaussian) and the symbol $\oplus$ means addition in quadrature, a is the sampling term (related to variations in sampling fraction and non-compensation), b is the constant term (due to inhomogeneities, losses and non-compensation) and c is the noise term (due to electronic and pile-up noise expected at low and high luminosity).

The three calorimeter systems in ATLAS: the electromagnetic, the hadronic and the forward calorimeters (FCal) are shown in Figure 2.5.

Electromagnetic calorimeters

The electromagnetic calorimeters are specialized in the detection and energy measurement of electrons and photons. Their design energy resolution is given by: $\frac{\sigma(E)}{E} = \frac{10\,\%\sqrt{\text{GeV}}}{\sqrt{E}} \oplus 0.7\,\%$ [21], where E is in GeV. When these particles interact with the material in the detector, so-called electromagnetic showers are produced. A generic drawing of the process is shown in Figure 2.6. For energies greater than about 20 MeV, most of the electron energy in materials such as copper is lost due to bremsstrahlung [25] (the radiation produced by a decelerating charged particle), as can be seen in Figure 2.7. An energetic photon is created and the electron loses some of its energy. Then the photon interacts with the nuclear electric field and an electron positron pair is created. The e^+ and e^- are free to reinteract with the material. This process continues until the electron energy reaches a threshold, E_c (the critical energy, defined as the energy at which the ionization and the bremsstrahlung energy loss are equal - see Figure 2.7), at which point the cascade stops propagating. The critical energy for electrons is about 7 MeV in lead, 20 MeV in copper and 37 MeV in liquid argon [25]. The energy dissipated through this process is proportional to the energy of the particle that instigated the shower.

The shower spreads both longitudinally and transversely. The longitudinal spread, radiation length X_0 (the distance for which the energy drops to 1/e of the initial value due to radiation losses only, in cm), is given by the formula [25]:

$$X_0 \simeq 180\,\text{cm}^4/\text{g}\,\frac{A}{Z^2}\,\rho \tag{2.5}$$

in which A is the atomic weight, Z is the atomic number and ρ is the material density (in g/cm^3). The transverse spread, Moliere radius ρ_m (defined such as 95 % of the

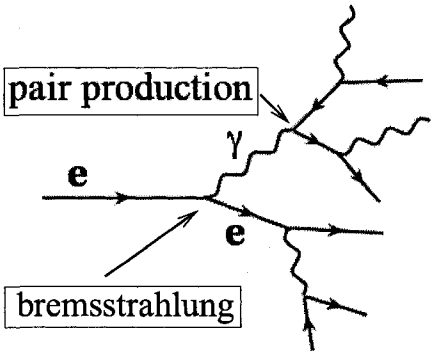


Figure 2.6: The progress of the electromagnetic shower: electrons - straight lines; bremsstrahlung radiation - wavey lines.

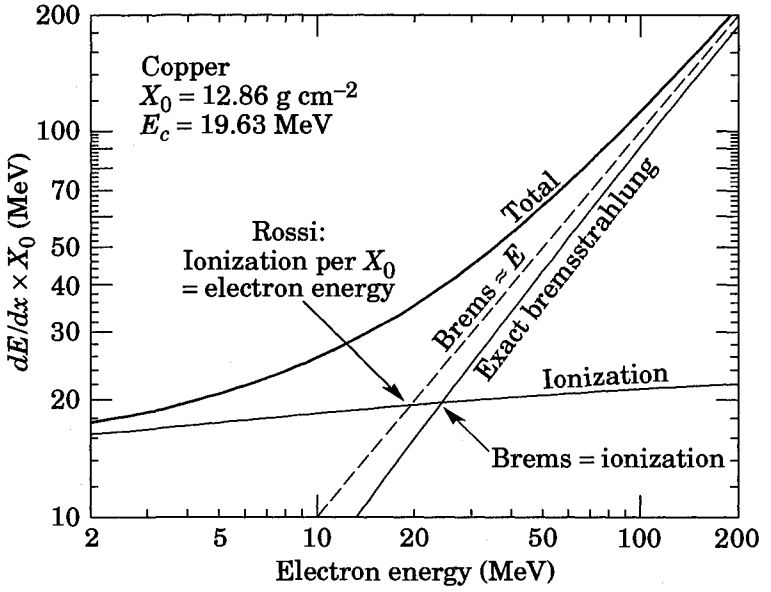


Figure 2.7: The energy lost by an energetic electron per radiation length in copper [25].

shower energy is contained in a $2\,\rho_m$ radius cylinder), is described by the formula:

$$\rho_m = \frac{21\,\text{MeV}}{E_c} X_0 \tag{2.6}$$

where E_c is the critical energy and X_0 is the radiation length previously defined. Some values for X_0 and ρ_m are respectively 6 mm and 16 mm for lead, 14 mm and 15 mm for copper and 140 mm and 80 mm for liquid argon [28].

Two types of electromagnetic calorimeters are used in ATLAS. In the central η region, $|\eta| < 1.475$, one finds the barrel calorimeter, while in the more forward η region, $1.375 < |\eta| < 3.2$ one finds the electromagnetic end-cap calorimeter (EMEC), shown in Figure 2.8. Liquid argon is used as the active material due to its radiation hardness, response stability of over a long period of time and linear behavior. Both have an accordion geometry of the lead absorbers and electrodes that provides complete ϕ symmetry without azimuthal cracks. Projective geometry of the readout is achieved with this arrangement. Several active layers are obtain in depth, as well as a constant LAr gap with depth. The dimensions of the barrel module calorimeter trigger towers and its cells η and ϕ granularity in each of the three layer are shown in Figure 2.8. Accurate position measurements are obtained by finely segmenting the calorimeters in η and ϕ directions. The depth of the ATLAS electromagnetic calorimeters is chosen to be greater than 24 radiation lengths in the barrel, and more than 26 X_0 in the end-caps, so that the electromagnetic showers are almost fully contained in these calorimeters.

Hadronic calorimeters

The hadronic calorimeters are specialized in the detection and energy measurement of hadrons, such as pions and protons. The required energy resolution is given by [21]:

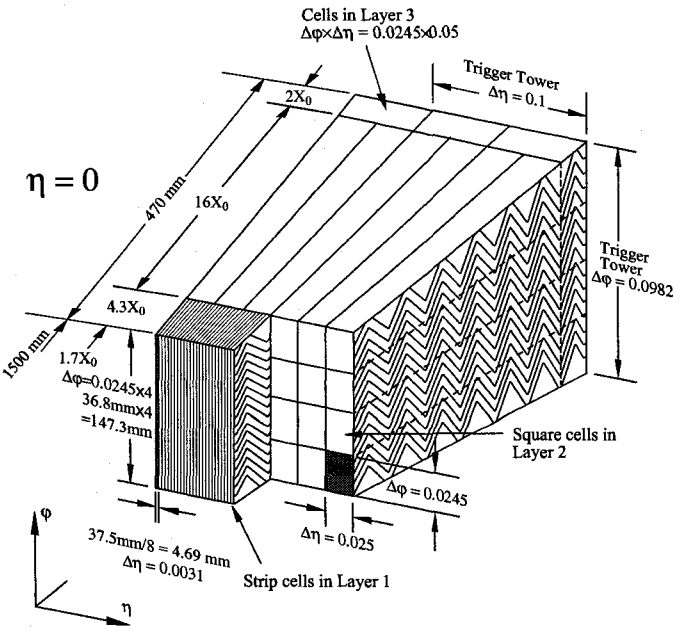


Figure 2.8: Diagram showing the accordion geometry of the electromagnetic calorimeters, the η and ϕ granularity and the projective trigger towers [24].

$\frac{\sigma(E)}{E} = \frac{50\,\%\sqrt{\text{GeV}}}{\sqrt{E}} \oplus 3\,\%$ for $|\eta| < 3.2$ and $\frac{\sigma(E)}{E} = \frac{100\,\%\sqrt{\text{GeV}}}{\sqrt{E}} \oplus 10\,\%$ for $3.1 < |\eta| < 5.2$, where E is in GeV.

The physical processes involved in the production of the hadronic shower are different from the electromagnetic ones. An energetic hadron interacts with the matter through nuclear processes and it creates new energetic hadrons, mainly pions, see Figure 2.9. These new energetic hadrons in turn interact with the material creating other hadrons. The process continues until a threshold is reached. The total energy deposited is proportional to the energy of the initial hadron. Due to the presence of neutral pions, π^0, in the hadronic showers (see Figure 2.9) and their subsequent decay to photons, there is also an electromagnetic component in the hadronic shower. For example, for 100 GeV pion showering in copper, about 60 % of the shower energy is deposited by the electromagnetic component (the fraction is in fact energy and ma-

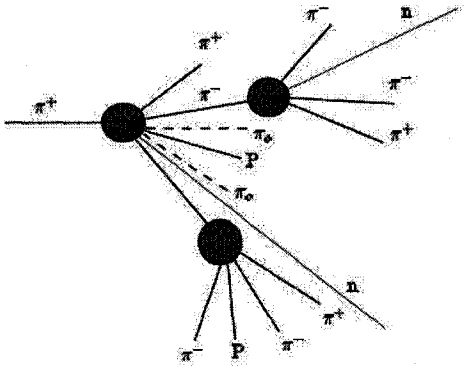


Figure 2.9: The development of the hadronic shower initiated by a pion. An electromagnetic component is present due to the decay of π^0.

terial dependent). Thus, the physics of the hadronic interactions is a convolution of strong interaction processes mentioned here and electromagnetic processes mentioned in the previous section.

The development of the hadronic shower is characterized by the nuclear interaction length λ_0 (the average distance a particle travels before strongly interacting with a nucleus, in cm) given by the formula:

$$\lambda_0 \simeq 35\,\mathrm{g/cm^2}\,\frac{A^{1/3}}{\rho} \tag{2.7}$$

where A is the atomic weight, and ρ is the material density (in $\mathrm{g/cm^3}$). The thickness of the hadronic calorimeters is chosen to be 11 interaction lengths (λ_0) at $\eta = 0$ to ensure a good containment for the hadronic showers and a good resolution for energetic jets. Some values for λ_0 are 170 mm for lead, 151 mm for copper and 837 mm for liquid argon [28]. Longitudinally, the energy deposition has a sharp peak near the first interaction point (due to the local deposition of the electromagnetic energy resulting from the neutral pions produced in the first interaction). As the shower develops, a maximum is reached at $[0.2\,ln\,(\,E\,/\,1\,\mathrm{GeV}\,) + 0.7]\,\lambda_0$ from the face of

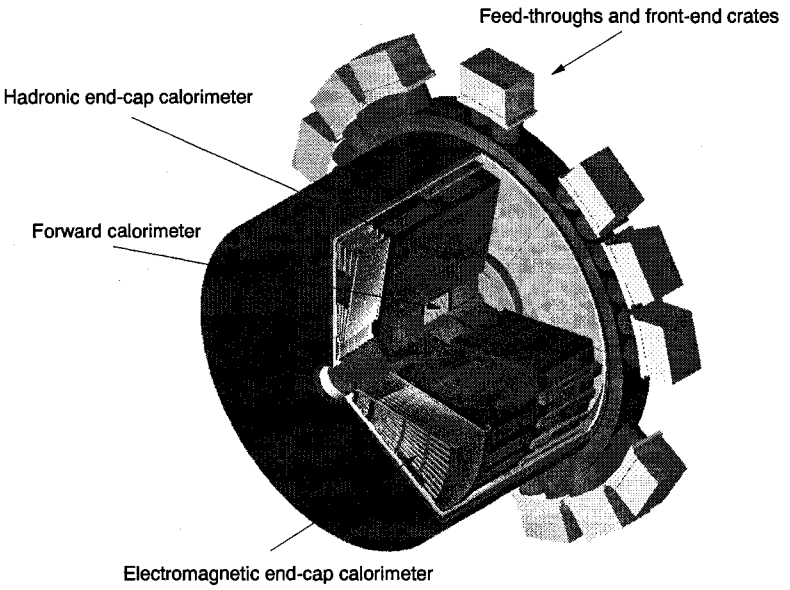
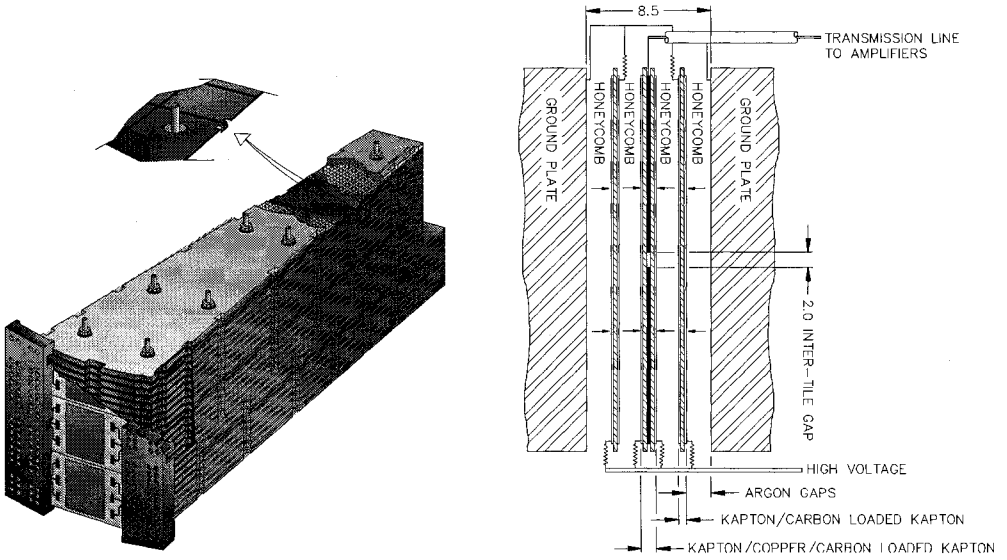


Figure 2.10: The three end-cap calorimeters (EMEC, HEC and FCal) are positioned in the same end-cap cryostat [24].

the calorimeter. Laterally, the hadronic shower is almost fully (95 %) contained in a cylinder with a radius equal to λ_0, but there are large fluctuations reported in the lateral spread.

Two different technologies are used in the construction of the two hadronic calorimeters. The barrel calorimeter uses steel absorber/scintillator tiles technology, and covers the range $|\eta| < 1.7$. The hadronic end-cap (HEC) calorimeter is a copper/liquid argon (LAr) sampling calorimeter with a flat plate design. It covers the pseudorapidity range $1.5 < |\eta| < 3.2$. The HEC is located in each of the two liquid argon end-cap cryostats, where the electromagnetic end-cap (EMEC) and forward (FCal) calorimeters are also found (see Figure 2.10). The HEC calorimeter is composed of a front, HEC1, and a rear, HEC2, wheel. Both HEC1 and HEC2 consist of two longitudinal sections.



(a) (b)

Figure 2.11: (a) HEC module; 32 such modules are assembled to create one end-cap wheel [24]. (b) The readout structure for the hadronic end-cap calorimeter [24].

A drawing of a HEC module is shown in Figure 2.11 (a). The 24 copper plates of the HEC1 are 25 mm each, with the first one being 12.5 mm. For HEC2 there are 16 copper plates of 50 mm each, plus the first plate being 25 mm. For both sections, the gaps between the plates are 8.5 mm in width. This arrangement results in a different sampling fraction for the two modules: 2.2 % for HEC1 and 4.4 % for HEC2.

In between the copper plates, there are three electrodes that divide the LAr gap into four drift zones, each having 1.8 mm. The schematical arrangement of the readout structure is shown in Figure 2.11 (b). A honeycomb sheet is used to maintain the spacing between the electrodes. High voltage is applied to the two lateral electrodes that have surfaces covered with a high resistive material. A pad structure covered by a highly resistive layer is mounted over the middle electrode in order to provide readout capabilities and to define the lateral segmentation of the calorimeter modules.

This ensemble is called a electrostatic transformer. When a charged particle passes through LAr, electron-ion pairs are created through ionization. By applying high voltages on the electrodes, the ionization electrons drift and are collected. A current is measured that is proportional to the incident particle energy. This principal has the advantage of offering a lower operation voltage of about 1.8 kV, compared to the situation in which only one gap of 7.2 mm were used, thus reducing the space-charge effects due to ion density for large η values. The electron drift time in the LAr drift gap is about 430 ns.

As in the case of the electromagnetic calorimeter, readout cells are defined by etching the central pads of the LAr gaps, resulting in a semi-pointing geometry. The granularity of the HEC readout cells in η and ϕ is 0.1×0.1 in the central region with $|\eta| < 2.5$, and 0.2×0.2 for regions with $|\eta| > 2.5$.

Forward calorimeters

The ATLAS forward calorimeter (FCal) is located in the same cryostat as the end-cap calorimeters and provides a coverage of $3.1 < |\eta| < 4.9$ (see Figure 2.10). It is very important for the measurement of energies of particles moving very close to the beam pipe, in the forward direction, where the radiation levels are very high. The close coupling between systems results in a hermetic design that minimizes energy loss across the calorimeter systems and limits the backgrounds that reach the muon system. It consists of three layers: FCal1 (made of copper) for electromagnetic calorimetry, and FCal2 and FCal3 (made of tungsten) for hadronic calorimetry. In each layer, the calorimeter consists of a metal matrix with regularly spaced longitudinal channels filled with concentric rods and tubes. A picture of the FCal structure is shown in Figure 2.12. The rods are at positive high voltage while the tubes and

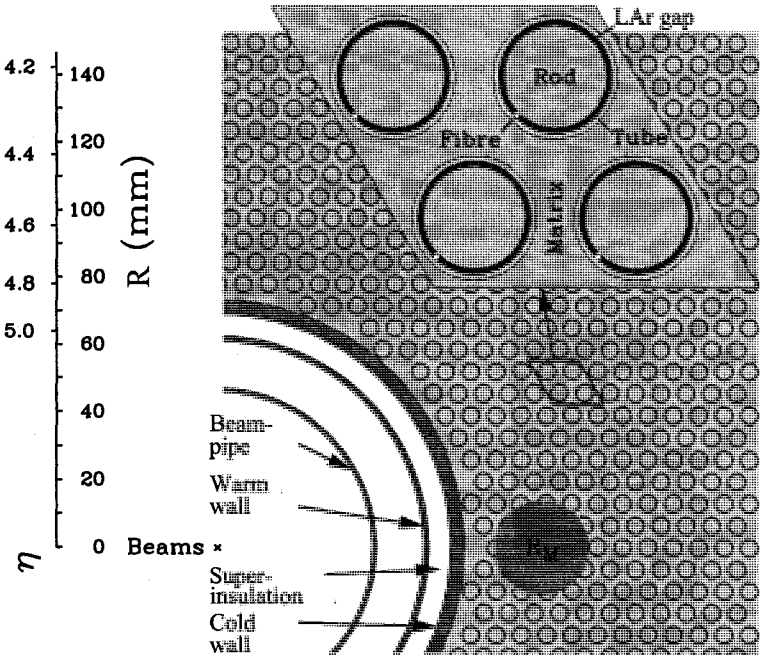


Figure 2.12: The FCal structure showing the copper plate matrix, copper tubes and the rods. The R_m circle represents the Molier radius [24].

matrix are grounded. The readout electrodes are hard-wired together using small interconnecting boards on the surfaces of the calorimeter modules. The liquid argon in the gap between constitutes the sensitive medium.

2.2.3 Muon spectrometer

The outer most part of the ATLAS detector is the muon spectrometer system [21]. It is instrumented with separate trigger and high-precision tracking chambers. In order to deflect the muons, the muon system is submerged in a magnetic field provided by three large air-core toroidal magnets (one barrel in the central region and two end-caps in the forward regions). The muon tracker is specialized in the detection and momentum measurement of muons and is divided into two regions: the central

region (barrel chambers - three cylinders concentric with the beam axis) for $|\eta| < 1$, and the forward region (end-cap chambers - four disks concentric with the beam axis) for $1 < |\eta| < 2.7$. The required p_T resolution is 10 % at $p_T = 1\,\text{TeV}$. The muon spectrometer components are shown in Figure 2.13.

Over most of the η range, a precision measurement of the track coordinates in the principal bending direction of the magnetic field is provided by monitored drift tubes, which have high measuring accuracy with mechanical rigidity and simplicity of construction. The monitored drift tubes provide a single-tube resolution of 80 μm when operated at an absolute pressure of 3 bars. At large pseudorapidities, cathode strip chambers with higher granularity are used in the innermost plane over $2 < |\eta| < 2.7$, because of their higher rate capability and time resolution. The resolution of the cathode strip chambers in the bending plane is 60 μm and about 5 mm in the transverse plane, the difference being due to different spacing of the readout channels.

The trigger system covers the pseudorapidity range $|\eta| < 2.4$. Resistive plate chambers are used in the barrel and thin gap chambers in the end-cap regions. The trigger chambers for the ATLAS muon spectrometer serve a threefold purpose. They provide bunch-crossing identification, well defined p_T thresholds, and the measurement of the second coordinate in a direction orthogonal to that provided by the precision chambers.

2.2.4 Triggering

At the LHC, the proton-proton interaction rate at the design luminosity of $10^{34}\,\text{cm}^{-2}\,\text{s}^{-1}$ will be about 1 GHz, while the bunch crossing rate will be about 40 MHz. In order to select and record only the events of interest, a trigger system is devel-

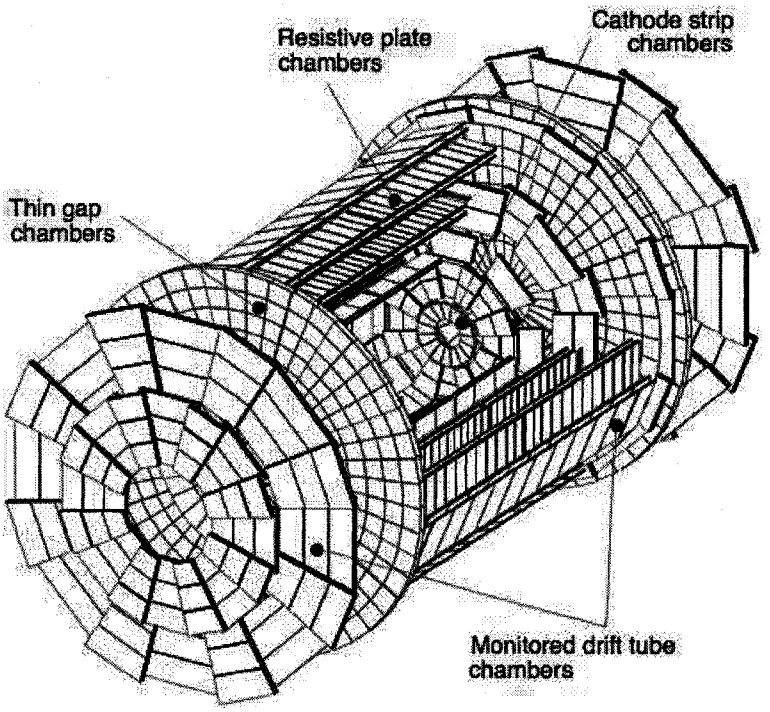


Figure 2.13: Drawing of the muon tracker system showing the various types of chambers used: MDTs (barrel and end-caps) and CSCs (end-caps) for precision measurement and RPCs (barrel) and TGCs (end-caps) for triggering [21].

oped [21]. Three levels of triggering are established for ATLAS: level one (L1), level two (L2), and event filter. The last two are grouped together into the high level trigger. A schematic drawing of the steps involved in the triggering process is presented in Figure 2.14. The level one trigger relies on custom-made electronic hardware, while the high level trigger uses commercially available computers and networking hardware.

The L1 trigger is hardware based and is used to search for high-p_T muons, electrons and photons, jets, and taus decaying into hadrons, as well as large missing and total transverse energy. At the L1 trigger level, two types of information are used: the calorimeters and the muon trackers, both with a latency of 2.5 μs (the time needed to form and distribute the decision of trigger). After the level one trigger, the frequency (rate) drops to 75 kHz and the events are fed into the high level trigger.

The L2 trigger system is software based. Only detector sub-regions (regions of interest) seeded by level one are processed. At L2, further processing is performed on the events that passed the level one trigger. For example the information about the electromagnetic clusters in the calorimeter is used to identify an electron/photon. This information is used to reconstruct the impact point on the face of the calorimeters. Then, one tries to match this impact point position to a track from the inner detector (by matching the cluster η and ϕ position with the η and ϕ values of candidate tracks extrapolated to the calorimeter surface). If one finds a match then the object is classified as an electron, and if not then is classified as a photon. After the L2 trigger system, the rate drops to 3.5 kHz, with a mean processing time for an event of about 40 ms.

At the event filter level, the events that passed the L2 trigger are further analysed using off-line type reconstruction software (unlike the L2 trigger, these tasks

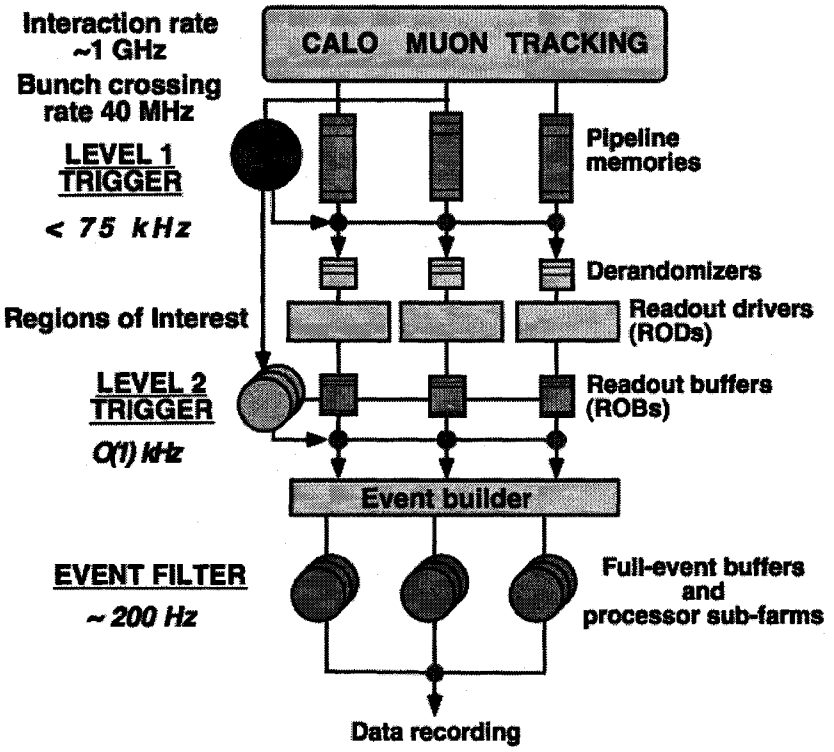


Figure 2.14: The trigger system in ATLAS, depicting the level 1, level 2, and event filter triggers [21].

are standard ATLAS event reconstruction and analysis applications). After the event filter, the event rate drops to 200 Hz with a mean processing time for an event of about 4 s and the events are recorded into CERN's storage facilities for detailed physics analysis.

Chapter 3

The 2002 beam test

Beam test measurements were performed at CERN on modules of the ATLAS calorimeters. These measurements provide the opportunity to analyze the performance of individual modules, test the energy reconstruction for different particle types and energies, and develop/use dedicated software for monitoring and analysis.

In the past, only stand-alone beam tests were performed in the end-cap region (EMEC and HEC modules tested separately). Several stand-alone beam tests were performed for the EMEC (1998, 1999, 2000 and 2001) and HEC (1998, 1999, 2000 and 2001). These calorimeter modules were tested to verify that they meet the design specifications, $\frac{\sigma(E)}{E} = \frac{10\,\%\sqrt{\mathrm{GeV}}}{\sqrt{E}} \oplus 0.7\,\%$ for the electromagnetic calorimetry and $\frac{\sigma(E)}{E} = \frac{50\,\%\sqrt{\mathrm{GeV}}}{\sqrt{E}} \oplus 3\,\%$ for the hadronic calorimetry [21], where E is in GeV, and to determine the energy scale of the calorimeter signals.

In the summer of 2002, the first combined EMEC/HEC beam test took place. Only combined tests provide calibration constants for single pions interacting with the ATLAS detector. This experiment studied the end-cap region corresponding to

the pseudorapidity interval $1.6 < |\eta| < 1.8$. The goals of the beam test were to obtain intercalibration constants for electrons and pions in the energy range $6\,\text{GeV} < E < 200\,\text{GeV}$, perform a detailed comparison with simulation to allow extrapolation to jets, study energy sharing between calorimeter technologies and obtain the energy scale for hadrons, and test methods and algorithms for optimal hadronic energy reconstruction in ATLAS.

In this chapter, the beam test is described, together with the analysis of data for beams of electrons and pions. In Section 3.1, the general test setup is discussed. A short description of the data collected during these tests is presented in Section 3.2. The electronic noise analysis is presented in Section 3.3. In Section 3.4, a clustering algorithm is presented that is used to reconstruct the particles' energy. Electron data analysis is discussed in Section 3.5 and pion results are presented in Section 3.6.

3.1 Test setup

The beam tests were performed at CERN in the H6 beam line of the SPS (see Section 2.1). The protons from the SPS collide on a target producing beams of various particle types (electrons, muons and pions) and energies (6 GeV to 200 GeV).

Figure 3.1 shows the beam test setup. The beam of particles comes from the right, passes through a number of monitoring detectors (scintillators for triggering and multiwire proportional chambers (MWPC) for charged particle tracking) and penetrates a cryostat filled with 90 K liquid argon in which the calorimeter modules are submerged.

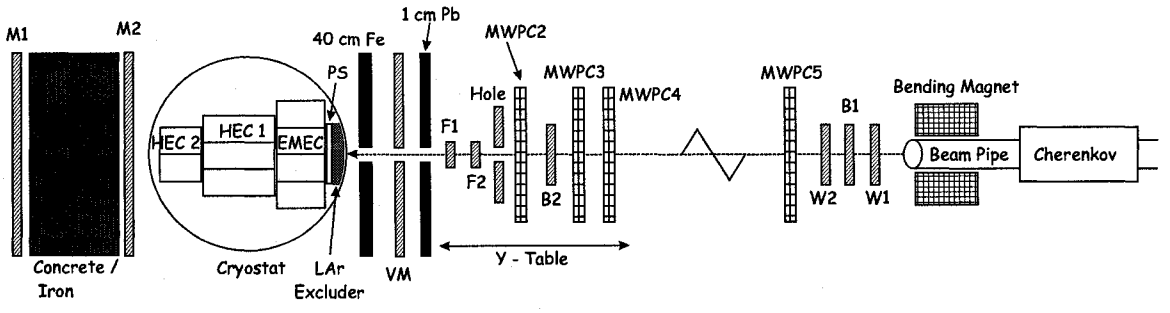


Figure 3.1: The 2002 beam test setup showing the monitoring instrumentation used to characterize the beam and the tested calorimeter modules in the cryostat. Beam direction is from right to left [37].

3.1.1 Trigger and tracking systems

In order to let the data acquisition (DAQ) system know when to read out the detector, various trigger detectors are used upstream and downstream of the cryostat. The locations of the trigger and the veto counters are shown in Figure 3.1. The trigger system is formed by the scintillation counter B1, three scintillation walls VM, M1, M2 and two scintillation counters F1, F2. Upstream of the bending magnet exists a Cherenkov counter that is used for beams with energies less than 80 GeV in order to separate electrons from pions. Four MWPCs are used to give information about the impact position and the angle of particles. Particles from the beam halo are rejected by the veto wall VM. The coincidence of the M1 and M2 scintillator signals is used for the identification of muons. Muons are the only particles that can penetrate through the concrete/iron block situated behind the cryostat.

3.1.2 Calorimeter modules tested

The setup of the calorimeter modules in the cryostat is shown in Figure 3.2. The beam penetrates the cryostat through a window of 60 cm in diameter. The window has a smaller thickness compared to the cryostat wall in order to reduce the amount of dead material in front of the calorimeter modules. A low density foam excluder was used to remove the liquid argon from the front of the modules and to limit the energy loss upstream the calorimeter. Next in the path of the beam is the presampler detector followed by a module of the EMEC. It is a sampling calorimeter in which lead plays the role of the absorber and LAr is the active layer as described in Section 2.2.2. The HEC modules lie behind the EMEC. HEC is also a sampling calorimeter, but has copper as absorbers. There are three HEC1 modules. Due to the cryostat size constraints, only half (in the longitudinal direction) of the normal HEC2 is used in the beam test. There are two such modules. Due to the reduced number of interaction lengths compared to the nominal HEC, longitudinal energy leakage is expected to occur for high energy hadrons.

Each calorimeter is segmented into readout cells. They are divided in longitudinal segments, along z beam axis, and lateral cells are defined in η and ϕ directions (see Section 2.2.2). This segmentation allows the calorimeters to be used also for determining the impact point of the beam particles. In the ATLAS experiment, the distribution of EMEC and HEC cells follows a pointing geometry along η - these cells are distributed along lines that originate at the interaction point. For the beam test setup, the beam is instead perpendicular to the surface of the calorimeter modules, as it can be seen in Figure 3.3 for the HEC. This non-pointing geometry has to be considered when translating beam test results into ATLAS performance.

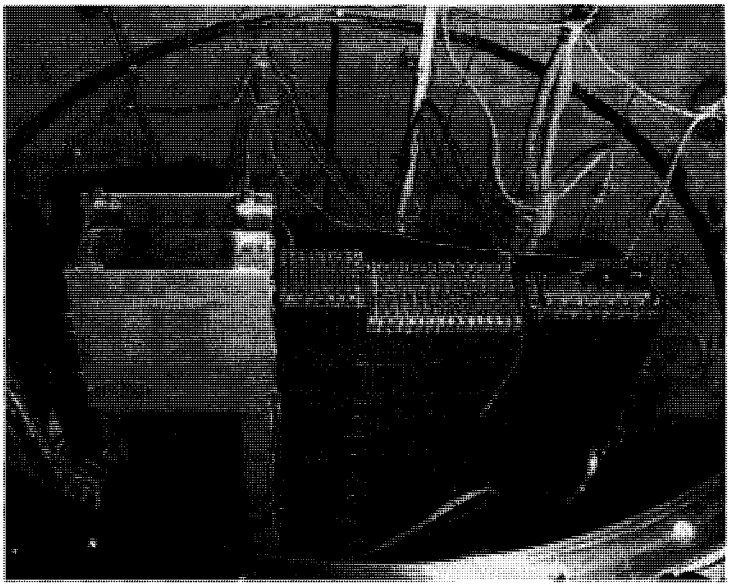


Figure 3.2: The 2002 beam test setup picture. Beam direction is from left to right. On the left is the presampler, then the EMEC module and to the right is the group of HEC modules - three HEC1 modules and two half longitudinal size HEC2 modules [37].

