Università degli studi di Pisa
Dipartimento di fisica
Scuola di dottorato in fisica “G. Galilei”
XVIII ciclo

Top quark mass determination at the ATLAS experiment using early data

Andrea Dotti

ADVISOR:
Prof. **Tarcisio Del Prete**

To my family

Contents

Introduction

The Large Hadron Collider (LHC), at present in an advanced construction phase at CERN will be the next generation accelerator that will allow to understand physics at a more fundamental level with respect to present knowledge.

The design center of mass energy is one order of magnitude larger than that of the most powerful existing hadron collider and the high luminosity will allow to study extremely rare events. Both experimental and theoretical physicists hope and expect that the LHC will reveal a lot of new physics.

The most important goals of the two general-purpose experiments at LHC (ATLAS and CMS) are the study of the spontaneous symmetry breaking mechanism and the discovery of possible extensions to the Standard Model of the Elementary Particles physics.

The LHC experiments will also play a fundamental role in precise measurements in the Standard Model field. These measurements are sensible to the presence of new particles and interactions that can reveal their presence through deviations of the measured quantities with respect to the predictions of the Standard Model.

In this sense the most promising field is the relatively recent top quark field. Its mass is close to the electro weak scale, ant it is a probe for extension of the Standard Model playing a fundamental role in many new physics models.

Both ATLAS and CMS have an extensive program to measure top quark properties with high precision. Top-antitop pair cross-section production and top quark mass measurement in the $t\bar{t}$ events are among the first measurements accessible at LHC.

In particular the top mass measurements, together with a precise W boson mass measurement can constrain the mass of the Higgs boson both in the Standard Model and in its Minimal Super Symmetric extension.

More challenging studies, such as the single top cross-section measurements, are directly sensitive to new heavy bosons predicted by extensions of the Standard Model.

The LHC "golden channel" for the top quark mass measurement is the semileptonic $t\bar{t}$ decay: $t\bar{t} \rightarrow \mathrm{W^+W^-b\bar{b}} \rightarrow \ell^{\pm}\mathrm{jjb\bar{b}}$. The top mass is extracted from the three jets (jjb) invariant mass of the hadronically decaying top.

The isolated lepton and the missing transverse energy revealing the presence of the neutrino, are used to select the events with high efficiency. The main backgrounds to this analysis are the direct production of W and Z bosons in association with jets and the QCD multi-jet production. Both can be drastically reduced requiring that two of the four jets in the final state are originating from b-quarks.

During the first period of data taking, even if the detector performances are not the final ones a preliminary top quark mass measurement can be performed, provided that the jets are reconstructed efficiently and their energy is measured with sufficient precision.

This thesis focuses on the top quark mass measurement with the first hundreds of pb^{-1} of collected luminosity, that can be achieved within some days of data taking. The top quark is reconstructed from the jets in the final state and emphasis is put in the jet reconstruction and calibration. This requires the precise knowledge of calorimetric performances and detailed studies on test beam data and on the detector simulations are discussed.

The first chapter provides and overview of the top quark physics. After a brief introduction to the Standard Model the role of top quark is presented, current knowledge on top quark physics, as achieved at the Tevatron experiments, is also summarized. Then the potential in this field of the LHC experiments is presented. For this discussion we will focus on the studies performed for the ATLAS detector.

The second chapter concerns details of the ATLAS detector. After a brief description of the LHC, we describe the inner detector, the calorimeters, the muon spectrometer, the trigger system and the computing model. ATLAS is in a well advanced status of construction and final testing, thus the current activities in the experimental hall are presented.

Since we are primarily interested in the energy measurements of hadrons the third chapter is dedicated to the ATLAS calorimeters. Details are given on both the electromagnetic and the hadronic calorimeters, the peculiarities of the interaction of hadrons with matter are presented as an introduction to the discussion of results obtained at the test beam programs of the ATLAS barrel hadronic calorimeter. We will present the calibration strategy for the barrel hadronic calorimeter and we will analyze the response to single electrons and pions. We will also start to compare the results with the prediction of the simulations.

The next chapter is entirely dedicated to the validation of the physics simulation with emphasis on the simulation of high energy hadrons interacting with the ATLAS calorimetric system. This is of fundamental importance since many calibration tools rely, at least partially, on the capability of the Monte Carlo to precisely describe real data. A detailed comparison between data and simulation can be performed with the data collected during the test beam activities. In particular during the ATLAS combined test beam, for the first time, an entire slice of the detector has been exposed to the different particles: single electrons, pions, muons. After presenting the experimental setup, the results obtained on single pions are compared to Geant4 Monte Carlo predictions.

Confident of the good level of agreement between data and simulation we present, in chapter five, a method that uses weights, extracted from Monte Carlo simulations, to calibrate jets. This method corrects for detector effects such as cracks and non linearity and, at the same time, improves the resolution on jet transverse energy measurement. The performance of the algorithm are tested on QCD and $t\bar{t}$ simulated events obtaining good results both on linearity and transverse energy resolution.

Finally, in chapter six, we discuss the top quark mass measurement with the very first collected data. The analysis concentrate on semi-leptonic events: $t\bar{t} \rightarrow \mathrm{W^+W^-b\bar{b}} \rightarrow \ell^{\pm}\mathrm{jjb\bar{b}}$. We will show that a preliminary top quark mass measurement is possible even without b-tagging and the role of the most important background, $W{+}4jets$, will be discussed. Top mass is measured from the three jets invariant mass and the jets are calibrated with the method that we developed. We will show that a good result is achieved and we will discuss a possible methodology to improve the signal over background ratio. The jj invariant mass will be used to attempt the reconstruction of the decay $W \rightarrow jj$. Even if, for a precise mass measurement and to deal with the combinatorial background, the use of the b-tagging is mandatory it is still possible to use the reconstructed W boson to improve the S/B ratio in the top sample.

Chapter 1

Phenomenological overview

The modern description of fundamental particles and their interaction is based on the gauge invariant theory known as Standard Model (SM) which has enjoyed a big experimental success. The missing ingredient of the SM, whose discovery and study is the main goal of near-future experiment, is the Higgs mechanism. The future experiments will also perform detailed studies of SM particles and phenomena to verify at a high precision level the SM predictions. In particular the top quark, due to its mass close to the electro weak scale, plays an important role in the SM and in any new physics scenario.
In this chapter, after an introduction on the SM, the top quark physics will be shortly reviewed, with particular interest for its mass measurement. The experimental potential in this field for the future ATLAS experiment will be presented.

1.1 The Standard Model

The modern description of the fundamental particles found in nature and of their interactions is based on a gauge invariant theory known as the *Standard Model* (SM) [1, 2, 3]. Since it was first proposed –more than 25 years ago– the SM has enjoyed a big experimental success and its experimental test became more and more precise over the last two decades. No significant deviation from the predictions of this theory was found [4].

In the SM the fundamental particles are divided into spin $\frac{1}{2}$ fermions and spin 1 bosons, which play the role of force carriers. There are three different interactions and three different kinds of vector bosons:

- **electromagnetic**, acting between charged particles. The carrier of this force is the neutral photon (γ). The theory of electromagnetic interactions (QED) is based on the $U(1)$ symmetry group;

- **weak**, acting between particles carrying weak isospin. The carriers of this force are three vector bosons W^+, W^-, Z^0. The theory of weak interactions is based on the $SU(2)_L$ group of symmetry. This symmetry is the same used in quantum mechanics to describe the spin of particles, and for this reason the word isospin is used. In particular the three force carriers constitute an isospin-1 triplet;

- **strong**, acting between particles carrying strong hyper-charge or colour. The strong interaction (quantum chromodynamics or QCD) is based on the $SU(3)_C$ symmetry group and the carriers of this force are 8 bosons called gluons. They carry colour charges themselves, and are thus self-interacting. This implies that the QCD coupling α_s is small for large momentum transfers but large for small momentum transfers, and leads to the confinement of quarks inside colour-neutral hadrons. Attempting to free a quark produces a jet of hadrons through production of quark-antiquark pairs and gluons.

All the particles are characterised using the “charges” they carry and, consequently, looking at the interactions they undergo. Fermions carrying a colour charge are called quarks, while those with zero colour charge are called leptons. All fermions are divided into isospin $\frac{1}{2}$ left-handed doublets and isospin 0 right-handed singlets (see table 1.1). Three quark doublets as well as three lepton doublets have been observed, while there are six quark singlets and three lepton singlets. Note that there are no right-handed neutrinos since there would be no way for them to interact with other particles. They would thus be undetectable.

The Standard Model gauge symmetry is:

$$G_{\mathrm{SM}} = SU(3)_{\mathrm{C}} \times SU(2)_{\mathrm{L}} \times U(1)_{\mathrm{Y}} \tag{1.1}$$

	Left handed doublets	**Right handed singlets**
Leptons	$\binom{\nu_e}{e}_L \quad \binom{\nu_\mu}{\mu}_L \quad \binom{\nu_\tau}{\tau}_L$	$e_R \; \mu_R \; \tau_R$
	$\lvert\vec{T}\rvert = \frac{1}{2}$, $Y = -\frac{1}{2}$	$\lvert\vec{T}\rvert = 0$, $Y = -1$
Quarks	$\binom{u}{d}_L \quad \binom{c}{s}_L \quad \binom{t}{b}_L$	$u_R \; d_R \; c_R \; s_R \; b_R \; t_R$
	$\lvert\vec{T}\rvert = \frac{1}{2}$, $Y = \frac{1}{6}$	$\lvert\vec{T}\rvert = 0$, $Y(u_r, c_r, t_r) = \frac{2}{3}$ $Y(d_r, s_r, b_r) = -\frac{1}{3}$

Table 1.1: *Classification of fermions predicted by the SM divided into leptons and quarks. The weak isospin $\vec{T}$ and hyper-charge Y are given for each group.*

A compact way to write the representations (table 1.1) of the SM gauge group, which describe quarks and leptons and include left-handed and right-handed fields, is the following:

$$Q_L^{fI}(3,2)_{+1/6}, u_R^{fI}(3,1)_{+2/3}, d_R^{fI}(3,1)_{-1/3}, L_L^{fI}(1,2)_{-1/2}, \ell_R^{fI}(1,1)_{-1}. \quad (1.2)$$

This notation means, for example, that Q_L^{fI} are the left-handed quarks, which are a triplet (3) of the $SU(3)_\mathrm{C}$ group, a doublet (2) of $SU(2)_\mathrm{L}$ and carry hyper-charge $Y = Q_\mathrm{EM} - T_3 = +1/6$. The index $f = 1, 2, 3$ is the *flavour* (or generation) index. The index I denotes *interaction eigenstates*.
The electromagnetic and weak interactions are unified in a single electroweak (EW) theory which describes the interactions among quarks and leptons mediated by the vector bosons. In the EW theory ($SU(2)_\mathrm{L} \times U(1)$) each particle is identified using the quantum numbers ($||\vec{T}||$, T_3) for the $SU(2)_L$ group and Y for the $U(1)$ group, $\vec{T}$ is called weak isospin and Y hyper-charge.
The EW theory is formulated using a Lagrangian invariant under transformations belonging to the $SU(2)_L \otimes U(1)$ symmetry group. The elementary fields described by the EW model are:

- 4 gauge boson fields, $W_\mu^i(x)$ (i=1,2,3) one for each generator of the $SU(2)_L$ group of symmetry and the boson field $B_\mu(x)$ for the $U(1)$ generator;
- left-handed quark fields $Q_L^f(x)$ which are doublets of the $SU(2)_L$ group, (f =1,2,3) is the *flavour* (or generation) index;
- right-handed quark fields $U_R^f(x)$ and $D_R^f(x)$, which are singlets of the $SU(2)_L$ group.

- left-handed lepton fields $L^f_L(x)$ ($SU(2)_L$ doublets), and right-handed ℓ^f_R lepton fields ($SU(2)_L$ singlets).

To simplify the notations we will concentrate on the quark sector; most of the considerations, anyway, apply also to lepton fields. The relevant part of EW Lagrangian is written as [5]:

$$\mathcal{L} = \bar{Q}^f_L \not{D} Q^f_L + \bar{U}^f_R \not{D} U^f_R + \bar{D}^f_R \not{D} D^f_R + \mathcal{L}_{\text{gauge}} \tag{1.3}$$

The covariant derivative $\not{D}$ is defined as:

$$\not{D} = \gamma^\mu D_\mu = \gamma^\mu \left(\partial_\mu + ig\vec{W}_\mu \cdot \vec{T} + ig'\frac{Y}{2}B^\mu \right) \tag{1.4}$$

Using this definition the first terms of Eq. (1.3) are written as:

$$\begin{aligned} \mathcal{L}(\text{quarks,gauge}) &= \bar{Q}^f_L \gamma^\mu \left(i\partial_\mu - g\frac{\vec{W}_\mu \cdot \vec{\tau}}{2} - \frac{1}{6}g' B_\mu \right) Q^f_L + \\ &\quad + \bar{U}^f_R \gamma^\mu \left(i\partial_\mu - \frac{2}{3}g' B_\mu \right) U^f_R + \\ &\quad + \bar{D}^f_R \gamma^\mu \left(i\partial_\mu + g'\frac{1}{3}B_\mu \right) D^f_R \end{aligned} \tag{1.5}$$

where $\vec{\tau}$ are the Pauli matrices and $\vec{T} = \vec{\tau}$ in the doublet representation, g and g' are the vector boson coupling constants. These three terms describe the interactions of vector bosons with left-handed and right-handed quarks respectively. As right-handed quarks have zero isospin ($\vec{T} = 0$) the $\vec{W} \cdot \vec{\tau}$ term disappears from the right side of Eq. (1.5).

The $\mathcal{L}$(gauge) term describes the evolution of the free gauge fields and is written as:

$$\mathcal{L}(\text{gauge}) = -\frac{1}{4}\vec{W}_{\mu\nu} \cdot \vec{W}^{\mu\nu} - \frac{1}{4}B_{\mu\nu}B^{\mu\nu} \tag{1.6}$$

with:

$$W^i_{\mu\nu} = \partial_\nu W^i_\mu - \partial_\mu W^i_\nu + g(\vec{W}_\mu \times \vec{W}_\nu)^i \tag{1.7}$$

$$B_{\mu\nu} = \partial_\nu B_\mu - \partial_\mu B_\nu. \tag{1.8}$$

The Lagrangian of Eq. (1.3) describes four massless vector bosons as well as massless quarks, hence this model cannot yet describe the real particles because there is only one massless vector boson observed in nature (the photon) and moreover the quarks have a mass.

The problem of the masses will be briefly discussed below. Let's now consider eq.1.5 and concentrate on the term describing the interaction of the $SU(2)_L$ gauge bosons (W^a_μ, $a = 1, 2, 3$) with quarks:

$$-\mathcal{L}_W = \frac{g}{2}\overline{Q^f_L}\gamma^\mu \tau^a Q^f_L W^a_\mu. \tag{1.9}$$

The 4×4 Dirac matrix γ^μ operates in Lorentz space and the 2×2 matrix τ^a operates in the $SU(2)_\mathrm{L}$ space. The coupling $\overline{Q_L^f} Q_L^f$ can be equivalently written as $\overline{Q_L^f} \mathbf{1}_{fj} Q_L^j$ where the 3×3 unit matrix $\mathbf{1}$ operates in flavour space and makes the so called "universality of the gauge interactions" manifest. The Standard Model gauge interactions do not distinguish between the different quark generations. The strength of the gauge interactions depends on the gauge quantum numbers given in 1.1 and not on the flavour index f.

Higgs mechanism

A solution to the mass problem has been given by P. Higgs [6] who proposed that the masses can be acquired by the interaction with another field. Fermions are then supposed to experience a further kind of interaction, called Yukawa interaction. Whereas in *gauge interactions* two fermions couple to spin 1 gauge bosons, which play the role of force carriers, as described by equation 1.5, in *Yukawa interactions* two fermions couple to a scalar field (the yet undiscovered Higgs).
The Lagrangian 1.3 has to be written incorporating this new field Φ as

$$\mathcal{L}(\mathrm{Higgs}) = D^\mu \Phi D_\mu \Phi + V(\Phi). \tag{1.10}$$

Under $SU(2)_L$, Φ is a doublet of scalar charged fields

$$\Phi = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1(x) + i\phi_2(x) \\ \phi_3(x) + i\phi_4(x) \end{pmatrix} \tag{1.11}$$

carrying hyper-charge 1/2. The potential is defined by

$$V(\Phi) = \mu^2 |\Phi|^2 + \lambda |\Phi|^4 \tag{1.12}$$

with $\lambda > 0$ and $\mu^2 < 0$, has a non trivial minimum for $|\Phi| \equiv vev = \sqrt{-\mu^2/2\lambda}$, called "vacuum expectation value". There is an infinite number of fundamental states since no direction is preferred. We can choose in an arbitrary way the "vacuum" state as

$$\langle \Phi \rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ vev \end{pmatrix} \tag{1.13}$$

which has the residual $U(1)$ invariance (corresponding to the only unbroken generator of the gauge symmetry), and develop the Higgs field around this value. Then the gauge fields has to be rewritten taking into account this change of variables. As a final result, looking at the quadratic terms in the gauge fields, we can see that the charged W_μ^1,W_μ^2 vector bosons have a non vanishing mass given by

$$M_W^2 = \frac{g^2 vev^2}{2} \tag{1.14}$$

whereas the mass terms for the W^3_μ and B_μ fields, is not diagonal:

$$M^2 = \frac{vev^2}{2}\begin{pmatrix} g^2 & -gg' \\ -gg' & g'^2 \end{pmatrix}. \tag{1.15}$$

Defining the "Weinberg angle" as $\tan\theta_W \equiv g'/g$, we can diagonalise the matrix using the orthogonal combination

$$Z^\mu = \cos\theta_W W_3^\mu - \sin\theta_W B^\mu \quad A^\mu = \sin\theta_W W_3^\mu + \cos\theta_W B^\mu. \tag{1.16}$$
$$\tag{1.17}$$

The massless field A^μ can be then identified with the vector field of the electromagnetic interactions, whereas the field Z^μ acquires a mass given by

$$M_Z^2 = vev^2\frac{g^2 + g'^2}{2}. \tag{1.18}$$

Three degree of freedom of the Higgs doublet are used to give mass to the vector bosons, while the fourth shows up as a new particle (the Higgs boson) with zero charge, 0-spin and mass related to the Higgs potential parameters (not predicted by the model) by the expression:

$$M_H = \sqrt{2}\mu = \sqrt{2\lambda}vev. \tag{1.19}$$

Fermions masses

The Higgs mechanism of spontaneous symmetry breaking gives rise to the quarks mass through the Yukawa coupling term:

$$-\mathcal{L}_Y = Y^d_{ij}\overline{Q^i_L}\Phi D^j_R + Y^u_{ij}\overline{Q^i_L}\tilde{\Phi}U^j_R + h.c. \tag{1.20}$$

where Φ is the Higgs doublet, and $\tilde{\Phi} = i\sigma_2\Phi^*$,is unitary equivalent to Φ but with a vacuum expectation value "rotated" in order to couple with the up quarks.

Y^d, Y^u are 3×3 complex matrices called Yukawa matrices. Since the Yukawa couplings may involve quarks of different generations the eigenstates of the EW interaction (*interaction basis*) do not have well-defined masses.

In order to transform to the mass basis, one has to take into account the spontaneous symmetry breaking, as discussed above. Replacing $\Phi \to \frac{1}{\sqrt{2}}\binom{0}{vev+H^0}$ and $\tilde{\Phi} \to \frac{1}{\sqrt{2}}\binom{vev+H^0}{0}$ and decomposing the $SU(2)_\mathrm{L}$ doublet into their components:

$$Q^i_L = \begin{pmatrix} u^i_L \\ d^i_L \end{pmatrix}, \tag{1.21}$$

the Yukawa interaction 1.20 gives rise to the mass terms:

$$-\mathcal{L}_M = (M_d)_{ij}\overline{d^i_L}d^j_R + (M_u)_{ij}\overline{u^i_L}u^j_R \tag{1.22}$$

where the mass matrices are defined as

$$M_f = \frac{vev}{\sqrt{2}} Y^f. \tag{1.23}$$

The mass eigenstates have, by definition, well-defined masses, and Yukawa interactions are diagonal in this *mass basis*. The mass basis corresponds, by definition, to diagonal mass matrices. We can always find unitary matrices V_{fL} and V_{fR}, such that

$$V_{fL} M_f V_{fR}^\dagger = M_f^{\rm diag}, \tag{1.24}$$

with mass matrix $M_f^{\rm diag}$ diagonal and real. The quark mass eigenstates are then identified as:

$$\begin{aligned} d_L^{i'} &= (V_{dL})_{ij} d_L^j, \quad d_R^{i'} = (V_{dR})_{ij} d_R^j, \\ u_L^{i'} &= (V_{uL})_{ij} u_L^j, \quad u_R^{i'} = (V_{uR})_{ij} u_R^j. \end{aligned} \tag{1.25}$$

The charged current interactions i.e. the interactions of the charged $SU(2)_{\rm L}$ gauge bosons $W_\mu^\pm = \frac{1}{\sqrt{2}}(W_\mu^1 \mp iW_\mu^2))$ described by 1.9 in the interaction basis, are now described by:

$$-\mathcal{L}_{W^\pm} = \frac{g}{\sqrt{2}} \overline{u_L^{i'}} \gamma^\mu (V_{uL} V_{dL}^\dagger)_{ij} d_L^{j'} W_\mu^+ + \text{h.c.} \tag{1.26}$$

where the 3×3 unitary matrix,

$$V_{uL} V_{dL}^\dagger = V_{\rm CKM} \tag{1.27}$$

is the CKM (Cabibbo,Kobayashi,Maskawa) mixing matrix for quarks [7].
Generally it would depends on nine parameters: three real angles and six phases. However, some of the phases of the $V_{\rm CKM}$ matrix are not physically meaningful. To see this, note that, apart irrelevant factor, V_{ij} is the amplitude for the quark transition $i \to j$ through W emission. Without changing the physics we can multiply each row and column for a phase factor removing five of the six phases. Within the SM case of three generation $V_{\rm CKM}$ has three mixing parameters and a single phase[1].
As a result of the fact that $V_{\rm CKM}$ is not diagonal, the $W^\pm$ gauge bosons can couple to quarks (mass eigenstates) of different generations. Within the Standard Model, this is the only source of *flavour changing* interactions.
In fact additional sources of flavour mixing in the lepton sector and in Z^0 interactions are not allowed within the Standard Model.

---

[1] This phase allows to accommodate CP violation in the SM, as was pointed out by Kobayashi and Maskawa in 1973 [8]

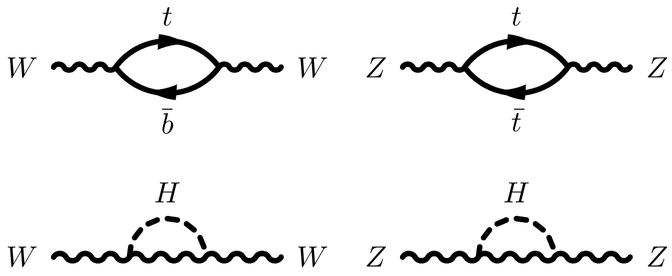


Figure 1.1: Contribution of self-coupling loops to higher order quantum corrections

In its simplest version, the model has 19 parameters: the three coupling constants of the gauge theory and the mass of the Z boson which sets the scale of weak interactions; the flavour parameters: three lepton and six quark masses and the four parameters which describe the mixing of the weak eigenstates in the mass basis. Between the two remaining parameters, the CP-violating parameter associated with the strong interactions must be very small. The last one is indeed associated with the mechanism responsible for the breakdown of electroweak symmetry $SU(2)_L \times U(1)_Y \to U(1)_e$. This can be taken as the mass of the Higgs boson, yet undiscovered.
Most experiments have provided tests of the gauge interactions at a very high accuracy level. In contrast, several flavour parameters are only known to an accuracy level of $O(30\%)$. Many rare decay processes that are sensitive to the flavour parameters have not been measured yet.

1.2 Top quark physics

The top quark is the heaviest elementary particle yet discovered. Its mass, of the same order as the electroweak scale, is about twice that of the W and Z bosons and about 35 times larger than its isospin partner, the b quark.
The top quark was first observed in 1995 at the Fermilab $p\bar{p}$ Tevatron collider by the two experiments CDF and DO [9, 10, 11], there were only about 100 total top candidates events found during the entire period of data taking which spanned 5 years known as Run I (1992-1996). Among the discovered particles top quark has a fundamental role also for new physics discoveries, not only because it is a fundamental background to all searches for new physics, but also because its precise mass measurement provides a constraint to the mass of the Higgs boson.
The mass of the top quark enters into calculations of higher-order (radiative) corrections, which connect electroweak processes, depending on the masses

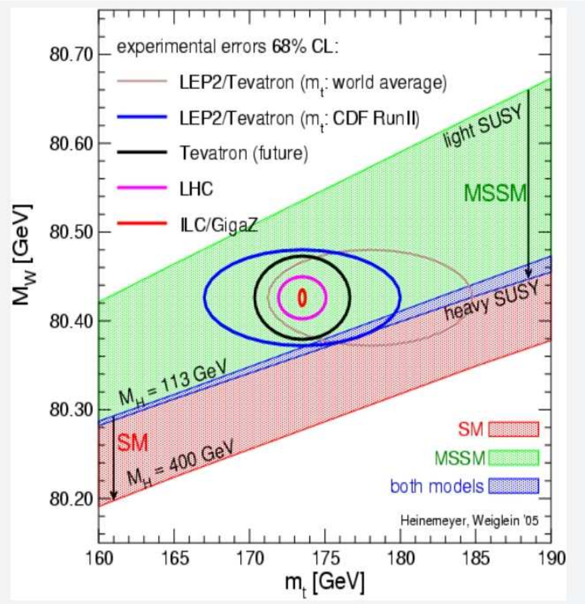


Figure 1.2: The W boson mass, top quark mass plane with the current measurements and predictions for future measurements. The SM and MSSM predictions for the Higgs boson mass are included showing the preference, from current, measurements of a light Higgs boson.

of the top quark and the Higgs boson via loop diagrams such as those shown in figure 1.1.

At one loop, for instance, the ρ parameter, which relates the W and Z boson masses and the weak angle,

$$\rho = \frac{M_W^2}{M_Z^2}(1 - \sin^2\theta_W) \equiv 1 + \Delta r, \tag{1.28}$$

gets a radiative correction

$$\Delta r = \frac{3G_F}{8\pi^2\sqrt{2}}M_{top}^2 + \frac{\sqrt{2}G_F}{16\pi^2}M_W^2\left[\frac{11}{3}\ln\left(\frac{M_H^2}{M_W^2}\right) + \cdots\right] + \cdots \tag{1.29}$$

which is quadratic in the top mass. Note, however, that the dependence on the mass of the Higgs boson is only logarithmic. Therefore, the mass of the top quark is the dominant term in the corrections for electroweak process. By measuring M_{top} very accurately, and using additional constraints from the large body of precise electroweak data we can test the consistency of the standard model and predict unknown parameters. One of such predictions is the mass of the Higgs boson, which can be constrained by the precise measurements of W boson and top masses.

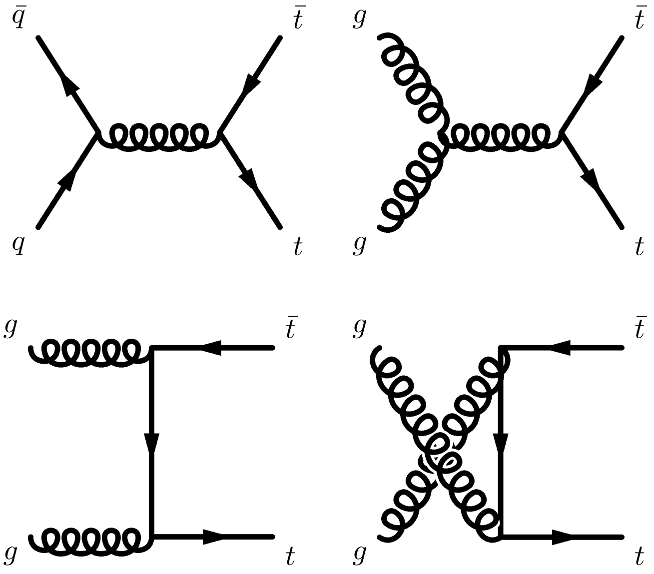


Figure 1.3: The lowest order diagrams for $t\bar{t}$ production

The direct measurements of W boson mass and top quark mass obtained directly from the Large Electron Positron collider (LEP) and the Fermilab Tevatron collider are shown in figure 1.2. The Higgs boson mass calculations in the standard model and in its Minimal Super Symmetric Model (MSSM) extension are also shown. This, together with the global electroweak fit, shows that a light Higgs boson is better in agreement with the known experimental data.

1.2.1 Top anti–top quark pairs production

In hadron collider top quarks are expected to be produced mainly as a pair by two distinct QCD processes: $q\bar{q}$ annihilation and gluon–gluon (gg) fusion. Figure 1.3 shows the Feynman diagrams at the leading order for these processes. At the Large Hadron Collider, the relative contribution from $q\bar{q}$ and gg are about 10% and 90% respectively[2].
The total cross section for the production of heavy quarks, at a pp machine

---

[2] This situation is opposite respect to the Tevatron collider, where the $q\bar{q}$ annhiliation process accounts for about 85% of the total cross section.

like LHC, is given by,

$$\sigma(pp \to t\bar{t}) = \sum_{i,j} \int dz_i dz_j f_i(z_i, \mu^2) f_j(z_j, \mu^2) \hat{\sigma}(ij \to t\bar{t}; \hat{s}, \alpha_s, \mu^2, M_{top}), \tag{1.30}$$

where the sum is over all partons: gluons, light quarks and anti-quarks. This formula expresses the total cross section in terms of the parton-parton process $ij \to t\bar{t}$. The parton distribution function (PDF) f_i corresponds to the probability density of finding a parton (of flavour i) with a given fraction of the proton momentum between z_i and $z_i + dz_i$, the short-distance cross section for the parton-parton subprocess is $\hat{\sigma}$. The center-of-mass energy of the $i-j$ parton system is given by $\hat{s}$ and it is related to the *pp* center-of-mass energy by $\hat{s} = z_i z_j s$. The parameter μ is the renormalization scale which is introduced to include the contribution from higher orders Feynman graphs. If the calculation could be carried out to all orders then the dependence on μ would vanish. The strong coupling constant is given by α_s and M_{top} is the mass of the top quark.

The current theoretical calculations (at Next-to-Next and Next-to-Next-to-Next Leading order) of the $t\bar{t}$ production cross section at the center-of-mass energy of 14 TeV, with a top quark mass of 175 GeV/c^2 , is 840 pb, with 5% error due to the extrapolation of the cross section at the LHC energy scale from Tevatron data and an extra 3% error due to uncertaininty on the PDFs [12].

1.2.2 Electroweak single top quark production

A single top can be produced via the electroweak interaction through three processes (the corresponding leading order Feynman diagrams are visible in figure 1.4):

- through the exchange of a space-like W boson and the promotion of a b quark to top, (a) and (b) diagrams of figure 1.4. This mechanism is referred to as t-channel;

- through the associated production channel (diagrams (c) and (d) in the figure) in which a single top is produced via weak interaction in association with a real W boson, one of the initial partons is a b quark in the proton sea, as in the t-channel;

- or through the last production mechanism: the s-channel (diagram (e) in the figure), the annihilation of a $q\bar{q}$ pair.

The cross section for all three processes is proportional to the matrix element $|V_{tb}|^2$ of the Cabibbo-Kobayashi-Maskawa matrix (CKM). Therefore, the measurement of the single top quark cross section provides a direct probe

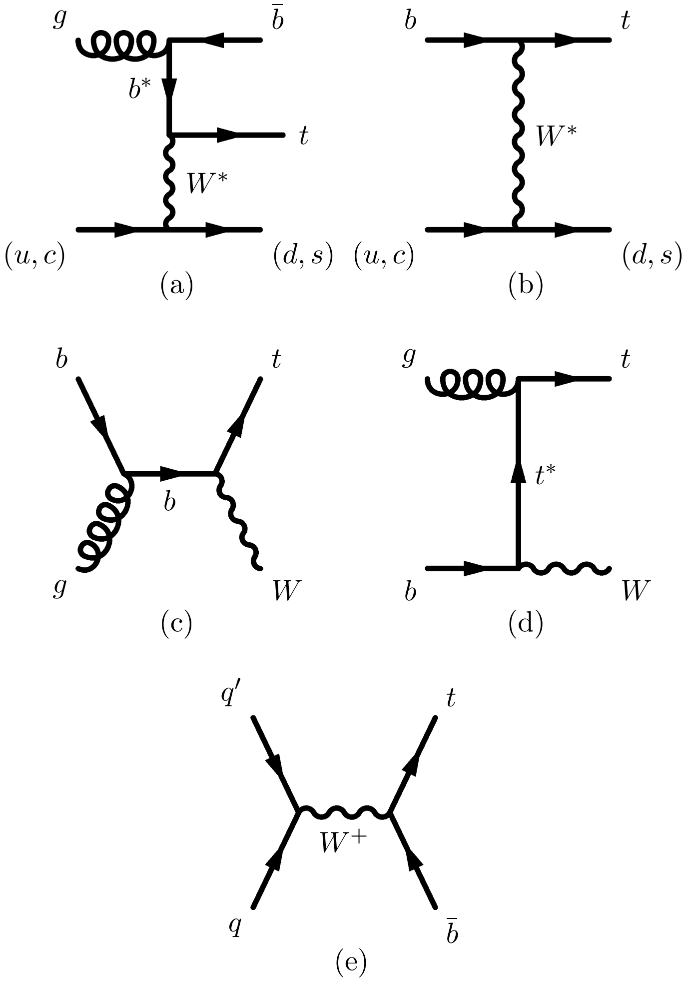


Figure 1.4: Leading order Feynman diagrams for electroweak production of single top quarks: (a,b) t-channel, (c,d) Wt production (e) s-channel.

of $|V_{tb}|$ and the weak tbW vertex in general. Next-to-leading order calculations for the t-channel at the LHC energy give $\sigma = 245 \pm 27$ pb [13], for the associated production the calculated cross section is $62.2^{+16.6}_{-3.7}$ pb [14] and for the s-channel calculations give 10.2 ± 0.7 pb [15].
The primary physics interest in single top production is its ability to directly determine the coupling strength for the $t - W - b$ vertex. The single top cross–section is unambiguously predicted by the SM, and it is important to cross check the three production channels separately. The various processes of single top production have different sensitivities to new physics. For example, the s-channel is sensitive to an additional heavy W' boson, since new s-channel diagrams in which the W' is exchanged would occur. In contrast, additional contributions to the t-channel would be suppressed by $1/m^2_{W'}$. Therefore, existence of a W' boson would be expected to produce an enhancement in both $\sigma(s - channel)$ and $\sigma(s - channel)/\sigma(t - channel)$. On the other hand, the t-channel is more sensitive to modifications of the top quark's couplings to the other SM particles. For example, a V+A contribution at the $t - W - b$ vertex could lead to both an increase in single top production and a modification of the decay angular distributions. Because it is an inherently weak production process, the W and top quark are produced in the appropriate mixture of helicities, as unambiguously predicted by the SM. A helicity analysis of top quark decay can check for new physics, such as right handed couplings, or an unexpected admixture of the left handed and longitudinal components for the W.

1.2.3 Top quark decays

The Standard Model predicts that top decays into W boson and a b quark with a branching ratio greater than 0.998. Other decays, such as $t \to Ws$ and $t \to Wd$ are also allowed but suppressed by a factor $10^{-3} - 10^{-4}$ by the square of the CKM matrix elements V_{ts} and V_{td}, whose values can be estimated, under the assumption of unitary of the three-generation CKM matrix, to be less than 0.042 and 0.014 respectively [16].
Top decay width is then:

$$\Gamma(t \to Wb) = \frac{G_F M^3_{top}}{8\pi\sqrt{2}} \left(1 - \frac{M^2_W}{M^2_{top}}\right)^2 \left(1 + 2\frac{M^2_W}{M^2_{top}}\right)^2 \left[1 - \frac{2\alpha_s}{3\pi}\left(\frac{2\pi^2}{3} - \frac{5}{2}\right)\right] \tag{1.31}$$

for $M_{top} = 175$ GeV/c^2 we have

$$\Gamma(t \to Wb) \approx 1.55 \text{ GeV/c}^2 \to \tau_{top} = \left(\frac{1}{\Gamma_{top}}\right) \approx 4 \cdot 10^{-25}\text{s} \tag{1.32}$$

Top decay width is smaller than the characteristic hadronization time of QCD ($\tau_{hard} \approx 28 \cdot 10^{-25}$ s) hence top decays before hadronizing.

Decay mode	Branching Ratio	Channel category
$t\bar{t} \to (q\bar{q}'b)(q\bar{q}'\bar{b})$	36/81	All-hadronic
$t\bar{t} \to (q\bar{q}'b)(e\nu\bar{b})$	12/81	Semi-leptonic
$t\bar{t} \to (q\bar{q}'b)(\mu\nu\bar{b})$	12/81	Semi-leptonic
$t\bar{t} \to (q\bar{q}'b)(\tau\nu\bar{b})$	12/81	τ channel
$t\bar{t} \to (e\nu b)(e\nu\bar{b})$	1/81	Di-leptonic
$t\bar{t} \to (e\nu b)(\mu\nu\bar{b})$	2/81	Di-leptonic
$t\bar{t} \to (e\nu b)(\tau\nu\bar{b})$	2/81	τ channel
$t\bar{t} \to (\mu\nu b)(\mu\nu\bar{b})$	1/81	Di-leptonic
$t\bar{t} \to (\mu\nu b)(\tau\nu\bar{b})$	2/81	τ channel
$t\bar{t} \to (\tau\nu b)(\tau\nu\bar{b})$	1/81	τ channel

Table 1.2: Branching ratios for $t\bar{t}$ decay modes in standard model coupling. q stands for u, d, c or s quark. Events are categorized into four channels: All-hadronic, Semi-leptonic, Di-leptonic and τ channels.

The W produced in the top decay, decays into lepton-neutrino pair or quark-antiquark pair. All the (anti)quarks present in the event will eventually hadronize and manifest themselves as jets: a b quark jet and, if present, the two light quark jets.

Table 1.2 lists the final states of $t\bar{t}$ production, according to the W decay modes. The events are classified as either an “all-hadronic”, “semi-leptonic” (or “lepton+jets”), “di–leptonic” or “τ channel”. An event is considered as an all-hadronic event when both W bosons decay to a quark-antiquark pair. Though this channel has the largest branching fraction, 44%, it suffers, at hadronic colliders like LHC, from a huge amount of background from QCD multijets production processes. An event is classified as a di-leptonic event when both W bosons decay leptonically to an e or a μ. Due to the two well isolated leptons and the missing energy due to the two neutrinos in the event this channel has the smallest background but also the smallest branching fraction, 5%. An event which contains the τ lepton can be classified separately. The τ lepton decays predominantly into charged and neutral hadrons and its signature is very close to the one of a jet.

The semi-leptonic events are characterized by one of the two W boson decaying in a e or μ lepton and neutrino while the other W boson decaying hadronically. This channel, with a branching fraction of 30%, has the advantage of presenting in the final state an isolated high-p_T lepton and high missing energy due to the presence of the neutrino. The final state is completed by 4 jets (of which two come from b quark). The Feynman diagram of the $t\bar{t}$ production from $q\bar{q}$ annihilation followed by decay in the semi-leptonic channel is presented in figure 1.5.

This channel presents in the final state four-jets a high $\mathrm{P_T}$ isolated lepton and E_T^{miss} , peculiar signature that can be used to efficiently distinguish the signal from the backgrounds. The hadronically decaying top can also be easily reconstructed measuring the jets in the event. For these reasons this channel is the preferred one for top mass precise measurements at the LHC and is often referred to as the “golden channel”.

1.2.4 Top quark mass measurement

Based on Run I Tevatron data, the CDF and DO collaborations have published several direct experimental measurements of M_{top} with all decay topologies arising from $t\bar{t}$ production. Tevatron Run I average claims for a top with mass 178.0±4.3 GeV/c^2 . On the other hand, based on Run II Tevatron data at the center-of-mass energy of 1.96 TeV, the top quark has been measured again. CDF most recent results are $M_{top} = 172.7 \pm$ 1.7(syst)±2.4(stat) GeV/c^2 .
The most precise single top mass measurement has been obtained in the semi-leptonic channel. The top mass is obtained from a χ^2 minimization procedure in which the top mass is a free parameter. To reduce the main systematic error, the jet energy scale, another free parameter is introduced in the χ^2 representing the uncertainty on the jet scale.
The results of the minimization procedure: top mass and jet energy scale are then compared to a set of Monte Carlo simulations in which the two parameters take different values. The data are fitted to this set of “templates” and the top mass is extracted.
This method allows, with increased statistic, to reduce the error on the jet energy scale, so far the dominant systematic error at the Tevatron, since the jet energy scale is extracted from data themselves. This latter is not a systematic, the errors will be reduced by increasing the statistics.
It is expected that, combining the results from the two general purpose experiments at the LHC, ATLAS and CMS, the error on top mass will be reduced at the level of 1 GeV/c^2 . This precise measurement together with the W boson mass measurements will better constrain the standard model.

1.3 Top physics at the ATLAS experiment

Study of the top quark may provide an excellent probe of the sector of electroweak symmetry breaking, and new physics may well be discovered in either its production or decay. At the LHC a very large variety of top physics studies will be possible with the high statistics samples which will be accumulated.
The LHC, already during the low luminosity period with $L = 10^{32}$ cm^{-2}s^{-1} at the nominal energy of 14 TeV, will produce about 8 millions of $t\bar{t}$ events. In only few days of data taking and assuming a selection efficiency of 1.5%,

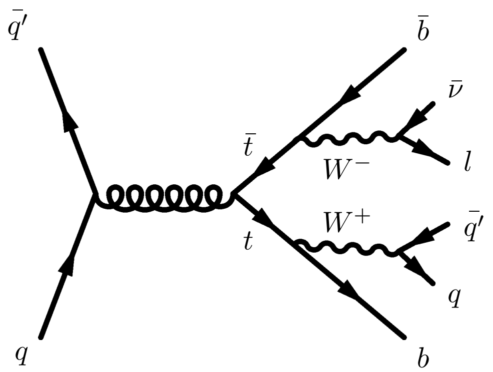


Figure 1.5: The tree-level Feynman diagram for $t\bar{t}$ production by $q\bar{q}$ annihilation and its decay in the semi-leptonic channel. Two b-jets, two light jets a lepton and a neutrino are present in the final state.

each LHC experiment will accumulate the same statistics as the Tevatron experiments up to now.

The Tevatron is providing both experiments with good luminosity and the precision on top mass will be improved down do few GeV/c^2 as it is shown in figure 1.6 where the total uncertainty on top mass is presented as a function of integrated luminosity. Of the order of 5-8 fb^{-1} of integrated luminosity are foreseen to be collected up to 2009, when LHC will begin to be competitive. The LHC experiments have the goal to decrease the top mass uncertainty to about 1 GeV/c^2 (3).

At LHC the statistics will not limit the precision of top mass measurement: already with an integrated luminosity of 10 fb^{-1} (one year of data taking at low luminosity) the statistical error will be about $\delta(M_{top}) = \pm 0.10$ GeV/c^2. The LHC experiments will have the real challenge in the understanding of the systematic uncertainties, mainly the light and b jet energy scale and the effect of initial and final state radiation.

In the remaining of this chapter, we will discuss the studies that will be done in ATLAS, with particular interest to the top mass measurement in the semi–leptonic channel.

This channel for the top mass measurement will be described in further detail in the last chapter of this work where the strategy to measure top quark mass with the very first data of ATLAS will be presented and where it will be discussed the role of the jet energy scale, the most important systematic

---

[3] The request to know the top quark mass at the level of 1 GeV/c^2 comes from models beyond the SM which attempt to explain in a more fundamental way the origin of mass and the observed fermion mass hierarchy, such as top-bottom-tau Yukawa coupling unification in supersymmetric models. To constrain the SM Higgs boson mass the main challenge comes from the precision measurement of W boson mass. Assuming m_W can be measured with a precision of ±20 MeV, a determination of m_{top} with a precision of $\delta(M_{top}) \leq$ 2 GeV/c^2 would be required to match that from m_W and from the current theoretical uncertainties.

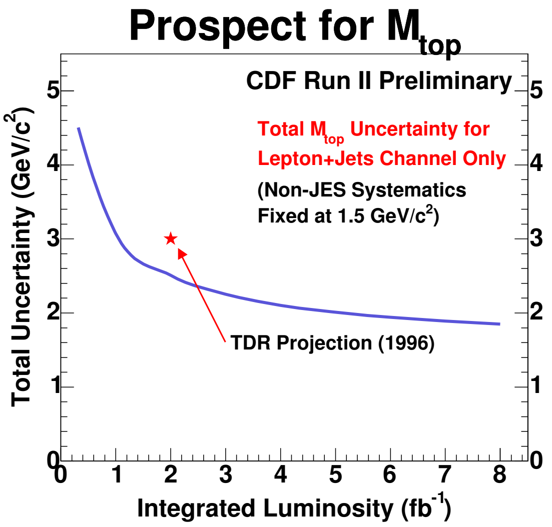


Figure 1.6: Extrapolation of total uncertainty on top quark mass measurement as a function of integrated luminosity for the CDF experiment. The best measurement achievable, for 8 fb^{-1} , has an error of about 1.8 GeV/c^2 . The systematic on the jet energy scale is fixed at 1.5 GeV/c^2 . Currently about 1 fb^{-1} of data has been collected.