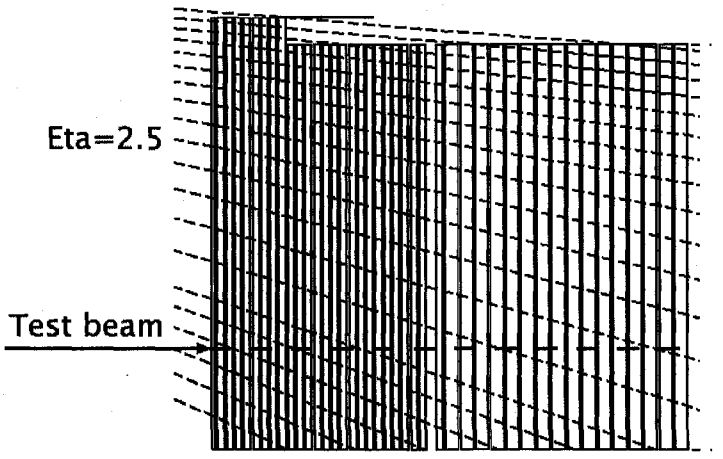


Figure 3.3: The distribution of HEC readout cells follows a pointing geometry along η - these cells are distributed along dashed lines that originate at the interaction point. For the beam test setup, the beam is perpendicular to the calorimeter's surface.

The positioning of the particle beam on the surface of the calorimeter is achieved by two different methods. The cryostat containing the modules was placed on a moving table allowing the motion of the system in the horizontal direction, while a magnet system (Bending Magnet in Figure 3.1) was used to deflect the beam in the vertical direction. A two dimensional scanning of the modules is thus possible.

3.1.3 Readout, calibration and signal reconstruction

When beams of particles of known energies (in GeV) interact with the active material in the calorimeters, signals (electric currents in nA) are produced in the calorimeters' readout cells. This signal is sent to the front end boards (FEBs) and read out by analog to digital converters (in ADC counts). The two types of radiation tolerant FEBs used for the EMEC and HEC readout channels (128 channels per FEB) are described in [31, 32, 35]. For the 2002 beam test, the total number of channels read out in the EMEC was 1152, but only 880 of them were physically connected to a calorimeter cell. The HEC had a total of 384 channels read out of which only 183 were physically connected to a calorimeter cell. The readout of the EMEC data was performed by nine MINI-ROD modules, previously used for the EMEC stand-alone tests [31]. The prototype of the ATLAS ROD module (ROD-demo [33]), designed to validate the final ATLAS LAr calorimetry readout system, was used for the HEC readout. As described in [37], the ROD-demo was a 9U VME motherboard with four mezzanine processing unit (PU) cards. Each PU processed the data from one half FEB, so in total six PU's on two ROD boards were used to read out the three HEC FEBs.

The ionization current measured in the calorimeter LAr gap has a triangular

shape [28]. Figure 3.4 shows the shape of physical, readout and sampled signal. The triangular shape is due the fact that at the beginning all ionization charges move with a constant drift velocity when an electric field is applied. As some of the charge arrives at the electrodes, the charge density diminishes linearly resulting in a linear decrease of the measured current. After about 400 ns, the signal decreases to zero. Because the LHC bunch crossing time is 25 ns, up to 16 bunches may produce signals during the interval of a single triangular current producing pile-up of the signal. The pile-up of the signal occurs when the charge collected from an event is contaminated by the charge produced by a previous or a following event. In order to reduce the pile-up of the signal and yield a fast response for energy measurements in calorimeter cells, the ionization signal is shaped as a bipolar pulse that yields a peak proportional to the peak drift current (and thus to the shower energy) in about 40 ns. A plot of the bipolar response is shown in Figure 3.4.

Known current pulses are injected in each readout cell in order to calibrate them (i.e. to extract the nA to ADC counts conversion). In this way, calibration parameters were obtained for both EMEC and HEC. An example is shown in Figure 3.5 for the HEC module. The left figure shows the calibration signal (points) fitted by the full electronics function (the convolution of the triangular ionization current with the electronics response function) (line). The right figure shows the corresponding residuals of the fit. The residuals are well within $\pm 1.5\,\%$, except for the signal rise, where the influence of small distortions of the calibration pulse shape is not taken into account. The conversion from nA to GeV is obtained by performing beam tests of calorimeter modules in beams of particles of known types and energies (see Section 3.5.2).

The analog signal is sampled at regular time intervals of 25 ns and the ampli-

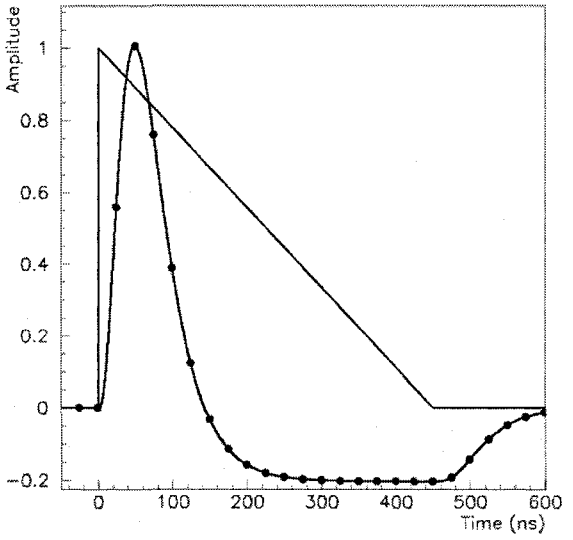


Figure 3.4: Signal shape: physical (triangular) and readout (bipolar) sampled every 25 ns (black dots). Normalized amplitude versus time corresponding to the input signal and the response of the bipolar shaper [21].

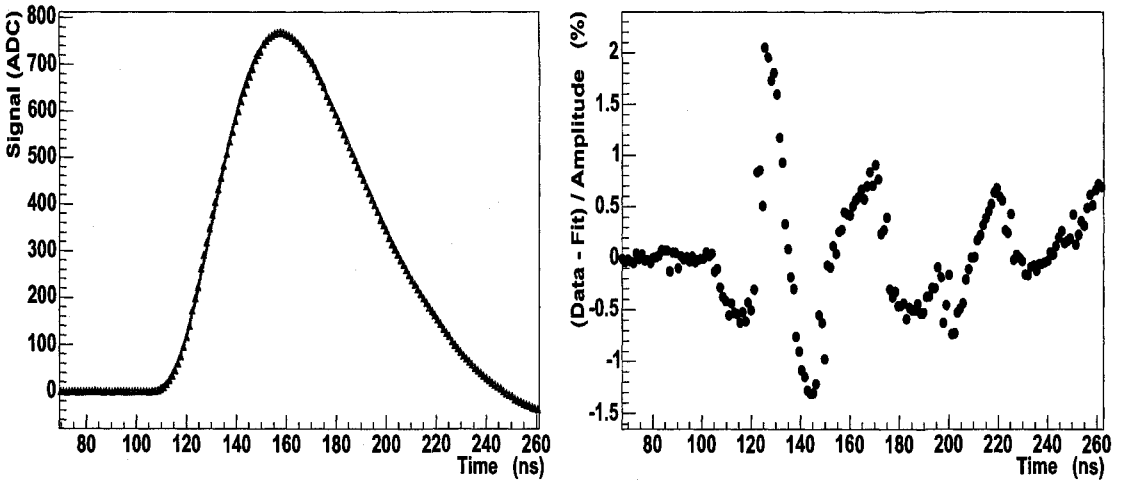


Figure 3.5: For a typical HEC readout channel. Left: the measured calibration signal (points) fitted by the full electronics function (line). The amplitude is shown in units of ADC counts. Right: the residuals of the fit [37].

tude of the signal for those time values are recorded in ADC counts. For the EMEC, typically 7 samplings have been read out. For the HEC, where due to the larger readout cell capacitances in comparison to the EMEC a larger noise is expected, usually 16 samplings were used. The signal usually peaks in sample three, while sample zero is used to calculate the electronic noise present in the cell (see Figure 3.4). The pedestal is the value of the signal generated in the cell even when no particle traverses the calorimeter. This value is subtracted from the measured signal amplitude in order to account for the electronic noise. The signal was reconstructed using the optimal filtering (OF) method [30] in a similar way to the method that will be used later in the ATLAS detector. This method allows the determination of the signal from the digitized signal values recorded at N time samples. The technique is an error minimization of the following two quantities:

$$S = \sum_{\mathrm{i}=0}^{N} a_{\mathrm{i}} \times (S_{\mathrm{i}} - Ped) \tag{3.1}$$

$$St = \sum_{\mathrm{i}=0}^{N} b_{\mathrm{i}} \times (S_{\mathrm{i}} - Ped) \tag{3.2}$$

where S is the signal per readout channel, S_{i} is the signal in time sample i, N is the number of time samples used in the OFC calculations (typically 5), Ped is the pedestal value, t is the time of the peak signal, a_{i} and b_{i} are the OF weighting factors. The two parameters, a_{i} and b_{i}, are thus extracted using information about the pulse shape, event phase (the position of the sample relative to the shape) and the time sample auto-correlation functions (the correlation of the noise contributions at two different times is expressed as a noise autocorrelation matrix) of each channel. The computation of the signal arrival time and of the signal maximum is done from a set of five measurements (samples) using the same constraints.

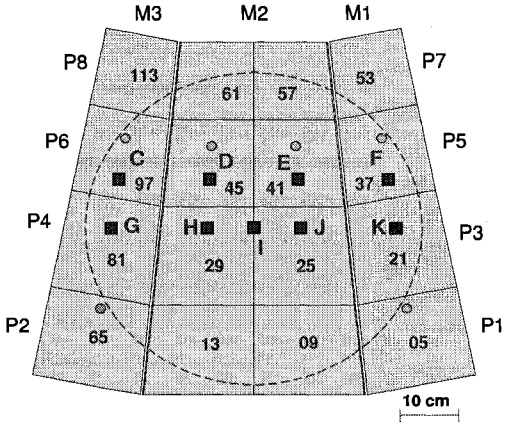


Figure 3.6: The positions of the 9 impact points (C, D, E, F, G, H, I, J and K) on the HEC1 surface. The dotted circle shows the outline of the cryostat window through which the beam passes before impacting on the calorimeter's surface [37].

3.2 Collected data

The data collection period extended from August 17, 2002 to September 15, 2002. Beams of electrons, pions and muons were used with energies between 6 GeV and 200 GeV. The beams were impacted at nine points (C, D, E, F, G, H, I, J and K) on the modules in order to study the response uniformity of the detector. The positions of the impact points on the HEC surface are shown in Figure 3.6. The impact point *I* is singular in the fact that it is the only one situated on the edge between two adjacent readout cells, so variations on the results are expected for this point. Most of the analysis in this thesis is performed for impact point *J*. Detailed horizontal and vertical beam scans were also taken during the beam test period in order to study the uniformity of the EMEC response. Eight surrounding points were chosen around each impact point as shown in Figure 3.7.

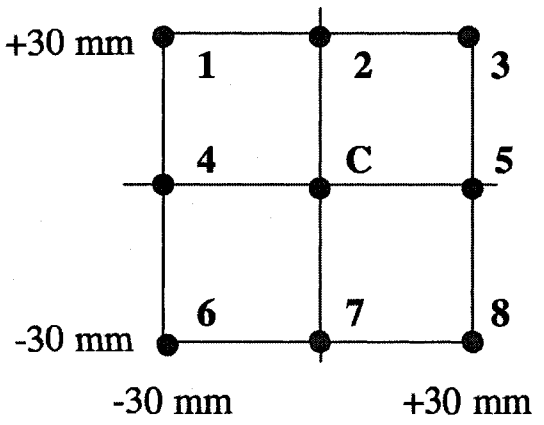


Figure 3.7: The positions of the eight surrounding points (1, 2, 3, 4, 5, 6, 7 and 8) and the centre, C, for a particular impact point [37].

Data were recorded in runs (collections of 10,000 to 20,000 events for a given particle type and energy). There were 743 runs collected.

3.3 Noise analysis for EMEC and HEC

For LAr technology based calorimeters, the readout signal is based on the collection of small amounts of electric charge. Due to the inherent capacitance of the LAr cells in the calorimeters and the instrumentation used to read out the detector, electronic noise contributes to the measured signal. This means that in the absence of particles traversing the detector, the collected charge (the signal) fluctuates from event to event. The standard deviation of these fluctuations, σ_{noise}, is measured and converted to an equivalent amount of deposited energy. Because of its effect on the calorimeter's performance, the electronic noise is studied for the calorimeter modules used in the beam test.

As mentioned in Section 3.1.3, the pedestal values of the signal can be used to measure the electronic noise in each cell. For this study, the first time sample (sample 0) of the signal (see Figure 3.4) was considered as the pedestal. The stability of the pedestals was studied approximately daily over a one month period of the 2002 beam test. Some of the runs used for the analysis and their characteristics are given in Table 3.1.

Figure 3.8 shows typical pedestal subtracted signal in ADC counts for four cells belonging to various layers in the EMEC (presampler, front, middle and an unconnected cell), for run 12100. For each cell, the mean and the RMS of the pedestal distribution for all events in the run are recorded (see for example Figure 3.9 showing the pedestal mean and RMS in ADC counts as a function of channel number in EMEC, together with their overall distributions). The average value for the pedestal mean over the 1152 EMEC channels is 1016 ADC counts with a RMS spread of the mean of 14.32 ADC counts.

The RMS distribution versus channel number in the EMEC shows a largely stable behavior for each EMEC layer with some features. A more detailed view of the top right plot of Figure 3.9 is presented in Figure 3.10, in which L0 is the presampler, L1 is the first longitudinal readout layer of the EMEC, L2 is the second readout layer and L3 is the third longitudinal layer. As it can be seen, steps are present at some boundary layers, e.g. L3→L2 at channel number 640. The steps are related to the structure of the EMEC that has different capacitances for each layer. Some channels that are connected to physical cells in L1 (1024-1056) and L2 (384-448) show some unstable behavior. Channels 48 (FEB 0, connector A) to 127 (FEB 0, connector B) are not connected to any physical cell in the EMEC (channels with a number less than 48 correspond to the presampler while channels with a number greater than 127

Day	Run	Date	Type
0	12100	17.08.2002	119 GeV e
1	12201	18.08.2002	119 GeV e
2	12276	19.08.2002	119 GeV e
3	12318	20.08.2002	100 GeV π
4	12340	21.08.2002	20 GeV π
4	12362	21.08.2002	180 GeV π
5	12390	22.08.2002	60 GeV π
5	12398	22.08.2002	60 GeV π
6	12410	23.08.2002	50 GeV π
6	12418	23.08.2002	50 GeV π
6	12443	23.08.2002	40 GeV π
7	12472	24.08.2002	100 GeV e
8	12495	25.08.2002	80 GeV e
9	12550	26.08.2002	40 GeV e
10	12616	27.08.2002	150 GeV π

Table 3.1: Partial list of runs used in this analysis. The entry "Day" represents the number of days since the reference day 17.08.2002

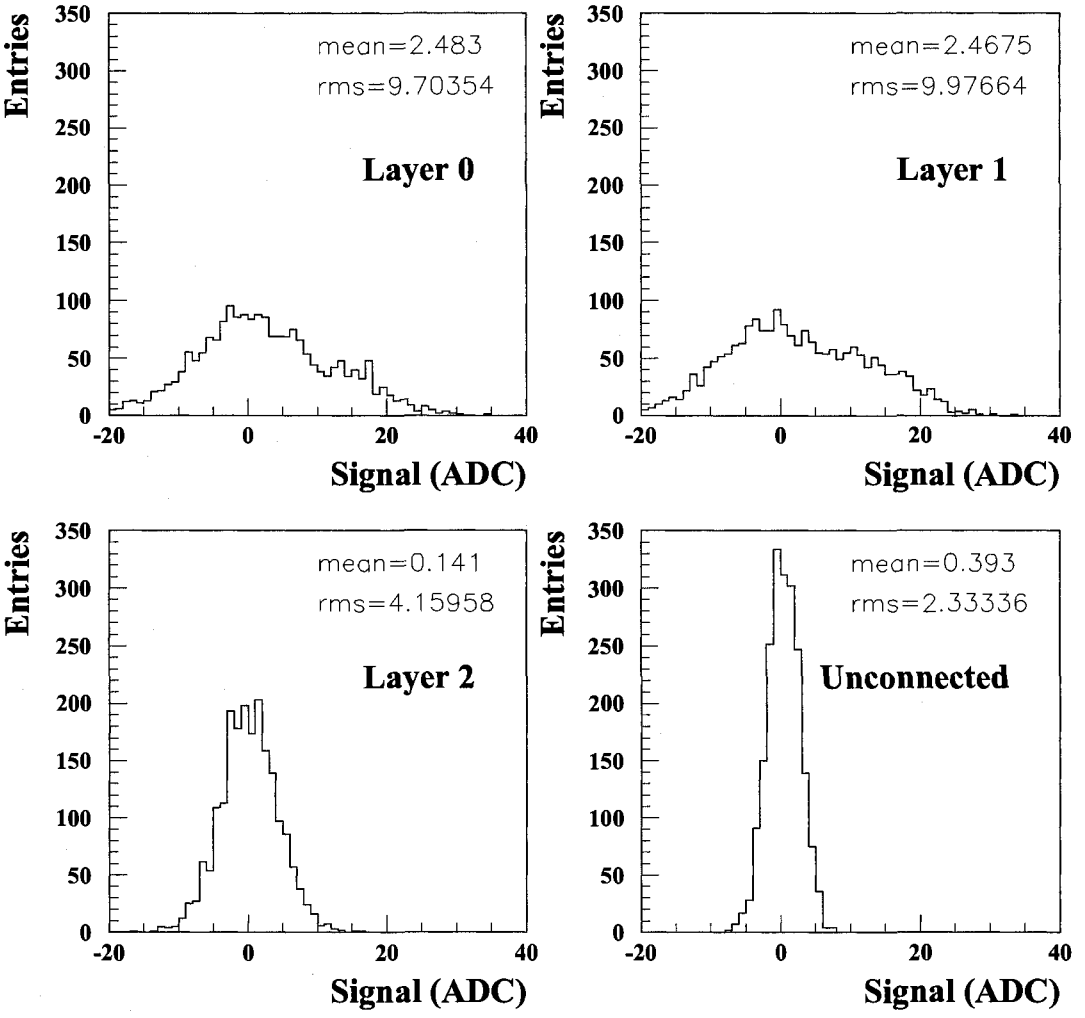


Figure 3.8: Signal (pedestal subtracted) from four cells in EMEC for run 12100 showing the distribution of the pedestal in ADC counts for 300 events in the run.

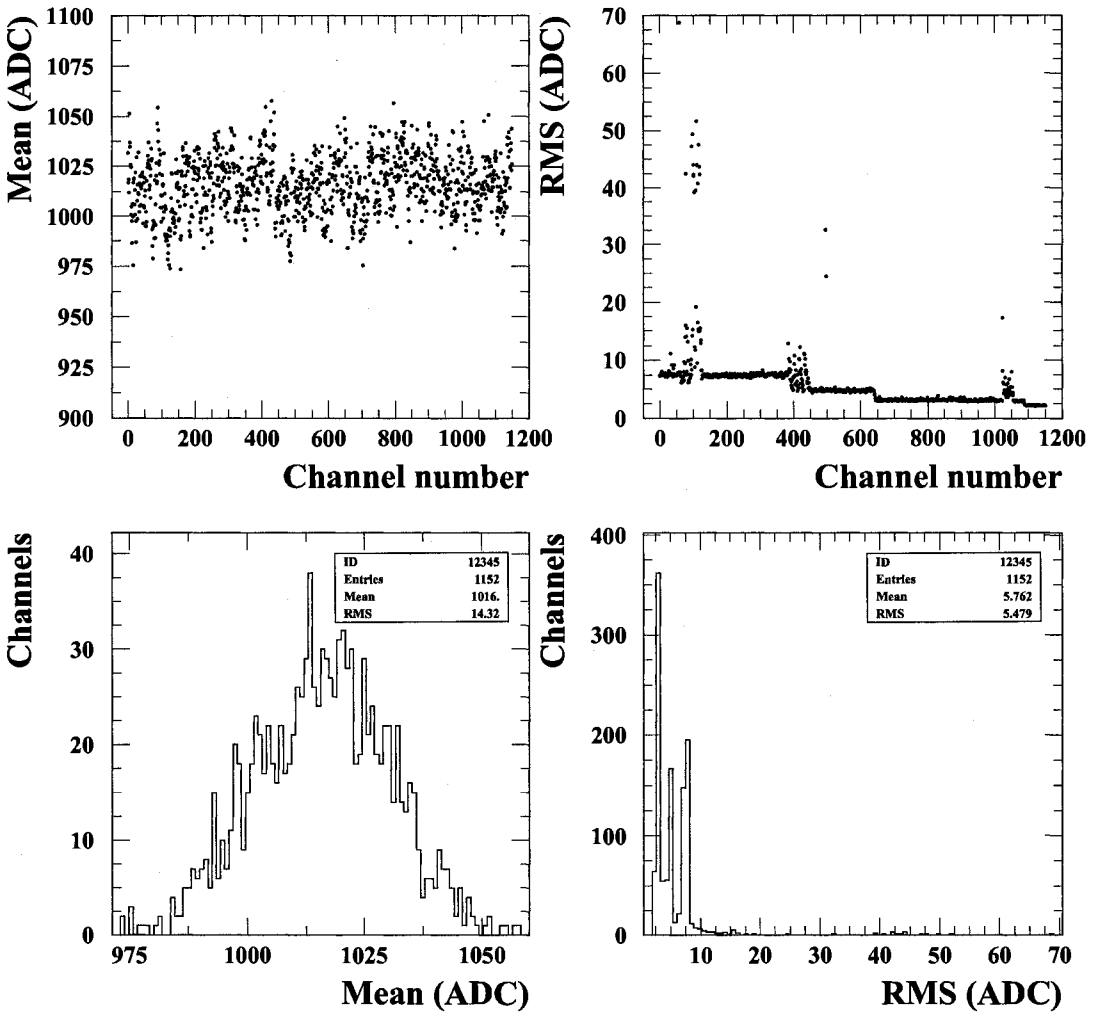


Figure 3.9: Pedestal mean vs channel (top left), pedestal RMS vs channel (top right) and overall channels mean (bottom left) and RMS (bottom right) for the EMEC (run 12100 at the beginning of the beam test period - 17.08.2002).

belong to the EMEC). Some of them (48-64 and 1056-1152) show stable behavior while some others (64-127) have unstable behavior.

Figure 3.11 shows the ADC counts output (pedestal subtracted signal) of four cells in various layers in the HEC, for run 12100. In Figure 3.12 the pedestal mean and RMS in ADC counts for all events in the run are plotted as a function of channel number in HEC, together with their overall distribution over the 348 HEC channels. It can be seen that the average value over all channels for the pedestal mean is 1018 ADC counts with a RMS spread of the mean of 14.93 ADC counts.

As is the case in the EMEC, for the HEC also the RMS distribution has a pattern. A more detailed view of the top right plot of Figure 3.12 is presented in Figure 3.13. The RMS distribution for the channels in the HEC shows a pattern related to the layer structure of the calorimeter, due to variations in cells capacitance. For the connected channels the highest RMS comes from the third (rear) layer, the middle values come from the second (middle) layer and the lowest values come from the first (front) layer of the HEC.

In order to study the time stability of the pedestals in the EMEC, run 12100 (17.08.2002) was chosen as the reference (see Figure 3.9). The same distribution is plotted for various runs during the beam test period. The stability is monitored by looking at the ratio of these distributions to the reference distribution, channel by channel. Figure 3.14 shows the average (top left) and the RMS (top right) over all channels of the ratio of the pedestal mean as a function of time. The same procedure is applied on the study the variation in time of the pedestal RMS for EMEC. The average (bottom left) and RMS (bottom right) over all channels of the ratio of the pedestal RMS as a function of time are shown in Figure 3.14. It can be seen that the RMS spread of the pedestal mean is stable to around 0.05 % (0.5 ADC counts), while

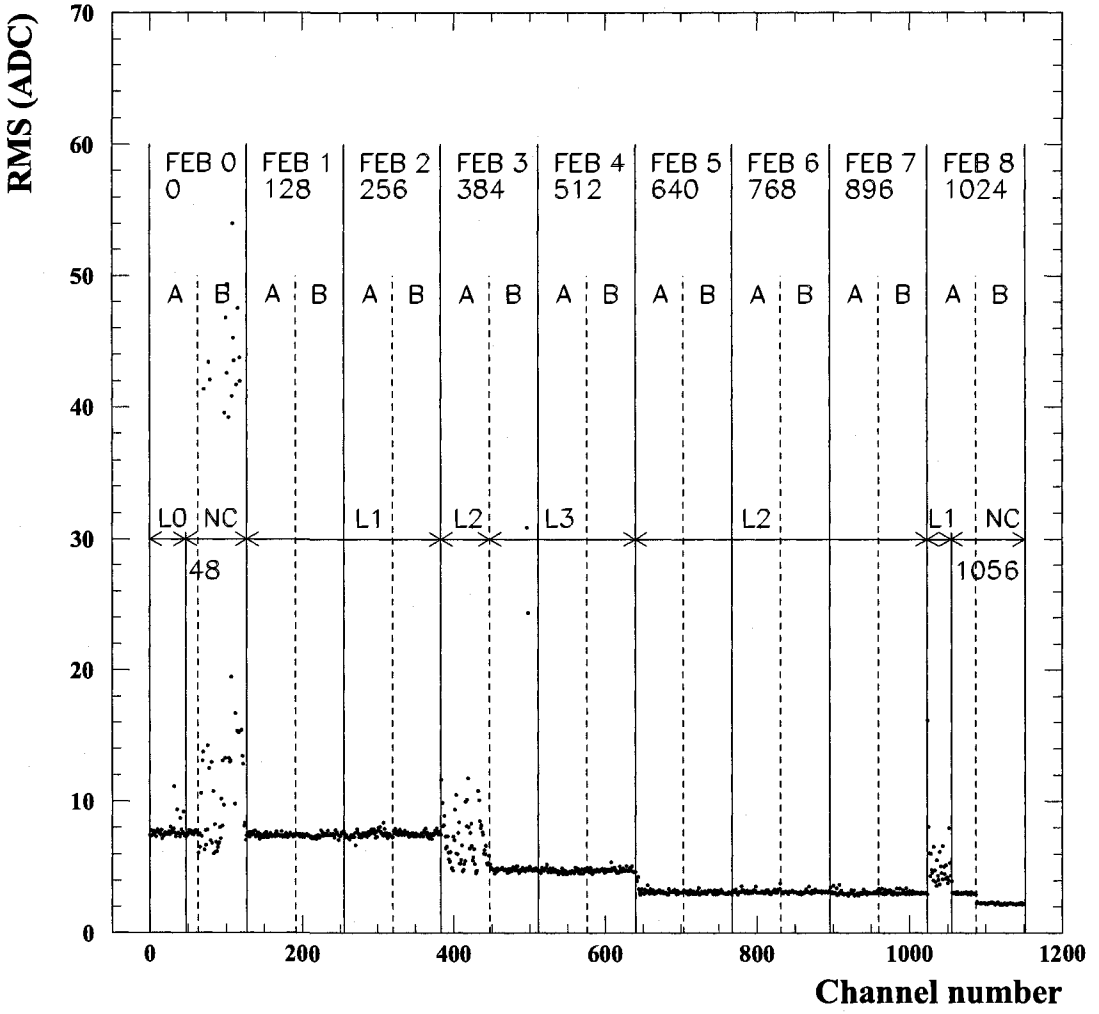


Figure 3.10: Pedestal RMS (ADC counts) vs channel for the EMEC (run 12100 - 17.08.2002). The EMEC layers (0, 1, 2 and 3) and the unconnected (NC) channels are delimited by arrows.

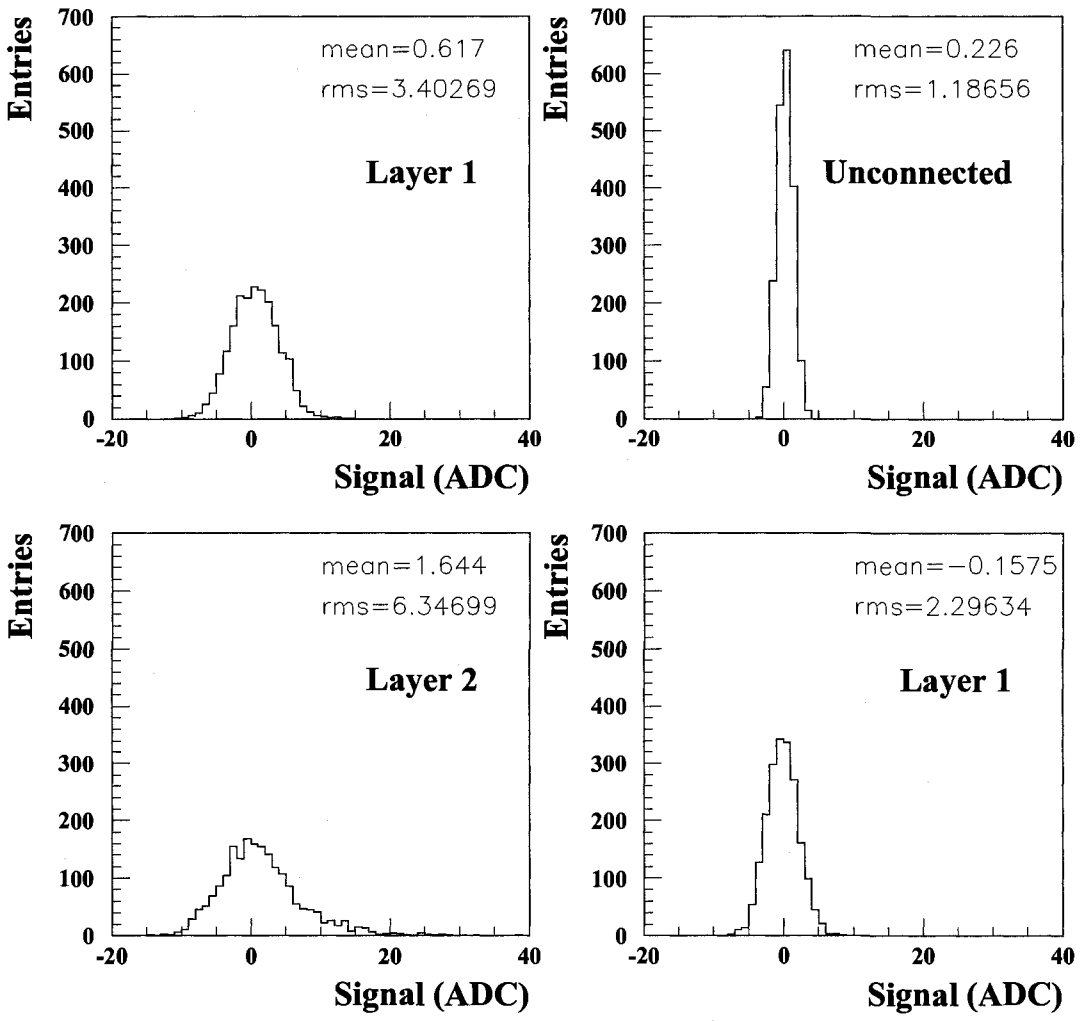


Figure 3.11: Signal (pedestal subtracted) from four cells in HEC for run 12100 showing the distribution of the pedestal in ADC counts.

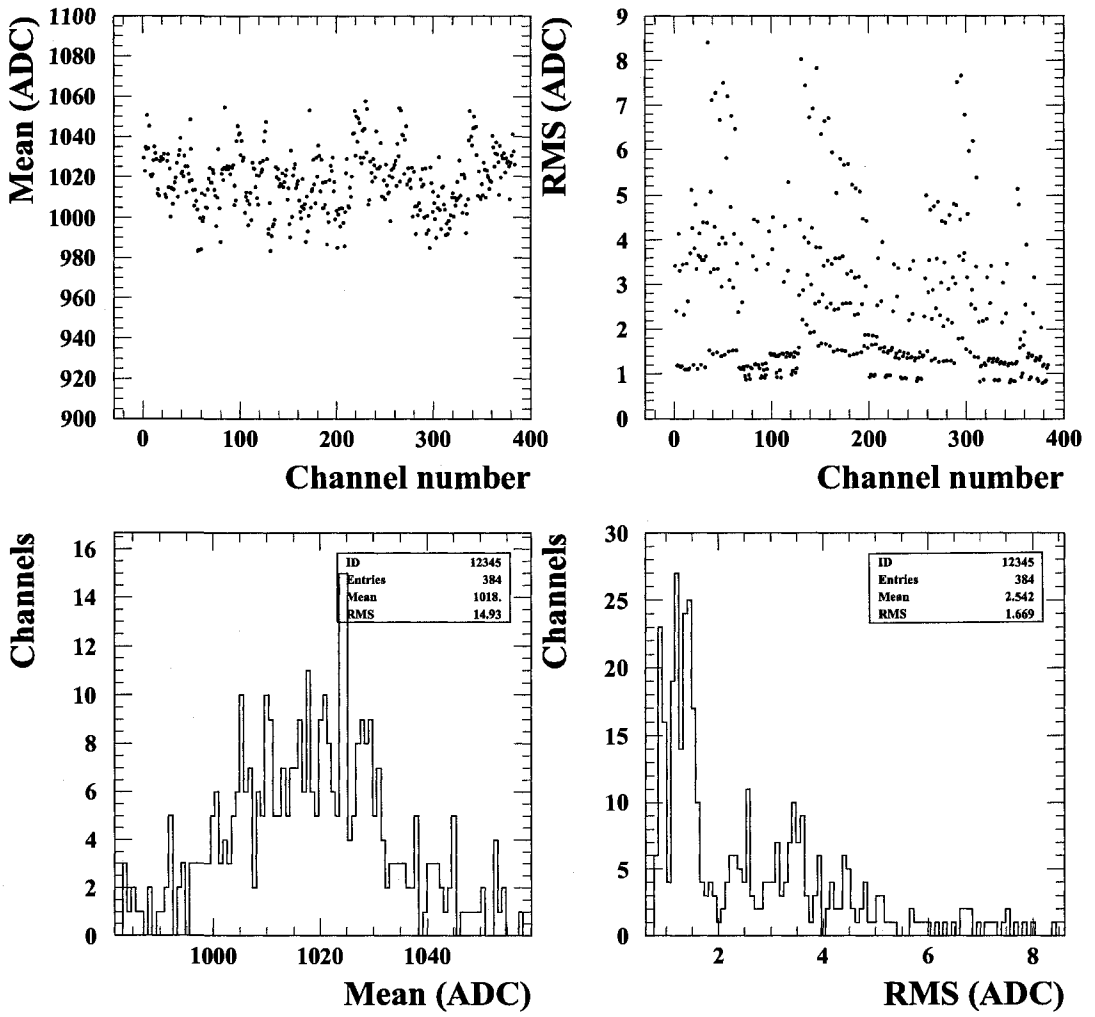


Figure 3.12: Pedestal mean for all events in a run vs channel (top left), pedestal RMS vs channel (top right) and overall channels mean (bottom left) and RMS (bottom right) for the HEC (run 12100 - 17.08.2002).

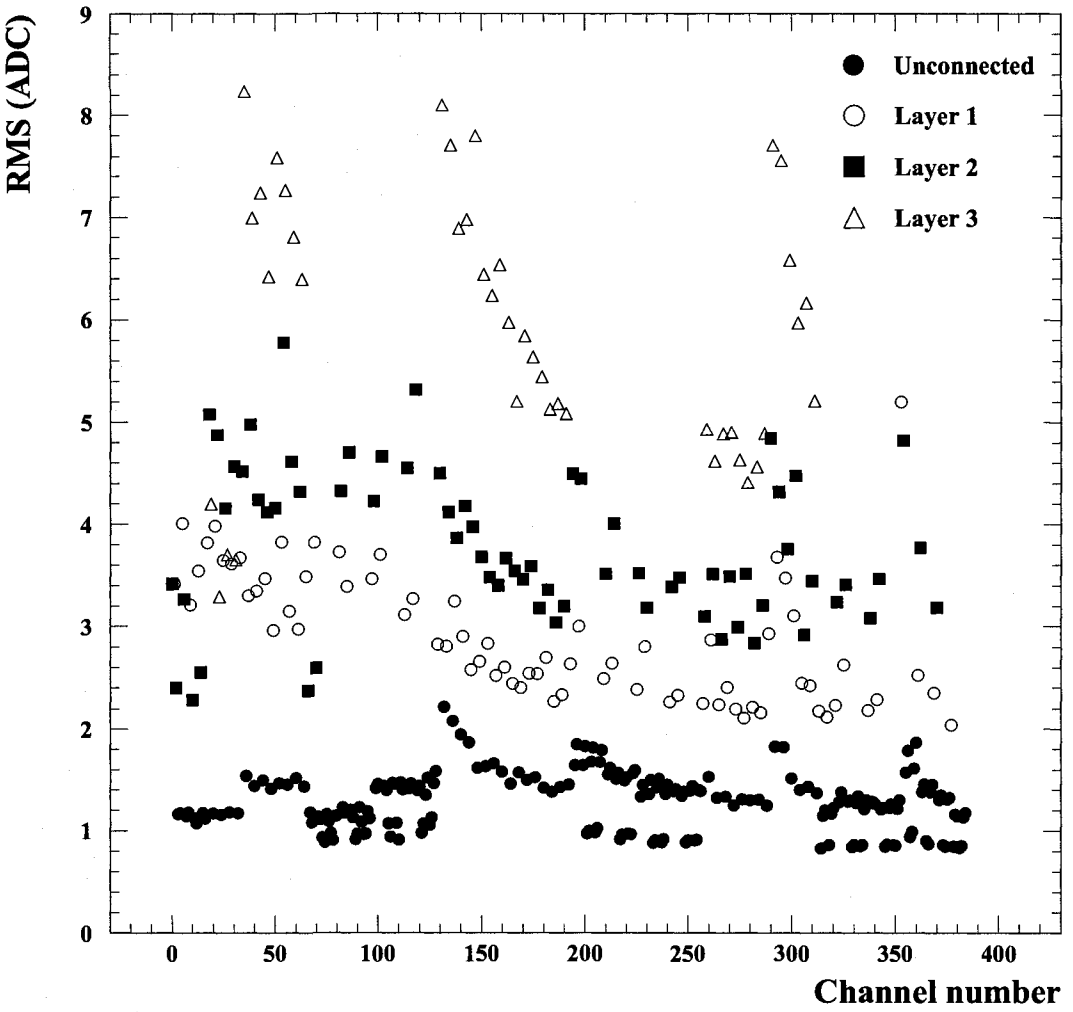


Figure 3.13: Pedestal RMS vs channel for connected (layers 1, 2 and 3) and unconnected channels in the HEC (run 12100 - 17.08.2002).

the RMS spread of the pedestal RMS is stable to around 3.5 % (0.2 ADC counts).

For the HEC, run 12100 (17.08.2002) was chosen as reference in order to study the time stability of the pedestals (see Figure 3.12). The same method described for the EMEC was applied, using all HEC channels. The time dependence of the pedestal mean and RMS ratios over the entire test period is shown in Figure 3.15. This figure shows the average (top left) and the RMS spread (top right) of the ratio of the pedestal mean, and the average (bottom left) and the RMS spread (bottom right) of the ratio of the pedestal RMS as a function of the day in the beam test. It can be seen that the RMS spread of the ratio of the pedestal mean is stable to around 0.03 % (0.3 ADC counts) while the RMS spread of the pedestal RMS is stable to around 3.9 % (0.1 ADC counts).

It can be concluded that the RMS spread of the pedestal mean is stable to $0.01\,\% - 0.05\,\%$ while the RMS spread of the pedestal RMS is stable to $2 - 4\,\%$.

3.4 Clustering algorithm

Each calorimeter readout cell (defined longitudinally - in z direction, and in the transverse plane - in η and ϕ) has an electronic noise, σ_{noise}. In addition, each cell may also have a signal due to energy deposit of particles traversing the cell. In each of these cells an electric current is produced, and it is read out by the electronics. The cells with a significant energy deposit above the contribution due to the noise σ_{noise} are used to reconstruct the total energy deposited by the particles. The cells that have a signal over a certain threshold, $E_{\text{cell}} > n\,\sigma_{\text{noise}}$ (n - a multiplicative coefficient to be determined), were chosen to belong to the cluster of energy deposit. The cluster energy is then computed by adding the signal contributions from each

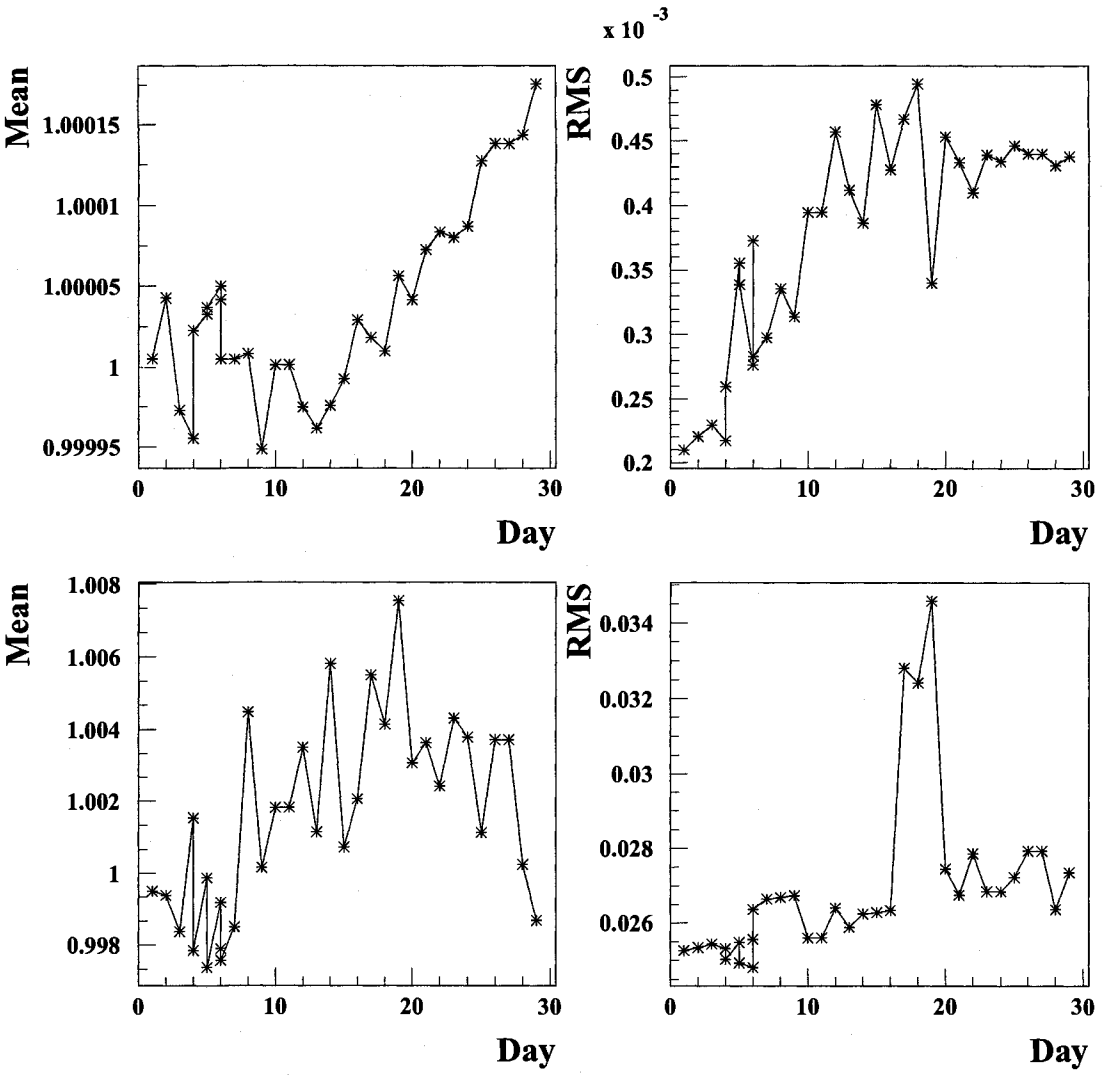


Figure 3.14: EMEC pedestal mean and RMS ratio with respect to reference run versus day in the beam test (reference run 12100 - 17.08.2002). Top left: average ratio over all channels of the pedestal mean, top right: RMS spread of the pedestal mean ratio. Bottom left: average ratio over all channels of the pedestal RMS, bottom right: RMS spread on the pedestal RMS ratio.

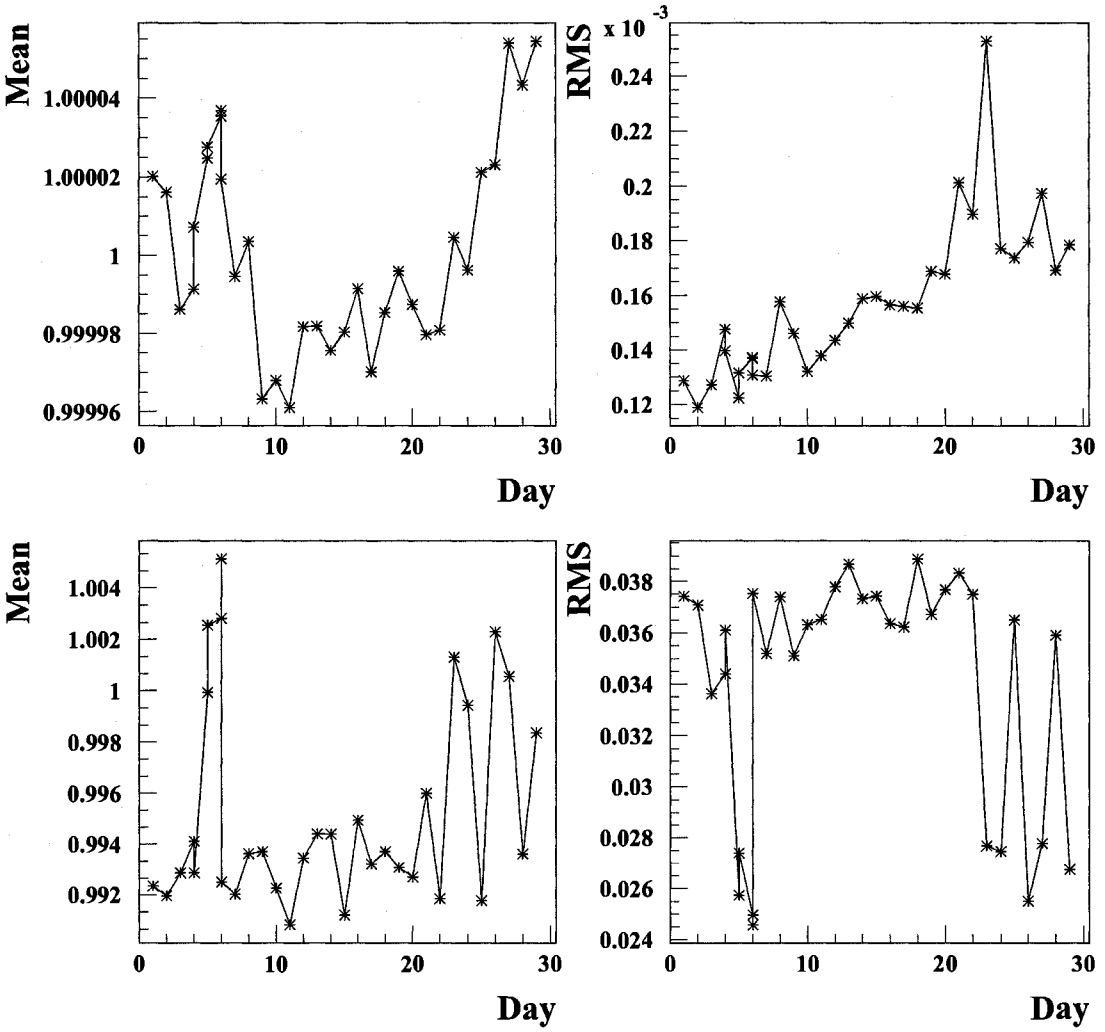


Figure 3.15: HEC pedestal mean and RMS ratio with respect to reference run versus day in the beam test (reference run 12100 - 17.08.2002). Top left: average ratio over all channels of the pedestal mean, top right: RMS spread of the pedestal mean ratio. Bottom left: average ratio over all channels of the pedestal RMS, bottom right: RMS spread on the pedestal RMS ratio.

cell in the cluster.

In order to determine an optimal value for the parameter n, the change in the number of cells in the cluster satisfying $E_{\text{cell}} > n\,\sigma_{\text{noise}}$ in EMEC was studied as a function of n (with n from 1 to 9). The results for electron beams at point J with 11 energies are shown in Figure 3.16. As expected, the number of selected cells in the cluster decreases with increasing n. For small electron energies, the number of cluster cells doesn't change significantly above $n = 2$. As the electron energies increase, the number of cells in the cluster becomes more sensitive to the variation of n, for n around 2. One would like to optimize the value of n in order to have as many cells with real signal in the cluster, but at the same time keep the added cluster noise as low as possible. A value of $n = 3$ seems to be a good first choice.

Let's consider $n = 3$ and look at the signal reconstruction for an electron beam of 119 GeV, using all EMEC clusters found by this method. The distribution obtained by summing the signals from all the cells in the clusters is shown in Figure 3.17. A Gaussian fit is done to the signal distribution to obtain the reconstructed signal, E, (the mean of the distribution) and the width σ_E of the distribution. The resolution of the signal is then calculated as σ_E/E. As can be seen, the signal follows well a Gaussian distribution.

One would like to minimize the signal resolution as a function of n. A study of the variation of the signal resolution as a function of n (with n from 1 to 9) was done for 11 electron beam energies, from 6 GeV to 148 GeV. The results are shown in Figure 3.18. As it can be seen, a minimum is obtained for $n = 3$ for most energies. A value of $n = 3$ is a good choice in order to reconstruct the signal accurately and minimize the amount of electronic noise in the cluster of energy. Similar values were used in the calibration of the H1 LAr detector at DESY, for both electrons and

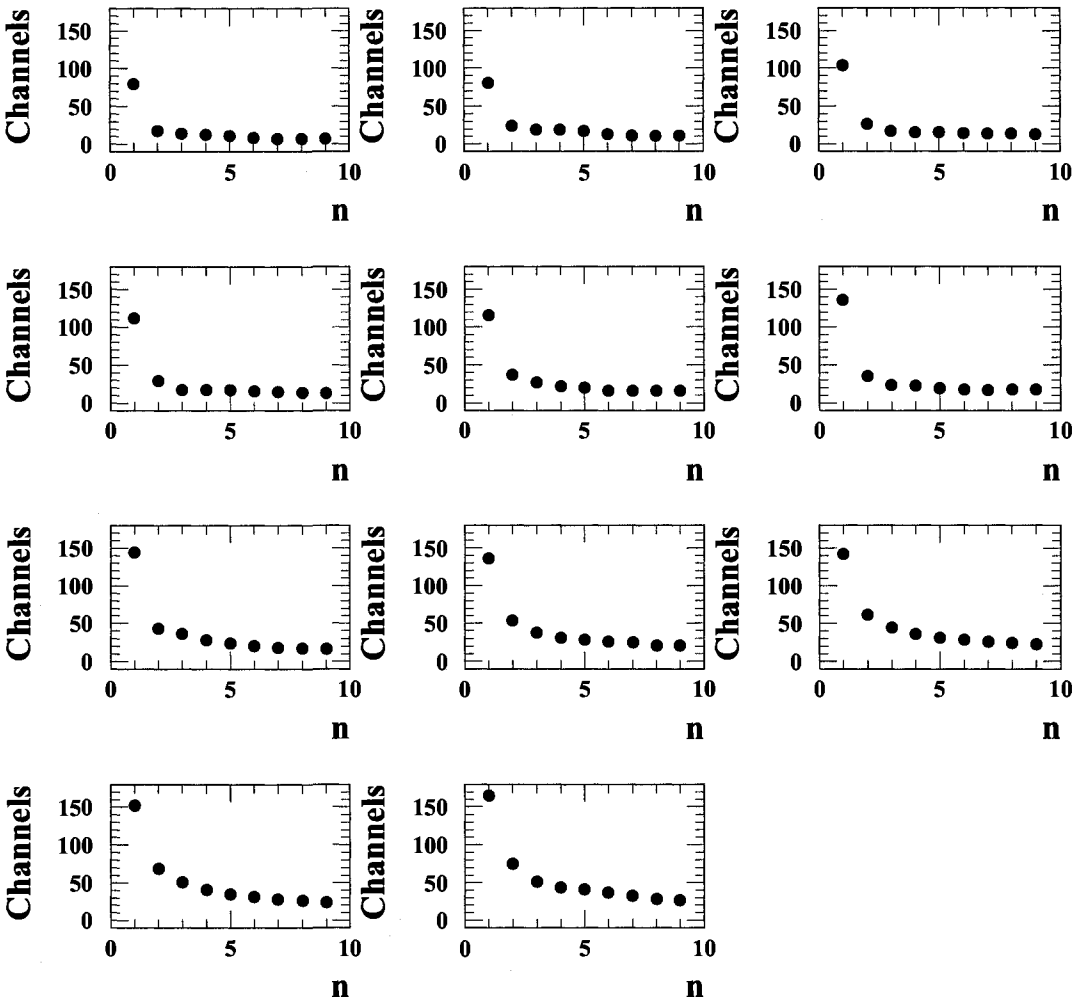


Figure 3.16: Variation of the number of cells in the cluster satisfying the requirement $E_{\text{cell}} > n\,\sigma_{\text{noise}}$, for electrons with 11 different energies, from 6 GeV to 148 GeV (6, 10, 20, 30, 40, 50, 60, 80, 100, 119 and 148 GeV) going from left to right, upper to lower rows, at impact point J.

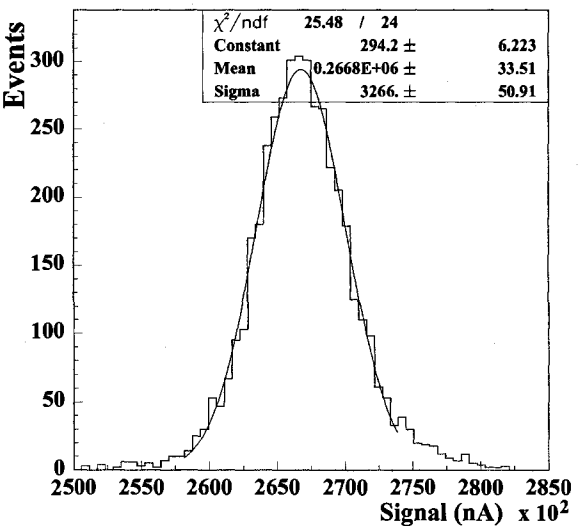


Figure 3.17: Reconstructed signal in nA for an electron beam of 119 GeV at impact point J. $E_{\text{cell}} > 3\,\sigma_{\text{noise}}$ used to reconstruct the signal.

pions [41].

Figure 3.19 shows the cells satisfying $E_{\text{cell}} > 3\,\sigma_{\text{noise}}$ cut of a typical 148 GeV electron event: in total there are three EMEC clusters (one per layer, with the exception of the PS) and no HEC clusters with 83 channels in total. The color code indicates the signal height of the individual readout cells. Due to the non pointing geometry of this beam test and the different clustering algorithm used, the cluster size obtained for electrons is bigger than the corresponding one used in the EMEC prototype module analysis which used a pointing geometry [31].

For a typical pion of 180 GeV, the corresponding distribution of selected cells is shown in Figure 3.20 and Figure 3.21: in total there are four clusters in the EMEC with 128 channels and three HEC clusters with 11 channels.

The clustering algorithm proves to be efficient for the reconstruction of the

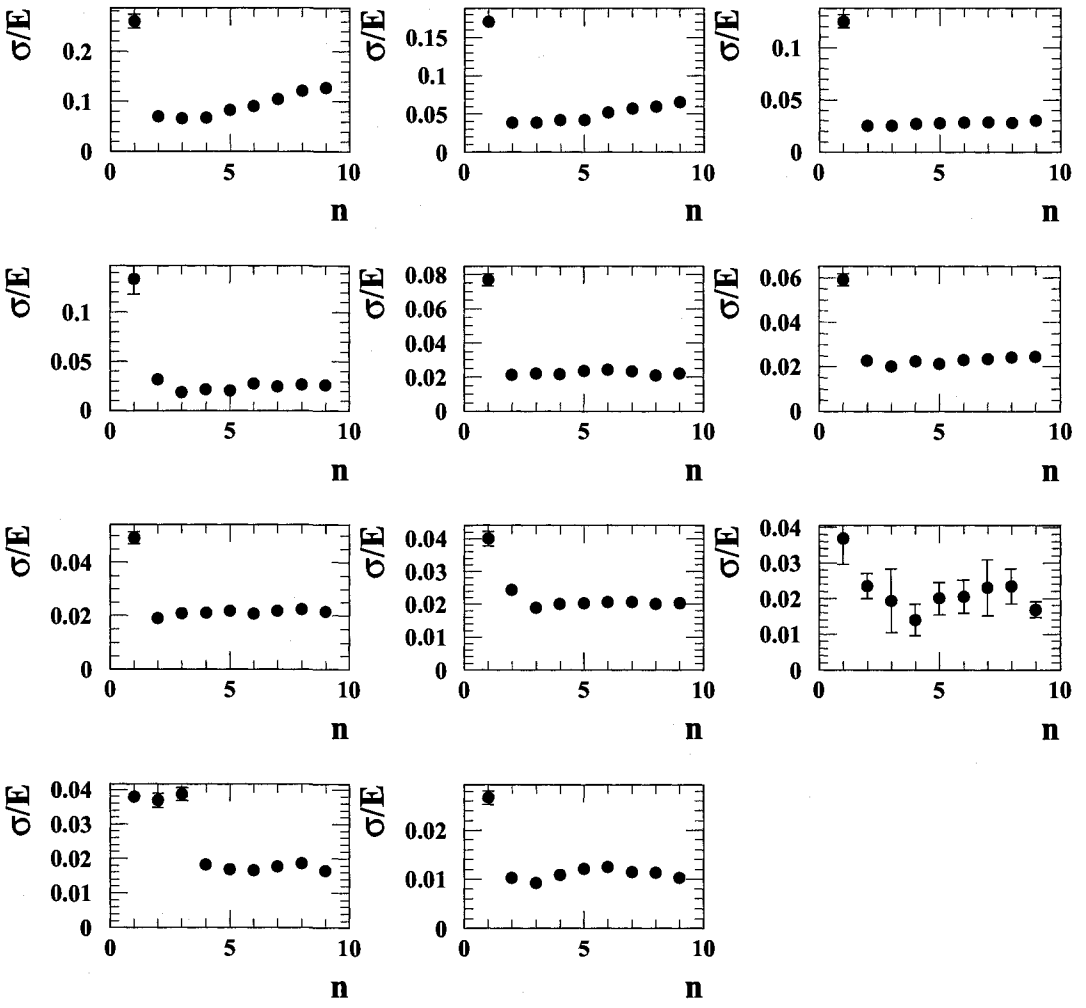


Figure 3.18: Variation of the resolution of the signal reconstructed in the cluster built by using the cells satisfying the requirement $E_{\text{cell}} > n\,\sigma_{\text{noise}}$, for electrons of 11 different energies, from 6 GeV to 148 GeV (6, 10, 20, 30, 40, 50, 60, 80, 100, 119 and 148 GeV) going from left to right, upper to lower rows, at impact point J.

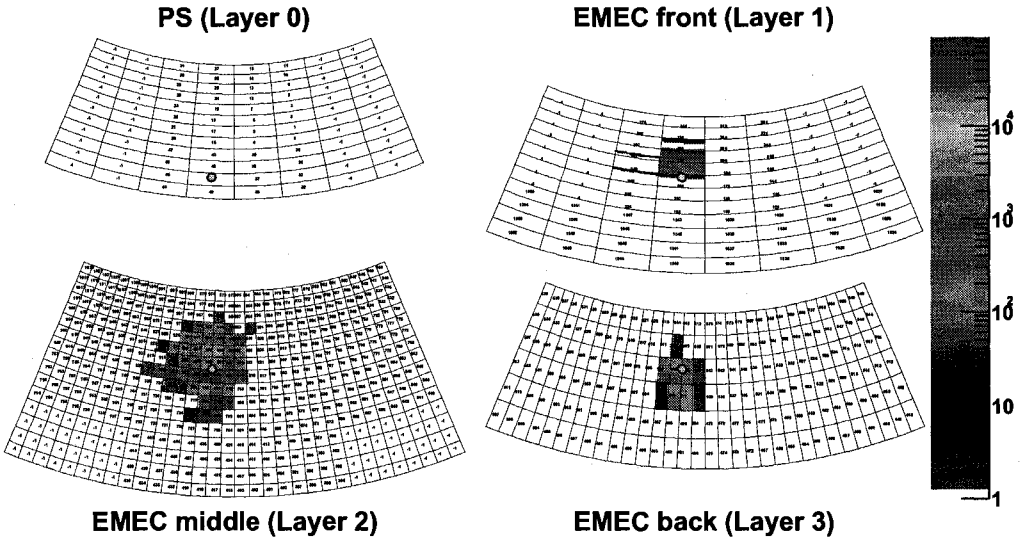


Figure 3.19: Reconstructed EMEC clusters for an electron of 148 GeV. The small circle shows the beam impact point; the scale is given in nA [37].

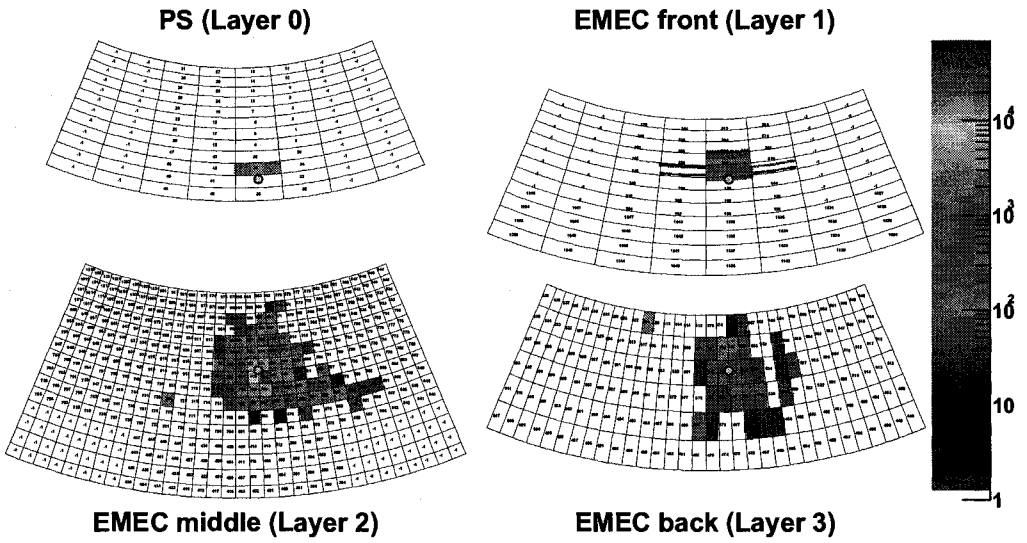


Figure 3.20: Reconstructed EMEC clusters for a pion of 180 GeV. The small circle shows the beam impact point; the scale is given in nA [37].

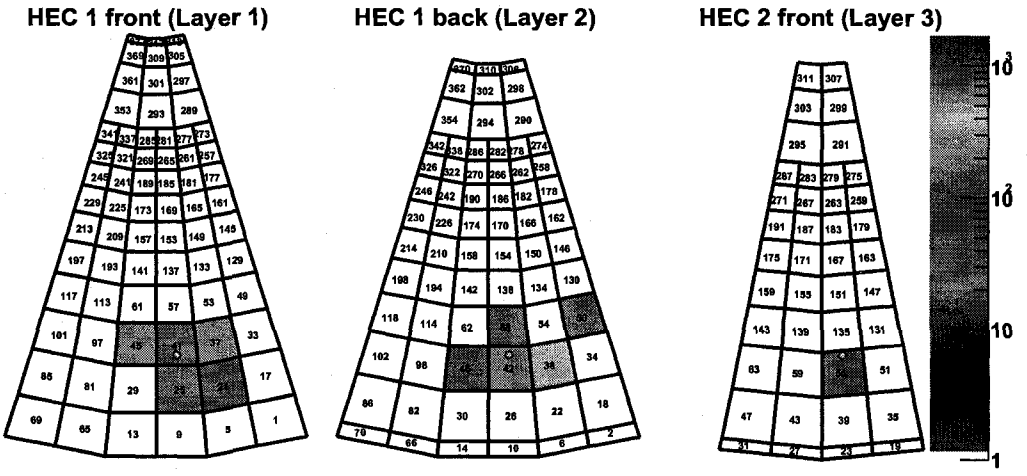


Figure 3.21: Reconstructed HEC clusters for a pion of 180 GeV. The small circle shows the beam impact point; the scale is given in nA [37].

particle energy deposited and minimizes noise in the reconstructed particle signal.

3.5 Electron energy reconstruction in EMEC

The understanding of the calorimeter response to electrons is crucial in many of the physics studies performed with the ATLAS detector. Available for this beam test were electrons with energies from 6 GeV to 148 GeV. By using the clustering algorithm described previously, one can measure the energy deposited in the calorimeter and hence reconstruct the initial energy of the electrons.

3.5.1 Corrections

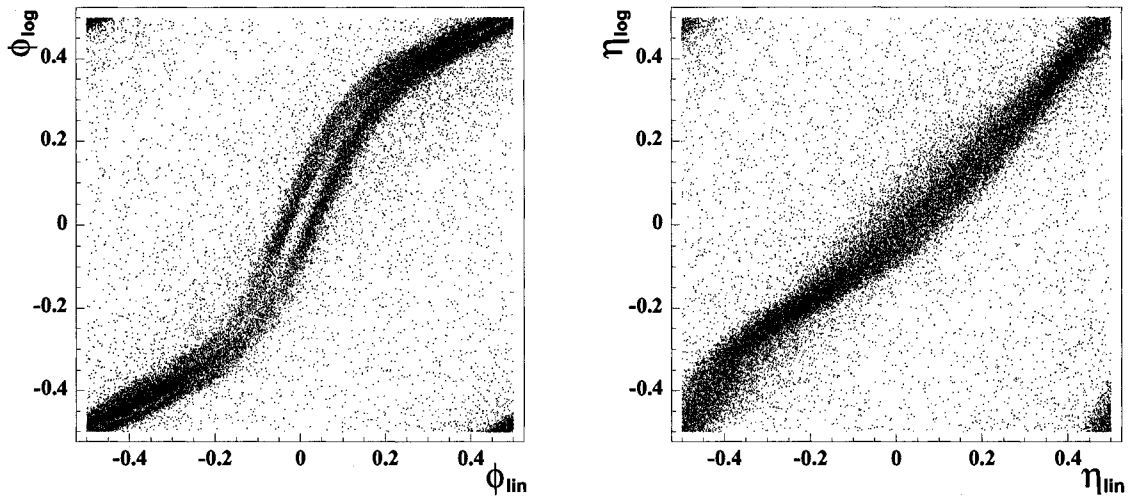
Due to the detector geometry, non-uniformities result in the energy reconstruction and need to be corrected. A significant ϕ variation of the signal response is expected due to the electric field and sampling fraction non-uniformities in the EMEC, and to the non-pointing geometry of the beam test. The influence of the point where the beam impacts on the readout cell has on the reconstructed signal in EMEC was assessed. For this purpose, the barycentre of the cluster was calculated by weighting the centre of each cell in the cluster by its signal contribution to the entire cluster. In this study only the cells in layer 2 of the EMEC were considered (most of the electron's energy is deposited in this layer). Linear and logarithmic weighting methods were used in order to calculate the weighted centre (given by ϕ_{lin} and ϕ_{log}), as described in Equations 3.3 and 3.4 [29]:

$$\phi_{\mathrm{lin}} = \frac{\sum_{\mathrm{i}=1}^{N} S_{\mathrm{i}} \phi_{\mathrm{c,i}}}{\sum_{\mathrm{i}=1}^{N} S_{\mathrm{i}}} \tag{3.3}$$

$$\phi_{\mathrm{log}} = \frac{\sum_{\mathrm{i}=1}^{N} [(ln \frac{S_{\mathrm{i}}}{S_{\mathrm{cluster}}} + 4) \phi_{\mathrm{c,i}}]}{\sum_{\mathrm{i}=1}^{N} (ln \frac{S_{\mathrm{i}}}{S_{\mathrm{cluster}}} + 4)} \tag{3.4}$$

where N is the number of cells in the cluster, S_{i} is the signal in cell i, S_{cluster} is the signal of the whole cluster and $\phi_{\mathrm{c,i}}$ is the ϕ of the centre of cell i. The offset of 4 in the weight was chosen to ensure that only cells which carry at least 2% of the cluster energy are included. Analogous formulas are used for the calculation of the η coordinate of the barycentre.

The dependence of the ϕ and η of the barycentre calculated using the logarithmic weight versus using the linear weight is shown in Figure 3.22. The cell units were obtained as $(\phi_{\mathrm{lin,log}} - \phi_{\mathrm{c}})/\Delta\phi$, where $\Delta\phi$ is the width of the cell, and the limits are ± 0.5 for edges, while 0 is for centre. This normalization eliminates the difference among the size of the cells. A typical "S" shape distribution is obtained [28].



(a) (b)

Figure 3.22: Correlation between ϕ (a) and η (b) of the energy barycentres using the logarithmic and the linear weighting method, as described in the text.

This type of shape is also found in other beam tests experiments when comparing the linearly reconstructed impact point positions with those measured more precisely using wire chambers upstream of the calorimeters [28]. The high concentration of reconstructed points around the cell's centre is due to the fact that for most of the impact points in between -0.25 and $+0.25$, the linear weighting scheme considers the deposited energy as being located at 0 and attributes the cell's centre the highest weight. This is less the case with the logarithmic weights, which makes their choice for the analysis preferable.

Figure 3.23 shows the variation of the reconstructed signal for 119 GeV electrons versus the normalized logarithmically weighted ϕ and η inside a cell. As can be seen, the distribution is not flat, but it has a number of minima and maxima. For example, the reconstructed energy variation in ϕ is on the order of $\pm 2\,\%$ for 119 GeV electrons, depending on the impact point inside the cell. The η distribution

is smoother with a variation smaller than in the ϕ distribution case, due to the limited η range accessible in the beam test setup and the non-pointing geometry (where the beam spread covers a larger η range than in the pointing geometry).