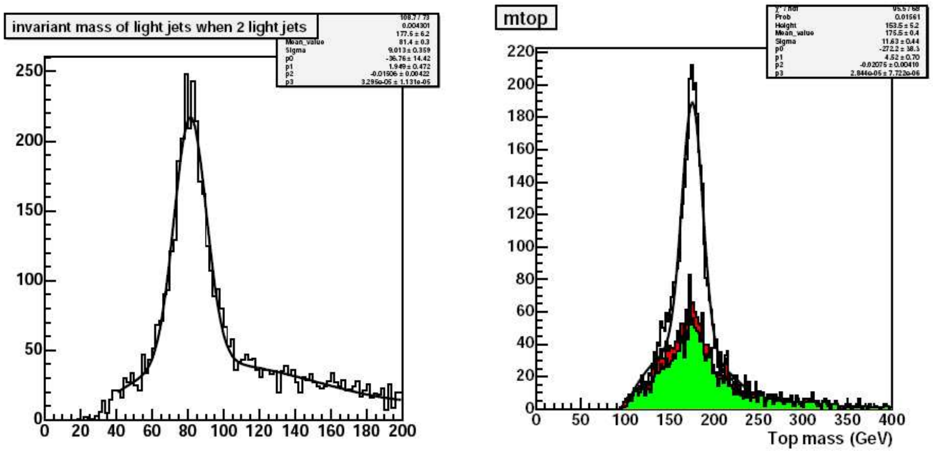
error.

1.3.1 Top quark mass measurement

At ATLAS the all-hadronic and τ channel of $t\bar{t}$ decays are difficult to isolate cleanly above the large QCD multi-jet background. For this reason ATLAS focus mainly on the channels where at least one W decays leptonically, i.e. the semi-leptonic and di-leptonic channel. The “golden channel” for top mass measurement is the semi-leptonic one.

The presence of a high p_T isolated lepton provides an efficient trigger. The lepton and the high value of $\mathrm{E}_{\mathrm{T}}^{miss}$ give a large suppression of background from QCD multi-jet and $b\bar{b}$ production.

The four momentum of the missing neutrino can be entirely reconstructed assuming $m_\nu = 0$, and $E_T(\nu) = \mathrm{E}_{\mathrm{T}}^{miss}$, and calculating $p_z(\nu)$, with a quadratic ambiguity, using the constraint $m_{l\nu} = m_W$. If one applies the further kinematic constraints that $m_{jj} = m_W$ and $m_{jjb} = m_{l\nu b} = m_{top}$, the top mass can be determined by a three-constraint fit. This techniques, used at Tevatron, gives the best determination of top mass, where statistics is limited. However, if the systematic errors are to be kept small, this method requires an excellent modelling and understanding of the $\mathrm{E}_{\mathrm{T}}^{miss}$ distribution and res-



(a) (b)

Figure 1.7: (a): Invariant mass of light jets m_{jj} in $t\bar{t}$ events with only 2 reconstructed light jets. (b): m_{jjb} invariant mass. Light jets have been calibrated as described in the text. Coloured histograms represents combinatorial background, green from wrong W reconstruction, red from wrong b-jet association.

olution.

An alternative approach is the one that uses the isolated lepton and the large $\mathrm{E}_{\mathrm{T}}^{miss}$ to tag the event, and the value of m_{top} is extracted as the invariant mass of the three jet system arising from the hadronic top quark decay [17]. With this strategy the events are selected requiring an isolated lepton with $p_T > 20$ GeV and $|\eta| < 2.5$ and $\mathrm{E}_{\mathrm{T}}^{miss} > 20$ GeV, and at least four jets. Jets are reconstructed with a fixed cone algorithm with $\Delta R = 0.4$. At least four jets with $p_T > 40$ GeV and $|\eta| < 2.5$ are required, and at least two of the jets are required to be tagged as b-jets[4] .

For accepted events, the decay $W \to jj$ is reconstructed among jets that are not tagged as b-jets. Figure 1.7 (a) shows the m_{jj} invariant mass of the light jets in the events accepted with the selections described above where only two light jets are reconstructed.

In order to reduce the importance of light–jet miscalibration (mainly due to out of cone energy) on the precision of top mass measurement, so called *in-situ* calibration of light jets can be performed through a χ^2 procedure [18]. Event by event, for each light jets pair, the quadratic form

$$\chi^2 = \frac{(m_{jj} - m_W)^2}{\Gamma_W^2} + \frac{(E_{j1}(1-\alpha_1))^2}{\sigma_1^2} + \frac{(E_{j2}(1-\alpha_2))^2}{\sigma_2^2} \tag{1.33}$$

---

[4] Also other jet reconstruction algorithm have been studied for top studies in ATLAS, as the fixed cone with $\Delta R = 0.7$. More recently some studies using the K_T algorithm have started. However, for the moment, the fixed cone size of $\Delta R = 0.4$ gives the best results.

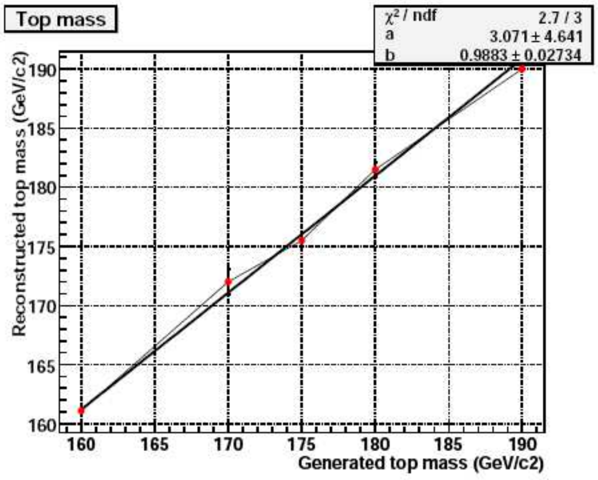


Figure 1.8: Reconstructed top mass as a function of the generated top mass.

is minimized respect to α_1 and α_2.

The first term contains the explicit constraint to the W boson mass. Among all jet pairs the one giving the smallest χ^2 is kept as the hadronic W candidate.

The two calibration parameters α_1 and α_2 are then fitted, as a function of the jet energy E_{jet}, to obtain the jet energy calibration function $\alpha(E_{jet})$ that is used for the final analysis jet-calibration.

The next step in the hadronic top reconstruction is the association of the hadronic W candidate with one of the b-tagged jets. In general there are ambiguities in the association. Two strategies are used both giving similar results: the selection of the b-jet giving the highest $\mathrm{P_T}$ to the reconstructed top and the selection of the b-jet closer to the hadronic W candidate.

The top mass is reconstructed as the invariant mass of the W+b-jet system and is shown in plot (b) of figure 1.7. The combinatorial backgrounds coming from an in-correct assignment of light jets to the reconstructed W boson (green histogram) and from the wrong b-jet association (red histogram) are shown.

A fit to the mass peak obtained for 500000 events corresponding roughly to six days of data taking at low luminosity gives $m_{top} = 175.5 \pm 0.4$ GeV/c^2, in good agreement with the mass used in Monte Carlo generation (175 GeV/c^2), the width of the distribution is equal to 11.6 ± 0.4 GeV/c^2. The stability of this method as a function of the value of the generated mass is shown in figure 1.8. The value of reconstructed top mass in in a good agreement with the top mass used in the generation of the Monte Carlo events.

Process	$p_T^l > 20$ GeV $\mathrm{E_T^{miss}} > 20$ GeV	As before, plus $N_{jet} \geq 4$	As before, plus $N_{b-jet} \geq 2$	Events per 10fb^{-1}
$t\bar{t}$ signal	64.7%	21.2%	5.0%	126000
$W^{+}jets$	47.9%	0.1%	0.002%	1658
$Z + jets$	15.0%	0.05%	0.002%	232
WW	53.6%	0.5%	0.006%	10
WZ	53.8%	0.5%	0.02%	8
ZZ	2.8 %	0.04%	0.008%	14
Total Background				1922
S/B				65

Table 1.3: Efficiencies for the inclusive $t\bar{t}$ semi-leptonic signal and for background processes, as a function of the selection cut applied. The last column gives the equivalent number of events for and integrated luminosity of 10 fb^{-1} .

Backgrounds to the semi–leptonic channel

In addition to the combinatorial, other sources of background are present in the semi–leptonic channel. The most severe are: the QCD multi-jets and $b\bar{b}$ production, the $W + jets$ production with $W \to l\nu$ decay and $Z + jet$ events with $Z \to ll$.
While the presence of the isolated lepton and the high value of the missing transverse energy suppresses the QCD background the b-tagging performances are crucial to suppress the $W + jet$ and $Z + jet$ backgrounds.
Potential backgrounds from ZZ, WW and WZ gauge boson pair production are reduced to a negligible level after all analysis selections. With an integrated luminosity of 10 fb^{-1} , only 1922 background events survive the cuts to be compared to 126000 signal events, yielding a value of S/B=65. The details of the analysis selections on signal and backgrounds samples are reported in table 1.3.

Systematics

An important systematic effect on the top mass measurement is the knowledge of the jet energy scale. Many strategies have been studied to correct for it. Light jets can be calibrated at the parton level applying the constraint $m_{jj} = m_W$. The calibration of b-jets can be performed using $Z + b$ events with the subsequent $Z \to ll$ decay. The shift of the top quark mass due to miscalibration of the light quark jets and of b-jets are studied separately. A miscalibration of 1% in the light jet energy and b-jet scale corresponds respectively to a shift of 0.3 GeV/c^2 and 0.7 GeV/c^2 in the reconstructed top mass.
The fraction of the original b-quark momentum which is reconstructed depends on the fragmentation function of the b-quark. This function is usually

parametrized in terms of one variable ϵ_b, using the Peterson fragmentation function [19]. The value of ϵ_b used in the Monte Carlo is $\epsilon_b = -0.006$ with an uncertainity of $\delta\epsilon_b = 0.0025$. We have compared the reconstructed top mass from Monte Carlo samples with $\epsilon_b = -0.006$ and $\epsilon_b = (-0.006 + 0.0025) = -0.0035$. The difference between the values of m_{top} obtained with these samples is taken as the systematic error, $\delta m_{top}(\epsilon_b)$, due to uncertainties in the knowledge of ϵ_b and amounts to another 0.3 GeV/c^2 .

The presence of initial state radiation (ISR) and final state radiation (FSR) have also effect on the measurement of m_{top}. A top mass shift due to ISR, Δm_{top}^{ISR}, is determined by artificially switching off the ISR in a Monte Carlo simulation and measuring the impact on the reconstructed top mass. Similarly we determined the shift due to FSR.

In case of FSR a large mass shift occurs, of about 10 GeV, for a jet cone size of $\Delta R = 0.4$. This mass shift is considerably reduced if the cone size is increased. The level of uncertainty on the knowledge of ISR and FSR is of the order of 10% and reflects the uncertainity on α_s. To estimate the systematic uncertainty due to FSR and ISR we have used more conservative systematic errors in m_{top}: 20% of the $\Delta m_{top}^{\mathrm{ISR}}$ and of $\Delta m_{top}^{\mathrm{FSR}}$ mass shift defined above is taken. The systematic errors estimated in this way are less then 0.3 GeV/c^2 for ISR, while 1-2 GeV/c^2 error results from effects due to FSR.

An alternative approach uses the measured jet multiplicity to search, event-by-event, for the presence of hard gluon radiation. This approach was adopted at the Tevatron [20, 21]: the mass shift would be defined by the difference, Δm_{top}, between the value of m_{top} determined from events with exactly four jets and that determined from events with more than four jets. Such a method shows systematic errors of the order of 0.4-1.1 GeV/c^2 for FSR, smaller than the more conservative approach adopted here.

Uncertainties in the size and shape of the background, which is dominated by the wrong combinations of jets in $t\bar{t}$ events, also affect the value of the reconstructed top mass. The resultant systematic uncertainty on m_{top} is estimated by varying the assumptions about the background shape. A conservative conclusion is a systematic error on m_{top} of 0.2 GeV/c^2 .

Also the structure of the underlying event can influence the top mass reconstruction. However it is possible to estimate and correct for this effect using data themselves by using the calorimeters cells not associated with the products of the top quark decay. The underlying event contribution is calculated as the average $\mathrm{E_T}$ deposited per calorimeter cell, averaged over those cells which are at least a distance ΔR (depending on the position of the cell ΔR varies from 0.7 to 1.0) away from the impact points of the top quark decay products.

Given the large statistics available at the LHC, it is assumed that the residual uncertainty from the underlying event will be small compared to the other errors [22].

Source of uncertainty	Comment on method	$\lvert\Delta m_{top}\rvert$ GeV/c^2 $\Delta R = 0.4$ (0.7)	$\lvert\delta m_{top}\rvert$ GeV/c^2 $\Delta R = 0.4$ (0.7)
Light-jet energy scale	1% scale error	0.3 (0.3)	0.3 (0.3)
b-jet energy scale	1% scale errror	0.7 (0.7)	0.7 (0.7)
b-quark fragm.	$(\epsilon_b = -0.006) - (\epsilon_b = -0.0035)$	0.3 (0.3)	0.3 (0.3)
Initial State Radiation	ISR ON - ISR OFF	0.2 (1.3)	0.04 (0.3)
Final State Radiation	FSR ON - FSR OFF	10.3 (6.1)	2.0 (1.2)
Background		0.2 (0.2)	0.2 (0.2)

Table 1.4: Top mass shift Δm_{top} and resulting systematic error on m_{top}, δm_{top} due to the various source of systematic errors.

The individual contribution to the systematic error of $t\bar{t}$ sample are summarised in table 1.4. In summary the jet energy scale and FSR dominate the systematic errors.

Complementary methods

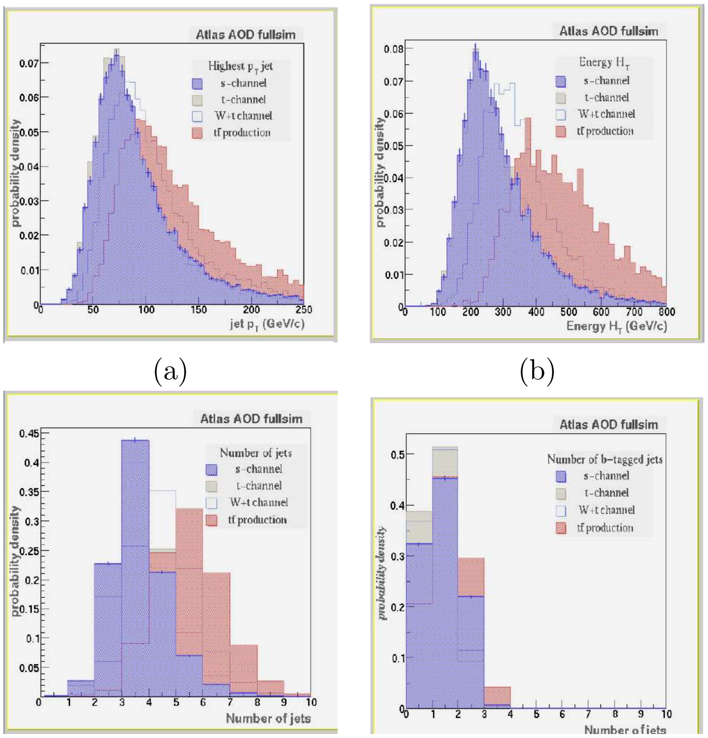
Although the semi–leptonic channel gives the most precise single measurement of the top quark mass, to reach the goal of measuring top mass with a precision of $\delta(m_{top}) = 1 - 2$ GeV/c^2 the systematic errors must be kept well under control. It is important to compare results obtained from different channels where systematic uncertainties play different roles.
Here only two methods among several will be shortly revised.
The high $\mathrm{P_T}$ channel method selects events in the semi-leptonic data sample with a high momentum top and reconstructs its mass from a large calorimeter cluster around the jets from the hadronic top decay. This method, after a proper treatment of the underlying event, can reduce the main systematic coming from the FSR contribution to a level well under 1 GeV/c^2 [22].
The analysis proceeds in two steps. First a jet cone of $\Delta R = 0.4$ is used in the lepton hemisphere, where, apart from radiation effects, only one *b*-jet is expected. In the hemisphere opposite to the lepton, where the three jets from the hadronic top decay are expected, a smaller cone size ($\Delta R = 0.2$) is used. To take into account the contribution of out of cone energy the selection cut on jets is lowered to $p_T > 20$ GeV.
For accepted events the two highest p_T non *b*-tagged jets are combined with the highest *b*-jet candidate in the hemisphere opposite to the lepton to form candidate for the jjb hadronic top decay. The selected combination is required to have $p_T > 150$ GeV and $|\eta| < 2.5$. Finally, after determining the top direction from the jet momenta, a large cone of $\Delta R = 1.3$ is used around this direction and the top mass is reconstructed from energy and momenta of the calorimetric cells in the cone.
In the absence of underlying event, and for cone sizes which are sufficiently large to contain all three jets from the hadronic top decay, the reconstructed mass is independent of the cone size. Hence, a method has been developed to subtract the contribution from the underlying event, by using the calorimeter

cells not associated with the products of the top quark decay. This method has shown to give promising results, at present the reconstructed top mass is found larger than the generated one of about 1%.
Also the $t\bar{t}$ di–lepton channel can provide a measurement of the top quark mass complementary to that obtained from the single lepton plus jets mode. The signature of the di–lepton event consists of two isolated high p_T leptons, high $\mathrm{E}_{\mathrm{T}}^{miss}$ due to the neutrinos, and two jets from the fragmentation of b-quarks. Measurement of m_{top} using di–lepton events is complicated by the fact that one cannot fully reconstruct either of the top quarks, due to the undetected neutrinos in the final state.
It is possible to take advantage of the fact that the kinematic distributions of the top decay products depend on m_{top}, and attempt to obtained the most likely top mass for a set of events. The mass determination depends on the assumption that the kinematic distributions for top production are well reproduced by the Monte Carlo simulation.
Of the many possible kinematic variables which could be studied, preliminary analyses [23] of three have been performed: the mass m_{lb} of the lepton-b-jet system, the energy of the two highest E_T jets, and the mass m_{ll} of the di–lepton system formed with both leptons coming from the same top decay (*i.e.* $t \rightarrow l\nu b$ followed by $b \rightarrow l\nu c$).

1.3.2 Other top studies

Search for single top

Single top production is an interesting channel as it allows the precise determination of the properties of the $W - t - b$ vertex, and the associated coupling strengths.
In order to reduce the enormous QCD multi-jet background, as well as provide high p_T lepton for trigger purpose, the leptonic decay of the W boson originated from top quark, is considered. The initial preselection cuts require the presence of at least one isolated lepton with $p_T > 20$ GeV, at least two jets with $p_T > 30$ GeV, and at least one b-tagged jet with $p_T > 50$ GeV. After these cuts, the dominant backgrounds are from processes with a real W in the final state, namely $t\bar{t}$ and Wjj (and in particular $Wb\bar{b}$) production.
Kinematics can be used to separate the various single top processes from these backgrounds and from each other.
Distributions of four quantities ($\mathrm{P_T}$ of the leading jet, H_T quantity, multiplicities of light and b-jets) are presented in figure 1.9. The $\mathrm{P_T}$ distributions of the jets in the events can be used to separate the single top signal from $t\bar{t}$ background: distribution (a) shows that the transverse momentum of the leading jet in $t\bar{t}$ events is harder than the one in the single top events. The H_T quantity, defined as the sum of the E_T values of all jets and leptons in the event, can be used to distinguish the signal from the background and



(a) (b)

(c) (d)

Figure 1.9: Distributions for quantities used to separate the signal from the $t\bar{t}$ background and to distinguish the different production channels: (a) Leading jet P_T , (b) H_T distribution, (c) number of jets and (d) number of b—-jets in the event.

to separate t-channel and s-channel from the $W + t$ associated production channel, as it is shown in distribution (b) of figure 1.9. Also the number of jets and especially the number of b-jets can be used to select the events (distributions (c) and (d)).

The direct reconstruction of the top mass from the lepton, the neutrino (reconstructed from E_T^{miss} and W decay with quadratic ambiguity) and one of the b-jets, in a window around the nominal mass is used to separate the signal from non top background events.

For an integrated luminosity of 30 fb^{-1} (three years of data taking at low luminosity) 1658 events of single top are expected (1106 from s-channel single top production, 510 from t-channel and 42 from associate production) while about 1844 events of backgrounds (of which about 1300 are from $t\bar{t}$ events) survive all the selections.

The relative statistical errors on the production cross-sections of the single top processes would be 0.71% for the t-channel, 2.8% for the associated

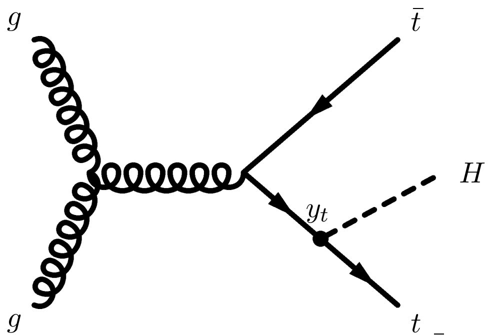


Figure 1.10: A lowest order Feynman diagram for $t\bar{t}H$ production.

production and 5.4% for the s-channel. These results imply statistical uncertainties on the extraction of V_{tb} of 0.36% for the t-channel, 1.4% for the associate production and 2.7% for the s-channel. The errors in the calculation of V_{tb} would be dominated by uncertainties in the theoretical predictions of the cross–sections, larger of about an order of magnitude. These arise from uncertainties in the parton distribution functions, uncertainty in the scale used in the calculation, and the experimental error on the mass of the top quark.

Top quark Yukawa coupling

In the SM, the mass of the top quark is due to its Yukawa coupling (y_t) to the Higgs boson. The values of the Yukawa couplings of the fundamental fermions are free parameters of the Standard Model. The measured value of m_{top} implies a value of the top quark Yukawa coupling of approximately unity. Alternative theories, such as Topcolor [24, 25] explain the large top mass as arising, at least in part, from some new strong dynamics. Clearly, measuring independently the value of the Yukawa coupling would provide important information on the mechanism of fermion mass generation.