In order to correct for this effect, two methods were investigated. First, an analytical fit (a combination of three sinusoidal functions) was performed on the data. The second method used a bin by bin numerical correction for the signal. For each ϕ bin, a multiplicative constant was calculated as the ratio of the bin value to the average signal over the entire η range. This constant then was used to multiply the signal for each event falling in the respective ϕ bin. For example, after applying a bin by bin correction for the logarithmic method the fluctuation is greatly reduced as it can be seen in Figure 3.24.

The effect the two corrections (for each of the two types of correction schemes) have on the signal resolution was studied. As it can be seen in Figure 3.25, using the corrections applied to the logarithmic weighting method gives better resolutions for all ϕ bins. The bin by bin correction improves the signal resolution by as much as 10 % and is therefore used in this analysis.

3.5.2 Calculation of the electromagnetic constant α_{em}

Because the signal in the LAr calorimeters is a current in nA, one needs to convert it to the energy of the particle that generated that signal (in GeV). This conversion factor, α_{em} in GeV/nA, can be obtained in beam tests in which the particle types and their energies are well known.

For the current study, electrons with 11 energies between 6 GeV and 148 GeV at impact point J were used to find the EMEC electromagnetic constant, α_{em}^{EMEC}. For

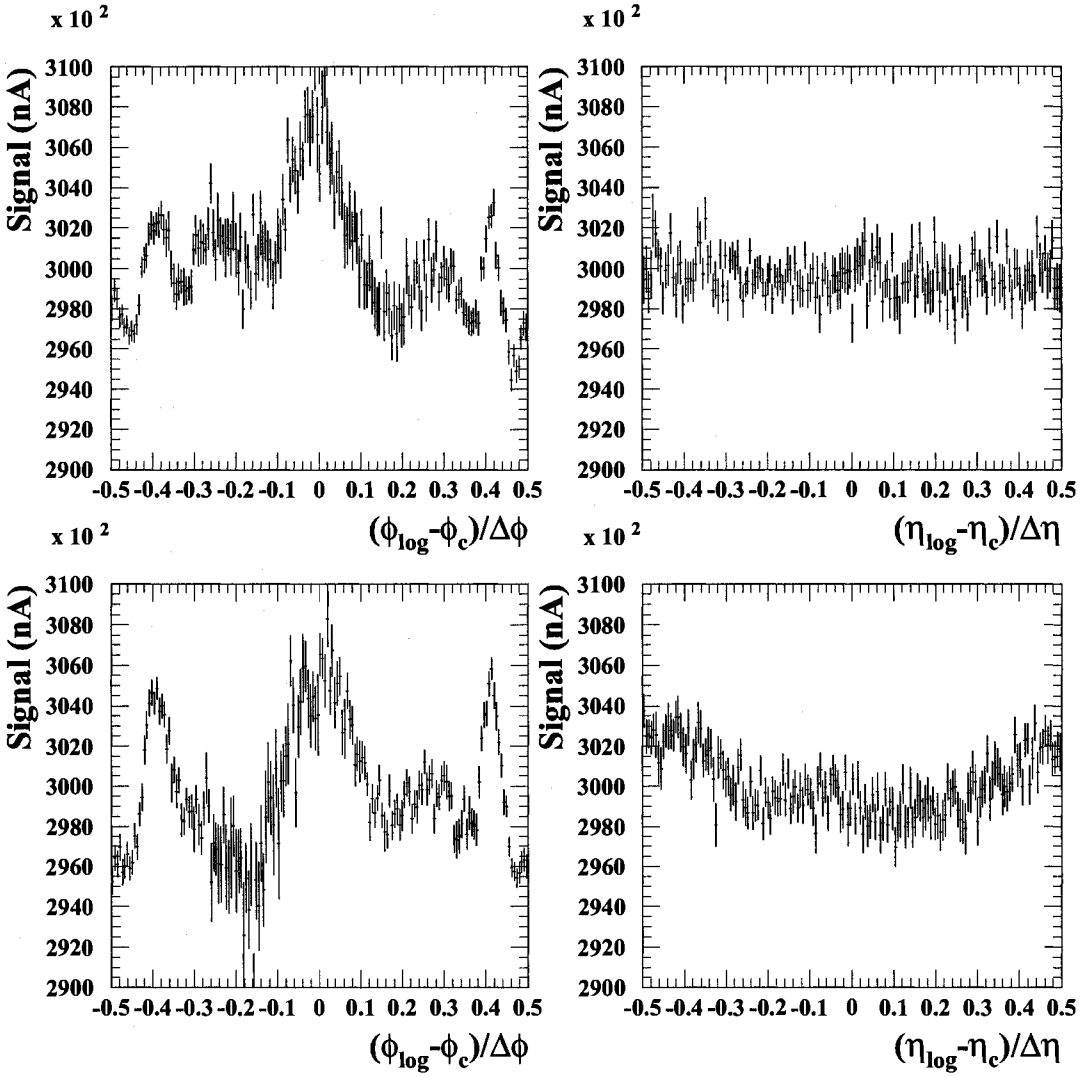
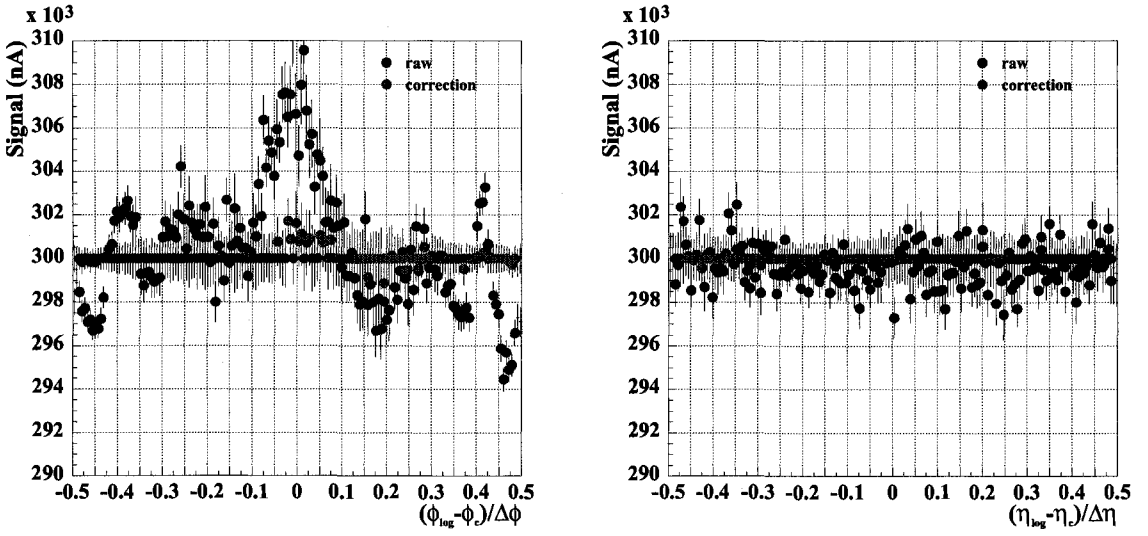


Figure 3.23: Signal distribution of an electron beam of 119 GeV versus the normalized logarithmic weighted ϕ and η for impact point J. The upper plots are averaged over impact points: C, D, E and F, while the lower plots are averaged over G, H, I, J and K (see Figure 3.6).



(a) (b)

Figure 3.24: Signal distribution versus the normalized ϕ (a) and η (b) before and after the bin-by-bn corrections are applied for the logarithmic weighting method - 119 GeV electrons at impact point J.

each beam energy, the signal in nA was reconstructed using the clustering algorithm previously described (Section 3.4), after the corrections discussed in Section 3.5.1 were applied. The electromagnetic constant is calculated as the beam energy divided by the reconstructed ionization current in the cluster. Figure 3.26 (a) shows the variation of the α_{em}^{EMEC} with the beam energy. The weighted average of α_{em}^{EMEC} over the entire beam energy range is 0.449 MeV/nA. For lower values of the beam energy α_{em}^{EMEC} has larger values than for the higher energies. This effect can be explained by the fact that at lower beam energies, the proportion of the ionization current in the cluster from the total ionization current is smaller than for higher energies. As it can be seen in Figure 3.26 (b), for lower energies, the reconstructed energy represents about 85 % of the beam energy. Considering that for higher energies most of the beam energy is reconstructed in the cluster, the last two energies, 119 GeV and 148 GeV were used

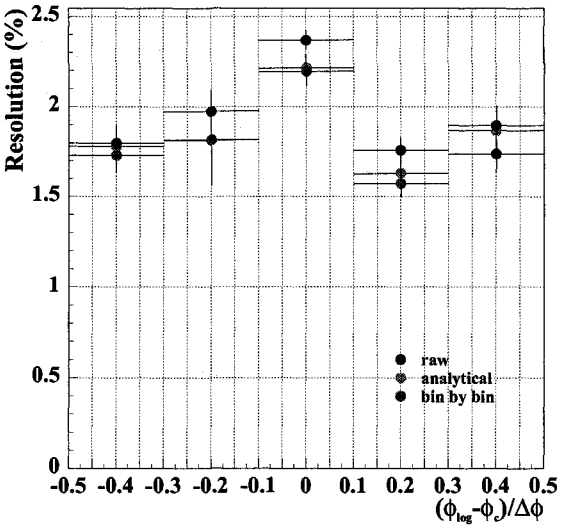
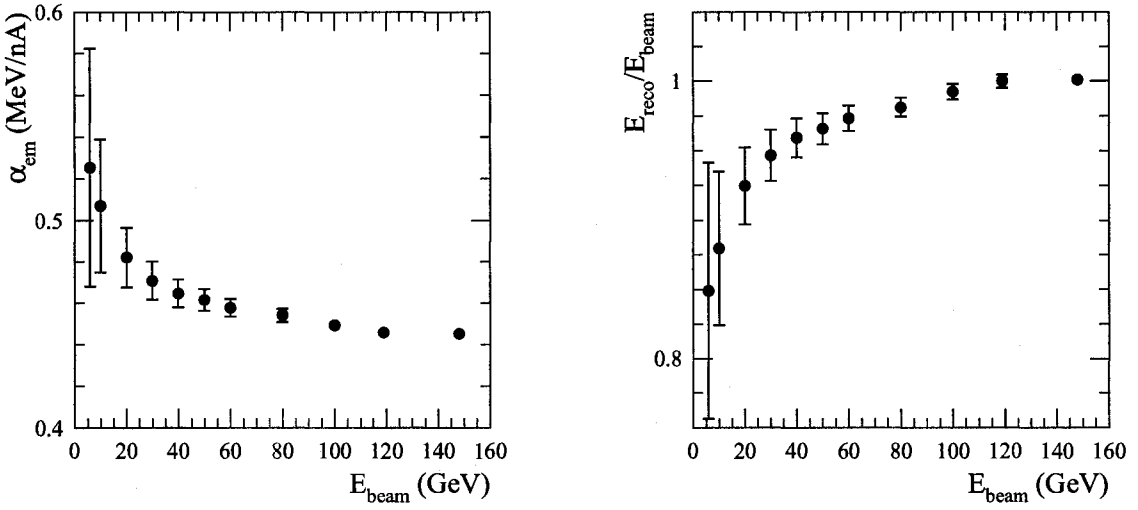


Figure 3.25: Resolution is improved in all ϕ bins after the bin-by-bin and analytical corrections are applied for the logarithmic weighting methods - 119 GeV electrons at impact point J.



(a) (b)

Figure 3.26: (a) Dependence of the α_{em}^{EMEC} on the beam energy at the impact point J. The α_{em}^{EMEC} at 119 GeV and 148 GeV are used to compute the overall $\alpha_{em}^{EMEC} = (0.446 \pm 0.009)$ MeV/nA. (b) Deviation from 1 of the ratio between the reconstructed energy using the electromagnetic constant α_{em}^{EMEC} and the known beam energy. This deviation is attributed to the energy leakage outside the reconstructed cluster.

to calculate the overall $\alpha_{em}^{EMEC} = (0.446 \pm 0.009)$ MeV/nA. The error is dominated by the systematic contribution evaluated at 2% from the spread of the distribution of α's for the full energies range. The obtained value of the α_{em}^{EMEC} is in agreement with the value published in [37] of (0.430 ± 0.009) MeV/nA and is compatible with the value obtained from EMEC stand alone beam tests of 0.426 MeV/nA [34].

3.5.3 Energy resolution

As described in Section 2.2.2, the energy resolution is an important characteristic of any calorimeter. To investigate the energy resolution of the EMEC, elec-

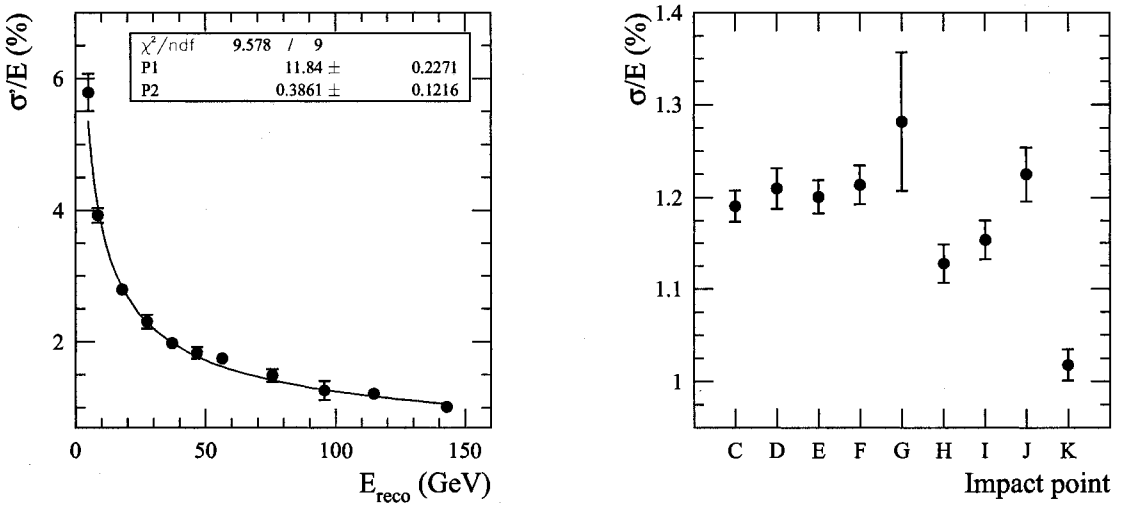
trons/positrons with energies between 6 GeV and 148 GeV at impact point J were used. Also, various impact points were used in order to study the spatial uniformity of the energy resolution. The clustering algorithm previously described was used to calculate the total signal collected in each EMEC layer. After the corrections in Section 3.5.1 were applied, the resulting signal (in nA) was converted in deposited energy (in GeV) by using the electromagnetic constant calculated in Section 3.5.2.

A Gaussian fit to the reconstructed energy distribution for all events in a run is performed to obtain the beam reconstructed energy, $E_{reco} = E_{peak}$, (the mean of the Gaussian) and the width $\sigma(E_{reco})$. The fit is performed for the range $E_{peak} - 2\sigma$ to $E_{peak} + 2\sigma$ and the resolution is calculated as $\sigma(E_{reco})/E_{reco}$ for each beam energy, E_{beam}. Figure 3.27 (a) shows the EMEC intrinsic energy resolution (noise subtracted in quadrature from σ to obtain the intrinsic resolution) versus the beam energy. The shape is described well by the general formula:

$$\frac{\sigma'(E_{reco})}{E_{reco}} = \frac{\sigma(E_{reco})}{E_{reco}} \ominus \frac{\sigma_{noise}(E_{reco})}{E_{reco}} = \frac{a}{\sqrt{E_{reco}}} \oplus b \qquad (3.5)$$

where σ_{noise} is calculated from the noise contributions of the cells included in the cluster. By fitting the distribution with this function, the following values are obtained for the coefficients: $a = (11.8 \pm 0.2)\,\%\sqrt{\text{GeV}}$ and $b = (0.4 \pm 0.1)\,\%$. These results agree well with the results published in [37] of $a = (12.1 \pm 0.2)\%\sqrt{\text{GeV}}$ and $b = (0.4 \pm 0.1)\,\%$, and with the values obtained previously in a pointing geometry stand alone EMEC test that produced a sampling term of about $a = 11.9\,\%\sqrt{\text{GeV}}$ and a constant term of about $b = 0.4\,\%$.

The variation of the energy resolution with the impact points from C to K (see Figure 3.6) was also studied, as a source of systematic uncertainty due to the positioning of the beam on the calorimeter modules. As it is shown in Figure 3.27 (b), the resolution is uniform to about 0.1 % for all impact points, when considering an



(a) (b)

Figure 3.27: (a) Electron intrinsic energy resolution σ'/E versus reconstructed energy for impact point J. The continuous line represents the fit with $\frac{\sigma'(E)}{E} = \frac{a}{\sqrt{E}} \oplus b$. (b) The electron energy resolution versus impact point for an electron beam of 119 GeV is uniform to about 0.1 %.

electron beam of 119 GeV (except at point K, for which some of the cells surrounding it were not connected to the readout, resulting in lateral energy leakage outside the cluster).

3.6 Pion energy reconstruction in EMEC and HEC

In order to get a measure of the total energy deposited by hadrons, besides an electromagnetic calorimeter one also needs a hadronic calorimeter. In the past, beam tests were performed on the EMEC and the HEC separately, to understand their response to electrons and pions. Their properties were studied considering that the electromagnetic/hadronic showers were almost fully contained in the tested modules. However, in the case of high energies hadrons in the end-caps of ATLAS, after depositing a significant part of their energy in the EMEC, they deposit the rest of their energy in the HEC. A typical distribution of energies of pions in the EMEC and HEC is presented in Figure 3.28 which shows the sharing of energy amongst these two calorimeters. The procedure of combining the two energy depositions is not trivial. In this section, a method of combining the signals from the two components is investigated, that provides the hadronic energy scale.

3.6.1 HEC high voltage corrections

One of the 4 high voltage lines, which are feeding the four sub-gaps of the second longitudinal segment of the central HEC1 module, had to be disconnected because of a short circuit. As the signal was still measured correctly in three out of the

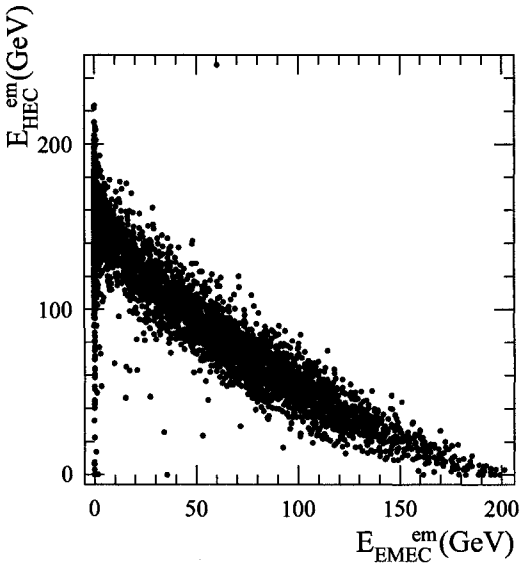


Figure 3.28: Sharing between EMEC and HEC of the reconstructed energy at the electromagnetic scale for 200 GeV pions at impact point J.

four sub-gaps, a correction factor of 4/3 has been applied to those data. The main consequence due to this high voltage short circuit is a reduced signal to noise ratio.

There were also some problems with the high voltage supply to the second HEC layer, that became obvious when analyzing the ratios of the energies deposited in the three layers of the HEC when pions scans were performed. In particular, it was noticed that pions that started showering early in the detector gave a smaller than usual signal. This effect can be seen in Figure 3.29 when looking at the ratio of the average signal collected in the second longitudinal layer of the central module over the one in the outer (in ϕ there are 3 HEC modules) modules (denoted R in the figure) as a function of the asymmetry of the total energy deposited in the first and last longitudinal segments (denoted D in the figure). This problem could be partially traced back to bad ground connections of the electrostatic transformer boards in this region of the central HEC module. In order to correct for this problem, a software

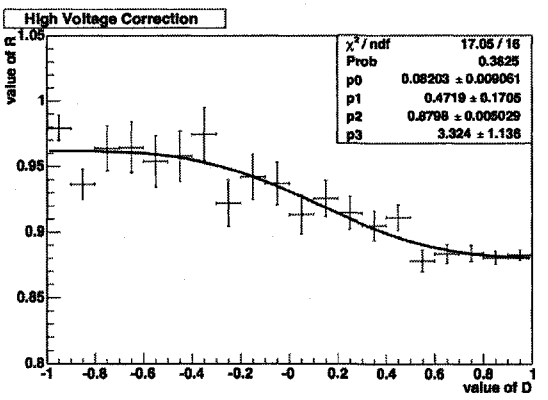
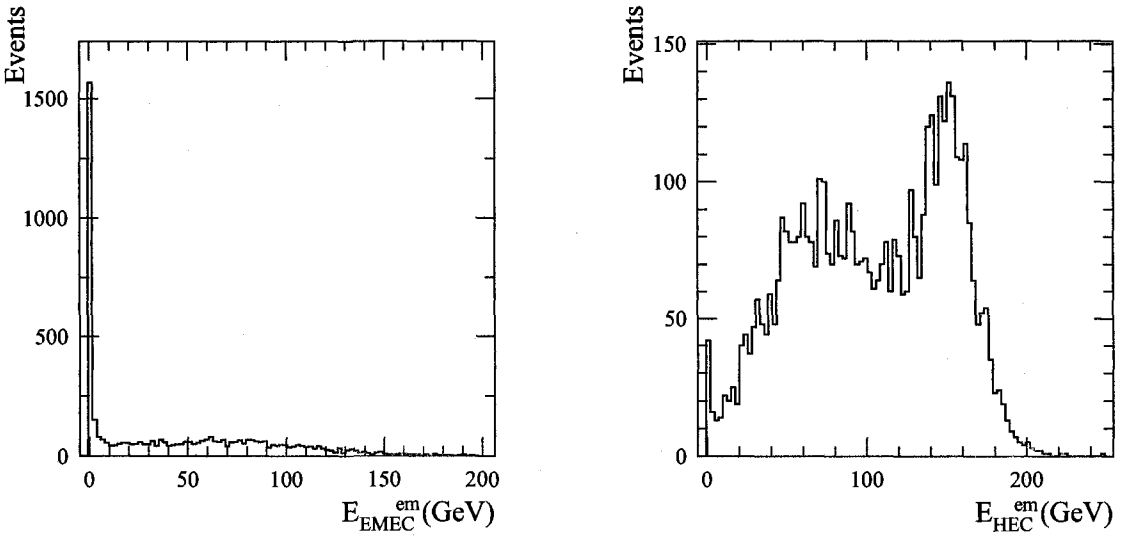


Figure 3.29: Ratio of the average signal collected in the second longitudinal layer of the central module over the one in the outer modules (R) as a function of the asymmetry of the total energy deposited in the first and last longitudinal segments (D) [37]. A third degree polynomial fitting is done and the resulting function is used to correct for the signal in the affected HEC cells.

correction is implemented. First, the distribution is fitted with a third degree polynomial and the fitting parameters are obtained. Then, event by event, the signal of cells in the central middle (second) layer in HEC are weighted by the 1/R fitting function at the point D calculated for that event.

3.6.2 Electromagnetic scale

The calorimeters respond differently to pions and electrons due to different physics processes involved in the electromagnetic and hadronic interactions (Section 2.2.2). The EMEC response to electrons was studied in Section 3.5.2 and an electromagnetic constant $\alpha_{em}^{EMEC} = (0.446 \pm 0.009)\,\mathrm{MeV/nA}$ was obtained. Previously, stand alone HEC beam tests with electrons calculated a $\alpha_{em}^{HEC} = (2.93 \pm 0.03)\,\mathrm{MeV/nA}$ [35]. These earlier tests used different calibration pulses/electronics. Therefore, a correc-



(a) (b)

Figure 3.30: Reconstructed energy at the electromagnetic scale in the EMEC (a) and HEC (b) for 200 GeV pions at impact point J.

tion for the new calibration pulse shape had to be applied for the 2002 beam test, yielding a value of $\alpha_{em}^{HEC} = (3.27 \pm 0.07)$ MeV/nA [36] (the error includes both statistical and systematic errors). If one tries to reconstruct the energy of pions in the same way as for electrons then one gets an energy that is much smaller than the beam's energy. In effect, the hadronic interactions are reconstructed as if they were electromagnetic, called the electromagnetic (em) scale. As an example, Figure 3.30 (a) and (b) show the electromagnetic scale (em-scale) reconstructed energy in the EMEC and HEC respectively, for 200 GeV pions at impact point J. After summing this em-scale signal in the EMEC and HEC, event by event, a Gaussian like distribution is obtained as it can be seen in Figure 3.31. The reconstructed energy at the electromagnetic scale is about 150 GeV for beams of 200 GeV pions. In order to properly reconstruct the hadrons' energy, one needs to correct to the hadronic scale. This procedure will be discussed in the next section.

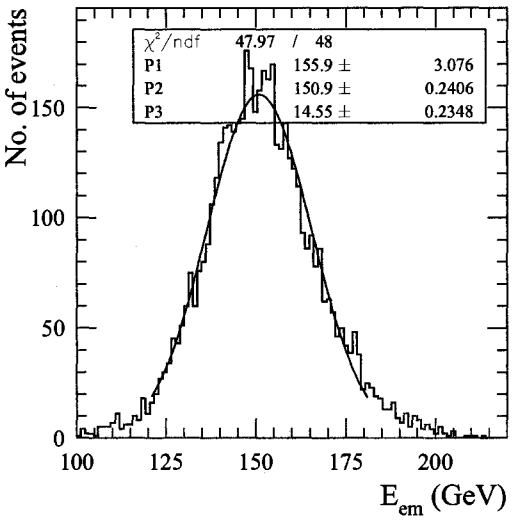


Figure 3.31: Reconstructed energy at the electromagnetic scale in EMEC and HEC for 200 GeV pions at impact point J. The summation is done event by event.

3.6.3 Hadronic scale

In order to correct for this discrepancy between the em-scale and the beam energy, as shown in Figure 3.31, a weighting method was applied to the energy at the electromagnetic scale. The total energy of the particles is calculated as the weighted sum of the em scale energy deposited in the EMEC and HEC as in the following formula:

$$E^{\text{reco}} = E^{\text{reco}}_{\text{EMEC}} + E^{\text{reco}}_{\text{HEC}} = w_{\text{EMEC}} E^{\text{em}}_{\text{EMEC}} + w_{\text{HEC}} E^{\text{em}}_{\text{HEC}} \tag{3.6}$$

where $w_{\text{EMEC,HEC}}$ are the EMEC and HEC weights.

The simplest method is to consider the weights as constants; $w_{\text{EMEC}} = C_E$, $w_{\text{HEC}} = C_H$. In this situation one can write:

$$E^{\text{reco}} = C_E E^{\text{em}}_{\text{EMEC}} + C_H E^{\text{em}}_{\text{HEC}} = C_E (E^{\text{em}}_{\text{EMEC}} + \frac{C_H}{C_E} E^{\text{em}}_{\text{HEC}}) \tag{3.7}$$

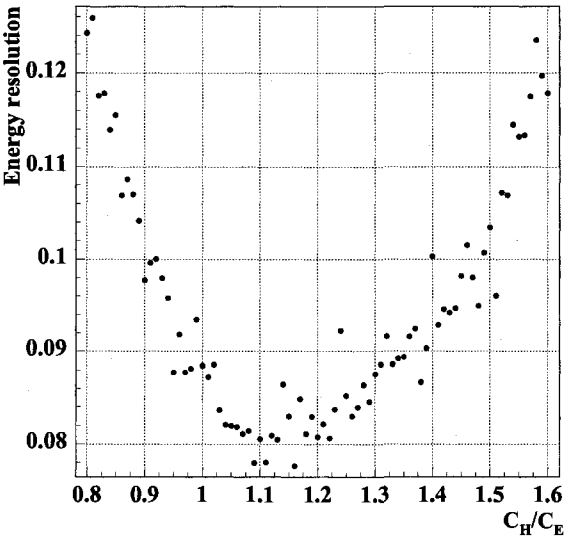


Figure 3.32: Variation of the reconstructed energy resolution with the ratio $\frac{C_H}{C_E}$ (as given in Equation 3.7) for pions of 200 GeV at impact point J.

where C_E and C_H are two constants that are obtained by minimizing the resolution of the reconstructed energy. Figure 3.32 shows the variation of the energy resolution with $\frac{C_H}{C_E}$ for 200 GeV pions. A minimum is obtained for $\frac{C_H}{C_E} \approx 1.15$ and a value of $C_E = 1.38$ is found. The calibration weights found in this way depend on the type and energy of the particle, which is not desirable for a system that has to measure a large range of energies and particle types.

Due to the different physics processes involved in the development of the electromagnetic and hadronic showers, the local density of energy deposits is different (high for the electromagnetic and lower for the hadronic component). The energy density in the EMEC and HEC clusters (in GeV/cm^3) is computed as the electromagnetic scale energy, E^{em},deposited in the cluster (in GeV) over the total volume, V, of the cells in the cluster (typically $20,000\,\text{cm}^3 = 0.02\,\text{m}^3$ for EMEC clusters and

about 10 times more for HEC clusters): $\rho_{\mathrm{EMEC,HEC}} = E^{\mathrm{em}}_{\mathrm{EMEC,HEC}}/V_{\mathrm{EMEC,HEC}}$. Using these properties, one can use weights that depend on the energy density measured in clusters of energy in both EMEC and HEC. It was observed in previous experiments [41] that the weights depend on the energy density in the cluster and have the following form:

$$w_{\mathrm{EMEC}} = C_1 \exp\left(-C_2 E^{\mathrm{em}}_{\mathrm{EMEC}}/V_{\mathrm{EMEC}}\right) + C_3 \tag{3.8}$$

$$w_{\mathrm{HEC}} = C_4 \exp\left(-C_5 E^{\mathrm{em}}_{\mathrm{HEC}}/V_{\mathrm{HEC}}\right) + C_6 \tag{3.9}$$

where $E^{\mathrm{em}}_{\mathrm{EMEC,HEC}}$ refers to the EMEC, HEC cluster energy at the electromagnetic scale, V to the corresponding cluster volume and C_{1-6} are six parameters that have to be determined.

Using these parametric forms for the two weights, a procedure is needed to extract the six parameters. A χ^2-fit was used to determine the parameters C_i:

$$\chi^2 = \sum_{\mathrm{events}} \frac{\left[E_{\mathrm{beam}} - E_{\mathrm{leak}} - E^{\mathrm{reco}}_{\mathrm{EMEC}}(C_1, C_2, C_3) - E^{\mathrm{reco}}_{\mathrm{HEC}}(C_4, C_5, C_6)\right]^2}{\left(\sigma^{\mathrm{reco}}_{\mathrm{noise}}\right)^2 + \left(\sigma^{\mathrm{leak}}_{\mathrm{noise}}\right)^2} \tag{3.10}$$

where E_{leak} is the energy leaked outside the reconstructed cluster, $\sigma^{\mathrm{leak}}_{\mathrm{noise}}$ is the noise contribution from the readout cells that are left outside the cluster and $\sigma^{\mathrm{reco}}_{\mathrm{noise}}$ is the integrated noise in the related cluster. The quantity $\sigma^{\mathrm{leak}}_{\mathrm{noise}}$ is given by the quadratic sum of the noise values on the electromagnetic scale of the calorimeter readout cells not included in the clusters. This value has been found to be rather independent of the cluster size and of the beam energy and therefore it was set to the average value of 5 GeV [37]. The contribution from $\sigma^{\mathrm{reco}}_{\mathrm{noise}}$, calculated as the quadratic sum of the noise of the cells included in the cluster, is instead dependent on the cluster's size. The values for E_{leak} were taken from the [37], e.g. for 200 GeV pions there is a 5 GeV lateral leakage outside the cluster and a 10 GeV longitudinal leakage outside the detector boundaries.

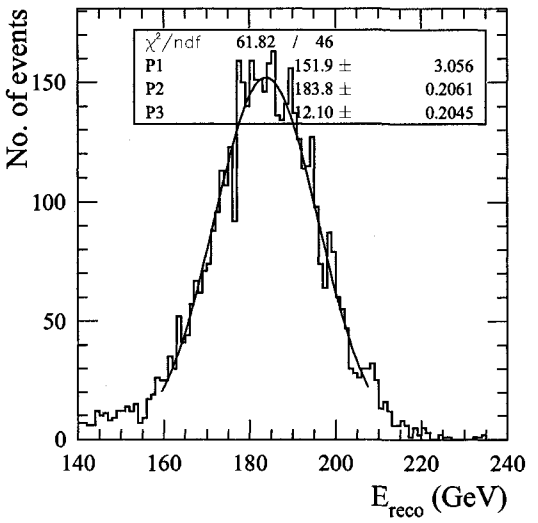


Figure 3.33: Reconstructed energy using the described weighting scheme for pions of 200 GeV at impact point J.

By minimizing the χ^2, typical values for the coefficients are obtained e.g.: $C_1 = 0.70$, $C_2 = 1000\,\mathrm{cm}^3/\mathrm{GeV}$, $C_3 = 0.98$, $C_4 = 0.83$, $C_5 = 1550\,\mathrm{cm}^3/\mathrm{GeV}$ and $C_6 = 1.02$. After that, the energy is reconstructed event by event using the calculated coefficients as described in Equation 3.6, where the weights are calculated from Equations 3.8 and 3.9 for each event.

As it can be seen in Figure 3.31, the electromagnetic scale reconstructed energy is about 150 GeV for pions of 200 GeV at impact point J. The result of applying the weighting scheme described above is shown in Figure 3.33. As can be seen, the reconstructed energy is now about 185 GeV, excluding the 15 GeV due to the leakage outside the cluster and the detector modules.

In order to assess the behavior of the weights obtained from the minimization procedure, one can also calculate the weights as the ratio of the beam energy minus

the leakage and the reconstructed energy in the other module (EMEC or HEC) over the reconstructed energy on the em scale, as in:

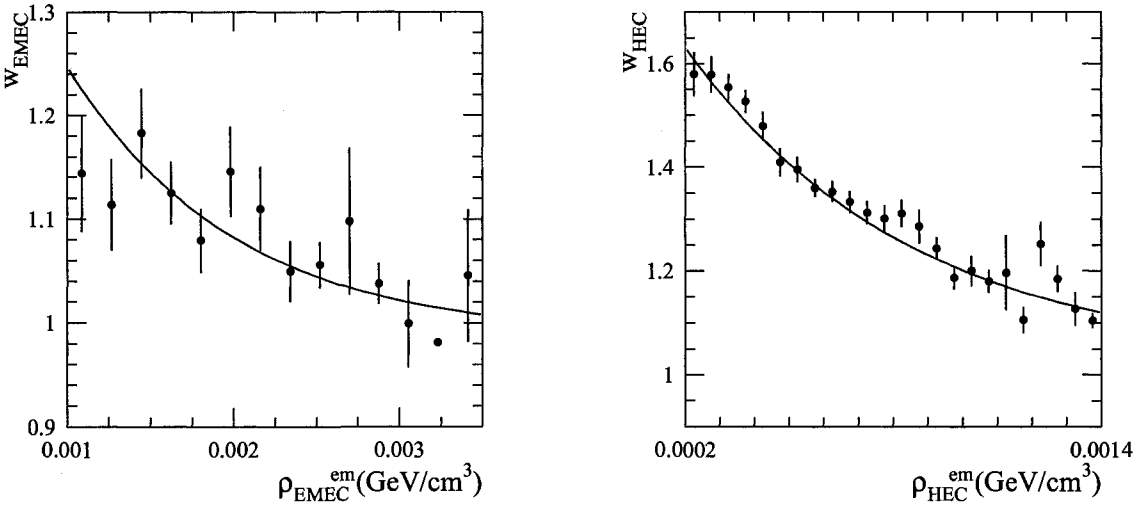
$$w_{\mathrm{EMEC}} = \frac{E_{\mathrm{beam}} - E_{\mathrm{leak}} - E^{\mathrm{reco}}_{\mathrm{HEC}}}{E^{\mathrm{em}}_{\mathrm{EMEC}}} \tag{3.11}$$

$$w_{\mathrm{HEC}} = \frac{E_{\mathrm{beam}} - E_{\mathrm{leak}} - E^{\mathrm{reco}}_{\mathrm{EMEC}}}{E^{\mathrm{em}}_{\mathrm{HEC}}} \tag{3.12}$$

with $E^{\mathrm{reco}}_{\mathrm{EMEC,HEC}}$ calculated using the $C_1 \rightarrow C_6$ parameters determined earlier. Figure 3.34 (a) shows the distribution of w_{EMEC} (calculated as the energies ratio) as a function of the cluster energy density in the EMEC. As can be seen, the curve obtained using the parameterization of Equation 3.8 overlaps well over the data obtained using the energies ratios (Equation 3.11). The same behaviour is seen in Figure 3.34 (b) for the w_{HEC} weight as a function of the cluster energy density in the HEC.

Because of the rather strong correlation between C_1 and C_2 in the EMEC and C_4 and C_5 in the HEC, the influence of varying the C_2 and C_5 coefficients (appearing in the exponential term of the weights) on the reconstruction energy was studied. By keeping the other coefficients fixed, the values of C_2 and C_5 were scanned (using MINUIT [42]) so that the change in χ^2 is no more than 1 unit. Figure 3.35 and Figure 3.36 show the ranges over which the two coefficients can be changed, while keeping the change in χ^2 under 1. As it can be seen C_2 is between 800 and 1200, while C_5 is between 1490 and 1590. The C_2 and C_5 parameters have been fixed to constant values: $C_2 = 1000\,\mathrm{cm}^3/\mathrm{GeV}$ for EMEC and $C_5 = 1500\,\mathrm{cm}^3/\mathrm{GeV}$ for HEC.

The cluster weighting method described above works well for beam tests in which single particles of known type and energy are used. In the case of the ATLAS detector, the calibration of the calorimeters based on jet interactions is required for the hadronic calibration. Due to the complex interaction environment in which many particles overlap, the difference between the electromagnetic and hadronic en-



(a) (b)

Figure 3.34: (a) Comparison for the EMEC weights versus the energy density in the EMEC clusters obtained with Equation 3.11 (dots) and Equation 3.8 (line) for pions of 200 GeV at impact point J. (b) Comparison for the HEC weights versus the energy density in the HEC clusters obtained with Equation 3.12 (dots) and Equation 3.9 (line) for pions of 200 GeV at impact point J.

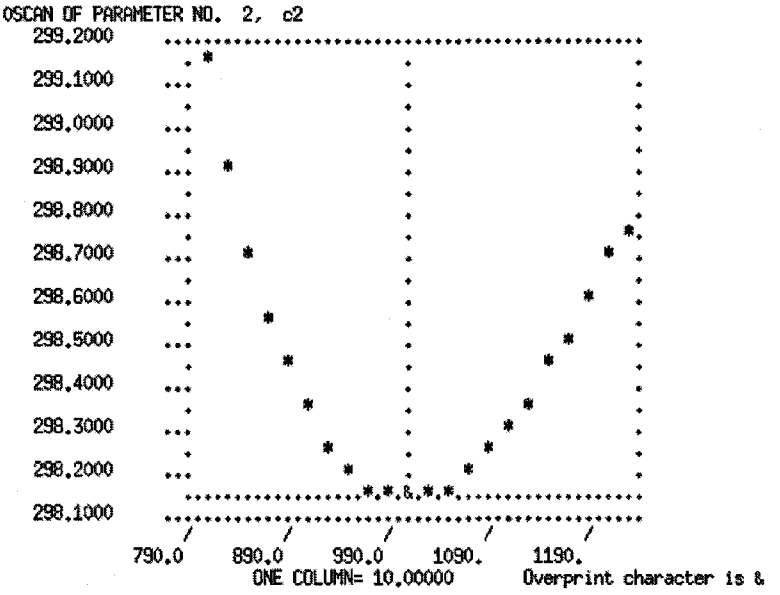


Figure 3.35: χ^2 values as the C_2 coefficient is scanned from 800 to 1200, while keeping the other 5 coefficients fixed.

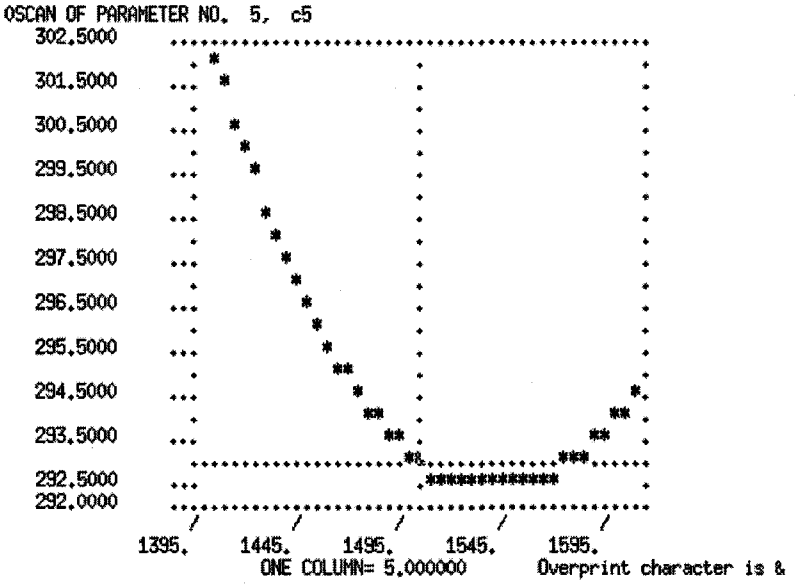
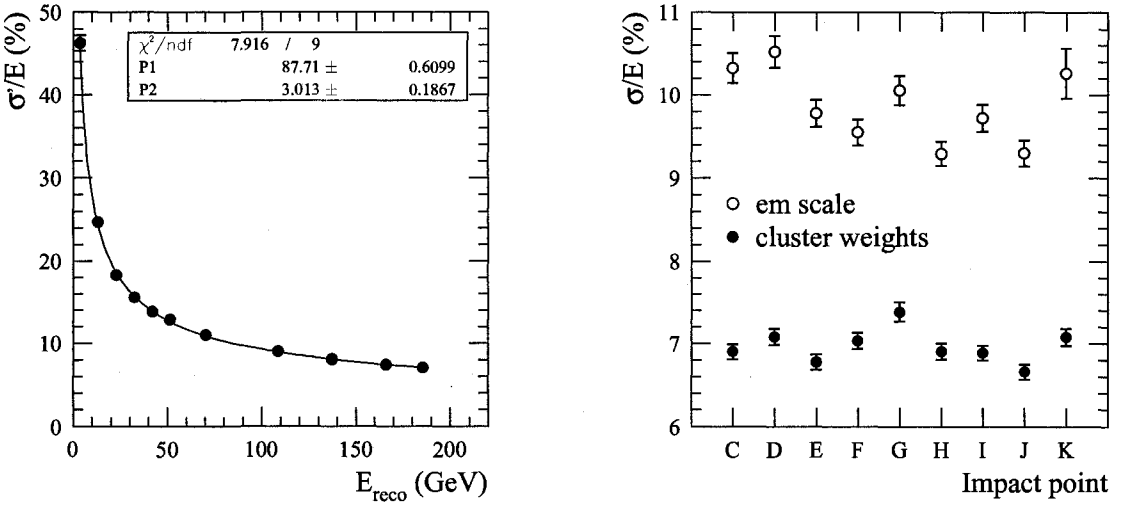


Figure 3.36: χ^2 values as the C_5 coefficient is scanned from 1400 to 1600, while keeping the other 5 coefficients fixed.

ergy deposition based on cluster energy density method is considerably reduced. To overcome this impediment, a new approach using the readout cell level weights is under development [37]. These weights are obtained from Monte Carlo simulations of the calorimeters for which the energy depositions in the active and inactive materials are known at the readout cell level.

3.6.4 Energy resolution

Using the cluster weighting scheme presented above, the reconstruction of energies for 11 pion beams of different energies was performed. The intrinsic energy resolution for pions was studied. The energy resolution for pions also follows the general form: $\frac{\sigma'(E)}{E} = \frac{a}{\sqrt{E}} \oplus b$. By fitting the distribution with this function, the following values are obtain for the coefficients: $a = (87.7 \pm 0.6)\,\% \sqrt{\text{GeV}}$ and $b = (3.0 \pm 0.2)\,\%$ (see Figure 3.37 (a)). These results are higher than the values published in [37] of $a = (84.1 \pm 0.3)\,\% \sqrt{\text{GeV}}$ and $b = 0\,\%$ within errors (a different clustering algorithm was used - the topological clustering uses three threshold values for the cell signal in order to define seeds, neighbors and overall selected cells). The variation of the energy resolution with the impact points from C to K (see Figure 3.6) was also studied, as a source of systematic uncertainty due to the positioning of the beam on the calorimeter modules. The resolution is uniform to about 0.3 % (with an average resolution of 7 %) for different scanning points and an improvement from the electromagnetic scale (about 10 %) is seen in Figure 3.37 (b) for beams of pions with 200 GeV.



(a) (b)

Figure 3.37: (a) Energy intrinsic resolution $\frac{\sigma'}{E}$ versus the reconstructed energies for pions at impact point J. The continuous line represents the fit with $\frac{\sigma'(E)}{E} = \frac{a}{\sqrt{E}} \oplus b$. (b) Energy resolution versus impact point for pions of 200 GeV. The resolution improves from 10 % for the em-scale to about 7 % after the weighting method is applied and is uniform to about 0.3 %.

Chapter 4

Monte Carlo simulation and object reconstruction

4.1 MC simulation

As the computing power continuously increased over the last few decades, the Monte Carlo simulation technique has become the preferred instrument for estimating the feasibility of potential physics analyses, modeling the particles interactions, studying the detector performance and comparing theoretical predictions with the data. Furthermore, for many parts of the simulation process, there are no analytical solutions, so MC methods have to be used. A multitude of simulation packages are presently available that incorporate our best knowledge of the parameters and physics of the Standard Model to model the LHC particle production and the detectors used for their study, such as the ATLAS detector.

Data collection at the LHC will start in 2008. In order to develop analysis tools

for physics analysis with ATLAS, data were generated using Monte Carlo techniques. Figure 4.1 shows the various stages in the generation of MC data.

The first step of the process is the *Generation* of primary particles from the hard scattering of the two colliding protons. At this level, the hard scattering processes are simulated, together with the decays of the primary particles and the hadronization of quarks and gluons into stable mesons and baryons. Next, at the *Simulation* level these primary generated particles interact with the detector and their interactions are simulated. At the *Digitization* level, the electronic response of the detector is simulated. The next step, *Reconstruction*, involves the use of dedicated software to reconstruct basic objects, such as: leptons, neutrinos (E_T^{miss}), jets etc... After some preliminary filtering of these objects, they are stored in various formats, such as ESD (Event Summary Data) and AOD (Analysis Object Data), to be used later for further, more detailed, physics analysis. Ntuples derived from AODs were used in this thesis.

4.1.1 Generation

Various generator level software packages are available for high energy particle physics. There are two main types of packages: more specialized ones (used for tree-level matrix element calculations of the hard scattering process) and general purpose ones (used for parton showering, underlying event, and hadronization).

Figure 4.2 shows the various steps involved in the event generation process. First, quarks and gluons in the protons carry a fraction of the proton momentum described by the parton density functions (PDFs), given as $f(x, Q^2)$ in Figure 4.2. One of the partons in each of the colliding protons undergoes a hard scattering pro-

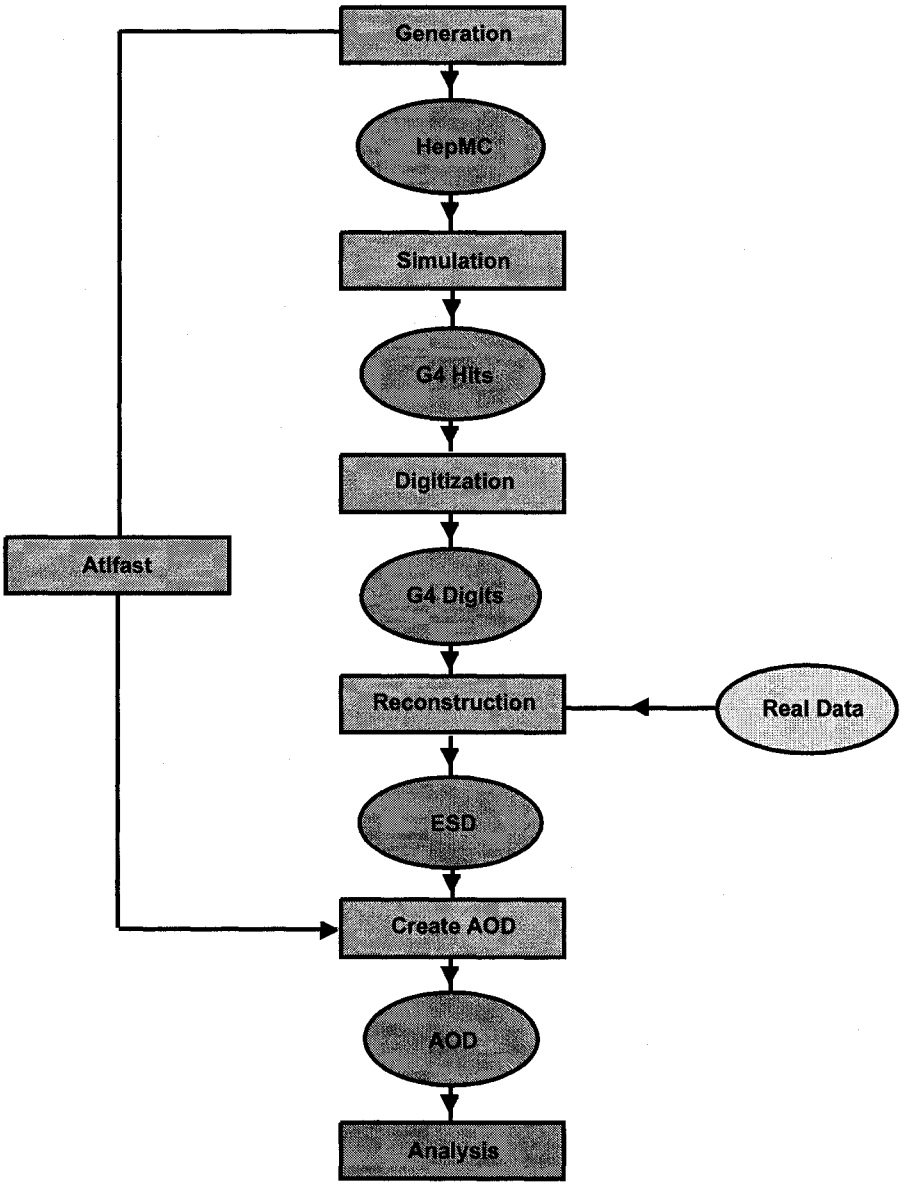


Figure 4.1: Schematic representation of the full chain Monte Carlo production [43]. For the fast simulation branch, the generated events are smeared with Atlfast and directly stored in AOD's for further analysis.

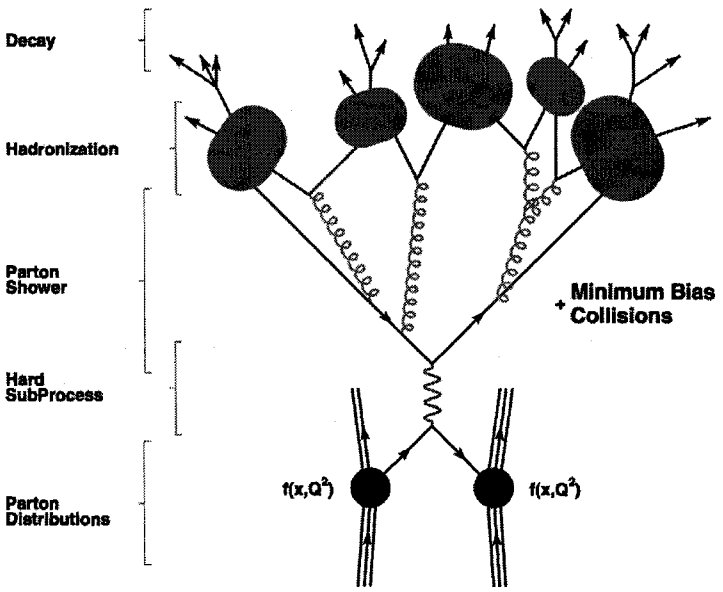


Figure 4.2: Schematic view of the steps involved in the generation of MC events following the collision of two protons [44].

cess in which particles with large mass or high p_T are produced. This part of the event is called the hard sub-process and it is described by the first type of generators, the matrix element calculators. Then, the resulting particles create a shower of new partons (quarks branch into qg pairs, while the gluons may branch into $q\overline{q}$ or gg pairs). The resultant partons may also branch, resulting in further cascades of partons (parton showers). Showering of the initial state partons (before the hard scattering takes place) can also be included (not shown in the figure). At this level, the event consists of a number of elementary particles such as quarks, anti-quarks, and gluons which due to color confinement are not allowed to exist in isolation. Next, the colored partons are grouped into color-singlet composite hadrons using a hadronization model (phenomenological), containing several cutoff parameters which cannot be computed from first principles but are tuned using available experimental data.

The final step of the chain is represented by the decay of the short-lived resonances created at the hadronization level. In addition, the generators can also add in features of the underlying event (due to beam remnants), multiple parton-parton interactions (interactions of partons belonging to the same interacting protons) and pile-up (other proton-proton collisions in the same bunch crossing).

The output of the *Generator* level is stored into a HepMC common format (an event record for high-energy physics Monte Carlo generators). HepMC stores the output from an event generator as a graph of particles and vertices, where the vertices maintain a listing of the incoming and outgoing particles and the particles point back to their production vertices.

The MC generators used for the production of the samples used in this thesis are: AcerMC [51] (provides a library of the matrix-element-based generators for several processes) and Pythia [45] (used for initial and final state parton showers, beam remnants and underlying events, and hadronization (fragmentation and remaining decays)) for the single top channels, MC@NLO [49] (matrix elements) and Herwig [46] (parton shower and hadronization) for the $t\bar{t}$ samples, AlpGen [52] (matrix elements) and Herwig (parton shower and hadronization) for W+jets, and Herwig for the QCD $b\bar{b}$ background.

In the AcerMC generator, the t-channel process (Figure 4.3) is modeled as the combination of the $qb \rightarrow q't$ and $qg \rightarrow q'tb$ W boson exchange processes. The double counting between the initial state radiation (ISR) $g \rightarrow b\bar{b}$ splitting and the next-order α_s process $qg \rightarrow q'tb$ needs to be removed. The procedure includes part of the NLO corrections, by removing the collinear singularities (described in [53]). This approach is different from the one used in the TopReX [50] generator (used for the generation of older single top samples in ATLAS) in which the combination of the

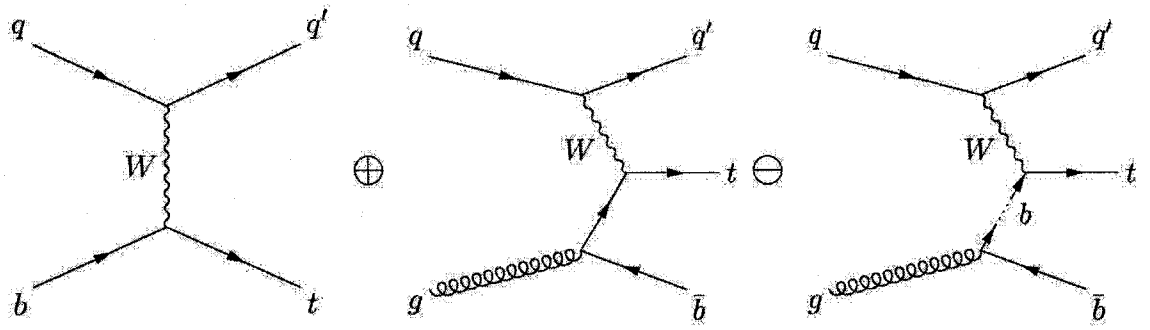


Figure 4.3: Feynman diagrams for the t-channel single top production for order $\alpha_s{}^{(0)}$ (left), order $\alpha_s{}^{(1)}$ (centre) and order $\alpha_s{}^{(1)}$ subtraction term (right) [53].

LO and NLO diagrams for the t-channel were done based on the p_T distributions of the additional b-quark defined in the initial pp reference frame (for details see [54]).

The MC@NLO event generator incorporates the next-to-leading-order calculations of rates for QCD processes and is capable of describing well multiple soft/collinear radiations, giving a final state description at the hadron level. Relevant for the top-anti-top and single top production simulation is the inclusion in the package of the decay angular correlations due to spin correlations.

Alpgen is a generator specialized for the generation of hadronic Standard Model events that have a large number of jets in their final states. The masses of the b-quark and t-quark, together with the t-quark and gauge boson decays with helicity correlations, are used in the program to extract the matrix elements at leading order.

The Pythia and Herwig programs are general purpose generators frequently used for hadronic event modeling and simulation in high-energy physics at particle accelerators, including pp, e^+e^- and ep colliders. The programs implement a vast diversity of hard scattering, $2 \rightarrow n$, processes, complemented by initial and final state

parton showers, hadronization and decays. The underlying event is added to complete the event generation simulation and analysis tools are also available with the packages.

4.1.2 Simulation

At the *Simulation* level, the particles created at the *Generator* level interact with the detector. Two types of detector simulations are available: full simulation (FullSim) and fast simulation (FastSim). The Full simulation uses a full Geant4 [55] detector description for ATLAS, while the fast simulation (Atlfast [56]) uses a smearing of the quantities resulted after the *Generation* stage. When very large numbers of events need to be simulated, the fast simulation is used due to its speed, 4-5 orders of magnitude faster than the full simulation. However, one needs to validate the fast simulation by comparing it with the full simulation results.

Geant4 is a package used for the simulation of the interaction of particles with detectors. It contains information about the geometry of the detector, the properties of the materials used in the components and the physics phenomena involved in the propagation of particles through matter. In addition, it has the capability to produce particles, track them through the detector and measure and store events for visualization and/or further, more in detail, analysis. Both, electromagnetic and hadronic physics processes are implemented, within a range of energies from 250 eV to the TeV level. The package is written in C++ object oriented programming language, giving it flexibility in interfacing with the other software modules in the Monte Carlo chain. The output of the *Simulation* level (from the FullSim - Geant4) is stored into Geant4 hits.

The ATLAS fast simulation program (Atlfast) simulates ATLAS physics

events, including the effects due to detector response and the software reconstruction chain through the application of a resolution function to the particle true properties (e.g. particle true energy). The resolution function is modeled after the resolution obtained using either test beam results or the full detector simulation. The input to the program is the collection of four-vectors for a physics event, usually provided by a physics event generator. The results produced by Atlfast are ready to be used for physics analysis.

4.1.3 Digitization

The hits produced by Geant4 (or after they were sent first to a pile-up algorithm in order to add the effects of the events pile-up and produce merged hits) need to be translated into the output mimicing that produced by the ATLAS detector. The propagation of charges (as in the tracking detectors and the liquid argon calorimeter) or light (as in the case of tile calorimeter) into the active media has to be considered, as well as the response of the readout electronics. This is a very detector specific task. The final output of the digitization step are raw data objects that should have the same format as the real detector data. As well, simulation data objects are created in addition to raw data objects, to save some simulation information that can be used by a downstream user. The output of the *Digitization* level is stored into a G4 Digits format similar to the G4 Hits but including the information about digitization.

4.1.4 Reconstruction

The ATLAS detector will produce approximately 3,000 terabytes of raw data per year. This is a very large amount of information which is difficult to distribute to the

institutes involved in data analysis around the world. To enable access from remote sites, several types of data sets (corresponding to different stages of reconstruction) are produced. The reconstruction step derives physics objects (needed for analysis) from the stored raw data. These objects are: photons, electrons, muons, tau-leptons, jets, missing transverse energy and primary vertices. At the end of the reconstruction step, two types of data sets are created: the Event Summary Data (ESD) which contains a detailed output of the detector reconstruction, and the Analysis Object Data (AOD) which is a summary of the reconstructed event, and contains enough information for common analyses. The target size for the ESD is 500 kB per event, while for the AOD it is 100 kB per event.

4.2 Reconstructed objects

In order to perform any physics analysis, one needs to be able to identify and reconstruct basic physics objects. For top physics, these objects are: leptons (e, μ, and τ), neutrinos (missing transverse energy - E_T^{miss}) and quarks (jets). Dedicated software tools have been developed for this purpose.

4.2.1 Electrons

For the identification and reconstruction of electrons, the EGamma algorithm is used [21]. It uses information from both the calorimeters and the inner tracking system in order to define discriminating variables used for separating the electrons from jets. In order to maintain a high electron efficiency, certain threshold values are used for the cuts on these variables. Only the objects for which the discriminating

variables pass the threshold cuts are selected as electrons.

The offline electron/jet separation procedure consists of using the calorimeter information (the hadronic leakage and the use of the first and second compartments of the electromagnetic calorimeter) to select events containing a high-E_T electromagnetic shower. Then, a good quality track is selected and required to match the electromagnetic calorimeter shower in (η, ϕ) as well as in the energy-momentum reconstruction. Finally, the TRT information is used to improve the identification of electrons.

4.2.2 Muons

The information collected from the muon spectrometer is combined with that from the inner detector and the calorimeters in order to reconstruct muons. In the inner detector and outer spectrometer, isolated tracks are identified as potential muon candidates. Two algorithms are used to reconstruct the muon objects for the samples used in this thesis: STACO and MuID [21].

The STACO (STAtistical COmbination) algorithm uses a statistical method to merge the track in the muon spectrometer with the one in the inner detector. The covariance matrices of the two tracks are combined together in order to fully exploit the fact that the two tracks are measured independently. The MuID algorithm uses a global combined fit of all measurements in the muon spectrometer, the inner detector and the calorimeter.

4.2.3 Missing transverse energy E_T^{miss}

Because the neutrinos cannot be detected directly in the ATLAS detector, one has to look for alternative ways to infer their momentum. One of the possibilities is to look for the transverse energy that is missing in an event. The E_T^{miss} of an event is defined as the negative of the vector sum of the transverse energy over all calorimeter cells, $-E_i \cdot sin\theta_i \cdot n_i$, where E_i is the energy deposited in the i-th calorimeter cell, θ_i is the polar angle of the i-th cell, and n_i is the unit vector in the azimuthal plane pointing from the interaction point to the i-th cell [21]. The sum is computed over all calorimeter cells. Since muons leave only a small amount of energy in the calorimeter, their energy does not get accounted for in the E_T^{miss} sum and so for these cases a correction must be applied. A correction is also applied for the energy lost in the cryostat (dead material in general). The electronic noise suppression in calorimeters is performed by applying a cell energy threshold in terms of the number of sigma noise (e.g. $|E| > 2\,\sigma_{noise}$), or using only cells in topological clusters (the selected cells in the topological cluster are already satisfying a noise threshold condition).

4.2.4 Jets

Jets are initiated by quarks and gluons ("parton jets") produced in the proton-proton collisions that spray (hadronize) into collimated groups of hadrons moving in the direction of the original partons. Before entering the detector, the jets are called "particle jets". When the hadrons in jets interact with the detector they produce clusters of energies in the calorimeters. The reconstructed jets in the calorimeters are called "calorimeter jets" or "reconstructed jets". The evolution of the three jet types is shown in Figure 4.4.

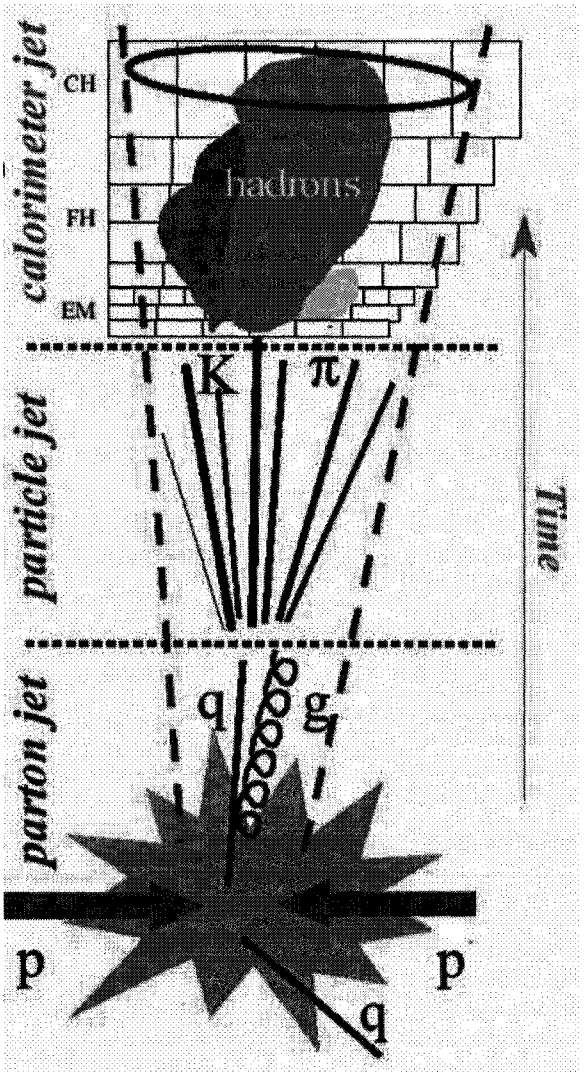


Figure 4.4: Schematic view of the evolution of the three types of jets. The parton jets are generated immediately after the proton-proton interaction and they consist of quarks and gluons [63]. The particle jets are produced through the hadronization of the quarks and gluons in the parton jets and they consist of mesons and baryons. The calorimeter jets are produced by the hadrons interacting with the materials in the calorimeters.

The same algorithms (jet finders) are used to reconstruct jets at both particle and calorimeter level. For jet reconstruction in ATLAS, the most commonly used jet finders are: the K_T algorithm and the cone algorithm [21]. There are two implemented K_T algorithms: $FastK_T$ and K_T algorithm. There are two types of cone algorithms implemented, one that requires a seed (particle, cell, cluster or tower with an energy greater then a threshold - faster) and other that does not (slower, but theoretically more accurate). Both cone algorithms pass through a split-merge tool to define non overlapping exclusive jets in the event.

A cone jet algorithm forms jets by associating particles (or calorimeter towers/clusters) within a circle in (η, ϕ) of specific radius $R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$, where $\Delta\eta$ is the distance in the η direction between the jet centre and the outer cone radius and $\Delta\phi$ is the distance in the ϕ direction between the jet centre and the outer cone radius. The algorithm starts by looking for seed objects with $p_T > 1\,\text{GeV}$. All objects around the seed axis (the direction defined by the beam interaction point and the seed) are added together and an energy-weighted centroid is calculated. Taking this axis (beam interaction point - energy-weighted centroid) as the new axis, the process is repeated and a new centroid recalculated. The cone centre therefore moves around until a stable solution is found in which the weighted centroid coincides with the cone axis. A process called "split-merge" is implemeted after all jet candidates were identified. When two cones are overlapping and the energy shared between them is more than 50 % of the less energetic jet, then the jets are merged. If the shared energy is less than 50 %, then the two jets are split into two, with the shared energy being attributed to the nearest jet.

There are two types of input used for the jet finder algorithms at the calorimeter level. The traditional input is based on the calorimeter towers. In this method all

calorimeter cells with signal are collected into projective towers. In an ideal detector geometry the towers span a grid defined by $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$. Due to the projective nature of the towers there is no explicit longitudinal (along the η direction) readout granularity used in this jet finder. No energy calibration has been performed at this level, therefore the signal is at the electromagnetic scale. In order to cancel part of the noise, protojets are formed by adding two near-by towers with positive and negative (due to negative fluctuation from noise) energies until the resulting protojet has a positive energy (all cells are kept). At this point the jet finding algorithms are run over the collections of protojets. The resulted jets are then calibrated at the particle level and finally an in-situ calibration is performed for calibration at parton level (e.g. the W boson with well known mass decaying into jets, or the Z boson decaying leptonically balanced by one high-p_T jet). The tower input is used for the jets in the samples used in this thesis.

The second type of input is based on topological clustering of the cells in the calorimeters [37]. Clusters of cells are formed by selecting cells with signals above certain thresholds (typically three thresholds are used) related to the noise in the cells. The clusters thus formed have a signal at the electromagnetic scale. These clusters are further classified and a hadronic cluster calibration is performed. After the clusters have been properly calibrated, the jet finder algorithm is run to build calorimeter jets. The resulting jets are then calibrated at the particle level and finally an in-situ calibration is performed for calibration at parton level.

Two values, 0.4 and 0.7 (optimized in the CDF and D0 experiments [64]), are generally used for the radius of the cone [65]. In the samples used for this analysis, the cone $R = 0.4$ algorithm was used (a smaller jet cone size is preferred for final states with a large number of jets).

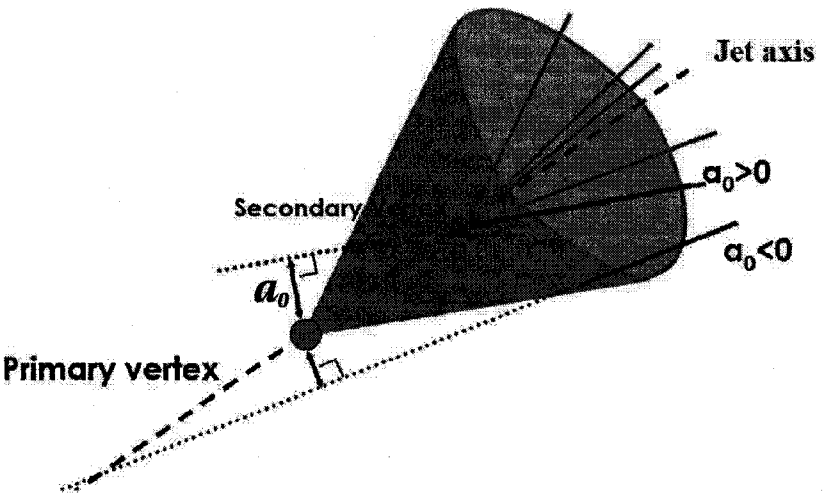


Figure 4.5: Schematic view of the method used for b-tagging in ATLAS [62]. The distance of closest approach of a track to the primary vertex position (a_0) is signed positive if the track crosses the jet axis in front of the primary vertex, and negative otherwise.