The value of the top quark Yukawa coupling can be accessed experimentally by searching for $t\bar{t}H$ production (figure 1.10). The analysis [26] requires one of the top quarks to decay leptonically and the other hadronically. Since $t\bar{t}H$ production has significant cross-section only for relatively light Higgs masses ($m_H < 150$ GeV/c^2), the Higgs boson is detected through its decay $H \to b\bar{b}$, the dominant decay channel for the m_H range of interest. Thus the final state contains an isolated lepton, missing p_T, two light quark jets, and a total of four b-jets. The resulting large combinatorial background is dealt with by first reconstructing both top quark decays. The combination which better satisfies both t and $\bar{t}$ mass constrains is used to assign jets to the top decays. A search is then made for a $H \to b\bar{b}$ signal using only the remaining unassigned b-jets.

Once the Higgs boson rate in a given decay channel is measured, an accurate theoretical prediction for the Higgs boson production cross-section

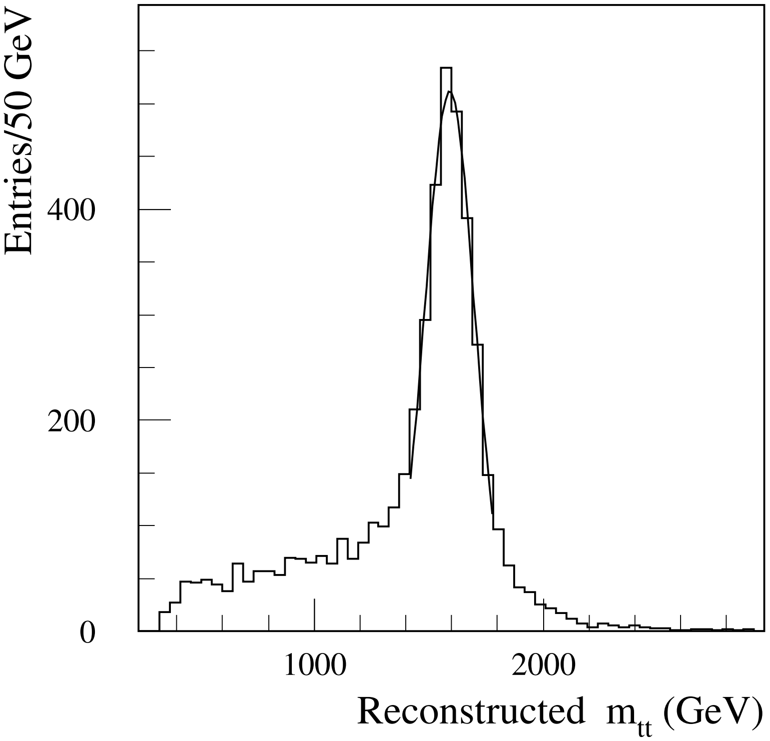


Figure 1.11: Measured $t\bar{t}$ invariant mass distribution for reconstruction of a narrow resonance of mass 1.6 TeV/c^2 decaying to $t\bar{t}$

would allow a measurement of the branching ratio for the decay in that channel. Without theoretical assumptions, one can only measure ratios of rates for different channels, which in turn provide ratios of couplings and branching ratios. By performing these measurements for several channels, one can obtain several constraints on the Higgs boson couplings to fermions and bosons, which can be used to test the theory.
For an integrated luminosity of 30 fb^{-1} and an Higgs boson mass of 100 GeV/c^2 , the Yukawa coupling can be measured with a relatively statistical error of 11.9%, improving to 9.2% for an integrated luminosity of 100 fb^{-1}

Search for $t\bar{t}$ resonances

A number of theoretical models predict the existence of heavy resonances which decay to $t\bar{t}$. In the case of MSSM if $m_A, m_H > 2m_{top}$ and $\tan\beta \approx 1$, then both A and H would decay almost totally in $t\bar{t}$ pairs [27, 28]. The possible existence of heavy resonances decaying to $t\bar{t}$ arises also in thechnicolor models [29, 30] as well as other models of strong electroweak symmetry breaking [31, 32]. Because of the large variety of models and their param-

eters, a study was made of the sensitivity to a “generic” narrow resonance decaying to $t\bar{t}$. Events of the semi–leptonic channel $t\bar{t} \rightarrow \mathrm{W^+W^-b\bar{b}} \rightarrow \ell^{\pm}\mathrm{jjb\bar{b}}$ are selected requiring a high isolated p_T lepton, missing transverse energy and at least four jets, with at least one tagged as *b*-jet. After this selection, the background to $t\bar{t}$ resonant production is dominated by continuum $t\bar{t}$ production (see figure 1.11).
For example, with 30 fb^{-1} , a narrow resonance X of mass 500 GeV decaying to $t\bar{t}$ could be discovered provided its $\sigma \times \mathrm{BR}$ is at least 2560 fb. This value decreases to 830 fb for $m_X = 1$ TeV, and 160 fb for m_X=2 TeV [33].

$t\bar{t}$ spin correlations

The SM prediction of the top quark width, given the large value of m_{top}, is $\Gamma_{top} \sim 1.5$ GeV/c^2 . Thus the top quark lifetime is very short in comparison with the hadronization time ($\sim 1/\Lambda_{qcd}$), and the top quark decays as a “bare quark” before hadronizing. In addition, the top quark decays before the strong interaction has time to depolarise its spin. As a consequence, the spin orientation of the top quark should be preserved in its decay. The weak decay of the top quark implies the daughters in the decay chain can be used to analyze its spin orientation.
To the lowest order, top quarks produced via the strong process $gg/q\bar{q} \rightarrow t\bar{t}$ are unpolarised, and the transverse polarisation effects due to loop diagrams are predicted to be very small. However, the spins of the t and $\bar{t}$ are correlated. At the LHC, the top and anti–top quarks tend to be produced with the same helicity, thus favouring the production of “Left-Left” (LL) or “Right-Right” (RR) $t\bar{t}$ pairs. For example, for $m_{t\bar{t}} < 500$ GeV, about 80% $t\bar{t}$ pairs are predicted [34] to be produced with either LL or RR helicities. This fraction falls slowly to a little under 70% for $m_{t\bar{t}} < 1000$ GeV. A measurement of this spin correlation would check wether the top quark does indeed decay before the strong interaction has time to depolarise its spin, and thereby would allow a lower limit to be set to Γ_{top}. Furthermore, new physics, such as large CP violation in the top system, could alter the spin correlations predicted by the SM.
In ATLAS with integrated luminosity of 100 fb^{-1} measuring the azimuthal and opening angle difference between the two leptons arising from the di–lepton events would allow to see differences, at the level of a few percent, between different models of correlation of $t\bar{t}$ spins [35].

Rare top decays

With its large mass, the top quark will couple strongly to the sector of electroweak symmetry breaking (EWSB). Many models of physics beyond the SM include a more complicate EWSB sector, with implications for top quark decays. Examples include the possible existence of charged Higgs bosons, or possibly large flavor changing neutral currents (FCNC) in top decays. ATLAS sensitivity to some of these scenarios have been studied, here they will be only listed.

- $t \to H^+b$. If a sufficiently light charged Higgs boson exists, the decay $t \to H^+b$ could compete with the SM decay mode $t \to W^+b$. Such a possibility could be seen by looking inclusively at the ratio of di-lepton and single lepton $t\bar{t}$ events [36].

- Favor Changing Neutral Currents (FCNC). Within the SM, FCNC decays of the top quark are highly suppressed, and so many observation of FCNC top decays at the LHC would be an indication of new physics. While the MSSM does enhance the branching ratios, they would be still too small to be observable. However, other extensions of the SM, including models with new dynamical interactions of the top quark, with multiple Higgs doublets, and with new exotic fermions, can lead to very significant enhancements of FCNC top decays. The ATLAS sensitivity to some channels have been studied, this includes: $t \to Zq$ [37], $t \to \gamma q$ [38], $t \to gq$ [39].

- $t \to WbZ$. The "radiative" top decay $t \to WbZ$ has been suggested as a sensitive probe of the top quark mass, since the measured value of m_{top} is close to the threshold for this decay. Withing the uncertainty of $\delta(m_{top}) \approx 5$ GeV/c^2, the predicted branching ratio varies by a factor of three. A measurement of BR($t \to WbZ$) could, therefore, provide a strong constraint on the value of m_{top} [38].

Chapter 2

The ATLAS detector

The stringent requirements for new physics discovery and precise measurements in the SM sector, calls for new experiments with extreme performance. The new LHC machine, under construction at CERN, will provide pp collisions at $\sqrt{s} = 14$ TeV *with a design luminosity of* $L = 10^{34}$ cm^{-2}s^{-1}.
The present chapter describes the ATLAS experiment, which is one of the two general purpose detectors under construction along the LHC ring. Each detector will be described and the status of installation will be shortly presented.

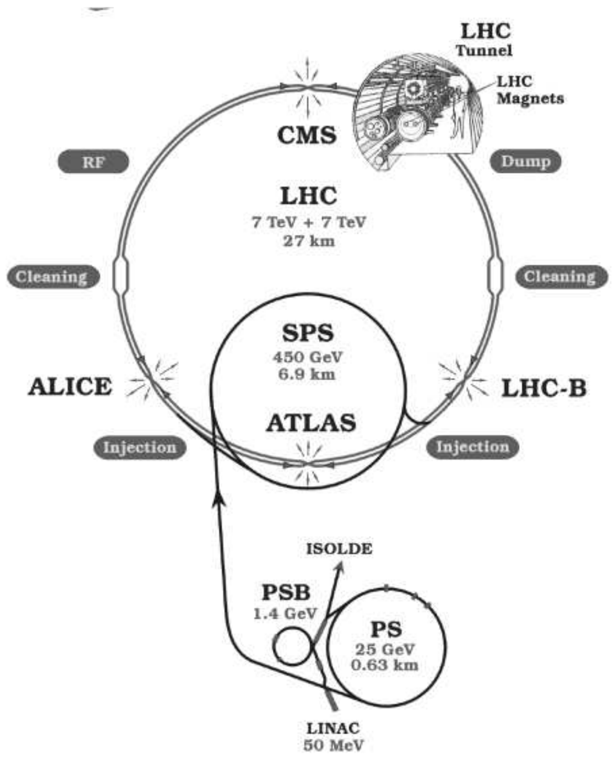


Figure 2.1: Layout of the LHC.

2.1 The Large Hadron Collider

The Large Hadron Collider (LHC) [40] is a superconducting proton Synchrotron which is being constructed at CERN inside the 27 Km tunnel used in the past years by LEP. LHC will accelerate from 450 GeV to 7 TeV two beams of protons (or heavy ions) traveling in opposite directions. The existing machines at CERN will do the first stages of acceleration (fig. 2.1): first the protons are accelerated up to 50 MeV in the proton linac, then the Proton Synchrotron Booster (PSB) will let them reach 1.8 GeV. The Proton Synchrotron (PS) will accelerate them up to 26 GeV. Finally, the SPS will be used to inject into the LHC 450 GeV protons.

The LHC dipole magnets will house in one single twin bore magnet (with the same yoke and cryostat) two different magnetic channels. The magnets will provide a magnetic field of 8.36 T, which allows the colliding beam particles to reach the design energy of 7 TeV. The design luminosity of 10^{34} $\mathrm{cm}^{-2}\mathrm{s}^{-1}$ is expected to be reached a few years after the first collisions. After the commissioning of the machine, it is foreseen to have one year at the so called low luminosity ($2 \cdot 10^{33}$ $\mathrm{cm}^{-2}\mathrm{s}^{-1}$), which should provide enough statistics to perform the first physics measurements.

Parameters	p–p	Pb^{82+}–Pb^{82+}
Beam energy (TeV)	7.0	7.0
Center of mass energy (TeV)	14	1262
Injection energy (GeV)	450	190.6
Bunch spacing (ns)	25	124.75
Particles per bunch	1×10^{11}	6.2×10^{7}
R.M.S bunch length (m)	0.075	0.075
Number of bunches	2835	608
Initial luminosity ($cm^{-2}s^{-1}$)	10^{33}	1.95×10^{27}
Luminosity ($cm^{-2}s^{-1}$)	10^{34}	1.8×10^{27}
Luminosity lifetime (h)	10	10
Dipole field (T)	8.3	8.3

Table 2.1: The LHC parameters.

The design luminosity for the *pp* collisions will be reached with 2835 bunches crossing at 25 ns intervals, corresponding to a spatial separation of 7.5 m. The number of protons per bunch will be 10^{11}.
At the LHC energies, the total inelastic non–diffractive *pp* cross section is about 70 mb. Since the interesting processes have cross sections that are several orders of magnitude lower (the cross section for $t\bar{t}$ pairs production production is $\sigma_{t\bar{t}} \simeq 800$ pb, the inclusive Higgs production is well below 1 nb for any Higgs mass), a very selective trigger system is required.
The main parameters of the LHC for *pp* and heavy–ion collisions are shown in tab. 2.1.

Along the perimeter of the LHC, four experiments are being built: two general–purpose experiments (ATLAS and CMS [41]), one experiment dedicated to the study of heavy–ion collisions (ALICE [42]), and LHCb [43], dedicated to the study of B mesons decays.
The primary goal of LHC is the study of the origin of the spontaneous symmetry breaking. The discovery (and study) of possible physics beyond the standard model is the other major scientific goal: the possible existence of supersymmetry, fermion compositeness and extra-dimensions can all be investigated with the LHC at the TeV scale.

2.2 The ATLAS detector

The ATLAS (A Toroidal Lhc ApparatuS) detector [44] is being installed at Point 1 of LHC. A picture of the status of the installation on October 2005 is shown in fig. 2.2. The eight superconducting coils forming the toroidal magnetic system are well visible together with, on the background, barrel calorimetric system in the parking position.

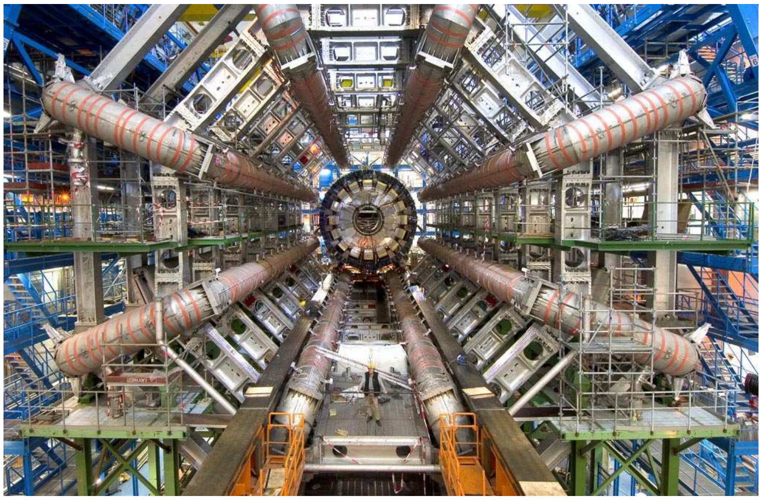


Figure 2.2: The ATLAS detector as it was on October 2005.

The calorimetry has been designed on the basis of the following criteria:

1. Very good identification and measurement of the four–vectors of electrons and γ's
2. Full hermeticity, both for the EM and hadronic sections
3. Accurate measurement of jets and missing transverse energy.

The coverage in the azimuthal angle ϕ is almost complete while the geometrical acceptance in pseudorapidity is $|\eta| < 5$. The external muon spectrometer allows high–precision measurements of the muon transverse momentum, even in stand–alone operating mode. This guarantees good $\mathrm{P_T}$ measurement also at high luminosity. The tracker has been designed to have high efficiency for non–isolated tracks and to provide electron, photon, τ and b identification.

2.2.1 Overview

Figure 2.3 gives an overall schematic view of the detector. The two independent magnetic systems (the solenoidal in the inner detector and the toroidal one in the muon spectrometer) are obtained with a thin superconducting solenoid surrounding the inner detector and with 8 independent coils arranged with an eight–fold symmetry outside the calorimeters, respectively. The central solenoid envelopes the Inner Detector and provides a 2 Tesla field oriented along the beam axis. The inner detector makes use of three different technologies, at different distances from the interaction point. Three inner

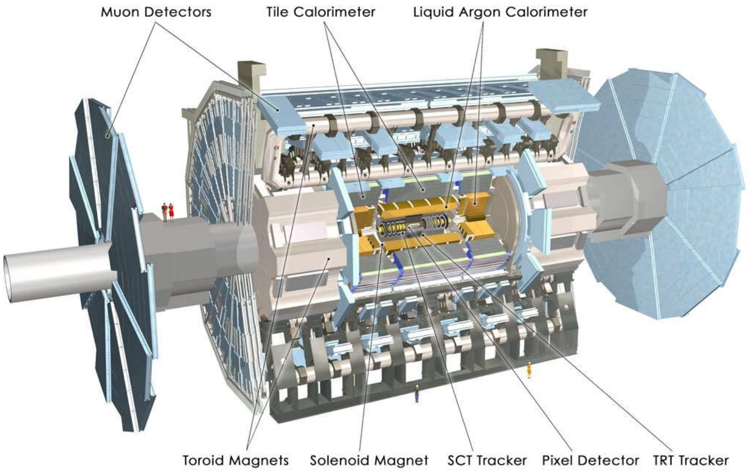


Figure 2.3: Overall view of the ATLAS detector.

layers of pixels allow good secondary vertex identification and, together with the four layers of silicon micro-strips, good momentum measurements. The tracking is then completed by continuous straw–tubes detectors with transition radiation detection capability in the outer part.

The calorimetry uses radiation–hard liquid argon (LAr) technology for the EM barrel and end–cap, for the Hadronic End–Cap (HEC) and for the Forward (FCAL) calorimeters. In the barrel region the cryostat is shared with the superconducting solenoid, while the EM end–cap, the HEC and the FCAL share the same cryostat in the forward region. In the barrel region the longitudinal hermeticity is provided by the Tile calorimeter (TileCal). Scintillating tiles are used as active material, while the passive material is iron. The TileCal is subdivided in a barrel ($|\eta| < 1$) and an extended barrel ($1 < |\eta| < 1.7$) region. The gap between the two is covered by the ITC and the Intermediate Gap Scintillators, which allow the recover part of the energy lost in the gap.

Outside the calorimeters there is the muon spectrometer. The magnetic field is provided by the 25 m long coils in the central region. The coverage at small angles is completed by two end–cap toroids. The magnetic field bends the particles inside the open structure that constitutes the support for the muon chambers. The multiple scattering is therefore minimized. This allows a very good measurement of the muon momentum with three stations of high precision tracking chambers. The muon detector includes fast response trigger chambers, which operate in coincidence to provide a fast trigger decision on the muon $\mathrm{P_T}$.

The total radius of the ATLAS experiment, from the interaction point to the last muon chamber, is about 11 m. The total longitudinal size is about

46 m, the overall weight is about 7 KTons. Almost 90% of the total ATLAS volume is occupied by the toroids and by the muon spectrometer.

2.2.2 The magnetic system

As already mentioned, two different magnetic fields are present within the volume of the detector: the central one, provided by the solenoid, and the outer one, produced by the toroids [45].

The central superconducting solenoid provides a central magnetic field of 2 T, while the peak value (at the superconductor face) is 2.6 T. In order to obtain the desired calorimetric performances, in particular for photon and electron energy measurements, a careful design to minimize the amount of dead material in front of the calorimeters has been done: the solenoid is placed inside the same vacuum vessel of the LAr calorimeter. The amount of dead material due to the solenoid and the cryostat wall is about one radiation length.

The magnetic field for the muon spectrometer in the barrel region is provided by a system of 8 coils assembled radially with an eight fold symmetry. The magnetic field in the forward region is obtained with the end–cap coils system, which is rotated by 22.5^o with respect to the barrel coils to provide radial overlap and to optimize the bending power in the interface regions of the two coil systems. The peak magnetic field obtainable in the barrel region is about 4 T. The coils of the barrel are 25 m long and their height is 4.5 m. One cryostat is present for each coil. In the end–cap region there is only one cryostat within which the coils (5 m long and 4.5 m tall) are housed.

2.2.3 Inner detector

The strategy used for the ATLAS tracker [46] is to combine few high precision measurements close to the interaction point with a large number of lower precision measurements in the outer radius. The inner detector is embedded in the 2 T magnetic field provided by the central solenoid. The structure of the inner detector is shown in fig. 2.4. Within a radius of 56 cm from the interaction point, pixel and silicon micro-strip technologies offer a fine–granularity, thus a high precision of the measurement. A track typically hits three layers of pixels (which measure both $R-\phi$ and z coordinate) and 8 layers of strips (SCT), for a total of 7 tracking points. Two SCT layers form one stereo strip (the angle between them is 40 mrad), allowing the measurement of the three coordinates. In the barrel region (which covers up to $|\eta| = 1$ for a total length of 160 cm), the pixels and SCT are arranged in concentric cylinders around the beam axis, while in the end–cap (up to $|\eta| = 2.5$) they are arranged in disks perpendicular to the beam axis.

A large number of tracking points (36) is provided by the Transition Radiation Tracker (TRT) that also can give e/π separation identifying the

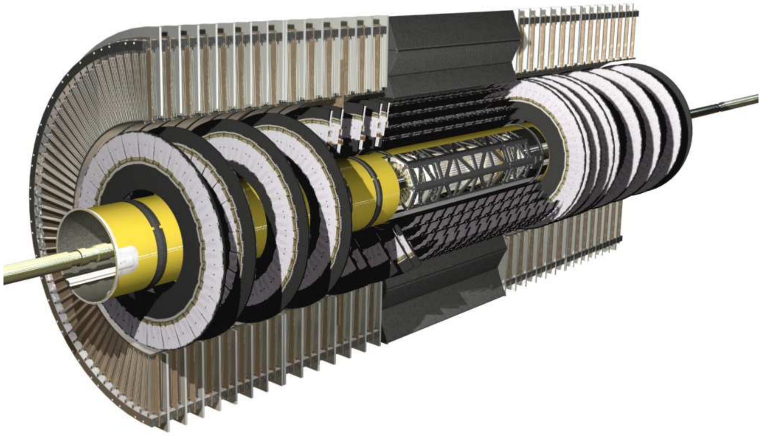


Figure 2.4: Overall view of the inner detector.

transition radiation emitted by electrons travelling at high speed. It consists of straw tubes arranged parallel to the beam axis in the barrel region and in wheels around the beam axis in the end–cap. The reduced resolution with respect of the silicon detectors is compensated by the higher radius and by the number of points measured. Therefore, the relative precisions of the measurements of the TRT and pixels/SCT are comparable. The TRT detector is intrinsically radiation hard.