4.2.5 B-jets

Two b-tagging methods are available for the ATLAS experiment. The first one, the vertex method, exploits the relatively long life time of b-hadrons (about 1.5 ps), while the second one considers the leptonic decays of the b-quarks into low-p_T electrons or muons [21].

For the first method, the ratio of the distance from the primary vertex position (a_0) to the considered track divided by its error (ϵ_{a_0}) is used to define a track impact parameter significance. The impact parameter definition is illustrated in Figure 4.5 and for b-hadrons it has an absolute value greater than zero, on the order of 400 μm.

The tracks belonging to the light jets have a track impact parameter significance much smaller than for b-jets. The significance distributions are used to create

weighting probabilities (probabilities that the track originated from a light quark jet or a b-jet). These weighting probability distributions for light, $P_l(\frac{a_0}{\epsilon_{a_0}})$, and b-jets, $P_b(\frac{a_0}{\epsilon_{a_0}})$, are used to define a weight for the jet, w_{jet}, given by the equation:

$$w_{jet} = \sum_{i \in jet} ln\,[P_b(\frac{a_0^i}{\epsilon_{a_0}^i}) \,/\, P_l(\frac{a_0^i}{\epsilon_{a_0}^i})] \tag{4.1}$$

where the sum is over all tracks i in the jet. This weight is used for the separation between light and b-jets. The procedure can be performed both in the longitudinal and transverse planes to the jet axis and the b-tagging performance can be improved by adding these two components of the significance (implemented in the so-called IP3D method). Characteristics of the secondary vertex (e.g. mass, fraction of energy) can be used to enhance the performance of the vertex method, the secondary vertex (so-called SV1 method). In ATLAS the two methods are combined in the IP3D+SV1 method, which is the method used for the analysis in this thesis.

4.3 Data: AODs and Ntuples

The MC data samples used for the analysis in this thesis were produced centrally at CERN in the Computing System Commissioning production. These data are used for validating new software tools, exercising the ATLAS computing model, and for preparing physics analyses with real data. Following the general procedure described previously in Section 4.1, AOD files are created for all channels of interest and their backgrounds. Furthermore, a special tool called TopView [58] was used in order to create common root ntuples to be used by various physics working groups in their physics analyses. The TopView package is fully based on the official ATLAS software Athena. It uses the AODs as input and performs tasks such as: overlap removal between different particle types within the AOD (by inserting in a given

order the reconstructed AOD objects) and application of pre-cuts (to reduce the size of the information to be stored in computer files) for further analysis. The insertion of the AOD defined physics objects is done in the following order: Muon, LooseMuon, Electron, LooseElectron, Photon, TauJet, ParticleJet. Overlap with muon and particle jet is not removed.

For the Single Top Working Group, the common ntuples were produced with various versions of TopView. For this analysis, samples were produced with version V1213. The AOD used for this TopView version were produced using the Athena version 12.0.6. The root ntuples (and the AODs) are available from the LHC computing grid (a computing network for data storage that maintains an analysis infrastructure for the entire physics community that will use the LHC).

The three single top MC samples were produced using the AcerMC event generator, simulated with the Geant4 (the so-called full simulation) and digitized and reconstructed using the Athena version 12.0.6.

The $t\bar{t}$ background sample was generated with the MC@NLO event generator, simulated with Geant4, and digitized and reconstructed using the Athena version 12.0.6.

For the W+jets background four separate sub-samples were generated corresponding to W+0jets (exclusive - exactly 0 jets), W+1jets (exclusive - exactly 1 jet), W+2jets (exclusive - exactly 2 jets) and W+3jets (inclusive - 3 or more jets). The samples were generating with the Alpgen event generator, simulated with Geant4, and digitized and reconstructed using the Athena version 12.0.6. In these samples also, the events contain only the W bosons decaying leptonically. Due to the particularity of producing these events with Alpgen, the cross section, σ_{GEN}, has to be corrected

with some efficiencies: the matrix-element hard partons and shower-generated jets matching (MLM matching) efficiency (MLM_{Eff}) and the filter efficiency ($Filter_{Eff}$) (to save computing time, even at the generation level, filters (cuts) are applied on quantities such as p_T and η). The values for the W+njets ($n = 0, 1, 2, 3$) cross sections are summarised in Table 4.1 (from [61]). In order to obtain the complete W+jets background one needs to add the four sub-sample contributions weighted by their cross sections.

	σ_{GEN}(pb)	MLM_{Eff}	$Filter_{Eff}$	σ(pb)
W + 0 parton	36833	0.8351	0.8	24607
W + 1 parton	16621	0.48	0.861	6869
W + 2 parton	8390	0.3533	0.888	2632
W + 3 parton	3787	0.606	1	2295

Table 4.1: Alpgen cross sections for the W+jets background samples used in the analysis: $\sigma = \sigma_{GEN} \cdot MLM_{Eff} \cdot Filter_{Eff}$.

The QCD $b\bar{b}$ background sample was generated with Herwig and Atlfast (private production by the Grenoble LPSC top group [59]).

Chapter 5

Reconstruction performance studies

The reconstruction performance of the objects relevant for the single top quark analysis was studied using data samples obtained for various production channels, reconstruction software versions and full (Geant4) and fast (Atlfast) detector simulations. When dealing with a procedure of reconstructing various objects, one would like to know how well the reconstructed objects describe the original objects. It is particularly important to compare the full and fast simulation results due to the fact that the top channels were simulated using the full simulation, while the W+jets and QCD $b\bar{b}$ backgrounds were simulated with Atlfast in order to obtain a data sample with large statistics. Studies are currently under way to compare full and fast simulated W+jets samples [66].

5.1 Definitions

For the characterization of the reconstruction performance, three quantities are defined: reconstruction efficiency, reconstruction purity and reconstruction resolution. In what follows, *objects* refer to electrons, muons and jets. Let us suppose that T objects are present at the generation level (the original, truth objects). After the reconstruction procedure, R objects are reconstructed. The first thing to see is how many of the reconstructed objects match back to the truth objects. The matching is defined only in terms of the angular distance $R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$, where $\Delta\eta$ and $\Delta\phi$ are the distances in η and ϕ between the directions of the truth and reconstructed objects. Two objects are considered as matching if this angular distance is smaller than a certain cut-off value (e.g. 0.05 for electrons and muons, and 0.3 for jets, values obtained by studying the distribution of the angular distances [58]). In the case in which more than one object is matched to the truth, the one that has the lowest angular distance is chosen. Thus, the reconstructed objects that matched to truth objects, $R^{matched}$, are a subset of all reconstructed objects, R. Next, a histogram of various parameters (p_T, η or ϕ) of the reconstructed matched objects is created. The same type of histogram is also created for all of the T original, truth objects.

In order to calculate the reconstruction efficiency for the object, a bin by bin division of the two corresponding histograms is performed. Generically, the reconstruction efficiency can be written as:

$$Efficiency^{reco}_{x,i} = \frac{x_{R^{matched},i}}{x_{T,i}} \tag{5.1}$$

where x can be any of the following: p_T, η or ϕ, i is the bin number, $x_{R^{matched},i}$ is the number of events in bin i of the x histogram of the reconstructed matched objects and $x_{T,i}$ is the number of events in bin i of the x histogram of all truth objects. The

purity of the reconstructed objects is defined as:

$$Purity_{x,i}^{reco} = \frac{x_{R^{matched},i}}{x_{R,i}} \tag{5.2}$$

where x_R is the x histogram of all reconstructed objects.

For the matched objects the resolution is defined event by event for each object in the set as:

$$\frac{\Delta x}{x} = \frac{x_{R^{matched}} - x_{T^{matched}}}{x_{T^{matched}}} \tag{5.3}$$

where $x_{R^{matched}}$ is the x value of the reconstructed matched object and $x_{T^{matched}}$ is the x value of the truth object to which the reconstructed object was matched. A histogram of the resolution values is then obtained, when considering all objects. The resolution can be plotted for any of the x variables: p_T, η or ϕ.

A truth object can be reconstructed as multiple reconstructed object (they are overlapping). In order to remove this overlap, the reconstructed objects were selected for analysis in the following precedence order: muons, electrons, photons, tau jets and particle jets [58]. If an object was reconstructed as a certain type then the other instances were not considered anymore.

5.2 Data

Various sets of MC data were used for the reconstruction performance analysis. These data were generated with two Athena software release versions (11 and 12) and processed with different TopView versions (v11, v12, v1212 and v1213). Unless otherwise specified, the results (graphs) are for the t-channel data and full simulation (FullSim) refers to FullSim(v1213), while fast simulation (FastSim) refers to FastSim(v12).

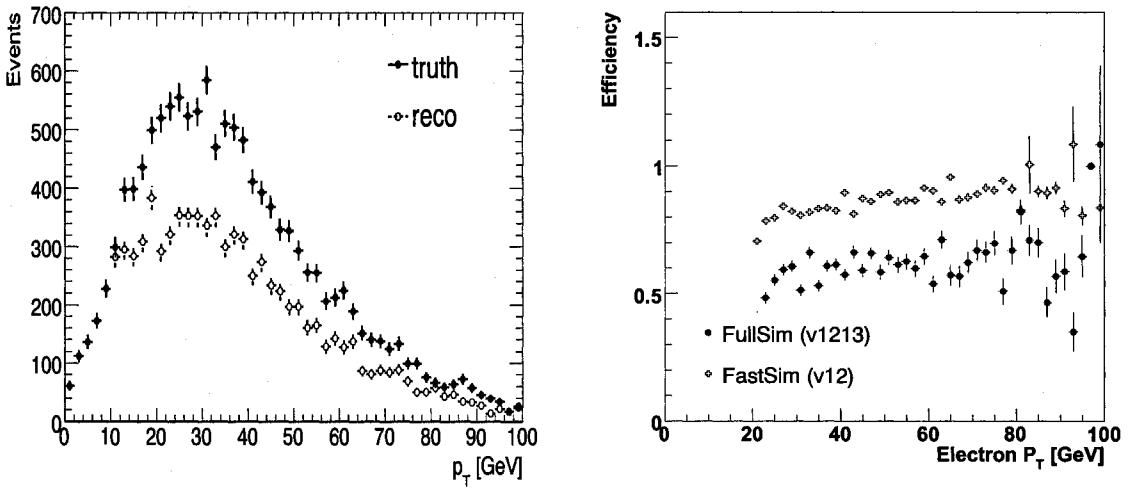
For this analysis only leptons originating from the W boson decay are analyzed. The electrons were reconstructed with the Egamma algorithm (Section 4.2.1) and have a minimum $p_T = 10\,\text{GeV}$, while the truth electrons have no minimum p_T cut. The muons were reconstructed using the STACO algorithm (Section 4.2.2) and they have a minimum $p_T = 20\,\text{GeV}$. A $|\eta|$ cut of ≤ 2.5 is applied for both reconstructed electrons and muons, as well as a isolation $E_T < 6\,\text{GeV}$ in a $R = 0.2$ cone around the lepton axis. The jets were reconstructed with the cone 0.4 algorithm (Section 4.2.4) using calorimeter towers (towers of calorimeter cells along a specific η) and they have a minimum $p_T = 30\,\text{GeV}$. A $|\eta|$ cut of ≤ 5 is applied for the reconstructed jets.

5.3 Reconstruction efficiency and purity

5.3.1 Electron reconstruction

The p_T distributions of the truth electrons and of the matched reconstructed electrons are shown in Figure 5.1 (a). Dividing bin by bin the two distributions (after the 10 GeV p_T cut was applied to both truth and reconstructed electrons), as mentioned above, one obtains the reconstruction efficiency for electrons in bins of p_T, as shown in Figure 5.1 (b). As can be seen in the figure, the reconstruction efficiency for electrons increases with increasing p_T up to about 25 GeV, after which it reaches a plateau of about 60 %. The average efficiency for electrons with $p_T > 25\,\text{GeV}$ is 60.8 %. The efficiency has a value greater than one in the last bin due to bin migration (events that belong to a bin in the truth histogram can belong to a different bin once the reconstruction is performed).

Figure 5.1 (b) shows the comparison of the efficiencies between the full and



(a) (b)

Figure 5.1: (a) The p_T distribution of all truth electrons and of the matched reconstructed electrons; (b) electron reconstruction efficiency versus p_T for the full and fast detector simulations.

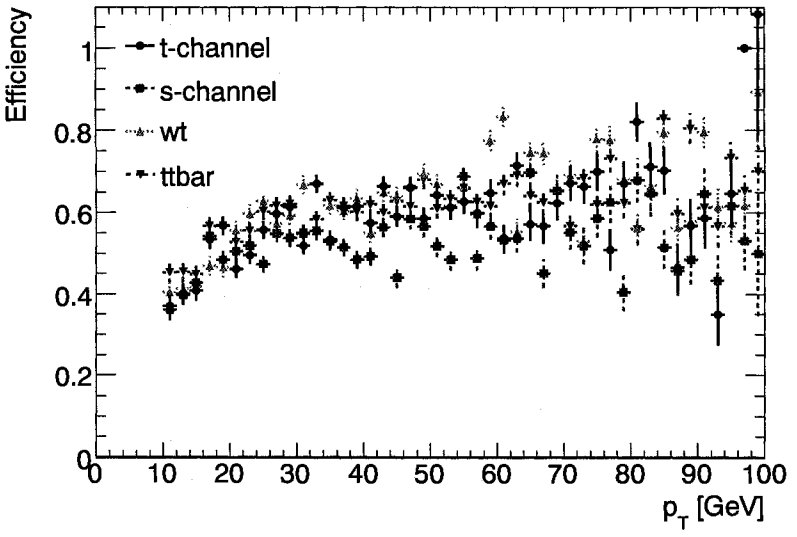
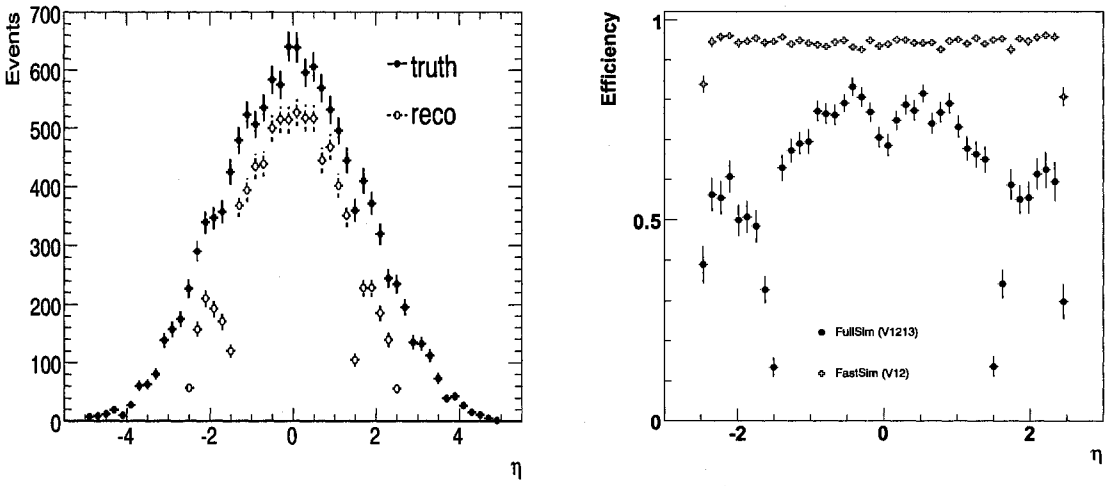


Figure 5.2: Reconstruction efficiency for electrons in bins of p_T for the t-channel, s-channel, W+t-channel and $t\bar{t}$ (full simulation).

fast detector simulations. The efficiency for Atlfast is higher over the full p_T range than the one for the full simulation, with an average value of 90 %. This effect is due to the fact that Atlfast does not smear the particles' tracks (therefore, no quality requirements on track and shower shapes are being considered).

A comparison is done for the electron efficiencies obtained for the t-channel, s-channel, W+t-channel and $t\bar{t}$. The distributions are presented in Figure 5.2 and the resulting efficiencies are summarized in Table 5.1. There are no major differences between the top quark channels with the exception of the s-channel that shows on average a lower electron efficiency.

Following the same procedure as for the p_T distributions, one can look also at the η and ϕ distributions. The η distribution of the truth electrons and of the matched reconstructed electrons are shown in Figure 5.3 (a). Dividing bin by bin

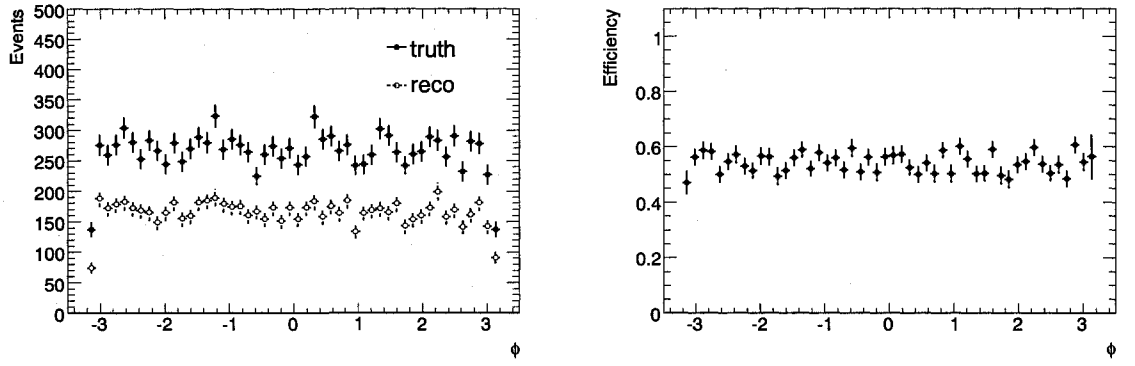


(a) (b)

Figure 5.3: (a) The η distribution of all truth electrons and of the matched reconstructed electrons; (b) electron reconstruction efficiency versus η for the full and fast detector simulations.

the two distributions, one obtains the reconstruction efficiency for electrons in bins of η, as shown in Figure 5.3 (b). It can be seen that the efficiency is higher in the more central region $|\eta| < 1.5$ where it reaches about 80 %. There is a sharp drop in the efficiency at $\eta = \pm 1.5$ to about 20 %, corresponding to the region where two major cracks are present in the detector. The reconstruction of electrons in these regions is quite poor due to problems in the electron isolation E_T calculated in the reconstruction when using the full detector simulation (the energy in one layer of the electromagnetic calorimeter in the crack region was not subtracted from the cone transverse energy resulting in a higher isolation E_T for electrons that were thus unable to pass the isolation $E_T < 6\,\text{GeV}$ criteria in this η region).

Figure 5.3 (b) also shows the comparison of the efficiencies between the full and fast detector simulations. The efficiency for Atlfast is higher over the full η range



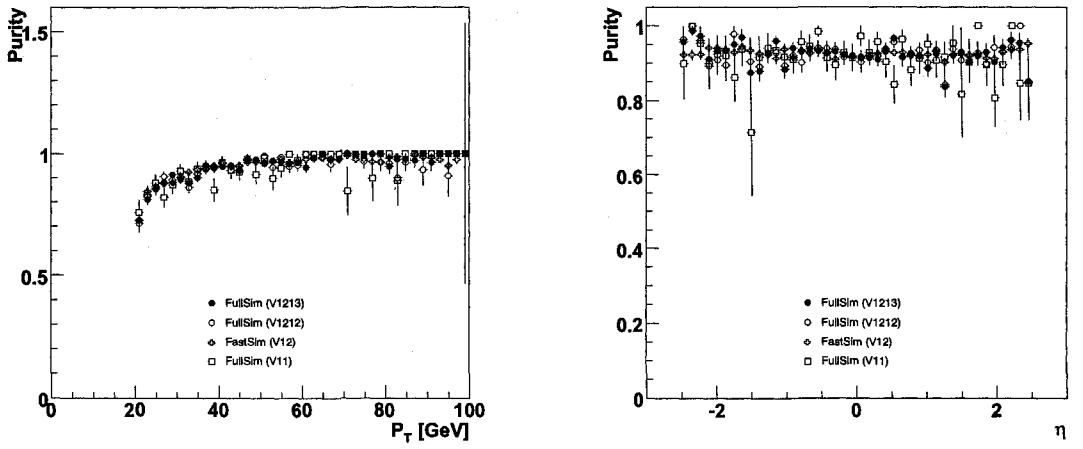
(a) (b)

Figure 5.4: (a) The ϕ distribution of all truth electrons and of the matched reconstructed electrons; (b) electron reconstruction efficiency versus ϕ for the full detector simulation.

than the one for the full simulation and it shows no decrease in efficiency in the crack regions. Therefore, a correction factor has to be applied to the fast simulated data in order to match the efficiency predicted by the full simulation.

The ϕ distribution of the truth electrons and of the matched reconstructed electrons are shown in Figure 5.4 (a). Dividing bin by bin the two distributions, one obtains the reconstruction efficiency for electrons in bins of ϕ, as shown in Figure 5.4 (b). The ϕ distribution of the electron reconstruction efficiency is uniformly distributed over the entire ϕ range from $-\pi$ to $+\pi$ and it has an average value of 55 %.

Another measure of interest is the reconstruction purity defined by Equation 5.2. Figure 5.5 (a) shows the comparison of the electron purity between the full and fast detector simulations as a function of p_T. The purity for Atlfast is close to the one for the full simulation over the entire p_T range. The reconstruction purity for



(a) (b)

Figure 5.5: Electron reconstruction purity versus p_T (a) and η (b) for the full and fast detector simulations.

electrons in bins of η, as shown in Figure 5.5 (b), is almost constant over the entire $|\eta|$ at about 95 %. Figure 5.5 also shows the comparison of the purity between the full and fast detector simulation. The purity for Atlfast is close to the one for the full simulation over the entire η range and it shows no decrease in purity in the crack regions.

The electron purity for the t-channel, s-channel, W+t-channel and $t\bar{t}$ (presented in Figure 5.6) increases with increasing p_T up to about 25 GeV, after which it reaches a plateau of about 95 %. The average purity for electrons with $p_T > 25$ GeV is 98 %.

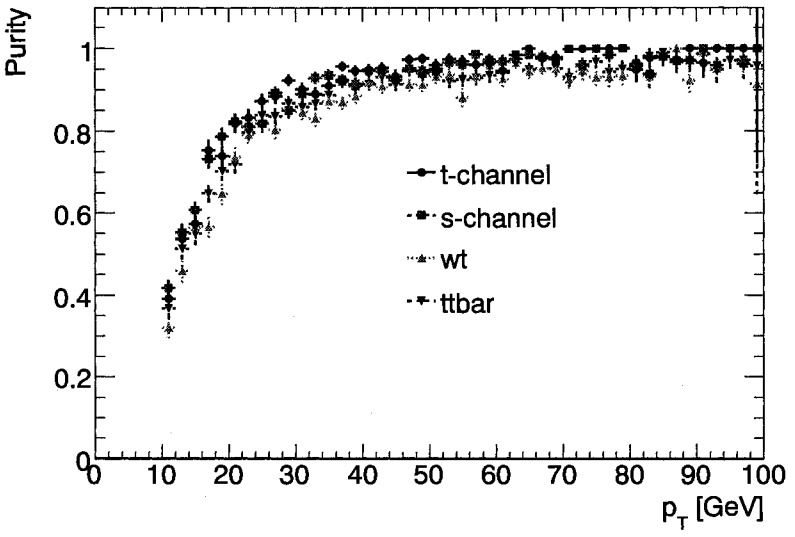
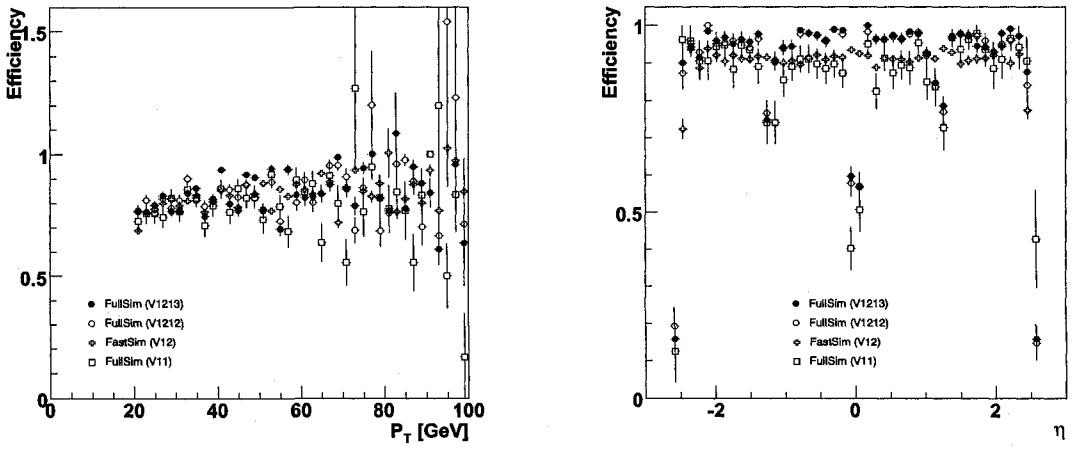


Figure 5.6: Reconstruction purity for electrons in bins of p_T for the t-channel, s-channel, W+t-channel and $t\bar{t}$ (full simulation).

5.3.2 Muon reconstruction

The reconstruction efficiency distribution for muons in bins of p_T, as shown in Figure 5.7 (a), is mostly a plateau of about 85 %, a value higher than for the electrons. The average efficiency for muons with $p_T > 25\,\text{GeV}$ is 86.8 %. Figure 5.7 (a) also shows the comparison of the efficiencies between the full and fast detector simulations, and various versions of the full simulation reconstruction software. The efficiency for Atlfast is close in value to the one for the full simulation over the full p_T range, a situation that is different from the electron's case. A comparison is done for the muon efficiencies obtained for the t-channel, s-channel, W+t-channel and $t\bar{t}$. The resulting efficiencies are summarized in Table 5.1.

The reconstruction efficiency for muons in bins of η, as shown in Figure 5.7 (b), is uniformly distributed almost over the entire $|\eta|$ region. There are though several

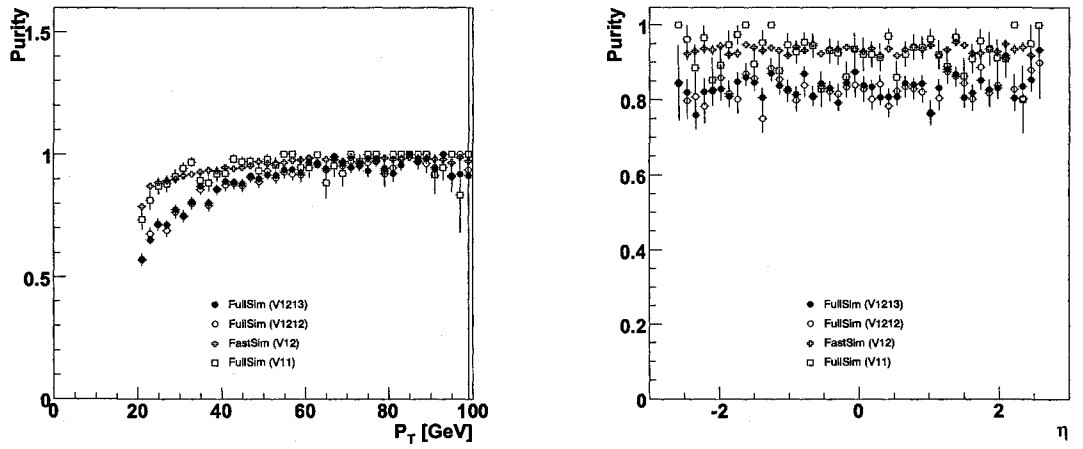


(a) (b)

Figure 5.7: Muon reconstruction efficiency versus p_T (a) and η (b) for the full and fast detector simulations.

noticeable drops in the efficiency. The one at $\eta = 0$ is related to the absence of the muon trackers in this region. The other one, at about $\eta = \pm 1.3$, corresponds to the region where extra material is present in the barrel-end-cap transition zone. Figure 5.7 (b) also shows the comparison of the efficiencies between the full and fast detector simulation. The efficiency for Atlfast is uniform and overall lower by a few percent than the one for the full simulation.

The reconstruction purity for muons in bins of p_T, as shown in Figure 5.8 (a), increases with increasing p_T up to about 25 GeV for FastSim and 40 GeV for FullSim, after which it reaches a plateau of about 95 %. The average purity for muons with $p_T > 25$ GeV is 98 %. Figure 5.8 (a) also shows the comparison of the purity between the full and fast detector simulations. The purity for Atlfast is higher than the one for the full simulation for small values of p_T. The reconstruction purity for muons in bins of η, as shown in Figure 5.8 (b), is almost constant over the entire $|\eta|$ at about



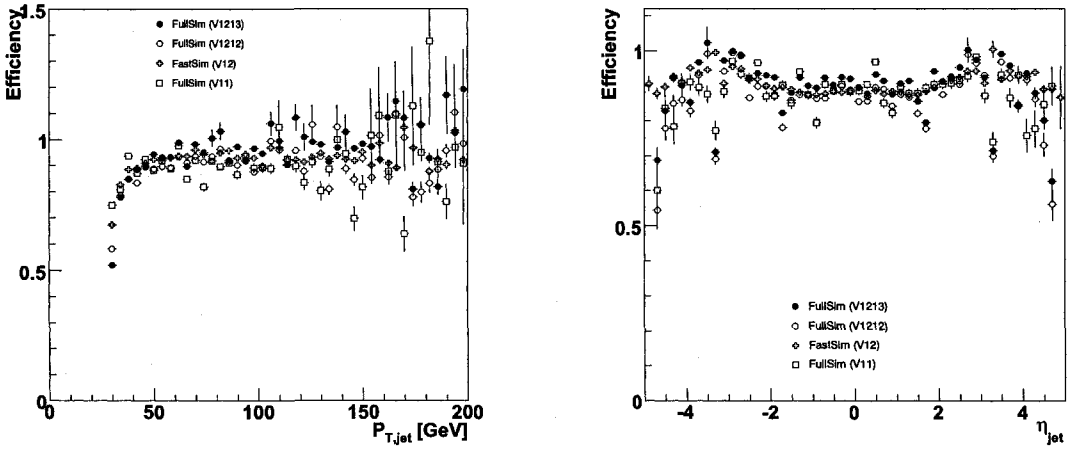
(a) (b)

Figure 5.8: Muon reconstruction purity versus p_T (a) and η (b) for the full and fast detector simulations.

85 % (for the full simulation). Figure 5.8 (b) also shows the comparison of the purity between the full and fast detector simulations. The purity for Atlfast is higher than the one for the full simulation over the full η range, except for the V11 full simulation data set.

5.3.3 Jet reconstruction

The reconstruction efficiency for jets in bins of p_T, as shown in Figure 5.9 (a), increases with increasing p_T up to about 50 GeV, after which it reaches a plateau of about 95 %. The average efficiency for jets with $p_T > 50$ GeV is 96.1 %. Figure 5.9 (a) also shows the comparison of the efficiencies between the full and fast detector simulation. The efficiency for Atlfast is close to the one for the full simulation over the full p_T range, with an average value of 96 %. A comparison is done for the jet efficiencies



(a) (b)

Figure 5.9: Reconstruction efficiency for jets in bins of p_T (a) and η (b) for the full and fast detector simulations.

obtained for the t-channel, s-channel, W+t-channel and $t\bar{t}$. The resulting efficiencies are summarized in Table 5.1.

The reconstruction efficiency for jets in bins of η, as shown in Figure 5.9 (b), is quite flat in the central region. There is a drop in the efficiency at about $|\eta| = 1.5$ and $|\eta| = 3.2$ (in the latter case the transition region between the end-cap and the forward calorimeters). Figure 5.9 (b) also shows the comparison of the efficiencies between the full and fast detector simulation. The efficiency for Atlfast is slightly lower than the one for the full simulation. The purity for reconstructed jets has a value around 90 % for both the full and fast simulation.

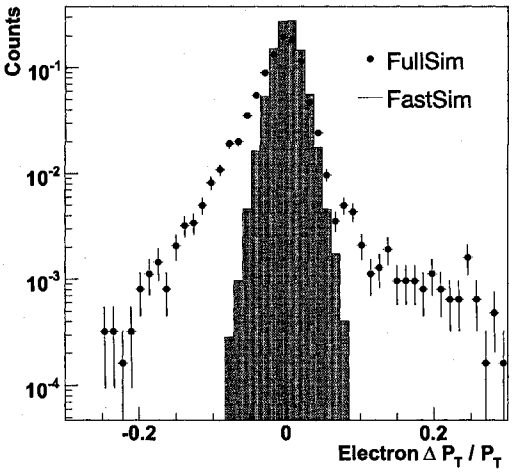


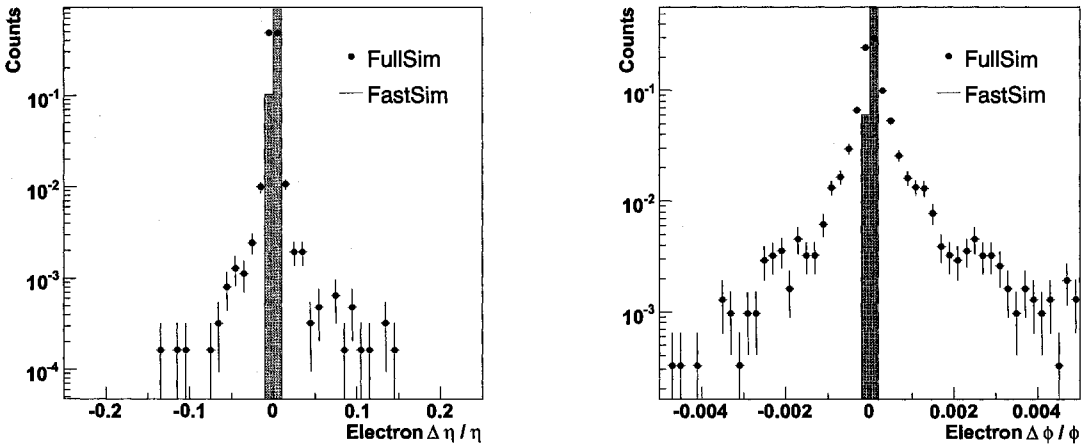
Figure 5.10: The electron p_T resolution for full and fast detector simulations.

5.4 Resolution

5.4.1 Electron resolution

The final reconstruction performance study is on the p_T, η and ϕ resolutions. For this part, the entire available p_T range for each object is considered. The comparison between the p_T resolution obtained for fast and full simulations is presented in Figure 5.10. It can be seen that the fast simulation predicts a narrower distribution of the electron resolution ($\sigma = 1\,\%$) than the full simulation ($\sigma = 2\,\%$). This effect can be attributed to the different treatment of the final state photon radiation in the fast simulation, where the photon (collinear with the emitting electron) deposited energy is not taken into account.