The outer radius of the inner detector cavity is 115 cm, while the total length is 7 m. The layout provides full tracking coverage within $|\eta| < 2.5$, including impact parameter measurement and vertexing for heavy flavors and τ tagging. The expected precision for the detector is

$$\begin{aligned} \sigma_{R-\phi}(\mu\mathrm{m}) &= 13 \oplus \frac{62}{\mathrm{P_T}\sqrt{\sin\theta}} \\ \sigma_z(\mu\mathrm{m}) &= 39 \oplus \frac{90}{\mathrm{P_T}\sqrt{\sin\theta}} \end{aligned} \tag{2.1}$$

While the radiation impact is low on the TRT detector, it is not in particular for the pixels, which are more exposed to the radiation since they are closer to the interaction point. The intrinsic radiation weakness of the silicon would probably impose their substitution after a few years of operation, depending on the luminosity profile.

2.2.4 Calorimeters

ATLAS calorimetric systems differ in technology and materials depending on the pseudorapidity region. Liquid Argon (LAr) technology is used as active material for the electromagnetic calorimeters on all pseudorapidity

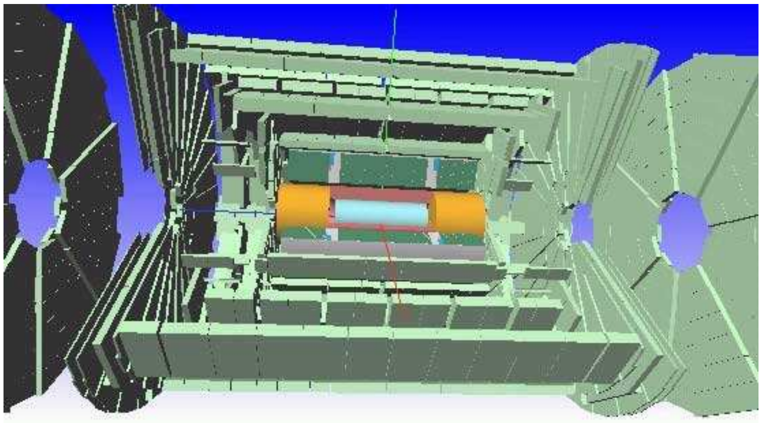


Figure 2.5: Overall view of the ATLAS muon detector.

range and for the hadronic calorimeter in the end-cap regions (HEC). In the end-caps regions the HEC and electromagnetic calorimeter are placed in the same cryostat.
Different absorbers are used in the different regions: lead in the barrel up to $|\eta| < 1.7$ and the end-caps ($1.5 < |\eta| < 3.2$), copper for the HEC.
An homogeneous LAr presampler detector is placed just behind the cryostat wall in the region up to $|\eta| = 1.8$.
In the barrel region ($|\eta| < 1.7$) the hadronic calorimeter is composed of a iron-scintillating tiles calorimeter (TileCal) subdivided into three parts: the central barrel covers up to $|\eta| \simeq 1$, while the two extended barrels cover up to $|\eta| \simeq 1.7$.
In the very forward region, up to $\eta \simeq 5$, the system is completed by a very dense LAr calorimeter consisting of rod-shaped electrodes in a tungsten matrix.
Since the importance of the calorimetric system for the topic of this work, a detailed description of the systems and results from the hadronic barrel calorimeters will be presented in the next chapter.

2.2.5 Muon spectrometer

One of the most important features of the muon spectrometer [47] is the possibility of a precise standalone measurement of the muon momentum. The magnetic field provided by the superconducting air–core toroid magnets deflects the muon trajectories that are measured by high precision tracking chambers. The magnetic field in the $|\eta| < 1.0$ range is provided by the barrel toroids, while the region $1.4 < |\eta| < 2.7$ is covered by the end–cap. In the so called transition region ($1.0 < |\eta| < 1.4$) the combined contributions

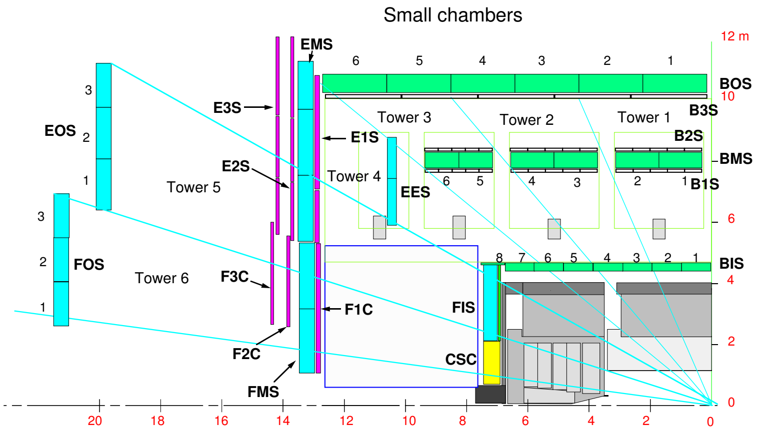


Figure 2.6: A view of one ϕ plane ($\phi = 22^o$) of the chamber layout.

of both the barrel and end–cap provide the magnetic field coverage. The magnetic field is mostly orthogonal to the muon trajectory in the covered pseudorapidity range, while the effect of multiple scattering is minimized.

In the barrel region, the muon chambers are arranged in three cylindrical layers (stations), while in the end–cap they form three vertical walls. The transition region is instrumented with one extra station.

The azimuthal layout follows the magnet structure: there are 16 sectors. The so called Large sectors lie between the coils, and they overlap with the Small sectors, placed in correspondence with the coils themselves. The $R-z$ layout of the chambers is shown in fig. 2.6, while fig. 2.5 offers a three dimensional view of the spectrometer.

The choice of the different chambers technology has been driven by the particle fluxes foreseen in the different regions of the detector. Criteria of rate capability, granularity, aging properties and radiation hardness have been considered. Table 2.2 summarizes the chamber technologies used in the various pseudorapidity regions.

The measurement of the track bending coordinate (η) is provided (in most of the η region) by the Monitored Drift Tubes (MDT), while at large pseudorapidity, the higher granularity Cathode Strip Chambers (CSC) are used.

The requirements on the momentum resolution ($\Delta P_T/P_T \simeq 10\%$ at 1 TeV/c) call for an accuracy of the relative positioning of chambers traversed by a muon track that matches the intrinsic resolution and the mechanical tolerances of the precision chambers.

the knowledge of the chamber positioning with an accuracy of 30 μm is required within a projective tower. The accuracy required for the relative positioning of different towers to obtain adequate mass resolutions for multi-muon final states is in the millimeter range. This accuracy can be achieved

Region		station I	station E	station M	station O
Barrel	$\|\eta\| < 1$	MDT		MDT RPC	MDT RPC
End-Caps	$1 < \|\eta\| < 1.4$	MDT TGC	MDT		
	$1.4 < \|\eta\| < 2$	MDT TGC		MDT TGC	
	$2 < \|\eta\| < 2.4$	CSC			MDT
	$2.4 < \|\eta\| < 2.7$	CSC		MDT TGC	

	TRIGGER CHAMBERS		PRECISION CHAMBERS	
Technologies used	RPC	TGC	MDT	CSC
Number of channels	354K	440K	372K	67K
Area (m^2)	3650	2900	5500	27
Time resolution	$< 5\ ns$	$< 7\ ns$	500 ns	<7 ns
Spatial resolution	5-10 mm		80 μm	60 μm

Table 2.2: *Design parameters of the Muon spectrometer.*

by the initial positioning and survey of chambers at the installation time. The relative alignment of muon spectrometer, calorimeters and ID will rely on the measurement of the high-momentum muon trajectories.

The MDT chambers are equipped with a in–plane alignment system aiming at a measurement of the tube position displacements, with respect to their nominal positions at the assembly phase, with a precision better than 10 μm. To achieve this the spectrometer is equipped with a RASNIK system: a laser, mounted at one side of a chamber, projects a pattern to a CCD camera positioned at the other end of the camera. From the displacement of the pattern–figure respect to what is expected, corrections for chambers deformation can be computed.

The chambers for the LVL1 muon trigger system covers the region $|\eta| < 2.4$. Resistive Plate Chambers (RPC) are used in the barrel region, while the Thin Gap Chambers (TGC) are used in the end–cap. Their first task is to identify without any ambiguity the bunch crossing of the triggered event. This requires a time resolution better than 25 ns. Next, they have to provide a well defined $\mathrm{P_T}$ cutoff for the LVL1 choice. This is obtained considering a window of a size defined by the LVL1 $\mathrm{P_T}$ threshold considered on the second RPC (or TGC) station once a super–hit has been obtained in the first station. Finally, the trigger chambers measure the non–bending coordinate (ϕ), in a plane orthogonal to that measured by the precision chambers, with a typical precision of 5–10 mm.

2.2.6 Trigger, Data Acquisition and Control systems

The ATLAS trigger and data acquisition systems (TDAQ) and the detector control system (DCS) are responsible for the data-flow of data from detectors' front-end electronics modules to data storage at CERN computing center, for the selection of events and for the control and monitoring of the apparatus [48].

A schematic diagram is presented in figure 2.7 and can be broken into four

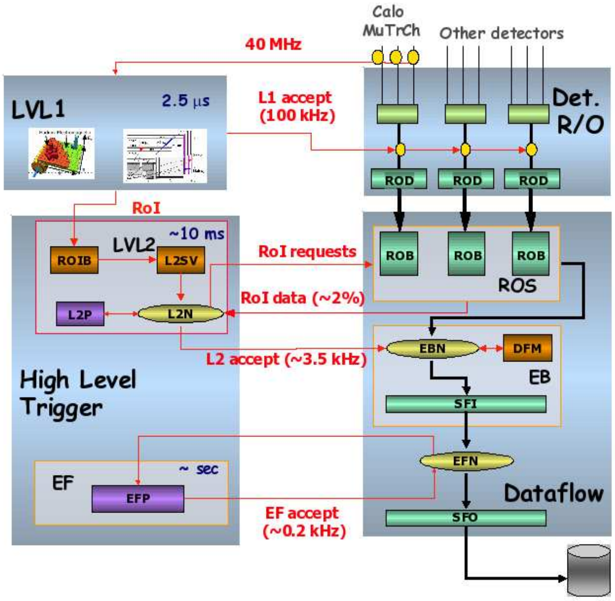


Figure 2.7: Principal components of the Data Flow (right) and trigger systems (left). Black lines represent the path of data, red lines movements of trigger data.

principal systems, namely:

1. The Data Flow System - responsible for the receiving the detector data, serving data to the trigger system, and transporting the data for selected events to mass storage
2. The Trigger System - subdivided into three layers. A first purely hardware level (LVL1) and two software levels (collectively called HLT)
3. The Online System - responsible for all aspects of experiment and TDAQ operation and control during data-taking, and during testing and calibration runs
4. The DCS - responsible for the coherent and safe operation of the ATLAS detector, as well as the interface with external systems and services including the LHC itself

The Online system is implicitly understood to be connected to all elements in figure 2.7, and the DCS to all hardware elements which need to be monitored and controlled.

In the right-half of figure 2.7 the data flow system is represented in the different components. The signals coming from the detectors (upper part), read by the input-output systems are managed by the back-end electronics boards: the so called Read Out Drivers (ROD). They are detector-specific and they have the responsibility of the reading of the data from the front-end electronics when a Level 1 trigger request is propagated to the system. The RODs, equipped with DSPs, have also the responsibility to make the first pre-formatting of the data to the ATLAS TDAQ final format and, in some cases, like for the calorimeters, apply some simple algorithms and performing calculations. From this point on, the components of the Data Flow are no more detector specific.

The data are then moved to the ROSes (Read Out System), normal PCs equipped with one or more buffer cards, the Read Out Buffers (ROBs). Here they will be further formatted and moved, after request, to the Event Building Network (EBN), passed to the last stage of trigger, and in case of acceptance, stored to a set of local disk pools (SFO: Sub Farm Output). The numbers of the TDAQ system are reported in table 2.3 where the numbers of RODs, ROSes and CPUs are reported for the different subsystems.

Starting from an initial bunch-crossing rate of 40 MHz, the rate of selected events must be reduced to about 100 Hz for permanent storage. The event rate is determined by the total *pp* cross section, and is about 10^9 evt/sec at the nominal luminosity. Hence an overall rejection factor of 10^7 against minimum bias events is required. This strong requirement must match the need of an excellent efficiency for the rare physics processes of interest. Therefore the principal requirement of the Level 1 (LVL1) trigger is that it identifies unambiguously the interesting physics events, while strongly reducing the overall rate.

The LVL1 trigger makes an initial selection based on a reduced granularity information from a subset of detectors. Objects searched by the calorimeter trigger are high P_T electrons and photons, jets, and τs decaying into hadrons, as well as large missing and total transverse energies. High and low transverse momentum muons are identified using only the muon trigger chambers.

The calorimeter selections are based on a reduced-granularity information from all the calorimeters. In the case of the electron/photon and hadron/τ triggers, energy isolation cuts can be applied. The missing and total scalar transverse energies used in the LVL1 trigger are calculated by summing over trigger towers. In addition, a trigger on the scalar sum of jet transverse energies is also available.

No tracking information is used at LVL1 due to timing restrictions and the inherent complex nature of the information from the inner detector.

The LVL1 trigger decision is based on a logical combinations of these objects. Most of the physics requirements of ATLAS can be met by using, at the LVL1 trigger level, fairly simple selection criteria of a rather inclusive nature. However, the trigger implementation is flexible and it can be programmed to select events using more complicated signatures.

The maximum rate at which the ATLAS front-end systems can accept LVL1 triggers is limited to 75 kHz (upgradeable to 100 kHz). The target rates es-

detector	RODs	ROSes
Pixel	120	11
SCT	92	8
TRT	192	20
LAr (EM barrel + EMEC + HEC + FCAL)	192	33
TileCal	32	8
Muon System (MDT + CSC + RPC + TGC)	296	26
LVL1	24	5
	PC	
LVL2 Processors	$\sim$ 500	
EF Processors	$\sim$ 1600	
DataFlow and Control Machines	$\sim$ 300	

Table 2.3: Size of different detectors TDAQ components: number or RODs and ROSes and, for the HLT and for the DataFlow system, the number of needed PCs.

timated in trigger performance studies, using trigger menus that meet the needs of the ATLAS physics program, are about a factor of two below this limit.

An essential requirement on the LVL1 trigger is that it should uniquely identify the bunch-crossing of interest. Given the short (25 ns) bunch-crossing interval, this is a non-trivial task. In the case of the muon trigger, the physical size of the muon spectrometer implies times-of-flight comparable to the bunch-crossing period. For the calorimeter trigger, a serious challenge is that the pulse shape of the calorimeter signals extends over many bunch crossings.

During this time, information for all detector channels are stored in pipeline memories. The LVL1 latency, measured from the time of the *pp* collision until the trigger decision is available to the front-end electronics, is required to be less than 2.5 μs. In order to achieve this, the LVL1 trigger is implemented as a system of purpose-built hardware processors.

Another important functionality of the LVL1 system is the identification of the Regions Of Interest (ROIs) representing the position of the triggering objects in the (η,ϕ) space. This is one of the main peculiarities of ATLAS triggering system.

This information is used to greatly reduce the needed computation time at the LVL2 triggering system and the size of data to be transferred in the system.

The LVL2 runs offline-like algorithms, optimized for the on–line use, using the full granularity information from the inner detector as well as from the muon detectors and calorimetry. However it is structured to process the data belonging only to a spatial window around the ROIs identified by the LVL1 trigger. Simulations showed that this corresponds to roughly 2-5% of the overall ATLAS data size. LVL2 has a maximum latency time of 10 ms, after this time the event is selected (and hence moved to the Event Filter system for further processing) or discarded and removed from the Data Flow chain (up to this moment the event fragments have been buffered in the ROBs. The final LVL2 rate is expected to be about 1-2 KHz.

After LVL2, the last stage of the on-line selection is performed by the Event Filter. It will employ offline algorithms and methods, slightly adapted to the on–line environment, and use the most up to date available calibration and alignment information and the magnetic field map. Complete event reconstruction is performed by the Event Filter, which will make the final selection of physics events to be written to mass storage for subsequent off-line analysis.

The time available for a decision at the event filter is 1 s. The output rate from LVL2 should be reduced by an order of magnitude, giving about 100 Hz. The final event size is expected to be 1 MBytes corresponding to an output data rate of about 100 MByte/s resulting in 10^{15} bytes of data per year.

2.2.7 Computing

The complexity and size of the ATLAS experiment imposes the use of new paradigms also in what the processing of the data is concerned once they are made available on mass storage. The events rate of 100 Hz, the size of the events (approximately 1 MB per event), the number of physicists involved in the analysis requires that the data distribution, processing and analysis is carried out according to a multi-tier schema, that is well suited to distribute the computing and storage loads among the different participating institutes. Even if similar strategies have been used in the past it is the first time that this kind of distributed analysis are performed on a ATLAS-size scale requiring the development of completely new tools extremely performing [49].

At the output of the event filter the raw data are transferred to the CERN's computing center, known as Tier-0, that is the first layer of the ATLAS analysis system. Here a complete copy of the raw data is stored and a first-pass reconstruction is applied producing ESD (Event Summary Data) and AOD (Analysis Object Data). The ESD data-format contains the reconstructed quantities measured by the detector (energy in the calorimetric cells, clusters information, tracks) as well as the reconstructed physics objects (electrons and gammas, jets, taus, muons). The small-sized data in AOD format are well suited for distribution to the physicists groups, to reduce their size only the physics objects are recorded. Each event can be characterized by few quantities, like, for example, the number of jets in the event, P_T of the leading jet, lepton multiplicity and so on, this information, produced by Tier-0 and stored in the TAGs and allows for a very fast filtering of the datasets.

Tier-0 has also the responsibility to run calibration and alignment algorithms that will be refined in future steps. Distribution of data to the community is done copying raw data, ESDs, AODs and TAGs to the Tier-1s. Tier-1s are big regional computer centers spread around the world (at the moment ATLAS foresees 10 of this centers). A copy of the raw data is divided among all the Tier-1s (each one having on average 10% of the entire raw data) while a complete copy of the ESDs, AODs and TAGs is distributed to each Tier-1. Tier-1s have also the responsibility to reprocess raw data performing more accurate reconstructions, updated version of ESDs, AODs and TAGs are therefore constantly produced and spread among the different computer centers.

Most of the physics analysis is performed at the Tier-2s center, an average of 5 Tier-2s are directly connected to one Tier-1 and typically receive a copy of one third of the most updated ESD and AOD data and a complete copy of the TAGs, they have the entire responsibility for the official Monte Carlo production (the simulated data are stored in the more reliable Tier-1s), the physics groups analysis and the development and refinement of calibration and reconstruction algorithms are also performed at the Tier-2 centers. The physics analysis will be performed mainly on the AOD data set (with the help of TAGs for pre-selections) or on even more compact derived formats like ntuples.

The multi-tier paradigm is deployed using *grid* technology and middle-ware that completely hides to the physicists the complex multi-tier structure [50].

2.3 Status of installation in the experimental cavern

LHC will start to provide interactions at the end of 2007 with reduced beam energy (E_b = 450 GeV) and initial low luminosity ($L = 10^{26/27}$ cm^{-2}s^{-1}). This pilot run will be followed by a slow ramping up of the luminosity and beam energy up to the nominal initial conditions. During this period the ATLAS detector will have the possibility to perform debugging of the detector and start detailed studies of its performances.

Prior to the data taking with real beams ATLAS will undergo intense testing of all the sub–detectors, both with calibration signals and with cosmic ray muons. Detectors will perform this tests at first in stand alone mode and they will later be integrated.

This long commissioning phase has already started for many sub–detectors and the analysis of first results is undergoing.

In this section an overview, updated to July 2006, of the current status of ATLAS installation and commissioning is presented, however due to the rapidly changing conditions only an overview, without details, is presented.

Figure 2.8: ID Barrel integration: SCT (left) being inserted in the TRT.

Inner Detector

The Inner Detector system is the last sub–detector that will be lowered in the pit and, due to the relative small size, it will be possible to install it already mounted (see figure 2.8).
Up to the first half of 2006 barrel inner detector have been integrated on the surface and have been taking cosmic rays together showing detection efficiencies as expected, they have been then un–cabled and prepared for the installation in the pit in the second half of 2006. Before this step the services (voltages, cooling) must be installed in inner detector cavity and a complete mapping of the magnetic field have been performed.
The end caps have been arrived at CERN and tests are undergoing before integration and tests.

Liquid Argon Calorimeters

Since July 2004 LAr calorimeters are undergoing a commissioning phase. In October 2004 the EM Barrel cryostat was lowered in the pit and positioned inside TileCal. The reached relative alignment with TileCal is of some millimeters fully compatible with specifications.
From that moment on the installation of services and front end electronics has started and at the end of 2005 it was moved, together with TileCal, at the final position at the interaction point. In June 2006 the cryostat has been filled with LAr and a cool-down phase has been done. At the moment EM Barrel is preparing for the first integration test with TileCal and for the combined calorimetric (Barrel TileCal and EM barrel) cosmic run foreseen for the end of the year 2006.
End-Caps cryostats construction have been completed and the calorimeters

Figure 2.9: TileCal barrel at the interaction point. The cryostat with the front-end electronics for EM calorimeter is visible in the inner part. TileCal is surrounded by the 8 superconducting magnets for the muon toroid system.

(EM End-Cap, HEC, FCAL) have been inserted. In December 2005 (side C) and April 2006 (side A) they have been lowered in the pit and positioned, with accuracy of few millimeters, inside Extended Barrels of TileCal. Installation of front-end electronics and services is ongoing. The cool-down is foreseen for middle Autumn of 2006 and commissioning with cosmic rays (combined data with TileCal) is foreseen for March 2007 on time for the first circulating beam in 2007.

Tile Calorimeter

TileCal was the first system to go for installation in the pit in 2004, at the moment the three cylinders are installed and the commissioning of the electronics services and the installation of services is ongoing (see figure 2.9).
For some ϕ sectors the electronics and the power supplies have been put in place and since March 2006 the final Read Out Drivers are available.
The final TDAQ systems (the so called pre-series) have been installed and used to test TileCal system. Currently the activity is ongoing in finishing the commissioning of the electronics system. This is done acquiring short runs of different type with the complete acquisition chain (Drawers - RODs - ROS - Event Building). TileCal is at the moment acquiring “noise” runs and calibration runs (CIS, Laser) to check the level of the electronic noise, the functionalities of the front-end and back-end electronics, with the final

setup.

A special setup to start looking at cosmic rays has been implemented. It uses trigger boards emulating the LVL1 system, this boards can read the analog signal from TileCal (that usually goes to the LVL1 calorimetric trigger system) and can produce a trigger signal in case a muon is observed in TileCal towers. Many cosmic muon runs have been collected since end of 2005 and they are at the moment under study. In the near future they will be compared with muons from test beams and from simulations.

Figure 2.10 shows some results obtained in a cosmic run. In the top picture a passing muon is displayed in ATLAS event display Atlantis, different views are visible showing the cells (in yellow) with a energy deposit compatible with a minimum ionizing particle.

On the bottom two histograms from TileCal online monitoring system are presented. The first (from top) plot shows the η distribution of the most energetic calorimetric tower for the upper half of TileCal (only $\eta < 0$ towers were read). Bottom plot shows reconstructed signal (in pC, the calibration constants are not applied in the online monitoring). The muon distribution is well visible at an energy deposit of around 2 pC. The electronic noise is well separated from the signal ($Q < 0.5$ pC).

This are the first cosmic rays runs done by ATLAS, in the Fall 2006 TileCal and EM LAr barrel calorimeter will be combined and will take data together.

Muon System

The muon chambers of the barrel toroid magnetic system, MDT and RPC, are undergoing installation in the pit. Before MDT chambers (together with RPC trigger chambers) are mounted and connected to the services (high voltage, gas–flow system) in the pit they undergo intense cosmic ray tests on the surface. After a validation of the chambers they are moved to final position and connected to services. At the end of July 2006 47% of all barrel chambers have been installed.

During July of the same year a first test combining TileCal and Muon system has been performed, acquiring cosmic ray muons with both detectors, achieving the goal to run together, for the first time in the pit, more than one detector.

The installation of the end–cap wheels has started recently, July 2006.

The rate of cosmic ray muons passing this chambers will be not enough to allow a rigorous commissioning with muons, at the moment is therefor under discussion the possibility to use the beam–halo muons, during the LHC pilot runs, to conclude the commissioning of this chambers in the pit with the final system.

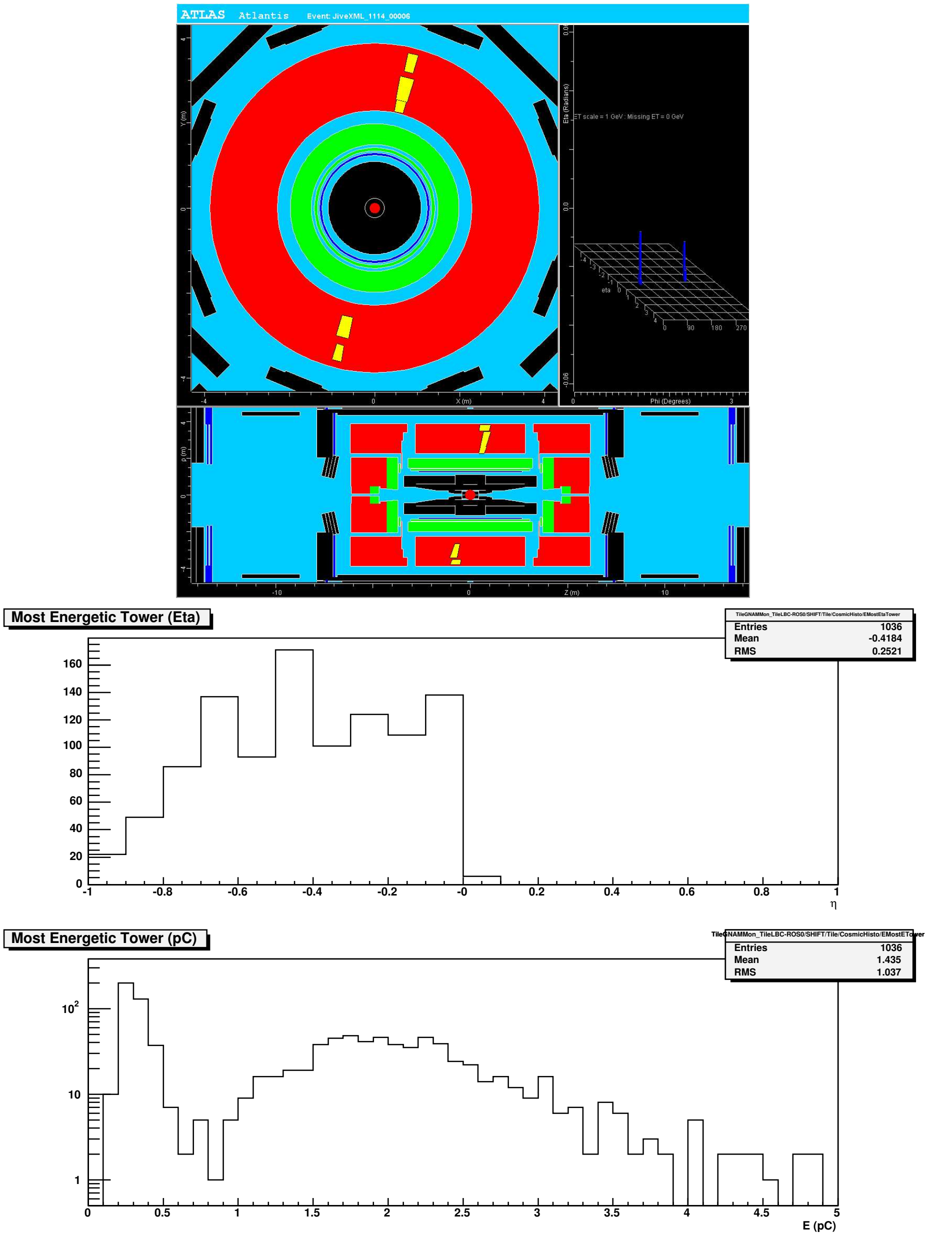


Figure 2.10: Cosmic rays muons from TileCal. Top: Event display showing a muon passing through TileCal in different views. Bottom: distribution for η of the most energetic tower in cosmic events (only one slice of the detector is being monitored) and energy distribution of the most energetic tower in cosmic events. The muons peak (around 2 pC) is well separated from the pedestal.

Chapter 3

ATLAS calorimeters and their response to hadrons: data from TileCal test beam

Different materials and technologies are used for the ATLAS calorimeters (LAr, scintillating tiles) for an efficient reconstruction of electromagnetic objects and of jets with a coverage up to $\eta = 5$*.*
In this chapter the ATLAS calorimeters will be presented in details. The response to hadrons of a non compensating calorimeter will be shortly reviewed and the ATLAS hadronic barrel calorimeter (TileCal) performance will be discussed in detail. Data from stand-alone test beams carried out in 2002 and 2003 will be presented and the response to electrons and pions of different energies will be discussed.

3.1 ATLAS calorimetry

3.1.1 Calorimeters requirements

The ATLAS detector is designed to operate in a challenging environment: LHC will have the highest luminosity and c.m. energy ever reached.

The calorimetric system will be used both to form the calorimeter triggers and in the measurement of jet momenta leading to complementary requirements.

The calorimeters will be used to form the trigger elements used by the first level of trigger (LVL1):

- **Electron/Photon trigger**. The electromagnetic calorimeter, thanks to the high granularity and to the good electromagnetic shower containment, will be used to identify single electrons and photons. The basic trigger elements will be used to trigger, for instance on $H \rightarrow \gamma\gamma$ events or the leptonic decay of W bosons.

- **Hadron/tau trigger**. The implementation of a Level 1 trigger based on the peculiar topology of τ decays (narrow jets with low multiplicity) can be used to improve the efficiency of triggering on events like $Z \rightarrow \tau^+\tau^-$ or on low-mass $A \rightarrow \tau^+\tau^-$. In coincidence with $\mathrm{E}_\mathrm{T}^{miss}$, it provides also a trigger on $W \rightarrow \tau\nu$. This trigger menu can also select high-E_T single hadrons for calibration studies in a way similar to what has been done at the Test Beam.

- **Jet trigger**. A trigger for high-P_T jets has proved extremely important at the past collider experiments (UA1/2, CDF/D0). At LHC we expect that this trigger will be essential in the study of QCD and helps in the study of new physics (SUSY, compositeness, etc.) as well as precise measurements of SM processes. A specific trigger on hadronic jets is therefor needed to select events with jets in the final state like.

- **Missing transverse energy and total transverse energy trigger**. These trigger menus are designed to identify, through the unbalanced energy measured in the calorimeters, final states with high energy neutrinos and particles that do not interact or interact weakly.

The trigger needs originating from the trigger needs outilined above and from the physics analysis require a good calorimetric system with the following requirements:

1. **Electromagnetic calorimeter requirements**
    - Good hermeticity with high granularity ($\Delta\eta \times \Delta\phi = 0.003 \times 0.1$) in the central region ($|\eta| < 2.5$) to have good acceptance on rare physics events ($H \rightarrow \gamma\gamma$, $H \rightarrow 4e$). In the forward direction

($2.5 < |\eta| < 5$) the main calorimeter task is the reconstruction of jets and the measurement of E_T^{miss} and a coarser granularity is acceptable.

- Capability to identify and measure electrons with energies above 1-2 GeV, since we need to reconstruct electrons produced in the b-mesons semi-leptonic decay, the Z/W leptonic decays and eventually new physics (Z'/W' decays).
- Excellent energy and angular resolution over the energy range 10-300 GeV. This is needed to achieve a mass resolution of $\backsim 1\%$ for the $H \rightarrow \gamma\gamma$ and $H \rightarrow 4e$ in the mass region 90-180 GeV.
- Total thickness of at least 24 radiation length at $\eta = 0$. This is required for full containment of electro magnetic shower hence to keep the contribution to energy resolution from longitudinal fluctuations of high-energy electrons ($E > 500$ GeV) showers to an acceptable level.

2. **Hadronic calorimeter requirements**

- Rapidity coverage. The main calorimeter task is the reconstruction of jets and the measurement of the missing P_T in the event: the jet detection has to cover up to $|\eta| \approx 5$ with very good hermeticity. This coverage is also required for an efficient tagging capability of the forward jets associated to the production of heavy Higgs.
- Granularity. The most stringent transverse granularity requirement comes from the $W \rightarrow jet + jet$ decays at high-P_T and requires, for $|\eta| < 3$, a granularity of $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$; at larger η regions a granularity of $\Delta\eta \times \Delta\phi = 0.2 \times 0.2$ is sufficient. The longitudinal segmentation is driven by particle identification, and by the possibility of achieving better energy resolution via weighting in a non-compensating calorimeter.
- Energy resolution. Resolution of $\Delta E = 50\%/\sqrt{E} \oplus 3\%$ in the central region is adeguate to the tasks of providing jet reconstruction and jet-jet mass reconstruction as well as missing P_T measurements.
- Energy linearity. The most stringent linearity requirements for the hadronic calorimeter come from the study of quark compositeness where the jet energy scale has to be linear within 2% up to transverse energy of 4 TeV. A 1.5% non-linear calorimeter can mask compositeness to a scale of $\Lambda = 30$ TeV.
- Total thickness. A total calorimeter (EM and hadronic) thickness of about 10 interaction lengths is required for shower containment, both for energy resolution reasons and for reducing the background in the muon chambers.

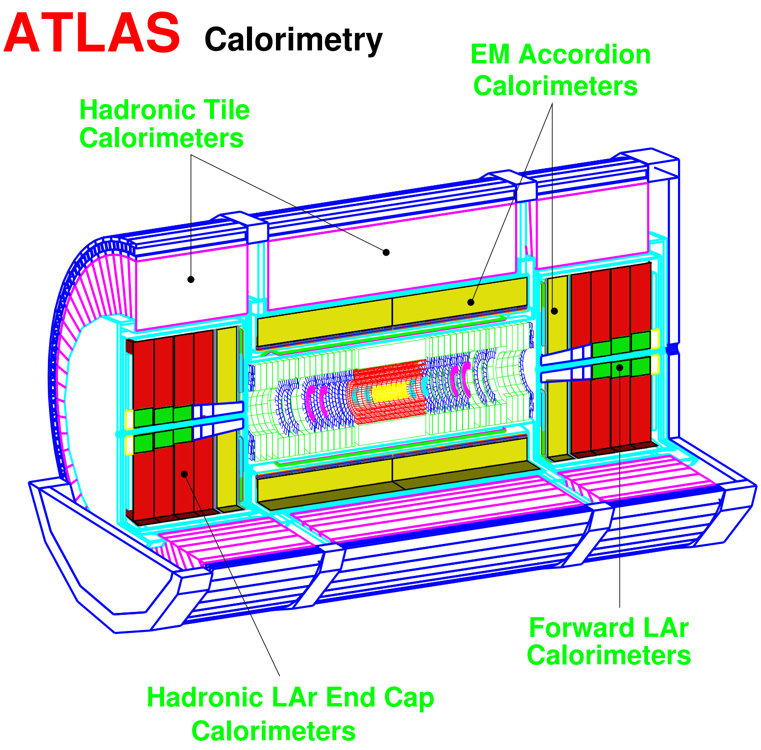


Figure 3.1: Overall view of the ATLAS calorimeters.

ATLAS calorimeters is designed keeping in mind the above requirements and will be able to cope with the physics goals.

3.1.2 The ATLAS calorimetric system

An overall view of the ATLAS calorimetric system is shown in figure 3.1. In this section we will describe the different technologies used in ATLAS calorimetric system.

The electromagnetic (EM) calorimeter of ATLAS uses liquid argon (LAr) as active material and lead as absorber. It covers up to $|\eta| < 3.2$: a high performance barrel calorimeter up to $|\eta| < 1.7$ and an end-cap section for $1.5 < |\eta| < 3.2$. In the region at $|\eta| < 1.8$ is located a fine granularity presampler just behind the cryostat wall. Given the large amount of material upstream the EM calorimeter (Inner Detector, cryostat, coil), the presampler is used to correct for the energy loss in the upstream material.

The hadronic barrel calorimeter (TileCal) is subdivided into three parts: the central barrel covers up to $|\eta| \simeq 1$, while the two extended barrels cover up to $|\eta| \simeq 1.7$. The gap between barrel and extended barrel, about 60 cm, necessary to accommodate the readout electronics and services of the EM calorimeter and the cabling of the inner detector, is instrumented with the ITC and gap scintillator. TileCal uses iron as absorber and scintillating tiles as active material. It is the mechanical support for the LAr cryostats, and it is the return yoke for the central magnetic field flux.

The hadronic calorimeter in the forward region uses LAr technology. The Hadronic End-Cap (HEC), which covers up to $|\eta| < 3.2$, is placed in the same cryostat of the EM end-cap and of the forward calorimeter (FCAL), which complete the pseudo-rapidity coverage up to $|\eta| \simeq 5$. The HEC uses copper as absorber, while the FCAL is a dense LAr calorimeter with rod-shaped electrodes in a tungsten matrix.

Table 3.1 shows the details of the segmentation of the calorimeters.

3.1.3 The LAr calorimeters

Though there are several differences in the absorber choice and in the read-out, the EM calorimeter, the HEC and the FCAL all use liquid argon as active material and are considered as a single sub-detector, the LAr calorimeter [51].

The EM calorimeters

The EM barrel consists of two identical half-barrels separated at $z = 0$ by a 6 mm gap, while each of the two EM end-cap are mechanically divided in two coaxial wheels. The outer one covers up to $|\eta| < 2.5$, while the inner one completes the EM coverage up to $|\eta| = 3.2$. The accordion shape of the electrodes (which envelop the lead absorber) provides complete ϕ symmetry without azimuthal cracks.

In the barrel region, the lead thickness changes, as a function of the pseudo-rapidity, to optimize the energy resolution while the LAr gap has a constant thickness of 2.1 mm. The geometry of the accordion becomes more complicated in the end-cap, where the amplitude of the accordion waves increase with the radius. Here the absorber has a constant thickness, therefore the size of the LAr gap increases with the radius.

Figure 3.2 shows a sketch of the EM structure in the barrel region. As shown, the modules are subdivided in three longitudinal samples. The outer wheel of the end-cap has also three samples, while the inner-wheel has only two. The first sample, which is 4.3 radiation lengths (X_0) long, has a fine segmentation in η (in the barrel $\Delta\eta = 0.003$), for a precise determination of the pseudorapidity of the impinging particle. Since the total amount of material in front of the EM calorimeter at $\eta = 0$ is about 1.7 X_0, there are 6 radiation length in front of the second sample, which is the largest one, with its 16 X_0's and a segmentation of $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$. The third sample has a coarser segmentation in pseudorapidity, and its thickness

EM Calorimeter	Barrel	End-Cap	
Coverage	$\|\eta\| < 1.475$	$1.375 < \|\eta\| < 3.2$	
Long. segmentation	3 samplings	3 samplings	$1.5 < \|\eta\| < 2.5$
		2 samplings	$1.375 < \|\eta\| < 1.5$
			$2.5 < \|\eta\| < 3.2$
Granularity ($\Delta\eta \times \Delta\phi$)			
Sampling 1	0.003×0.1	0.25×0.1	$1.375 < \|\eta\| < 1.5$
		0.003×0.1	$1.5 < \|\eta\| < 1.8$
		0.004×0.1	$1.8 < \|\eta\| < 2.0$
		0.006×0.1	$2.0 < \|\eta\| < 2.5$
		0.1×0.1	$2.5 < \|\eta\| < 3.2$
Sampling 2	0.025×0.025	0.025×0.025	$1.375 < \|\eta\| < 2.5$
		0.1×0.1	$2.5 < \|\eta\| < 3.2$
Sampling 3	0.05×0.025	0.05×0.025	$1.5 < \|\eta\| < 2.5$
PRESAMPLER	Barrel	End-cap	
Coverage	$\|\eta\| < 1.52$	$1.5 < \|\eta\| < 1.8$	
Long. segmentation	1 samplings	1 samplings	
Granularity ($\Delta\eta \times \Delta\phi$)	0.025×0.1	0.025×0.1	
TILE-CAL	Barrel	Extended Barrel	
Coverage	$\|\eta\| < 1.0$	$0.8 < \|\eta\| < 1.7$	
Long. segmentation	3 samplings	3 samplings	
Granularity ($\Delta\eta \times \Delta\phi$)			
Sampling 1 and 2	0.1×0.1	0.1×0.1	
Sampling 3	0.2×0.1	0.2×0.1	
Hadronic LAr		End-cap	
Coverage		$1.5 < \|\eta\| 3.2$	
Long. segmentation		4 samplings	
Granularity ($\Delta\eta \times \Delta\phi$)	0.1×0.1	$1.5 < \|\eta\| < 2.5$	
	0.2×0.2	$2.5 < \|\eta\| < 3.2$	
FCAL Calorimeter		Forward	
Coverage		$3.1 < \|\eta\| < 4.9$	
Long. segmentation		3 samplings	
Granularity ($\Delta\eta \times \Delta\phi$)		0.2×0.2	

Table 3.1: Design parameters of the ATLAS calorimeters.

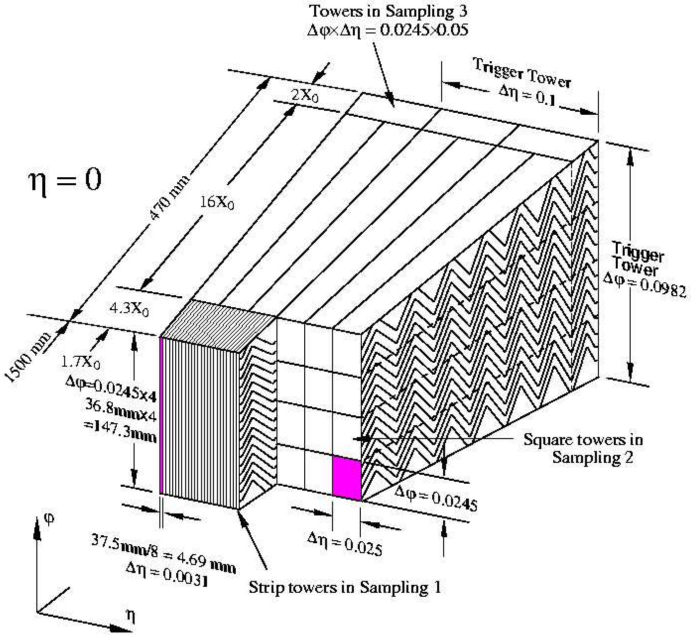


Figure 3.2: Sketch of the accordion structure of the EM calorimeter.

changes with η to have a total length of the EM calorimeter of 24 X_0. More details about the EM LAr segmentation can be found in table 3.1.

The analog signal from the EM calorimeters leave the cryostat through cold-to-warm feedthroughs. They are located in the outer faces of the cylinder, in the gap between the barrel and the extended barrel of the TileCal. The front-end electronics is also located here. The front-end electronics provides the digitization of the the analog signal (figure 3.3) with fast ADCs. As for any other sub-detector in ATLAS, the data are then stored in pipeline memories, waiting for a possible Level 1 Acceptance (L1A) signal. If the event is accepted by the Level 1 (LVL1), the corresponding signals (5 samples) are extracted from the pipelines and sent to the Read Out Drivers (ROD). Online analysis of the signal is performed using Optimal Filtering (OF) technique. The energy measurement in each channel is given by:

$$E_{cell}(\mathrm{MeV}) = F \sum_{i=1}^{5} a_i (ADC_i - P)$$

where F is a conversion factor between ADC counts and MeV (obtained from specific calibration runs), P is the pedestal in the cell (also obtained from specific runs) and a_i are the optimal filtering coefficients. A complete

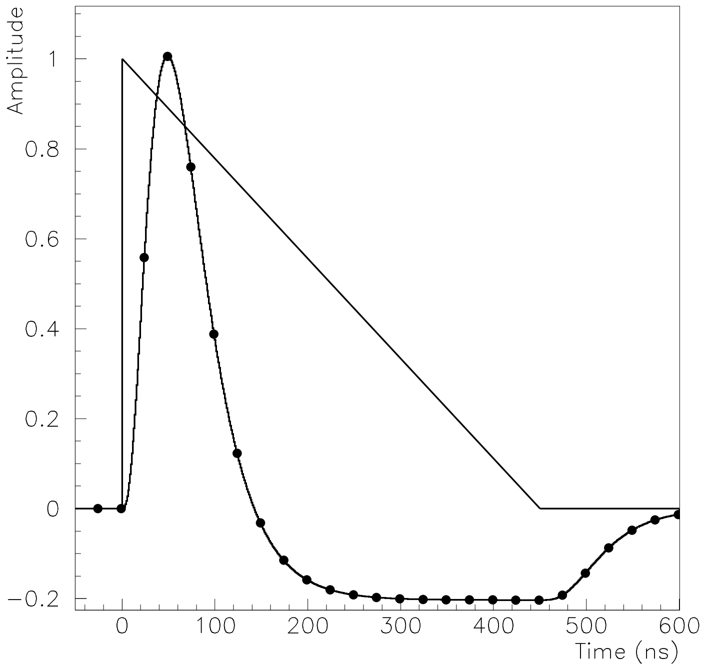


Figure 3.3: Signal shape of the EM calorimeter as produced in the detector (triangle) and after shaping (curve with dots). The dots represent the position of the successive bunch crossings.

discussion about the optimal filtering can be found in [52].
The linearity of the EM calorimeters has been verified (both for the barrel and for the extended barrel) with electrons up to 350 GeV at the test-beam obtaining a linearity better than 1%. The resolution found for the barrel at $\eta = 0.9$ is:

$$\frac{\sigma(E)}{E} = \frac{10\%}{\sqrt{E}} \oplus \frac{0.38\ \text{GeV}}{E} \oplus 0.3\% \tag{3.1}$$

which well suites the requirements of the physics.