The electron η and ϕ resolutions using the fast and full simulations are presented in Figure 5.11 (a) and Figure 5.11 (b), and show that Atlfast produces almost



(a) (b)

Figure 5.11: The electron η (a) and ϕ (b) resolution for full and fast detector simulations.

no smearing in the η and ϕ directions, while the full simulation shows a broader resolution ($\sigma = 0.06\,\%$ for both η and ϕ).

Figure 5.12 shows the electron p_T resolution for all studied top quark channels.

5.4.2 Muon resolution

For muons, the p_T resolution comparison between fast and full simulations, as presented in Figure 5.13, predicts a narrower muon resolution distribution for the fast simulation ($\sigma = 2\,\%$) than for the full simulation ($\sigma = 2.4\,\%$), although the difference is smaller than in the case of electrons. The distribution in the tails of the full simulations are also smaller than for the electrons.

The muon η and ϕ resolutions using the fast and full simulations are presented

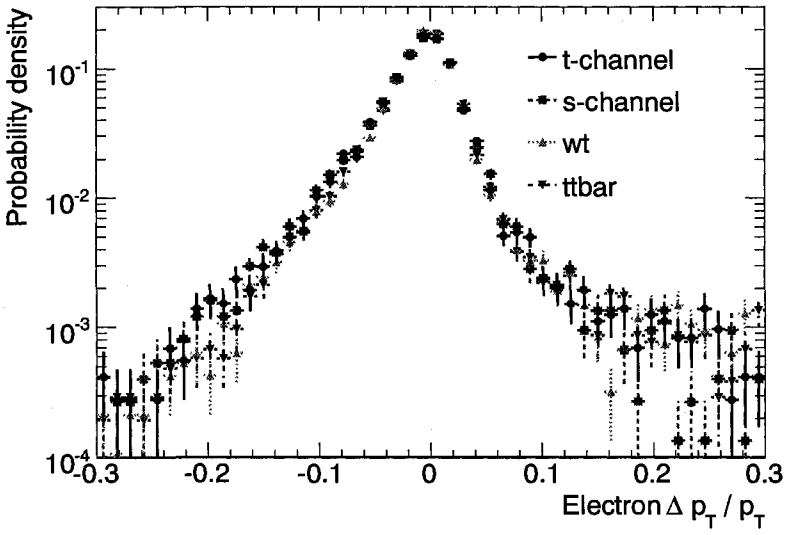


Figure 5.12: The p_T resolution for electrons in bins of $\Delta p_T/p_T$ for various top channels (full simulation).

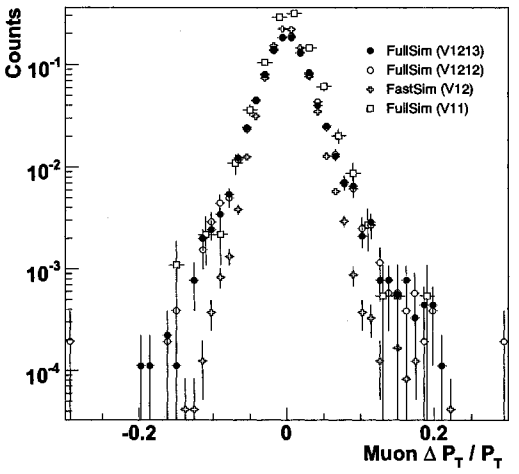
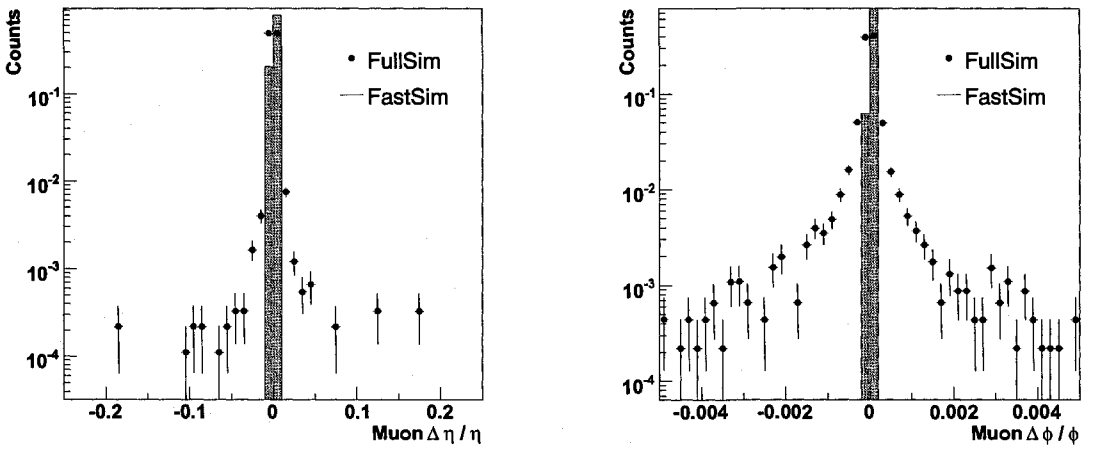


Figure 5.13: The muon p_T resolution for full and fast detector simulations.



(a) (b)

Figure 5.14: The muon η (a) and ϕ (b) resolution for full and fast detector simulations.

in Figure 5.14 (a) and Figure 5.14 (b). As in the electron case, it can be seen that Atlfast produces almost no smearing in the η and ϕ directions, while the full simulation shows broader resolutions ($\sigma = 0.05\,\%$ for both η and ϕ).

5.4.3 E_T^{miss} resolution

Due to the presence of neutrinos in the final state of the top quark leptonic decay, the study of the E_T^{miss} is also important. In the case of the E_T^{miss}, the resolution is defined as the difference between the truth and reconstructed quantities for an event. The E_T^{miss} resolution using the fast and full simulations, as presented in Figure 5.15, shows a narrower distribution of the E_T^{miss} resolution for the fast simulation ($\sigma = 11\,\text{GeV}$) than for the full simulation ($\sigma = 15\,\text{GeV}$). Both distributions, however, are well centered around zero.

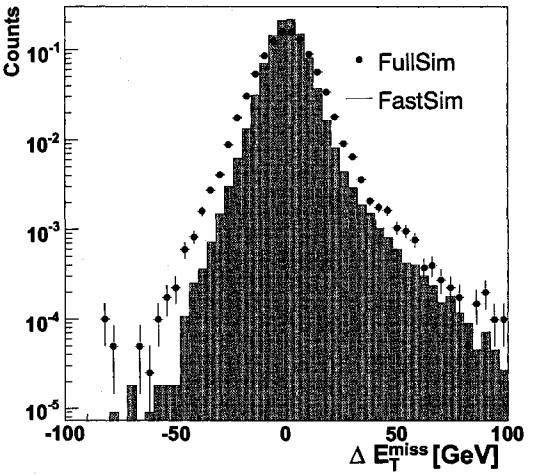


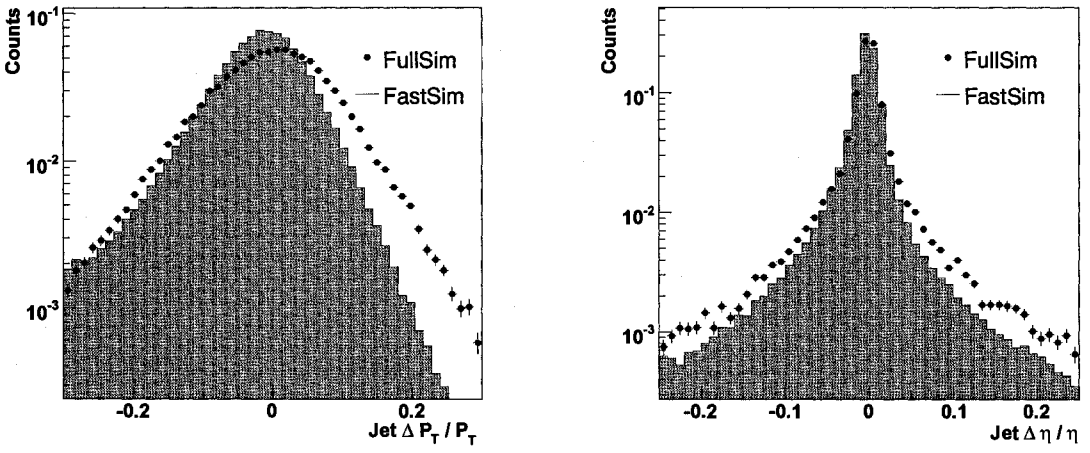
Figure 5.15: The E_T^{miss} resolution for full and fast detector simulations.

5.4.4 Jet resolution

The jet p_T and η resolutions using the fast and full simulations are presented in Figure 5.16. A negative shift in the position of the $\Delta p_T/p_T$ maximum is observed for the fast simulation (5 %). No correction has been applied for this effect. The fast simulation distribution ($\sigma = 7.7\,\%$ for the p_T distribution and $\sigma = 1.7\,\%$ for the η distribution) is also narrower than the full simulation one ($\sigma = 7.9\,\%$ for the p_T distribution and $\sigma = 1.8\,\%$ for the η distribution) in both cases.

5.5 Summary

The main results of the reconstruction performance study for the full simulation are summarized in Table 5.1. The table contains the p_T efficiencies and the standard deviation (a gaussian fit was performed around the pick of the distribution) of the



(a) (b)

Figure 5.16: The jet p_T (a) and η (b) resolution for full and fast detector simulations.

resolution for electrons, muons, jets and E_T^{miss}, for the t-channel, s-channel, W+t-channel and the $t\bar{t}$ channels together with their estimated statistical uncertainties.

The reconstruction performance study shows some differences between the full and fast simulations. Particularly, the electron reconstruction efficiency is largely overestimated by Atlfast. The η and ϕ resolution distributions point out the almost absent smearing of Atlfast on the directions of the particles. The shift in the jet p_T resolution is also an issue that needs further attention, as well as the overall resolution underestimation with Atlfast. In some situations, applying correction factors to scale to the values predicted by the full simulation is a first approach to the solution. Corrections (scaling factors) for electron efficiencies and b-tagging performance (described in [60]) were applied for the W+jets fast simulation samples. These corrections are important when studying the selection efficiencies for the top channels versus the selection efficiencies of the backgrounds (the ones generated using the fast

	Electron		Muon		Jet		E_T^{miss}
	Eff (%)	Res (%)	Eff (%)	Res (%)	Eff (%)	Res (%)	Res (%)
	±0.5	±0.03	±0.4	±0.03	±0.1	±0.1	±0.1
t-channel	60.8	2.01	86.8	2.41	96.1	7.9	17.7
s-channel	53.6	2.02	78.8	2.43	94.4	7.8	16.7
W+t	65.2	1.90	87.9	2.51	90.4	8.2	17.2
$t\bar{t}$	63.0	1.98	86.9	2.44	94.3	7.7	17.7

Table 5.1: Reconstruction efficiency (Eff) for electrons ($p_T > 25\,\text{GeV}$), muons ($p_T > 25\,\text{GeV}$) and jets ($p_T > 50\,\text{GeV}$), and standard deviation of the p_T resolution (Res) for electrons, muons, jets and E_T^{miss} (statistical uncertainties only - about 50,000 fully simulated events for each channel).

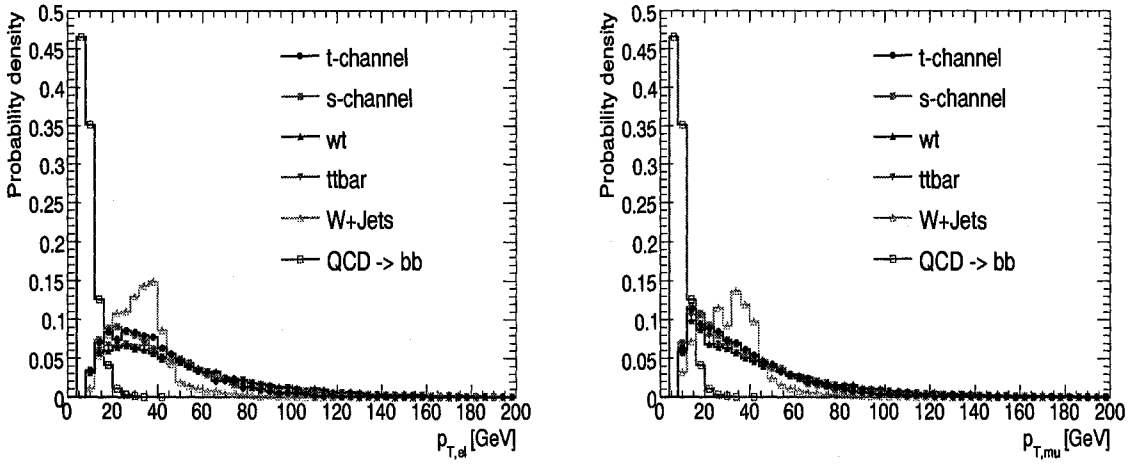
simulation) as is the case in the following section.

Chapter 6

Selection efficiencies for the t-channel and its backgrounds

6.1 Selection criteria

In order to isolate the t-channel signal from the background, one needs to apply well defined selection criteria (cuts). Various physics objects and quantities are studied in order to determine their possible use as discriminating variables for the single top quark t-channel signal versus background. Selection criteria based on lepton multiplicity, p_T and η; E_T^{miss}; jet multiplicity, p_T, η and flavor tagging information, will be applied sequentially, resulting in final efficiencies for the t-channel and its backgrounds.

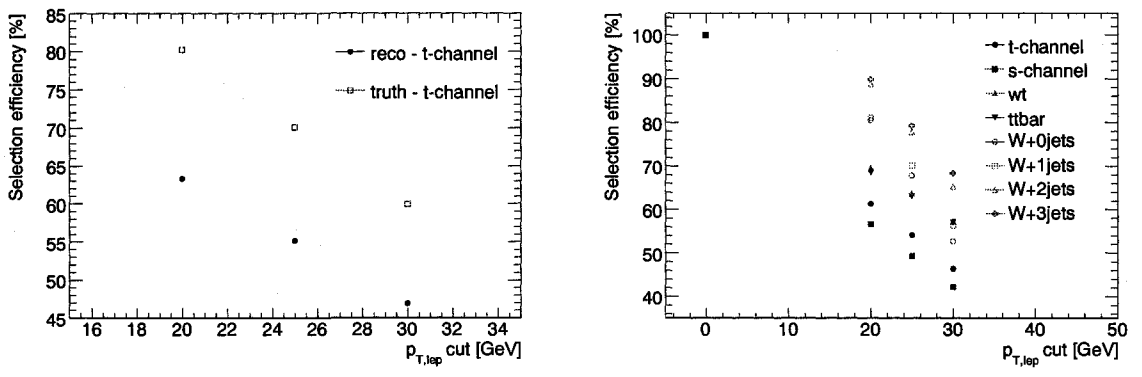


(a) (b)

Figure 6.1: Electron (a) and muon (b) transverse momentum distributions for the t-channel and the backgrounds.

6.1.1 Leptons (e and μ)

The first objects considered are the leptons. Due to the topology of the single top quark event in the t-channel, an energetic lepton originating from the W boson is expected. This lepton has a high transverse momentum, p_T, and is emitted at low pseudo-rapidity η. For this analysis, only final states containing electrons or muons are studied. Figure 6.1 (a) shows that the p_T distribution for the emitted electron has a maximum at about 30 GeV for the t-channel. The p_T distributions for backgrounds show a maximum at different p_T values. For muons, as can be seen in Figure 6.1 (b), the signal maximum is reached at lower values of p_T, about 15 GeV. Selecting only leptons with p_T greater than 25 GeV will reject mostly events coming from background processes like QCD, for which the selection efficiency becomes negligible, at about .02 %. This p_T cut value is also in the plateau region of the lepton reconstruction efficiency, as described in Section 5.3.



(a) (b)

Figure 6.2: The change in the selection efficiency with the lepton p_T cut for (a) signal at the truth and reconstruction level and (b) all channels (the QCD $b\bar{b}$ selection efficiency is on the order of .1 % (for 20 GeV), .02 % (for 25 GeV) and .006 % (for 30 GeV), and is not explicitly shown on the graph).

In order to study the effect of the lepton p_T cut on the selection efficiency for the signal and backgrounds, three values were chosen for the p_T cut: 20 GeV, 25 GeV and 30 GeV. The results for the signal selection efficiency at the reconstructed and truth level are presented in Figure 6.2 (a), while Figure 6.2 (b) shows the results for signal and backgrounds. For both signal and backgrounds, the selection efficiencies decrease almost linearly when increasing the p_T cut. The slope is almost the same for the signal and most backgrounds, therefore the lepton p_T cut is mostly effective for the rejection of the QCD $b\bar{b}$ events.

6.1.2 Missing transverse energy E_T^{miss}

Although neutrinos originating from the W boson decay are not detected by the detector, their transverse energy can be calculated as the missing transverse energy,

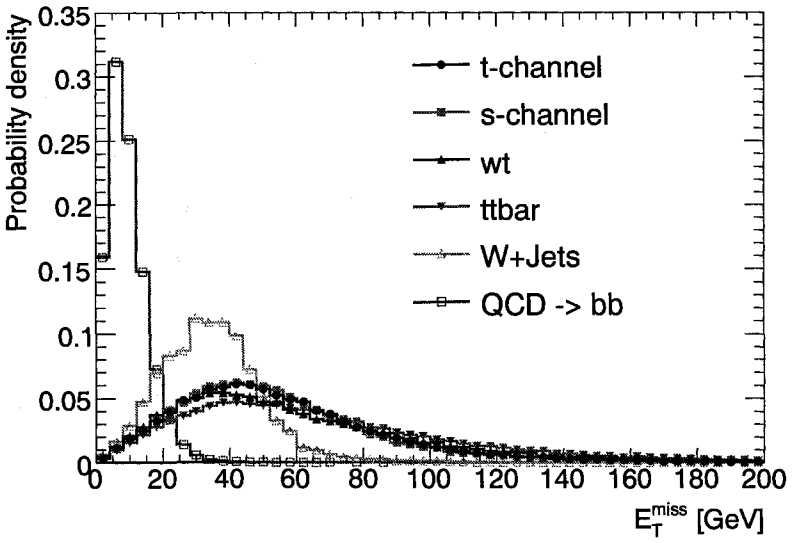
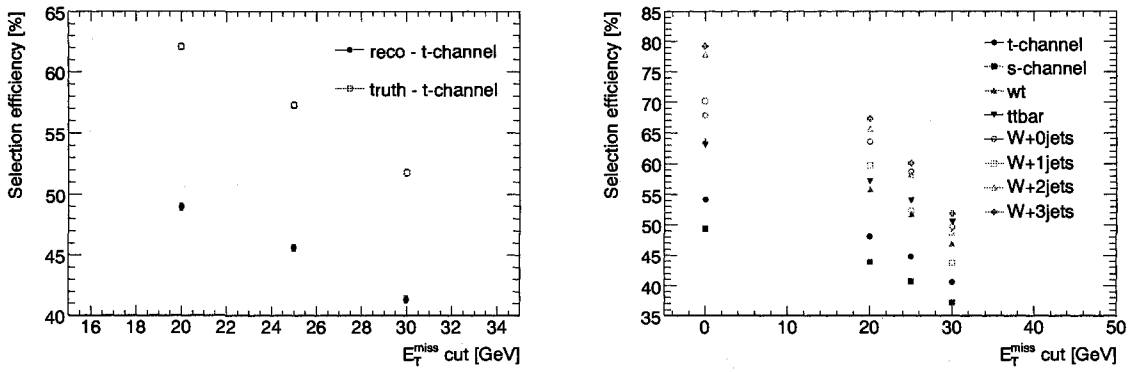


Figure 6.3: Missing transverse energy distributions for the t-channel and the backgrounds.

E_T^{miss}. This missing energy is associated with the neutrinos created in the event. In the t-channel only one neutrino is expected, associated with the lepton from the W boson decay. Figure 6.3 shows the distribution for the reconstructed E_T^{miss} in the t-channel and the backgrounds. Based on these distributions a E_T^{miss} cut of 25 GeV is chosen to reject the background events, mostly from W+jets (most of the QCD $b\bar{b}$ events were rejected by the lepton p_T cut, while the E_T^{miss} cut further reduces this background).

The effect of the E_T^{miss} cut on the selection efficiencies was studied by looking at three values of the E_T^{miss} cut: 20 GeV, 25 GeV and 30 GeV, while keeping the lepton p_T cut at 25GeV. The results are shown in Figure 6.4 (a) for the signal (reconstructed and truth level) and Figure 6.4 (b) for the signal and backgrounds. For both signal and backgrounds, the selection efficiencies decrease almost linearly when increasing



(a) (b)

Figure 6.4: The change in the selection efficiency with the E_T^{miss} cut for (a) signal at the truth and reconstruction level and (b) all channels (except QCD $b\bar{b}$).

the E_T^{miss} cut. The slope of all W+jets components is greater than the one for the signal, therefore the E_T^{miss} cut is effective for the rejection of the W+jets events.

6.1.3 Jets

The topology of the t-channel allows for the presence of two (or three) jets in the final state (Figure 1.8). They are produced by the hadronization of the b-quark and the forward quark (or the extra b-quark). The p_T distributions of the b-quark, forward quark and the $\bar{b}$-quark at the generator level are presented in Figure 6.5. As can be seen, both the b-quark and the forward quark have high p_T distributions, while the $\bar{b}$-quark has a much lower p_T. Due to its lower p_T, the third quark will mostly not be detected (or pass the basic selection cuts), leading to a low event selection efficiency. Complications with the three jets final state occur also due to the large jet multiplicity, therefore only two jets final states were considered in this thesis.

At the reconstructed level as well, the jets in the t-channel show large p_T

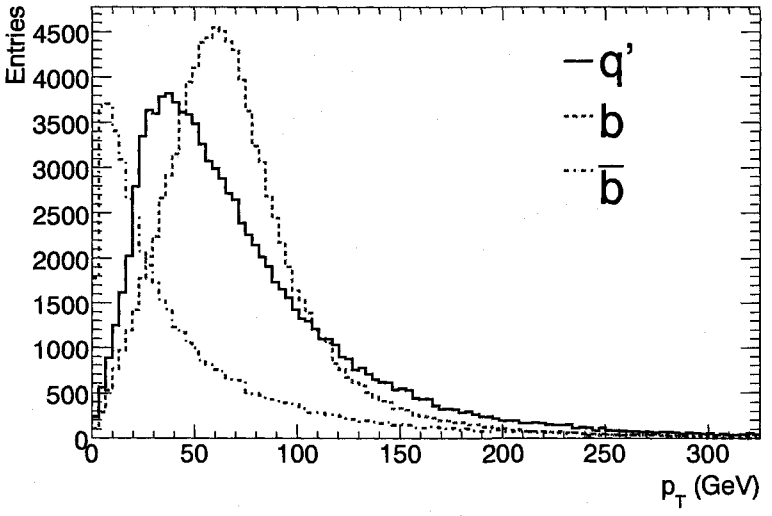
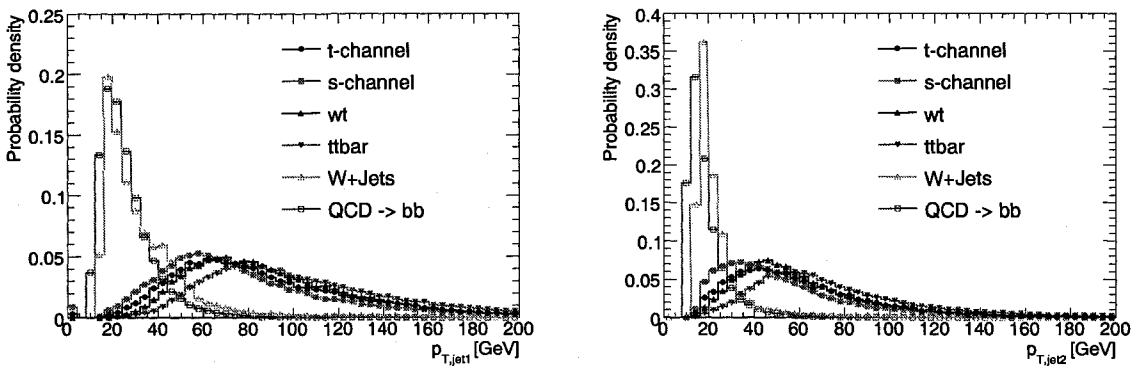


Figure 6.5: The p_T distribution for the b-quark, forward quark and the $\bar{b}$-quark at the generator level for the t-channel.

values. Figure 6.6 (a) shows the p_T distribution of the highest p_T jet for the t-channel and for the backgrounds, while Figure 6.6 (b) shows the same distributions for the second highest p_T jet. Based on these distributions a p_T cut of 50 GeV is chosen for the two jets. This cut has a high impact on the suppression of the QCD and W+jets backgrounds for which the jets are much softer. This p_T cut value is also in the plateau region of the jet reconstruction efficiency, as described in Section 5.3.

The effect of the jet p_T cut on the selection efficiencies was studied by looking at three values of the p_T cut: 40 GeV, 50 GeV and 60 GeV, while keeping both the lepton p_T and E_T^{miss} cuts at 25GeV. The results are shown in Figure 6.7 (a) for the signal (reconstructed and truth level) and Figure 6.7 (b) for the signal and backgrounds. For both signal and backgrounds, the selection efficiencies decrease when increasing the jet p_T cut, but the W+jets background is dramatically reduced (particularly in the case of the W+0jets, W+1jets and W+2jets for which the selection efficiency



(a) (b)

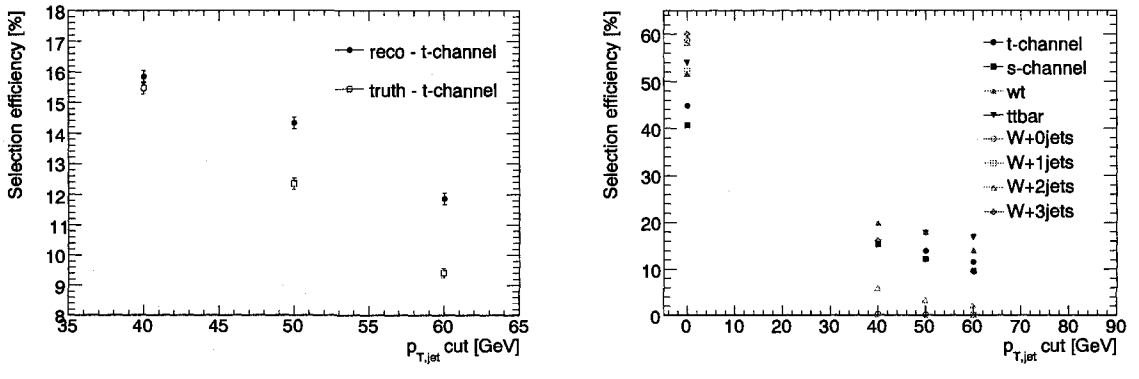
Figure 6.6: Distribution of the p_T for the highest (a) and second highest (b) p_T jets in bins of p_T for various channels.

drops to almost zero).

The generator level η distributions shown in Figure 6.8 (a) confirm the existence of a large η, forward quark, q' (it can also be seen that the b-quark has mostly a central direction). At the reconstruction level, shown in Figure 6.8 (b), a large numbers of high η jets is also observed.

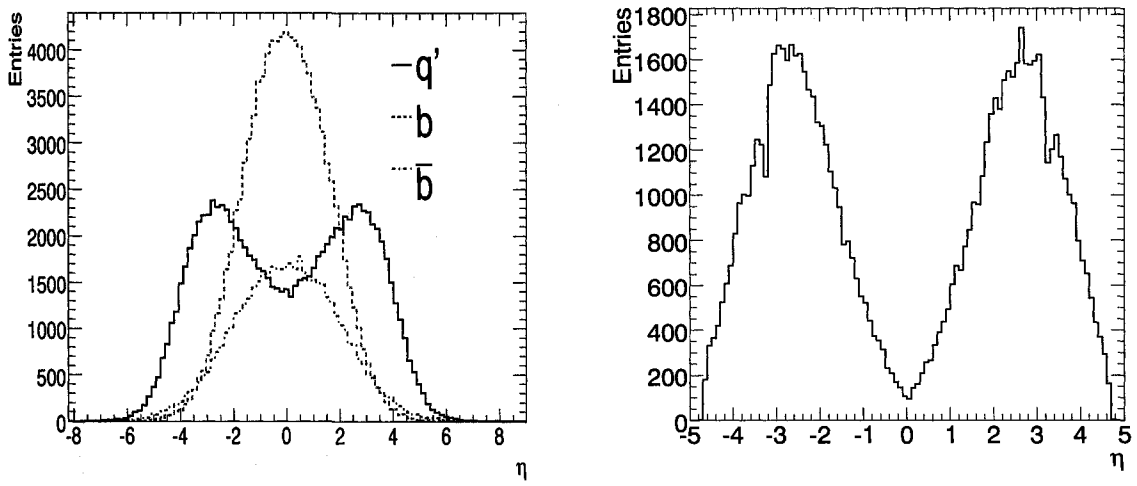
A central b-jet is expected in the t-channel together with a forward high p_T jet. Therefore the requirements of exactly one b-jet with $|\eta| < 2.5$ and one light jet with $2.5 < |\eta| < 5$ are implemented to enhance the signal events.

Various methods are implemented for the jet tagging in ATLAS. For this work the so-called IP3D+SV1 algorithm (described in Section 4.2.5) is used. Figure 6.9 shows the IP3D+SV1 distribution of the jet weight (as defined by Equation 4.1) for b-jets and light jets. In order to maintain a high b-jet tagging efficiency (about 60 %) and a low mis-tagging rate, a value of 7 was chosen for the weight cut. For the light jets, the requirement is to not be tagged as b-jets.



(a) (b)

Figure 6.7: The change in the selection efficiency with the jet p_T cut for (a) signal at the truth and reconstruction level and (b) all channels (except QCD $b\bar{b}$).



(a) (b)

Figure 6.8: The η distribution for: (a) the b-quark, forward quark and the $\bar{b}$-quark at the generator level and (b) the highest η jets at the reconstruction level.

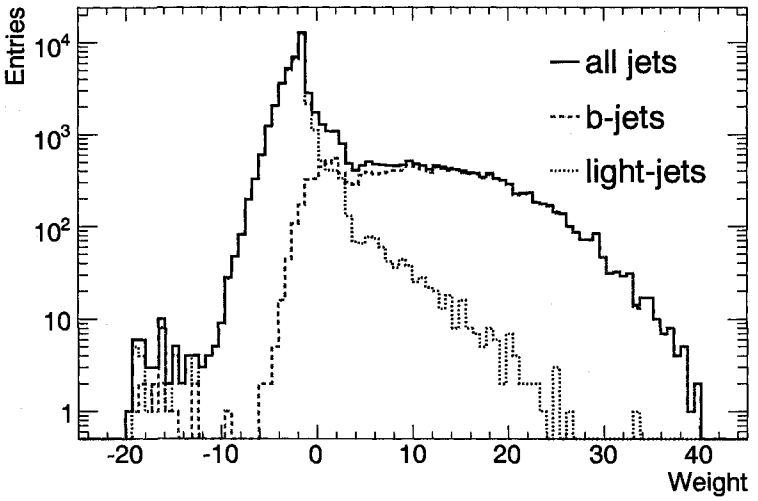


Figure 6.9: Weights for jet tagging using the IP3D + SV1 method.

6.2 Event selection efficiency and t-channel cross-section

Based on the study in the previous section, the following selection cuts are applied to enhance the selection of single top quark t-channel events with respect to the background:

- $n_{lepton} = 1$, $p_{T\,lepton} > 25\,\text{GeV}$ and $|\eta_{lepton}| < 2.5$ ($lepton = e$ or μ) (no more than one lepton in the event is required to have a p_T greater than 25 GeV and $|\eta| < 2.5$).

- $E_T^{miss} > 25\,\text{GeV}$ (the transverse missing energy of the event is required to be greater than 25 GeV).

- $n_{jets} = 2$, $p_{T\,jets} > 50\,\text{GeV}$ and $|\eta_{jets}| < 5$ (exactly two jets are required with a p_T greater than 50 GeV and $|\eta| < 5$). Further conditions on the two jets are

	Selection efficiency (%)				
	t-channel	s-channel	W+t	$t\bar{t}$	W+jets
$N_{lep} = 1$; $p_T > 25\,\text{GeV}$	54.2	49.3	63.6	63.1	72.4
$E_T^{miss} > 25\,\text{GeV}$	44.8	40.7	51.7	54.0	54.7
$N_{jet} = 2$; $p_T > 50\,\text{GeV}$	13.9	12.3	18.0	17.8	1.8
$N_{b-jet} = 1$; $\lvert\eta\rvert < 2.5$	2.86	1.05	0.59	1.23	0.026
$N_{light-jet} = 1$; $\lvert\eta\rvert > 2.5$	± 0.10	± 0.06	± 0.04	± 0.06	± 0.007

Table 6.1: Selection efficiencies after each cut for all channels. The statistical errors are given for the final selection efficiencies.

applied as follows:

- $|\eta_{jet1}| < 2.5$ (central) and tagged as b-jet (one jet, jet1, has to be tagged as a b-jet and be in the central region with $|\eta| < 2.5$).

- $|\eta_{jet2}| > 2.5$ (forward) and tagged as non-b-jet (the other jet, jet2, has to not be tagged as a b-jet and have the direction in the forward region of $2.5 < |\eta| < 5.0$).

These requirements are implemented into a ROOT macro and run over the ntuples containing the samples described in Section 4.3, both signal and backgrounds. The event selection efficiency (the ratio of the number of selected events after the cut was applied over the initial number of events, N_i) after each of the individual cuts are summarized in Table 6.1 (due to its low selection efficiency, see Section 6.1.1, the QCD $b\bar{b}$ background is not considered further in this analysis). The lepton cut has a significant impact on the number of events selected, eliminating about half of the candidates. Next, the E_T^{miss} cut decreases the number of selected events by 10 %. A

	t-channel	s-channel	W+t	$t\bar{t}$	W+jets	S/B
Initial events (e, μ)	54,200	2,200	17,800	307,333	6,546,667	0.008
Final events (e, μ)	1,550	23	105	3,780	1,702	0.28

Table 6.2: Initial and final number of events selected (for 1 fb^{-1} integrated luminosity), and the initial and final signal over background ratios, S/B. The numbers reflect the branching ratio of the leptonic (e or μ only) decay of the W ($\sigma_{e,\mu} = BR_{e,\mu} \cdot \sigma_{t-channel}$).

major effect is noticed after applying the two jet p_T cut. Only about 14 % of the events survive it. The final cut is represented by the b-tagging of the central jet and the non-b-tagging of the forward jet. This last requirement establishes the final selection efficiency for the signal events, $\epsilon = N_f/N_i$, at 2.86 % (where N_f is the final number of selected events). The background events' selection follow generally the same pattern.

The initial number of events for signal and backgrounds for 1 fb^{-1} integrated luminosity and the final number of events selected after all cuts have been applied are presented in Table 6.2. As can be seen, the major component of the backgrounds is represented by the $t\bar{t}$ events, followed by the W+jets events. This is expected due to the similar final state particles present in these channels and their large cross sections. Based on the final numbers of selected events, a signal to background ratio (S/B) of about 0.28 and a signal significance ($S/\sqrt{B}$ - an estimator of the deviation from the assumption that this is a fluctuation from purely the background [67]) of 21 are obtained (based on 1 fb^{-1} integrated luminosity).