The hadronic end cap calorimeter

Each one of the two Hadronic End-Cap consists of two independent wheels of outer radius 2.03 m, for a total coverage of $1.5 < |\eta| < 3.2$. Both wheels consist of an array of copper plates (their thickness is 25 mm in the first wheel, 50 mm in the second). The gap between the plates (8.5 mm) is split by three electrodes into 4 drift spaces of 1.8 mm. The readout electrode is the central one, while the side ones are HV carriers. The scheme is shown

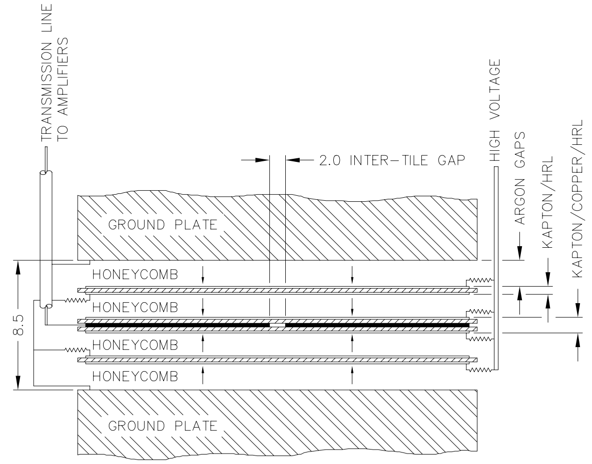


Figure 3.4: Structure of the readout gap in the Hadronic End-Cap.

in figure 3.4.

Each of the two wheels is composed by 32 identical modules, assembled with fixtures at the periphery and a central ring. Each wheel is divided into two longitudinal segments.

Primarily in order to limit the capacitance seen by a single preamplifier, and thus to allow for a faster response, only two gaps are ganged together at the pad level. Miniature coaxial cables running between the sectors carry signals to the pre-amplifiers boards located at the wheel periphery. Output signals from (typically) four pre-amplifiers are summed together on the same board. A buffer stage drives the output signal up to the cold-to-warm feedthroughs. Cells defined in this way are fully projective in azimuth, but only pseudo-projective in η. However, the detector envelope is cylindrical, for sake of mechanical simplicity. To minimize the dip in the material density at the transition between the end-cap and the forward calorimeter (around $|\eta| = 3.1$), the end-cap EM calorimeter reaches $|\eta| = 3.2$, thereby overlapping the forward calorimeter.

The HEC standalone resolution for single pions (6-200 GeV) was measured at the test beam. The results [53] is:

$$\frac{\sigma(E)}{E} = \frac{(70.6 \pm 1.5)\%}{\sqrt{E}} \oplus (5.8 \pm 0.2) \tag{3.2}$$

The e/h factor was obtained as well from test beam measurements of the e/π ratio. The result is $e/h = 1.5$.

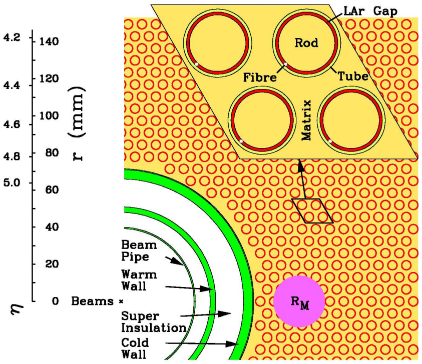


Figure 3.5: Sketch of the structure of the Forward Calorimeter.

The forward calorimeter

Because of the high level of radiation it has to cope with, the forward calorimeter (FCAL) is built using the intrinsic radiation-hard LAr technology. It is integrated in the forward cryostat together with the EM end-cap and the HEC, and its front face is at about 4.7 m from the interaction point.

In order to minimize the amount of neutron albedo in the inner detector cavity, the front face of the FCAL is recessed by about 1.2 m with respect to the EM calorimeter front face. This severely limits longitudinal space for installing about 9.5 active interaction lengths, and therefore calls for a high-density design, which also avoids energy leakage from the FCAL to its neighbors.

The FCAL consists of three sections: the first one is made of copper while the other two are made of tungsten. In each section the calorimeter consists of a metal matrix with regularity spaced longitudinal channels filled with concentric rods and tubes. The rods are at positive high voltage while the tubes and the matrix are grounded. The LAr in the gap is the sensitive medium. This geometry allows for an excellent control of the gaps which are as small as 250 μm in the first section. Figure 3.5 shows a sketch of FCAL structure.

3.1.4 The Tile calorimeter

The central part of the hadronic calorimetry [54] differs from the rest of the calorimetry because it does not use LAr as active material. Instead scintillating tiles within an iron structure are used. The structure is periodical along z and the tiles are oriented perpendicular to the beam axis. The tiles thickness is 3 mm while, for each period, the total iron thickness is 14 mm.

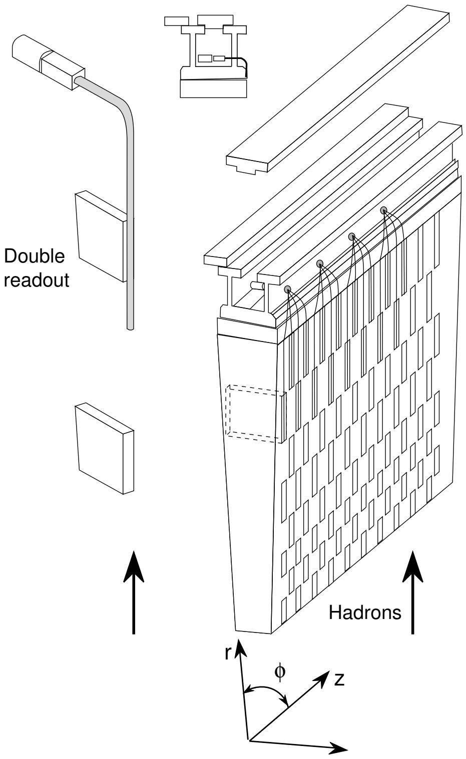


Figure 3.6: Optical layout of the Tile calorimeter

The tiles are read out by two wave length shifting (WLS) fibers, one for each side. The WLS fibers are grouped together in order to reach the desired granularity: three longitudinal samples, with a lateral segmentation of $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ in the first two samples,$\Delta\eta \times \Delta\phi = 0.2 \times 0.1$ in the last one. The light signal is readout by PMTs located, with the front-end electronics inside the *girder*: the iron structure which holds each azimuthal sector at the outer radius of TileCal. The girder provides also sufficient magnetic shielding for the PMTs.

The light emitted inside the tiles in the UV range is shifted to visible violet light by two tiles components: paraterphenyl (PTP) and POPOP. The wavelength of the fibres has been chosen so that to maximize the efficiency of the collection and detection of the light signal. A light mixer at the end of the fiber bunch makes the readout of the light more uniform over the PMT photocatode.

A sketch of the optical layout of the TileCal is given in figure 3.6.

TileCal is subdivided into one barrel region ($|\eta| < 1$) and two extended barrels ($0.8 < |\eta| < 1.7$) one on each side of the barrel. The gap between them (68 cm) accommodates the services for the inner detector and the front-end

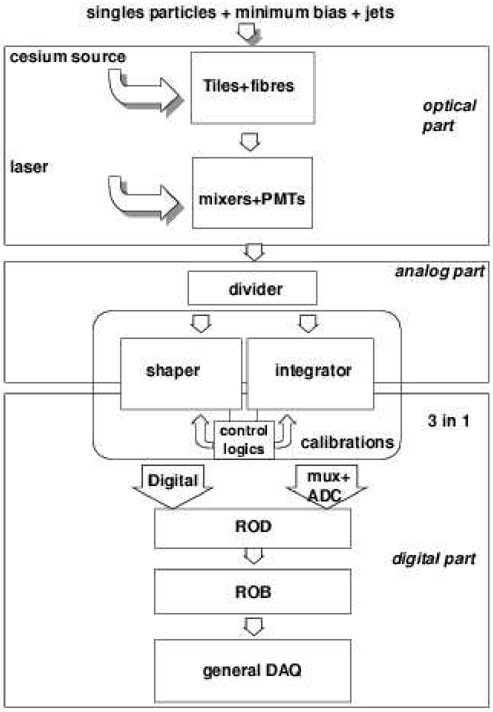


Figure 3.7: General Scheme of the Tile Calorimeter read-out.

electronics of the EM calorimeter. Both the barrel and the two extended barrels are subdivided in 64 modules, one for each ϕ slice ($\Delta\phi \simeq 0.1$). The extended barrel has the same azimuthal segmentation as the barrel, while the longitudinal segmentation differs in the second and third layer.

The gap between the barrel and extended barrel is partially instrumented by the Intermediate Tile Calorimeter (ITC). It is composed of two radial sections attached on the face of the extended barrel made, as the rest of TileCal, of scintillator in an iron matrix. The outer section, 31 cm thick, starts at the outer radius and covers 45 cm in radius. It is followed by the inner section which is 9 cm thick and extends over 45 cm to lower radii. The ITC is extended further inward by a scintillator sheet, covering the inner part of the extended barrel and extending to the region between the LAr barrel and end-cap cryostat over $1.0 < |\eta| < 1.6$. This scintillator samples the energy lost in the cryostat walls and dead material. It is segmented into three sections each covering a range of $\Delta\eta \simeq 0.2$.

The front-end electronics of each ϕ wedge of TileCal is placed inside the girder.

PMTs and front-end electronics are assembled in a modular structure called

drawer. A drawer houses 48 PMTs and their electronics: the digitizers, the interface cards and the ancillary components. Fibers and PMTs are light coupled with an air gap of about 1 mm. This mechanically decouples PMTs and drawers from the rest of the calorimeter for a fast replacement if needed.
Each PMT is equipped with a 3-in-1 card that provides three basic functionalities: pulse shaping and accommodation of the needed large dynamic range, charge injection calibration, and slow integration of the PMT signals for monitoring and calibration. The shaped analog signal is sent to two different amplifiers (whose relative gain is 64) and, at the same time, to the analog trigger summation board from where they reach the LVL1 signal on dedicated cables.
After the amplification the signal is sent to the *digitizers boards* for digitization. In each drawer there are 8 of such boards. Each one provides a double sampling (one for each of the two gains) using two separate 10.bits fast ADCs. The sampled signal is then stored in pipelines, waiting for the LVL1 trigger decision.
Once the LVL1 signal is received for one event, the corresponding 7 samplings are extracted from the pipelines and send to the RODs, where the signal is processed using Optimal Filtering techniques.
Figure 3.7 shows a scheme of the front-end electronics for the Tile Calorimeter.
The calibration system is also shown in the picture. It consists of three separate calibration streams.

- A hydraulic system can drive a ^{137}Cs source through each tile row, allowing a current measure of the response of all the optics and PMT chain. The readout is performed via a dedicated slow integrator on the 3-in-1 cards. In this way it is possible to check the calorimeter response (optics, PMTs) independently of the electronic chain.

- The laser system sends short light pulses directly to the PMTs. A system of clear optical fibres is used to split the signal between all the PMTs of each module. The laser calibration checks the linearity of the system and its stability. It will be used also to verify the equalization of the acquisition chain from PMT to electronics boards, and to set the timing of the PMTs inside a single drawer.

- Finally, the Charge Injection System (CIS) injects a well known amount of charge directly into the 3-in-1 card, and, thus, to check the stability of the electronics chain itself and to verify the gain of the two amplifiers.

The slow integrator of the 3-in-1 cards will also be used to read, independently from the digital data stream, the integrated signal over many bunch crossings. The relative stability of the curent monitored in this way from each TileCal cell is an indipendent check of the stability of the system and can be used to monitor the relative luminosity of the LHC.

3.2 Calorimeter response to hadrons

The raw response of hadron calorimeters needs a few corrections before it can be used in the analysis. These corrections concern the energy scale. The main effect to deal with is non linearity of the hadronic scale, that is a general problem when considering intrinsic non-compensating calorimeters as in ATLAS.

When a single hadron impinges on a material, a hadronic interaction can occur, giving rise to a hadronic shower. This shower is mainly composed by charged and neutral pions.

The π^0s decay into photons originating electromagnetic showers. Therefore, a hadronic shower, is composed of a pure hadronic component and a electromagnetic one.

The response of a calorimeter to the non electromagnetic component is smaller than the electromagnetic one due to several effects: the invisible energy, that is mainly the energy spent in breaking a target nucleus, invisible to the calorimeter, the decays of escaping muons.

The relative amount of energy carried by π^0s and the invisible energies fluctuates from event to event. Therefore, in general, calorimetric response to hadrons is smaller and broader with respect to that of electrons or photons of the same energy.

The calorimeter response to the pure hadronic component of a shower can be considered constant as a function of energy, but since the number of π^0 produced, and thus the EM fraction, is energy dependent, the response to single hadrons is non linear unless $e/h = 1$.

We can understand with a simple model how the fraction of electro-magnetic energy in a hadronic shower depends on the energy of the primary particle [55].

A hadronic shower has a rather complicated development and we can only describe it in an oversimplified way.

We shall describe the hadronic shower as made of n generations of interactions. In each interaction the interacting primary particle produces, in average, $< m >$ secondaries. The produced π^0 will undergo an electromagnetic cascade and their energy is considered electromagnetic component. The available hadronic energy is therefore reduced by the fraction of energy carried by the π^0. Let us call f_{π^0} the fraction of produced neutral pions, on average, in each interaction. If only pions would be produced, $f_{\pi^0} = n_{\pi^0}/ < m >= 1/3$ for isospin invariance, since also other particles are produced 1/3 is only an upper limit for f_{π^0} (Monte Carlo simulations sug-

gest values closer to 1/4). We shall assume that f_{π^0} and $< m >$ are constant along the shower development ($< m >$ depends only logarithmically on the parent's energy).
At the first interaction the fraction of hadronic component in the shower is:

$$f_h^1 = 1 - f_{em}^1 = 1 - f_{\pi^0} \tag{3.3}$$

At the second interaction the available hadronic component is further reduced of a factor $(1 - f_{\pi^0})$:

$$f_h^2 = (1 - f_{\pi^0}) f_h^1 = (1 - f_{\pi^0})^2 \tag{3.4}$$

At the n-th generation we will indeed have:

$$f_{em} = 1 - f_h^{(n)} = 1 - (1 - f_{\pi^0})^n \tag{3.5}$$

where f_{em} is the fraction of electromagnetic component (fraction of neutral pions) in the shower. Since we consider f_{π^0} independent of the energy, the energy dependency for f_{em} originates from the energy depenedence of the number of generations n.
T. A. Gabriel [56] deduced with a heuristic derivation the energy dependence of the electromagnetic component in a hadronic shower for a high energy ($E > 10$ GeV) primary hadron:

$$f_{em} = 1 - \left(\frac{E}{E_0}\right)^{k-1} \tag{3.6}$$

$$k = \frac{\ln 1/(1 - f_{\pi^0})}{\ln < m >} \tag{3.7}$$

where E_0 is a scale factor (typically E_0, that must be extracted from data, ranges from 1 GeV to 3-4 GeV) and can be considered as the extrapolated energy at which the cascade is entirely hadronic ($f_{em} = 0$ in equation 3.6), i.e. an effective turn-on energy for π^0 production.
For $f_{\pi^0} = 1/3$ and typical multiplicity $< m >= 5$ we obtain $k = 0.75$, with $f_{\pi^0} = 1/4$ we obtain $k = 0.82$. Changing to $< m >= 10$ and $f_{\pi^0} = 1/4$ we get $k = 0.87$. Hence for quite wide ranges of the parameters $< m >$ and f_{π^0}, the parameter k gets values ranging between 0.75 and 0.88.
The energy response to a pion can be written as:

$$\pi = f_{em} e + (1 - f_{em}) h \tag{3.8}$$

with e (h) is the calorimeter response to the electromagnetic (hadronic) component of the shower. This leads to:

$$\frac{e}{\pi} = \frac{e/h}{1 - f_{em}(1 - e/h)} \tag{3.9}$$

Since f_{em} is a function of the energy, e/π depends on the energy and introduces a nonlinearity in calorimeter response to hadrons (unless the intrinsic calorimeter parameter e/h equals 1). The ATLAS calorimetric system is non-compensating and a strategy to correct for the non linearity must be developed when measuring energy of hadronic objects.

In chapter 5 a method to correct for non-compensation will be presented and discussed, while the rest of this chapter we will discuss in detail the response of a module of TileCal to single particles.

The test beam operations are also simulated with detailed Monte Carlo programs. Test beam geometry is fully described as well as most of the physics processes that control the interactions of the particles with the calorimeter. As essential part of this analysis will concern the comparison between data and Monte Carlo. In the following chapters it will become clear why a Monte Carlo which describes well the performances of the calorimeter is essential to develop the tools to calibrate hadronic showers and to improve the performance of the calorimeter.

3.3 Results from TileCal stand alone test beams

3.3.1 Test beam setup for the tile calorimeter

During the period 2001 - 2003 an extensive test beam program has been carried out at the CERN SPS H8 beam line exposing 9 Barrel and 14 Extended Barrel modules of the Tile hadron calorimeter to electron, pion and muon beams of energies ranging between 10 and 180 GeV. In addition, during August 2003 data run, special Very Low Energy (1-9 GeV) beam data (VLE) has been collected.

From April 2004 a Combined Test Beam period has been also carried out: a complete slice of ATLAS experiment has been assembled and all the subdetectors took data together. Results from the Combined Test Beam will be presented in next chapter.

The results reported here refer to data taken during 2002 and 2003 periods in which TileCal took data in stand alone mode. Two extended barrel modules, one barrel module and the barrel module zero (see figure 3.8) were exposed to pions and electrons of energies between 20 and 180 GeV.

The modules were mounted on a rotating table to expose the calorimeter to different beam directions.

The beam line was instrumented with a set of ancillary detectors: one threshold Cerenkov counter used for beam particles identification, four delay-line wire chambers providing the coordinates in the plane transverse to the beam, and scintillating counters to build a trigger signal. A set of scintillators (muon wall) was placed just after TileCal to separate electrons and pions from passing muons, contaminating the beam.

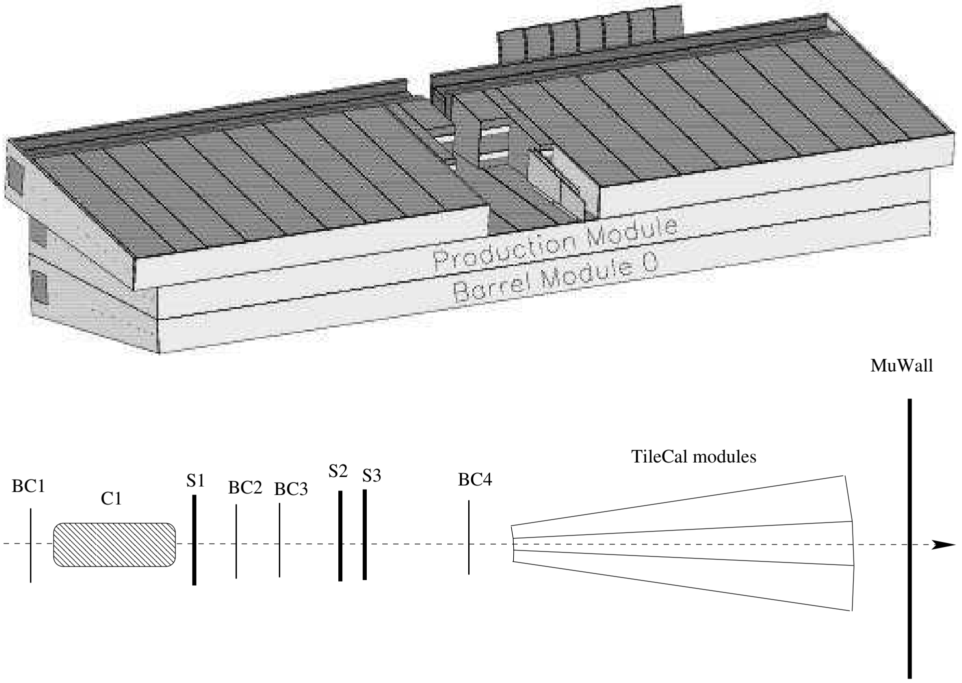


Figure 3.8: Schematic view of the ATLAS TileCal test beam set-up. In the upper plot the Geant4 simulation is shown with the two extended barrels on top of the module 0 and one barrel module. The Muon Wall is also visible in the back of the modules. In the bottom picture a schematic side view of the beam line is presented. C1 is a threshold Cerenkov counter, BC1-BC4 are the multi wire beam chambers to monitor the beam position, S1, S2 and S3 are the scintillators used to form the trigger. Downstream of TileCal modules, scintillators, used to identify the muons, are shown (Muon Wall).

Particle selection

The wire chambers are used to select the particles impinging the fiducial region at the center of a cell of the central module.

Pions and electrons are separated using the Cerenkov counters, which provide electron separation from pions up to $\sim 50\ GeV$. When the Cerenkov information is insufficient or unavailable, particles are selected exploiting the longitudinal shower profile.

Muons are identified using the muon wall information. When this is not possible (due to the limited coverage of the muon wall), the longitudinal shower profile is used again to identify muons. Events with a MIP in TileCal longitudinal samples are identified as muons.

Electron beams are used to calibrate the calorimeter modules. Mean charge deposit is normalized to beam energy obtaining the calibration factor pC/GeV. This factor is applied also to pion response, obtaining the hadronic response

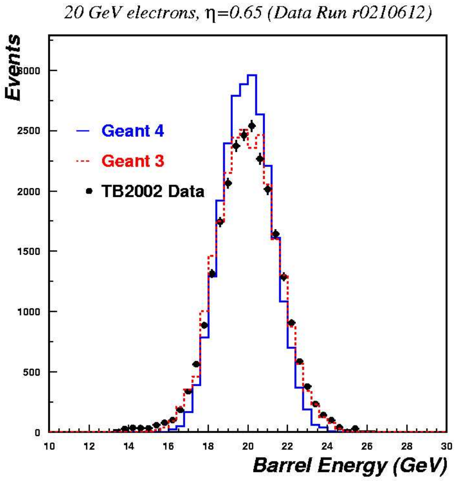


Figure 3.9: Distributions of electron response with (right) and without (left) adding the photo-statistic and electronic noise contributions.

at the electromagnetic scale.
Before proceeding with a discussion of the results obtained at the test beam, we describe the tools used to simulate the performances of the calorimeter. Indeed one of the outcomes of this work is to understand to what extend the simulation describes the performances of TileCal. As we will see later in this work the simulation will be needed to correct for the energy scale (hadronic) and for linearity of the ATLAS calorimetric system.