The statistical error for 1 fb^{-1} integrated luminosity is estimated at 2.5 % using the formula (binomial error):

$$\delta\,\epsilon_{Stat} = \delta\,(\frac{N_f}{N_i}) = \sqrt{\frac{\epsilon\,(1-\epsilon)}{N_i}} = \sqrt{\frac{\epsilon\,(1-\epsilon)}{L\,\cdot\,\sigma_{e,\mu}}} \tag{6.1}$$

The cross section, $\sigma_{t-channel} = \sigma_{e,\mu}/BR_{e,\mu}$, for the signal can be calculated from the selection efficiency using the equation:

$$\sigma_{t-channel} = \frac{N_{signal+background} - N_{background}}{BR_{e,\mu} \times Signal\,selection\,efficiency \times L} \tag{6.2}$$

where $N_{signal+background}$ is the number of detected events (signal and background together), $N_{background}$ is the number of estimated background events, $BR_{e,\mu} = 21.32\,\%$ is the branching ratio of the W leptonic (e or μ only) decay and L is the integrated luminosity of the data set.

6.3 Estimation of systematics

6.3.1 Top quark mass systematics

The value of the top quark mass has an impact on the selection efficiency for all single top quark channels. Jet p_T distributions are shifted towards higher values for larger top quark masses, implying a higher selection efficiency for all top quark events. The theoretical estimation of the relative error on the cross section due to a uncertainty in the top mass of $\pm\,2.5$ GeV is about $\pm\,2\,\%$ [15].

6.3.2 Jet energy scale systematics

Although the ATLAS goal is to reach a $\pm 1\,\%$ uncertainty on the jet energy scale, JES, its value will be initially known only from Monte Carlo, beam tests and early running data, to a level of a few percent. The main effect of JES is on the selection of the objects of interest that use a p_T cut as a discriminant, such as jets and E_T^{miss}. Figure 6.11 shows the change in the selection efficiency with the variation in the jet energy scale for both the light and the b-jets. For a $2\,\%$ miscalibration in the light energy scale combined with a $3\,\%$ uncertainty in the b-jet scale, the estimated effect on the t-channel selection efficiency using Atlfast simulated data is about $\pm 5\,\%$ [69] ($\pm 4\,\%$ calculated as the addition in quadrature of the light and b-jet contributions). Due to the correlation between the two energy scales, the actual value is expected to be smaller. The lower value for the selection efficiency reported in [69] and shown in Figure 6.10 is attributed to the differences in the MC data samples used and in the selection cuts applied.

6.3.3 B-tagging systematics

The effect of b-tagging performance on the t-channel study is usually assessed by changing the b-tagging efficiency by $\pm 5\,\%$ (this relative error on the b-tagging efficiency is expected for $100\,\mathrm{pb}^{-1}$ of data [68]). Due to the fact that the b-tagging only occurs in the last cut for the t-channel events selection, the efficiency change in the b-tagging efficiency will produce a equal change in the selection efficiency. Thus, the expected effect of this change on the t-channel selection efficiency is about $\pm 5\,\%$. This value is the same as the one estimated in a previous study using Atlfast data [69]. For the Atlfast study, the variation of the selection efficiency with the change in the

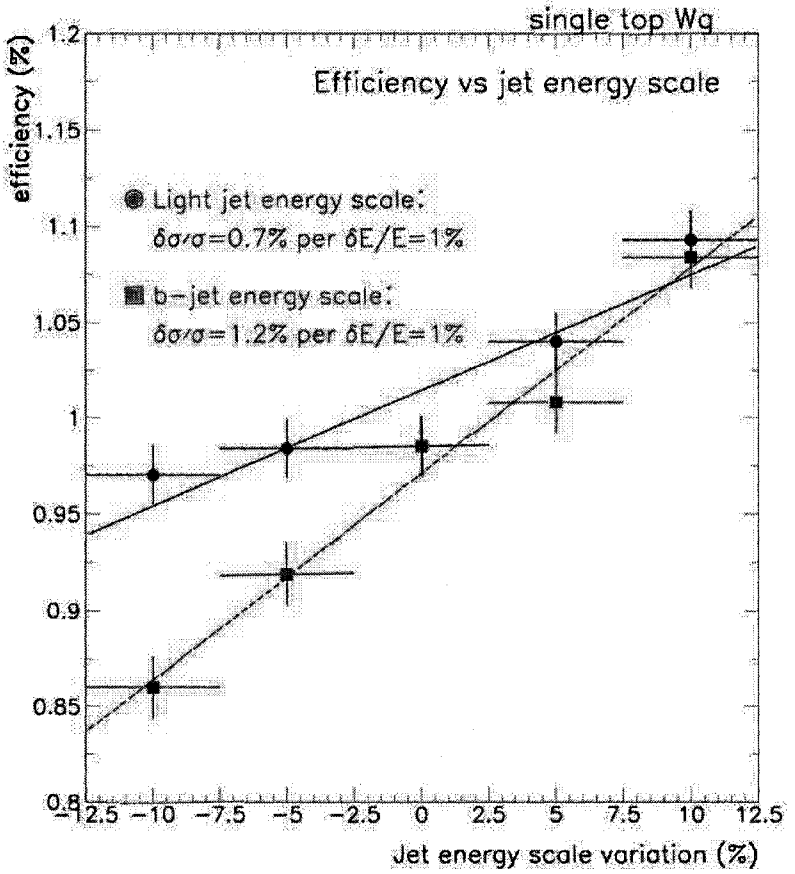


Figure 6.10: Variation of the t-channel selection efficiency with the change in the jet energy scale for Atlfast simulated data [69].

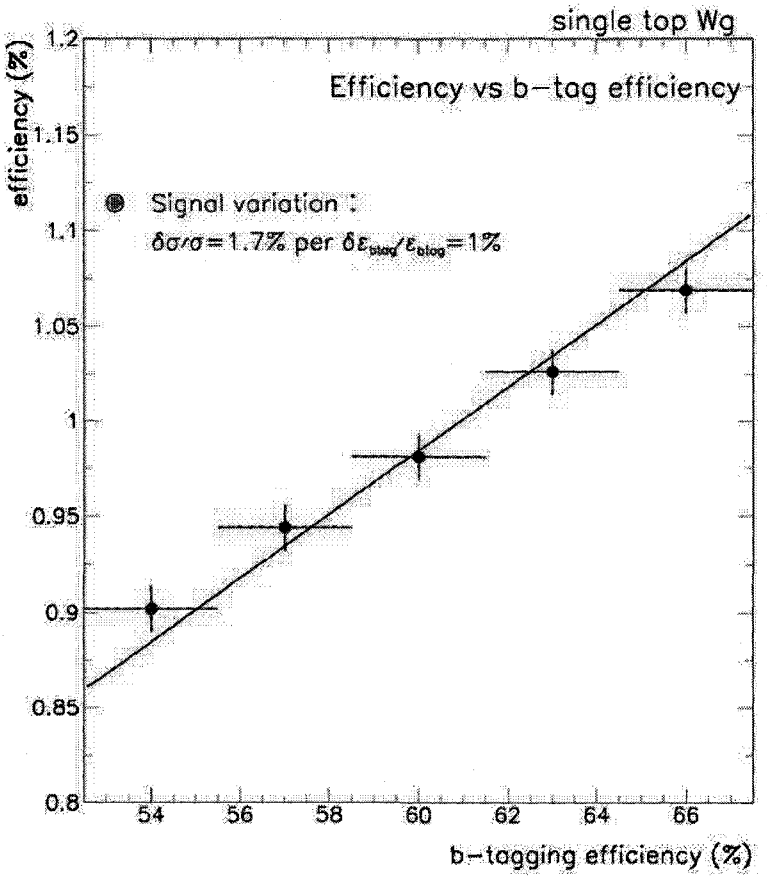


Figure 6.11: Variation of the t-channel selection efficiency with the change in the b-tagging efficiency for Atlfast simulated data [69].

b-tagging efficiency is shown in Figure 6.11. Due to the stability in the S/B ratio over the full b-tagging efficiency variation range, the quoted systematic error of this study is $\pm 5\,\%$ (about half of the value obtained form Figure 6.11).

6.3.4 ISR/FSR systematics

The events of interest at the LHC are produced through hard scattering of quarks and gluons. Before or after the hard scattering event, the involved partons can radiate

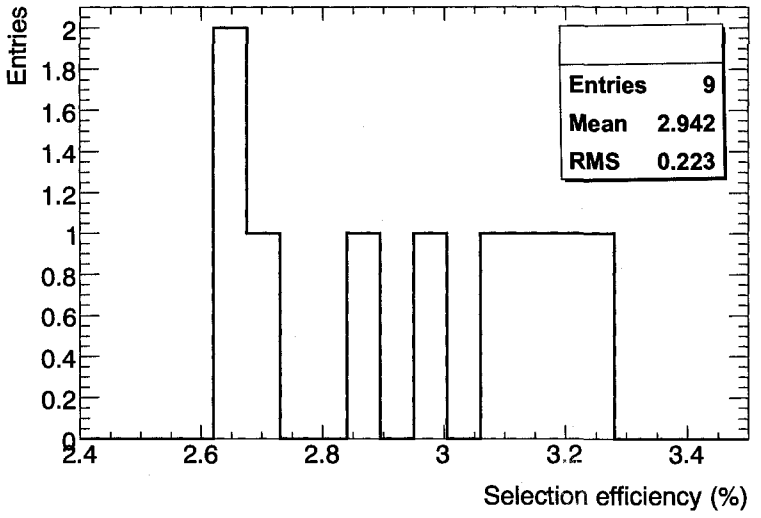


Figure 6.12: Variation of the t-channel selection efficiency for nine sets of ISR/FSR CTEQ data samples obtained when varying about 70 Pythia parameters related to ISR/FSR.

gluons, leading to the presence of initial state radiation (ISR) or final state radiation (FSR), respectively. These radiated partons manifest themselves as extra jets in the event's final state. The selection of events in the t-channel makes use of the number of jets, therefore it will be affected by these extra final state jets. In order to analyze the effect of the ISR and FSR on the t-channel cross section, the selection cuts were applied on nine samples generated by varying the corresponding default parameters in the Monte Carlo generators (about 70 variations related to ISR/FSR). More details about these samples are provided in [70]. The selection efficiencies obtained with these samples are presented in Figure 6.12.

As can be seen, the mean and the RMS of the distribution are 2.9 and 0.2, respectively. Considering the RMS as an estimate for the systematic uncertainty due to the ISR/FSR effects, the relative uncertainty on the t-channel selection efficiency

is about $\pm 8\,\%$. This value is very close to the one of $\pm 8.6\,\%$ estimated in a previous study using Atlfast data [69].

6.3.5 PDF systematics

The knowledge of the parton distribution functions in the proton plays an important role in the calculation of the production cross sections at the LHC. The effect of the PDFs on the t-channel selection efficiency was studied using two sets of distribution functions which allow estimates of the PDF uncertainties for the selection efficiencies. The first set of PDFs, CTEQ [71], includes a central value set, CTEQ6m, and 40 error sets (CTEQ6) obtained by varying each of the 20 eigenvectors in the parameter space by $\pm 1\,\sigma$ of their values. The second set of PDFs, MRST [72], contains the central value, MRST2001E, together with 30 error sets (MRST2001E) created by varying each of the 15 eigenvectors in parameter space by $\pm 1\,\sigma$.

The selection efficiency was calculated for each set using the same selection cuts as described in Section 6.2. The variation of the selection efficiency within the CTEQ set is shown in Figure 6.13. As can be seen, the selection efficiency fluctuates around $2.5\,\%$. Following the prescription described in [71] the uncertainty Δ for an observable S (e.g. selection efficiency) is calculated as:

$$\Delta S = \frac{1}{2}\sqrt{\sum_{i=1}^{N_{par}} (S_{+\sigma,i} - S_{-\sigma,i})^2} \tag{6.3}$$

where $S_{\pm\sigma,i}$ represents the calculated value of S when using the data sets in which the i-th PDF parameter was changed by $\pm\sigma$ and N_{par} represents the number of eigenvectors in the parameter space. This method yields a selection efficiency of $(2.5 \pm 0.5)\,\%$ (this value is lower than selection efficiency calculated in Section 6.2 of $2.86\,\%$, most likely due to the different data version used in the systematics study).

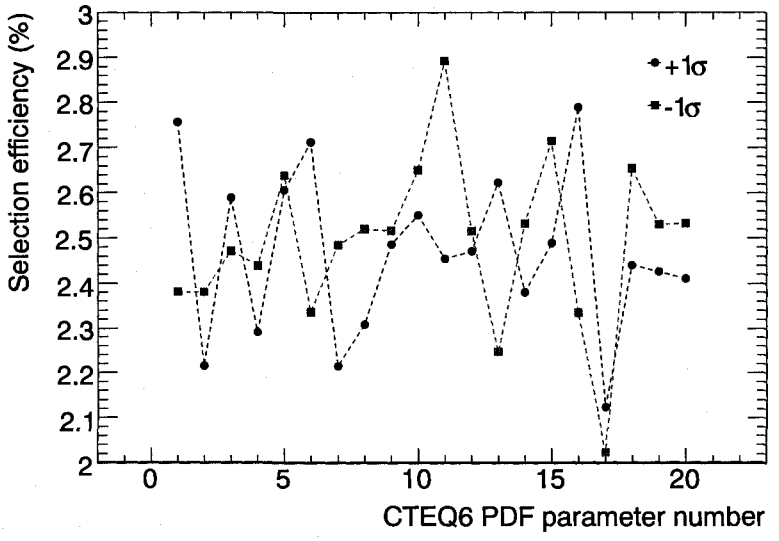


Figure 6.13: Variation of the t-channel selection efficiency for CTEQ data samples obtained when varying the 20 eigenvectors in the parameter space with $\pm 1\,\sigma$.

Due to low available statistics, the uncertainty is dominated by the statistical error, and the effect of the systematics is difficult to evaluate. A larger data sample would be necessary, but the computing resources are a limiting factor.

Similarly, Figure 6.14 shows the efficiency fluctuation for the MRST set. A selection efficiency of $(2.4 \pm 0.5)\,\%$ is obtained, limited by the low statistics as well. An alternative method is underway that uses a PDF re-weighting approach (explained in [73]). The benefit of this method is that it only requires large statistics for the central value PDF.

Thus the relative uncertainty (statistical and systematical combined) on the t-channel selection efficiency due to the PDFs is about $\pm 20\,\%$. Considering only the systematic effects (using the re-weighting approach) this value is expected to drop to about $\pm 2\,\%$ [73].

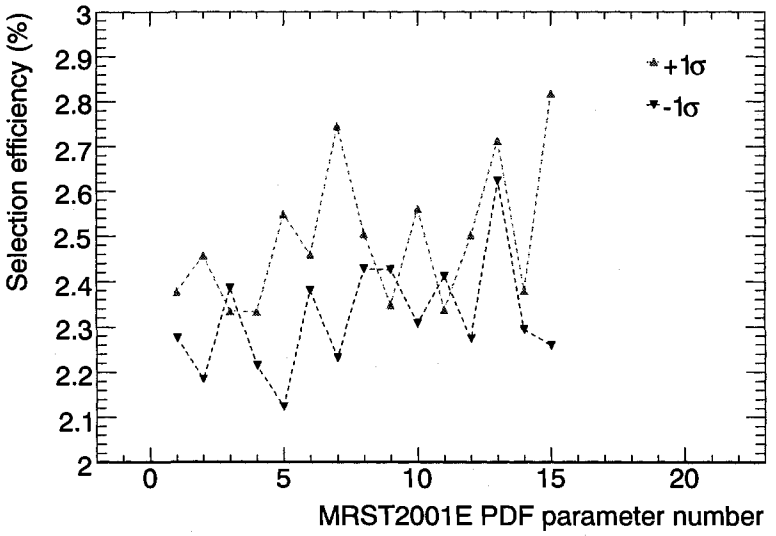


Figure 6.14: Variation of the t-channel selection efficiency for MRST data samples obtained when varying the 15 eigenvectors in the parameter space with $\pm 1\,\sigma$.

6.4 Summary

In this chapter, a selection method was developed to isolate the t-channel signal events from the background. The kinematic cuts applied to the various physics objects involved in the analysis were optimized to obtain a high signal significance.

Various systematic effects on the signal selection efficiency were considered. The final value for the estimated t-channel selection efficiency, ϵ (in percent), is:

$$\begin{aligned}\epsilon\,(\%) \;=\; & 2.86 \pm 2.5\%_{Stat} \\ & \pm 2\%_{TopMass} \pm 5\%_{JES} \pm 5\%_{B-tagging} \pm 8\%_{ISR/FSR} \pm 2\%_{PDF} \quad (6.4)\end{aligned}$$

The expected cross section for the t-channel at LHC is about 247 pb. The results given in Equation 6.4 feed into the total cross section uncertainty. The relative error in the cross section is estimated to be 44.7 % [74], leading to an uncertainty on the t-channel cross section value of approximately 110 pb for 1 fb^{-1} integrated luminosity.

Chapter 7

Conclusions

In the first part of this work, ATLAS end-cap-region calorimeter modules' response to electrons and pions of known energies was investigated in the 2002 beam test. The response of the EMEC to electrons over the energy range 6 GeV to 148 GeV was found to be linear and an electromagnetic calibration constant $\alpha_{em}^{EMEC} = (0.446 \pm 0.009)$ MeV/nA was calculated for this calorimeter. The study of the energy resolution for the EMEC showed that it can be parametrized as:

$$\frac{\sigma'(E_{reco})}{E_{reco}} = \frac{(11.8 \pm 0.2)\,\%\sqrt{\text{GeV}}}{\sqrt{E_{reco}}} \oplus (0.4 \pm 0.1)\,\% \tag{7.1}$$

where E_{reco} is in GeV.

For the response to pions, a weighting technique was developed that allowed the conversion of energy from the electromagnetic to the hadronic scale. It was shown that by applying this technique, the reconstructed energy for pions in the EMEC and HEC is consistent with the energy of the particles interacting in the calorimeters. The parametrization for the energy resolution of the hadronic response was found to

have the form:

$$\frac{\sigma'(E_{reco})}{E_{reco}} = \frac{(87.7 \pm 0.6)\,\%\sqrt{\mathrm{GeV}}}{\sqrt{E_{reco}}} \oplus (3.0 \pm 0.2)\,\% \tag{7.2}$$

where E_{reco} is in GeV. The energy resolution for both electromagnetic and hadronic response is consistent with the ATLAS detector expectation. Good energy resolutions are needed for the study of many physics channels at the LHC.

One of these channels is the single top quark t-channel. The second part of the thesis focused on the study of this channel, in particular by investigating the selection efficiency for the signal events over the background. Kinematic characteristics of physics objects were studied for their use as discriminant variables. After identifying these discriminants, a cut flow analysis was performed on both signal and backgrounds. It was found that the selection efficiency, $\epsilon\,(\%)$, for the t-channel is:

$$\begin{aligned}\epsilon\,(\%) \;=\; & 2.86 \pm 2.5\%_{Stat} \\ & \pm 2\%_{TopMass} \pm 5\%_{JES} \pm 5\%_{B-tagging} \pm 8\%_{ISR/FSR} \pm 2\%_{PDF}\end{aligned} \tag{7.3}$$

where the statistical error is evaluated for an integrated luminosity of $1\,\mathrm{fb}^{-1}$. For the same luminosity, a signal to background ratio (S/B) of 0.28 and a signal significance ($\mathrm{S}/\sqrt{\mathrm{B}}$) of 21 were obtained. The t-channel cross section value with this luminosity is estimated at $(247 \pm 110)\,\mathrm{pb}$.

The LHC is expected to start running in the fall of 2008. The study of the electroweak production of the top quark will be among the first physics studies that can be performed. Although the production of single top events was reported last year by the Tevatron, the high statistics available at the LHC will allow the study of top quark properties that can only be accessed directly in the single top channels.

The results of this thesis have provided valuable feedback to the improvement of the analysis of ATLAS data. Several aspects of the beam test studies of the

ATLAS end-cap calorimeters have evolved to become part of the standard calibration procedure of the calorimeters. The clustering algorithm developed in this thesis is the precursor of the topological clustering algorithm that is used in the ATLAS official software, while the electromagnetic and hadronic calibration methods are used for the ATLAS calorimeters calibration. The calibration constants derived from beam test data are being used both in Monte Carlo and in the commissioning of the calorimeters, while resolution results are used to tune the Monte Carlo.

In order to refine the accuracy of the fast Monte Carlo simulation, detailed comparisons of fast and full Monte Carlo's are needed. The performance reconstruction results are used to calibrate the Atlfast response and validate the Computing System Commissioning production. The t-channel selection cuts used in this thesis were extended and improved to develop the official ATLAS single top t-channel analysis strategy. These studies have lead to a realistic estimation of the identification of a top channel which can be accomplished with first data.

Bibliography

[1] *http://cpepweb.org*

[2] F. Abe *et al.*, The CDF Collaboration, *Observation of Top Quark Production in $p\overline{p}$ Collisions with the Collider Detector at Fermilab*, Phys. Rev. Lett. **74**, 2626 (1995).

[3] S. Abachi *et al.*, The D0 Collaboration, *Observation of Top Quark*, Phys. Rev. Lett. **74**, 2632 (1995).

[4] The CDF Collaboration, the D0 Collaboration and the Tevatron Electroweak Working Group, *Combination of CDF and D0 Results on the Mass of the Top Quark*, hep-ex/0703034 (2007).

[5] *http://www-cdf.fnal.gov/physics/new/top/2007/mass/tevcombination*

[6] R.K. Ellis, W.J. Stirling, and B.R. Webber, *QCD and Collider Physics*, New York: Cambridge University Press (1996).

[7] *http://www-cdf.fnal.gov/physics/new/top/2007/mass/topwidth/topwidth.html*

[8] V. M. Abazov *et al.*, The D0 Collaboration, *Experimental Discrimination between Charge 2e/3 Top Quark and Charge 4e/3 Exotic Quark Production Scenarios*, Phys. Rev. Lett. **98**, 041801 (2007).

[9] I. Borjanovic *et al.*, The ATLAS Collaboration, *Investigation of top mass measurements with the ATLAS detector at LHC*, Eur. Jour. Phys. C **39S2**, 47 (2004); hep-ex/0403021.

[10] *http://www-cdf.fnal.gov/physics/new/top/public_xsection.html*

[11] *http://www-d0.fnal.gov/Run2Physics/top/top_public_web_pages/top_public.html*

[12] J. M. Campbell, J. W. Huston and W. J. Stirling, *Hard Interactions of Quarks and Gluons: a Primer for LHC Physics*, hep-ph/0611148 (2006).

[13] *http://www-cdf.fnal.gov/physics/new/top/2007/singletop/ME1_5fb*

[14] V. M. Abazov *et al.*, The D0 Collaboration, *Evidence for Production of Single Top Quarks and First Direct Measurement of* $|V_{tb}|$, Phys. Rev. Lett. **98**, 181802 (2007).

[15] Zack Sullivan, *Understanding Single-Top-Quark Production and Jets at Hadron Colliders*, Phys. Rev. D **70**, 114012 (2004).

[16] John Campbell and Francesco Tramontano, *Next-to-Leading Order Corrections to Wt Production and Decay*, Nucl. Phys. B **726**, 109 (2005).

[17] Nikolaos Kidonakis and Ramona Vogt, *Next-to-Next-to-Leading Order Soft-Gluon Corrections in Top Quark Hadroproduction*, Phys. Rev. D **68**, 114014 (2003).

[18] CERN Communication Group, *CERN FAQ LHC the guide*, CERN-Brochure-2008-001-Eng (2008); *http://public.web.cern.ch/public/en/LHC/Facts-en.html*

[19] The ATLAS LARG Unit, *Liquid Argon Calorimeter Technical Design Report*, CERN/LHCC 96-41 (1996).

[20] The ATLAS Collaboration, *ATLAS Calorimeter Performance Technical Design Report*, CERN/LHCC 96-40 (1997).

[21] The ATLAS Collaboration, *ATLAS Detector and Physics Performance Technical Design Report*, Vol. I, CERN/LHCC 99-14 (1999).

[22] The ATLAS Collaboration, *ATLAS Detector and Physics Performance Technical Design Report*, Vol. II, CERN/LHCC 99-15 (1999).

[23] The ATLAS Computing Group, *ATLAS Computing Technical Design Report*, CERN/LHCC 2005-022 (2005).

[24] The ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, submitted to JINST (2007).

[25] W.-M. Yao *et al.*, *The Review of Particle Physics*, J. Phys. G **33**, 1 (2006). *http://pdg.lbl.gov*

[26] Richard Fernow, *Introduction to Experimental Particle Physics*, New York: Cambridge University Press (1986).

[27] D. J. Griffiths, *Introduction to Elementary Particle Physics*, John Wiley & Sons Inc. (1987).

[28] Richard Wigmans, *Calorimetery: Energy Measurement in Particle Physics*, New York: Oxford University Press (2000).

[29] B.B. Brabson *et al.*, *A study of two prototype lead glass electromagnetic calorimeters*, NIM A **332**, 419 (1993).

[30] W.E. Cleland and E.G. Stern, *Signal processing consideration for liquid ionization calorimeters in a high rate environment*, NIM A **338**, 467 (1994).

[31] B. Aubert *et al.*, *Performance of the ATLAS electromagnetic end-cap module 0*, NIM A **500**, 178 (2003).

[32] B. Aubert *et al.*, *Performance of the ATLAS electromagnetic barrel module 0*, NIM A **500**, 202 (2003).

[33] S. Simion, *Liquid Argon Calorimeter ROD*, 2nd ATLAS ROD workshop, October 2000, University of Geneva,
http : *//dpnc.unige.ch/atlas/rod*00*/transp/S_Simion.pdf*.

[34] O. Martin, E. Monnier and S. Tisserant, *Update of some Geometrical Parameters for the ATLAS E.M. End-Cap Calorimeter*, ATL-LARG-96-047 (1996).

[35] B. Dowler *et al.*, The ATLAS Liquid Argon HEC Collaboration, *Performance of the ATLAS Hadronic End-cap Calorimeter in Beam Tests*, NIM A **482**, 94 (2002).

[36] The ATLAS LAr Collaboration, *INTAS Final Report 2003: INTAS-CERN99-00278*, INTAS-CERN99-00278 (2003).

[37] C. Cojocaru *et al.*, The ATLAS Liquid Argon EMEC/HEC Collaboration, *Hadronic Calibration of the ATLAS Liquid Argon Endcap Calorimeter in the Region* $1.6 < |\eta| < 1.8$ *in Beamtests*, NIM A **531**, 481 (2004).

[38] D. M. Gingrich *et al.*, The ATLAS Hadronic End-Cap Calorimeter Group, *Construction, Assembly and Testing of the ATLAS Hadronic End-Cap Calorimeter*, JINST **2**, P05005 (2007).

[39] M. Vincter and C. Cojocaru, *Electronic Noise in the 2002 HEC/EMEC Test Beam*, ATLAS-HEC-Note-149 (2003).

[40] C. Cojocaru *et al.*, The ATLAS Liquid Argon EMEC/HEC Collaboration, *Muon Results from the EMEC/HEC Combined Run Corresponding to the ATLAS Pseudorapidity Region* $1.6 < |\eta| < 1.8$, ATL-LARG-2004-006 (2004).

[41] I. Abt *et al.*, *The tracking, calorimeter and muon detectors of the H1 experiment at HERA*, NIM A **386**, 348 (1997).

[42] F. James, *MINUIT Function Minimization and Error Analysis Reference Manual Version 94.1*, CERN Program Library Long Writeup D506 (1998).

[43] *https://twiki.cern.ch/twiki/bin/view/Atlas/WorkBookFullChain*

[44] M.A.Dobbs *et al.*, *Les Houches Guidebook to Monte Carlo Generators for Hadron Collider Physics*, hep-ph/0403045 (2004).

[45] T. Sjostrand *et al.*, *Pythia 6.4 Physics and Manual*, JHEP **5**, 26 (2006); hep-ph/0603175; *http://www.thep.lu.se/~torbjorn/Pythia.html*

[46] *http://hepwww.rl.ac.uk/theory/seymour/herwig*

[47] *http://www.hep.fsu.edu/~isajet*

[48] *http://projects.hepforge.org/sherpa/dokuwiki/doku.php*

[49] *http://www.hep.phy.cam.ac.uk/theory/webber/MCatNLO*

[50] *http://sirius.ihep.su/~spitsky/toprex/toprex.html*

[51] *http://borut.web.cern.ch/borut*

[52] *http://mlm.home.cern.ch/mlm/alpgen*

[53] B. P. Kersevan and I. Hinchliffe, *A Consistent Prescription for the Production Involving Massive Quarks in Hadron Collisions*, hep-ph/0603068 (2006).

[54] E.E. Boos, L.V. Dudko and V.I. Savrin, *"SingleTop" - An Event Generator for the Single Top Quark Production at the LHC. Part1.*, CMS NOTE-2000/065 (2001).

[55] S. Agostinelli *et al.*, *Geant4 a simulation toolkit*, NIM A **506**, 250 (2003).

[56] E. Richter-Was, D. Froidevaux and L. Poggioli, *ATLFAST 2.0 a fast simulation package for ATLAS*, ATL-PHYS-098-131 (1998).

[57] *http://atlas-computing.web.cern.ch/atlas-computing/computing.php*

[58] A. Shibata, *TopView - An AOD Analysis Package for ATLAS Top Physics Analysis*, ATL-SOFT-PUB-2007-002 (2007); *https://twiki.cern.ch/twiki/bin/view/Atlas/TopView*

[59] *http://lpsc.in2p3.fr/atlas/lucotte/4.GENERATOR/page_generator.html*

[60] A. Shibata and B. Clement, *Tagging Rate Function B-Tagging*, ATL-PHYS-PUB-2007-011 (2007).

[61] *https://twiki.cern.ch/twiki/bin/view/Atlas/TopNtupleAvailability*

[62] S. Kotov, *B-tagging in ATLAS*, MPI seminar, Munich, Germany, May 19, 2005 (2005).

[63] M. Lefebvre and P. Loch, *Introduction to Hadronic Calibration in ATLAS*, 3-rd ATLAS Hadronic Calibration Workshop, Milan, Italy, April 26-27, 2007 (2007).

[64] G.C. Blazey *et al.*, *Run II Jet Physics*, hep-ex/0005012 (2000).

[65] S. D. Ellis *et al.*, *Jets in Hadron-Hadron Collisions*, COM-PHYS-2007-085 (2007).

[66] A. Shibata, *Remaining Issues*, talk at the Single Top Group meeting - CSC T8, February 13, 2008 (2008).

[67] S.I. Bityukov and N.V. Krasnikov, *On the observability of a signal above background*, NIM A **452**, 518 (2000).

[68] *https://twiki.cern.ch/twiki/bin/view/Atlas/BTaggingRecommendationsForCSCNotes*

[69] A. Lucotte, A. Lleres and F. Chevallier, *Prospects for single top cross-section measurements in Atlas*, ATL-PHYS-PUB-2007-005 (2007).

[70] J. Cochran and K. Yamanaka, *ISR/FSR MC studies*, talk at the Single Top Group meeting - CSC T8, September 20, 2007 (2007).

[71] J. Pumplin *et al.*, *New Generation of Parton Distributions with Uncertainties from Global QCD Analysis*, hep-ph/0201195 (2002).

[72] A.D. Martin *et al.*, *Uncertainties of predictions from parton distributions. I: experimental errors*, hep-ph/0211080 (2002).

[73] G. Khoriauli and M. Cristinziani, *Study of Single Top Systematic Uncertainties*, talk at the Top working meeting, February 13, 2008 (2008).

[74] ATLAS Single Top Working Group, *Prospect for single top cross section measurements in ATLAS*, note in preparation (2008).

* All web pages were last accessed in March 2008.

Appendix A

Participation in the ATLAS experiment

My first involvement with the ATLAS experiment was through the participation in the beam tests for EMEC and HEC calorimeter modules at CERN. I was part of the teams that run the beam tests in 2001, 2002 and 2004. During the run shifts I monitored the data collection, inspected the infrastructure to insure the proper functioning of the experiment and, as a shift leader for some of the 2004 runs, I oversaw the activities of regular shift-crews. In the summer of 2002 I also helped in the assembling of the HEC wheels at CERN. The analyses described in this thesis is based on the data collected in the 2002 beam test. I started by studying the electronic noise in both the EMEC and HEC modules. The results on the noise analysis were presented at various LAr Group meetings at CERN and were published in a ATLAS internal scientific note. Next, I studied the response of the EMEC to electron beams. A clustering algorithm was developed that was used to reconstruct the energy of electrons in the EMEC. After a few corrections were applied, the linearity of the

detector response was tested using various beam energies and an electromagnetic conversion constant was calculated. For the response of the calorimeter modules to pions, a weighting method was implemented and tested. For both electron and pion responses, the energy resolution was investigated as a function of beam energy and impact point. The results of these studies were presented at various LAr Group meetings at CERN, at the 2005 CAP congress and were part of a paper published in NIM. Due to my involvement in the above mention activities, I am also a co-author on some other scientific publications.

The second part of my involvement with the ATLAS experiment was on the analysis of the top quark production at the LHC. I started by investigating the effects of various jet algorithms on the jet reconstruction performance for top channels and the mass of the top quark. Once the Computing System Commissioning (CSC) data became available, I worked on the reconstruction performance of physics objects relevant for the analysis of single top channels. I studied the reconstruction efficiency, purity and resolution of objects such as: electrons, muons, E_T^{miss} and jets. Next, I investigated the selection efficiency of the t-channel events and its backgrounds. The results of these analyses were presented at various Top Group meetings at CERN and North American ATLAS Physics meetings, and are part of the Top CSC note, currently in the process of being written.

Refereed Publications: [24], [37], [38].

Internal Scientific Notes: [36], [39], [40], [74].

Appendix B

Abbreviations

ADC	Analog to Digital Converter
AOD	Analysis Object Data
ATLAS	A Large Toroidal LHC ApparatuS for the LHC
CERN	the European Organization for Nuclear Research
CKM	Cabibbo-Kobayashi-Maskawa
CSC	Computing System Commissioning
EMEC	Electromagnetic End-cap Calorimeter
ESD	Event Summary Data
FCal	Forward Calorimeter
FEB	Front End Boards
FSR	Final State Radiation
HEC	Hadronic End-cap Calorimeter
ISR	Initial State Radiation

LAr	Liquid Argon
LHC	the Large Hadron Collider at CERN
LO	Leading Order
MC	Monte Carlo
MLM	Michelangelo Mangano
MuID	Muon IDentification
MWPC	Multiwire Proportional Chambers
NLO	Next to Leading Order
NNLL	Next to Next to Leading Logarithm
OF	Optimal Filtering
OFCs	Optimal Filtering Coefficients
PDF	Parton Distribution Function
PS	Proton Synchrotron
PSB	Proton Synchrotron Booster
PU	Processing Unit
QCD	Quantum ChromoDynamics
ROD	Read Out Driver
SPS	Super Proton Synchrotron
STACO	STAtistical COmbination
SUSY	SUper SYmmetry
TRT	Transition Radiation Tracker