Geant4 peculiarities for TileCal test beam

A special setup for Geant4 (FADS/Goofy) [57] has been developed for the test beam configuration.
Geant4 [58] is a C++ set of libraries and utilities to develop MC application and simulate the interaction, transport and digitization of particles in the matter. Geant4 is provided with a single particle generator and is interfaced to the physics generators (PYTHIA, MC@NLO, GENSER). In the test beam simulations the single particle generator is used. FADS/Goofy is the concrete implementation of the simulation code. The geometry of the setup is described and materials associated to volumes. It also provides a system to control general parameters using a macro file (like direction of beam line, beam energy, particle type).
FADS/Goofy is responsible to collect the Monte Carlo hits and build a data structure according to the cell structure of modules.
To take in account the saturation of the scintillating tiles used in TileCal the Birk's law [59] is applied to energy deposited in the cells before computing the light response.

Geant4 simulation of physics processes in action during the shower development is implemented using the so called physics-lists. The individual processes can be independently turned off and on.

FADS/Goofy uses as electromagnetic physics the Geant4 standard package, derived from the well known and validated Geant3 and EGS4 [60] package, while for what concerns the hadronic physics, we exploited mainly two out of the many available physics lists: LHEP 3.6 and QGSP 2.7. These physics lists seems to reproduce better the processes in action in the hadronic shower development.

In the first list the hadronic inelastic scattering cross sections are parametrized from real data (mainly the LEP and HEP parametrized models) [61]. The second list, QGSP, instead uses a theoretical model (the quark-gluon string model) to compute the cross sections of the reaction of energetic pions, kaons and nucleons [62]. In all the studies performed so far, results obtained from both physics lists have been compared to real data.

Electronic noise and the photo-statistic fluctuations are not yet included in FADS/Goofy simulation and have been added in the analysis step. Both have been measured from the data and added to the simulated response as a gaussian source of noise. For the photo-statistic fluctuations this is a reasonable approximation due to the high number of photo-electrons produced (54 pe/GeV deposited in each cell).

The comparison of Monte Carlo predictions to data has been made first applying the photo-statistic fluctuations to energy deposited in active material: for every event an additional term have been added distributed like a gaussian with $\sigma = \sqrt{E(\text{GeV}) * 54}$ and zero mean.

The data are then calibrated to the e.m scale: the mean energy deposit for electrons in the calorimeter have been made equal to the beam energy. This procedure is the same as with real data.

The noise is then added: a further gaussian term is added distributed according to the noise distribution obtained from test beam data looking to out of the beam signal.

This last point has been showed important to compare the Monte Carlo to data as is clearly visible in figure 3.9. The response to a 20 GeV electron beam is shown together with the Geant4 prediction. Electrons impinges at $\eta = 0.65$ in the central module. In the left plot, the data are superimposed to Monte Carlo predictions without noise contribution and photo-statistic effect, it is visible the poor agreement obtained, while adding the two contributions the Monte Carlo well describes real data (right plot).

Geant3 predictions have been analyzed with the same method and are also shown for comparison.

In the left plot Geant4 predictions show a poor agreement with data respect to Geant3 predictions. However after including the noise and photo-statistic the Geant4 predictions better describe the distribution of the data shape. The same distributions have been studied for different energies and different

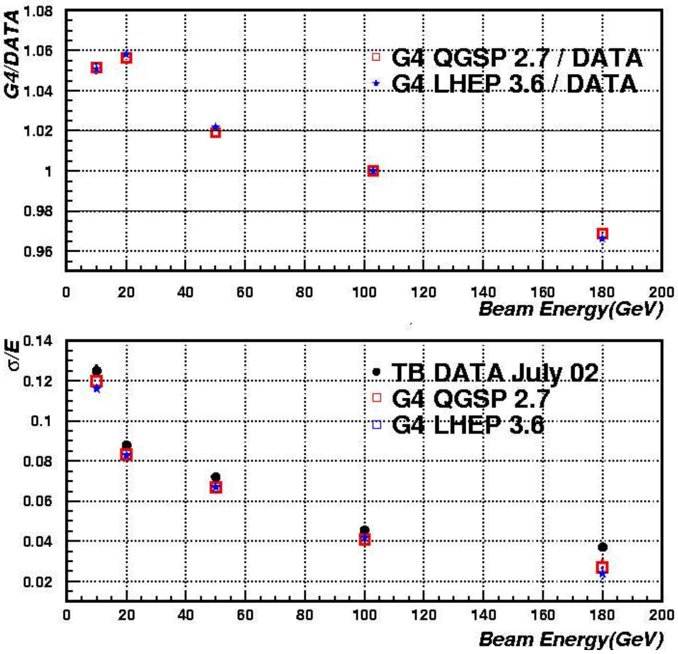


Figure 3.10: Top: Ratio between Geant4 and test beam data predictions for linearity Bottom: resolution. The points refer to electrons from 10 to 180 GeV at $\eta = -0.65$. The calibration constants are normalized to 100 GeV point.

impinging directions showing the same result. We can therefore conclude that Geant4 is in better agreement with data once that all the read-out effects have been taken into account.

A small discrepancy in the low energy tail between Monte Carlo and data is visible. This is probably due to the incomplete description of the material in front of the calorimeter setup in the Monte Carlo. In fact in the Monte Carlo the ancillary detectors and the beam line setup are not fully simulated.

For this study this discrepancy is a small effect. In fact the quantities that are used in the analysis (mean value and σ of distributions) are obtained with a gaussian fit of the signal peak, excluding the low-energy tails.

The analysis of simulated data is done in the same way as for the test beam data: simulated electrons are used to set the electromagnetic scale and the calibration factors are used also for pions.

3.3.2 Results for electrons

Linearity

In the test beam the energy deposited in the active material (scintillator) is converted into light, taken to the PMTs and transformed in measured charge (pC). The gains of PMTs have been set in order to have 1.2 pC per GeV deposited in the calorimeter (both in the active and material and in the absorber).

The mean calorimeter response to electrons is used to compute the electromagnetic energy scale. The charge measured in the active material is normalized to the beam energy, obtaining the calibration factor pC/GeV, which is expected to be energy independent. The linearity obtained from test beam data is better than 2% over a wide energy range [63] [1].

In Geant4 the energy deposited (measured in GeV) in the active material is normalized to the energy of the incident simulated electron, obtaining the calibration factor $\mathrm{GeV}_{active}/\mathrm{GeV}_{beam}$.

The ratio between the Geant4 and data electron calibration factors obtained at $\eta = -0.65$ are shown in the upper plot of figure 3.10 as a function of the beam energy. The data have been normalized at the 100 GeV point. As expected no differences are visible between the two physics lists as they differ only for the hadronic part.

The agreement between data and simulation is about ±2% above 20 GeV, linearity is consistent with the results obtained during different test beam periods.

A substantial worsening is visible below 20 GeV. In particular the response in data is lower than expected producing an enhancement in the Geant4/DATA ratio. The discrepancy found for these energy values is probably related to instrumental effect: for cell energy deposits around 20 GeV the high-gain readout amplifier saturates, this is the case for electron beams of energies 10 and 20 GeV. The different response is indicating that a more detailed study of this "transition" region should be performed, the inter-calibration of the bi-gain readout should be studied in more details. In addition to that, energy is reconstructed from energy deposits in all cells of the system. The preliminary energy reconstruction method (flat-filter) used for this data-set can introduce a bias in the energy measurement that can become not negligible at low energies.

The re-processing of the entire data-sets with more sophisticated energy reconstruction algorithms (fit method and optimal filtering, now standards tools), will be performed also on 2002 data to better understand this effects. Work in this direction is in progress.

Resolution

In the lower plot of figure 3.10 the measured and predicted resolution are shown. As it can be seen a good agreement between data and simulation has been obtained for both the physics lists. This level of agreement has been obtained once the noise and the photo-statistic effect have been included in

---

[1] Possible reasons to have an energy dependence effect are saturation of scintillators (Birk's effect), variation of the sampling fraction, geometry effects.

Data Set	a (GeV$^{1/2}$)	b
QGSP	$(34 \pm 1)\%$	$(0.9 \pm 0.2)\%$
LHEP	$(34.5 \pm 0.8)\%$	$(0.7 \pm 0.1)\%$
TB Data	$(35 \pm 1)\%$	$(1.1 \pm 0.1)\%$

Table 3.2: Values for the sampling and constant term for the electromagnetic resolution. Values are obtained from electrons from 10 to 180 GeV impinging at $\eta = -0.65$.

Data Set	a(GeV$^{1/2}$)	b
QGSP	$(58 \pm 3)\%$	$(2.9 \pm 0.5)\%$
LHEP	$(59 \pm 3)\%$	$(2.4 \pm 0.5)\%$
TB Data	$(53 \pm 2)\%$	$(2.0 \pm 0.2)\%$

Table 3.3: Values for the sampling and constant term for pions. Values are obtained from pions from 10 to 180 GeV impinging at $\eta = -0.65$. For VLE the data refer to pions from 1 to 9 GeV at $\eta = -0.35$.

the simulation.
The resolution scales, as a function of energy, as $1/\sqrt{E}$ and hence the test beam data and the simulation have been fitted using the formula:

$$\frac{\sigma}{E} = \frac{a}{\sqrt{E}} + b \tag{3.10}$$

In table 3.2 the results of the fit are quoted.
The sampling term obtained from data is $(35 \pm 1)\%$ in agreement with the simulation, also in agreement is the constant term b.
As expected, when considering electrons, the results obtained from the two physics list, that differ only in the hadronic part, are in agreement within error bars with each other.

3.3.3 Results for Pions

Resolution

The same analysis has been repeated also for pions. Once the calorimeter modules have been calibrated to electromagnetic scale, the response to pions has been measured and compared to Geant4 predictions. The pion energy distribution is not perfectly gaussian, this is due to events in which the hadronic component over-fluctuates or, for high energy beams, is due to leakage. This problematic events are excluded from the fit. A gaussian fit is performed in a restricted range: a first fit is made to obtain a $\pm 2\sigma$ range inside which a second gaussian fit is done to extract the needed parameters.

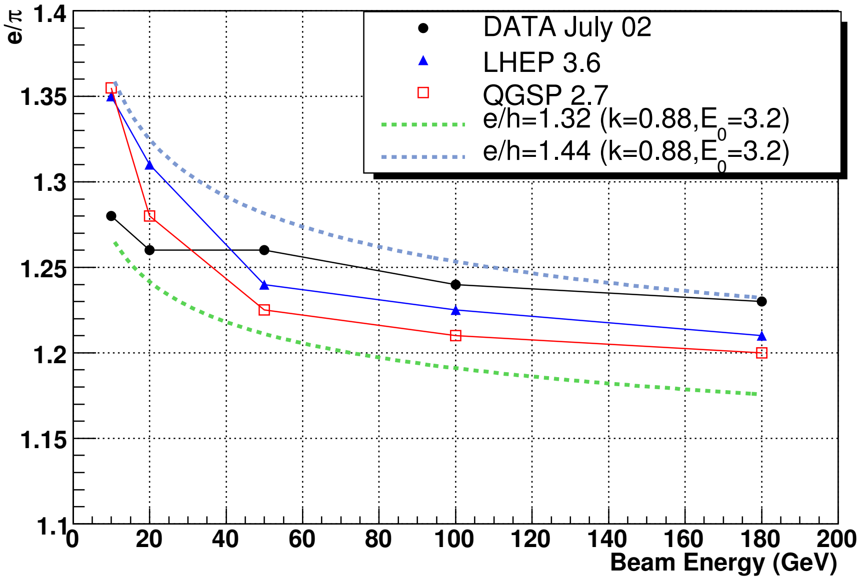


Figure 3.11: e/π ratio for pions (10-180 GeV) impinging at $\eta = -0.65$. Superimposed are Geant4 predictions for both physics lists. The two dashed lines represent the e/π values obtained from equation 3.9 for two different values of the e/h constant.

In table 3.3 the results from the fit are quoted. Geant4 simulation predicts a broader distribution compared to data (data and Geant4 predictions are anyway within a 2 σ interval), the two physics lists do not present considerable differences. The discrepancies between data and Monte Carlo are due to the less precise description of hadronic physics with respect to the e.m one.

e/π ratio

The e/π ratio has been obtained as the ratio of the response to electrons and pions at the same energy ($e/\pi = R_e/R_\pi$ with $R = Q$ for test beam data and $R = E_{active}$ in Geant4 simulation).
Results for experimental data and the simulation are shown in figure 3.11. Data and Geant4 predictions are within $2-3\%$ above 20 GeV, while for lower energies the agreement is worse. This problem has the same origin as the one discussed for electrons. For these two energies the response to electrons is lower than expected leading to a smaller e/π compared to Geant4 predictions.
In the same figure the expected value of the e/π ratio is also reported. Equation 3.9 has been used with two different choices of the e/h parameter.

Upper dashed line has been obtained with $e/h = 1.44$, while the bottom dashed line with $e/h = 1.32$. The other parameters are fixed to $E_0 = 3.2$ GeV and $k = 0.88$. The experimental and simulation points lie between the two values, an e/h value between 1.32 and 1.44 is compatible with the 1.34-1.36 value obtained on prototype modules [54].

Shower Profile

Using TileCal segmentation a coarse shower profile study can be performed. The fraction of energy deposited in each single longitudinal sample is plotted as a function of the beam energy in the first three upper plots of figure 3.12. Geant4 predicts longer showers with higher energy deposit in the second and third layers, This is true mainly for LHEP (points marked as stars). The fourth plot represents the transverse shower profile: the fraction of energy deposited outside the central module as a function of pion initial energy. Contrary to longitudinal shower profile, the simulation predicts more compact showers with less energy deposit in nearby modules. In the longitudinal profile QGSP list better reproduces the data, while in the transverse profile is LHEP that simulates better TileCal response. This is the only case, studied so far, in which the two physics lists have a significantly different behavior.

The correct simulation of shower shape (mainly the longitudinal shape) is particular important to develop realistic weighting methods for jet energy reconstruction. TileCal, an intrinsic non compensating calorimeter, can use the longitudinal segmentation to recover for $e/h > 1$. In this aspect more work is needed before being confident of Geant4 simulation.

Conclusions

The TileCal test beam data collected during 2002 have been analyzed to asses the quality of the data reconstruction and start a comparison with the Monte Carlo predictions.

The electron analysis has shown that to correctly compare data and Monte Carlo the electronic noise and the photo-statistic effect must be included in the simulation. Test beam results obtained over many periods show that an electron linearity of about 1-2% can be achieved over a wide energy range. We have compared the measured response to electrons to Monte Carlo simulation obtaining a good agreement for energies above 20 GeV. For lower energies the response is lower than expected and a more detailed treatment of the data should be performed as soon as new and more precise reconstruction strategies are made available.

Excluding the low energy regime Geant4 can predict reasonably well the electrons impinging in TileCal and we can be confident of the simulation.

The situation is complicated in the hadronic sector. Single pions can be

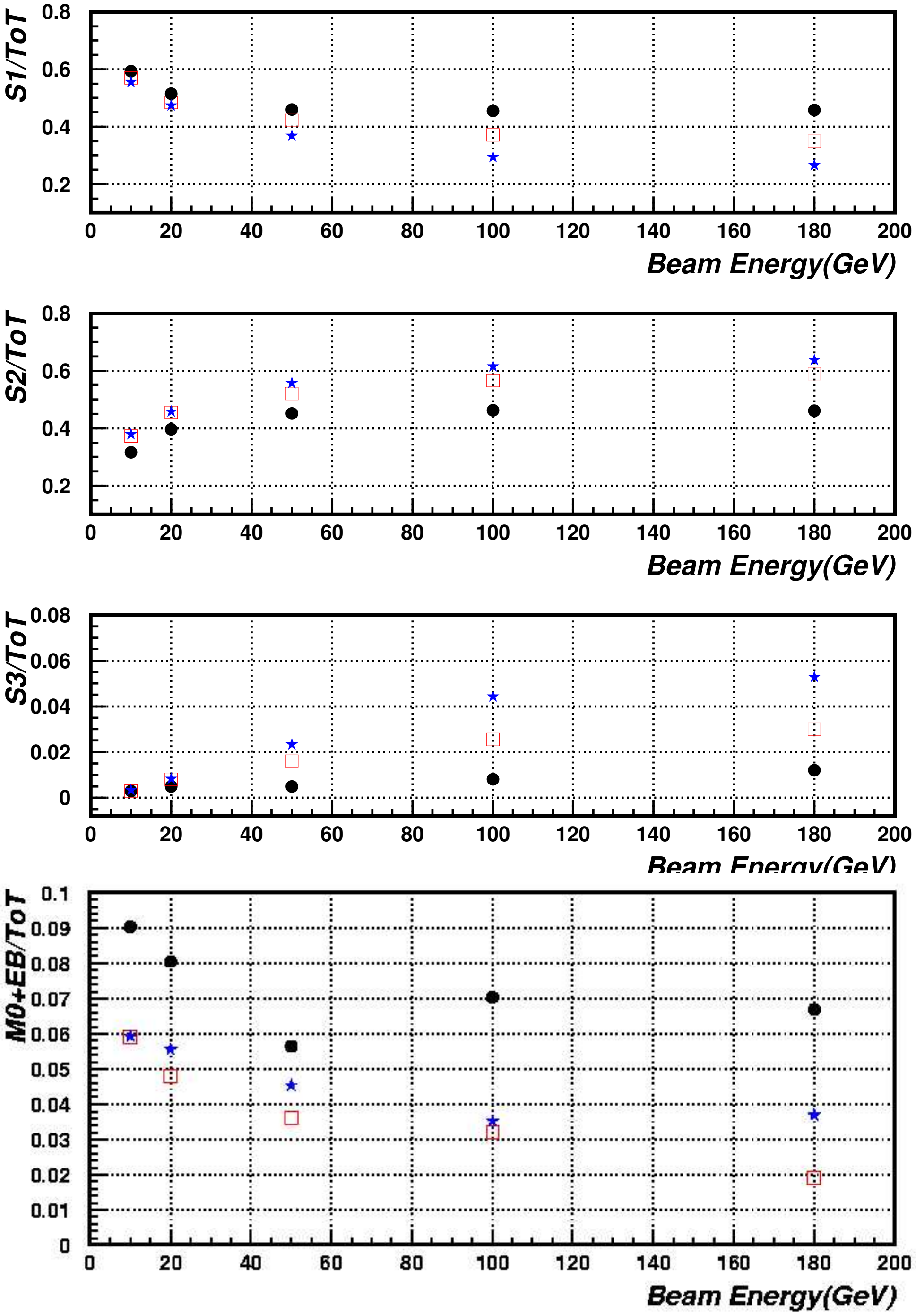


Figure 3.12: Longitudinal shower profile (first 3 plots) for pions (10-180 GeV) at $\eta = -0.65$. The partial transverse profile (last plot) is the ratio between the energy deposit out of the central module over the total energy deposit.

selected in data and compared to Monte Carlo predictions. Geant4 comes with different models to simulate hadronic interactions. This models and their parameters are collected in the so called physics lists. We concentrated on the two that are expected to predict better the data.

The agreement between resolution measured from real data and from simulation can be considered satisfying, while the e/π ratio is predicted at the level of $< 5\%$ again for energies above 50 GeV. In fact the poor agreement obtained at 10 and 20 GeV is due to the lower response of electrons.

It is important that Geant4 can predict the hadronic shower development since many hadronic calibration schemes in ATLAS use the calorimetric longitudinal segmentation to achieve compensation (see chapter 5).

We compared predictions with data obtained from single pions. Both physics list fail, at the level of some percent, to predict data. Between the two, QGSP is better simulating the longitudinal profile (while LHEP does better for the lateral profile).

Also using the results presented here the ATLAS calorimetry simulation has been tuned and a more precise version of the Geant4 QGSP physics list has been chosen for the new ATLAS simulations that will be discussed in the next chapter.

Chapter 4

Towards Geant4 hadronic physics validation

The stand alone test beams program performed by the ATLAS calorimeters in the past have been used to assess the electromagnetic scale and to study the detector performance in stand alone mode. In 2004 a combined test beam program has been carried out at CERN allowing, for the very first time, all the ATLAS detector to take data together. In addition to this, this data are of fundamental importance, especially for the calorimetric system, to assess the combined performance and to start a detailed comparison of the Geant4 predictions. Many hadronic calibration scheme indeed use Monte Carlo simulations to extract calibration constants for single hadrons and jets. It is therefor extremely important to have a simulation code that can describe the detector as precisely as possible.

In this chapter results from the 2004 ATLAS Combined Test Beam period will be presented. Preliminary comparison between data and simulation will be discussed for single pions at different impinging energies. After a first introduction to the experimental setup, the peculiarities of the Geant4 simulation program used for the combined test beam will be presented. The analysis procedure will be presented together with preliminary results for response, resolution and shower profile.

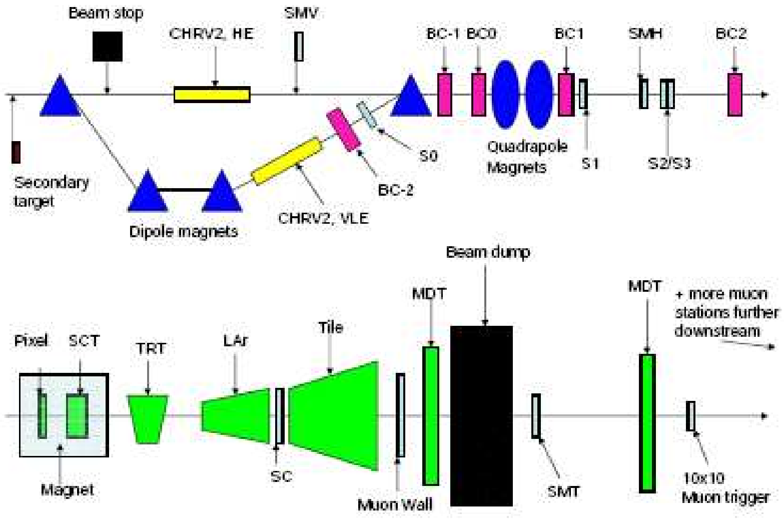


Figure 4.1: Schematic view of the beam-line instrumentation (top) and of the detectors setup (bottom). The meaning of the different acronyms is explained in the text.

4.1 The Combined Test Beam 2004

4.1.1 Introduction

From April to October 2004 an intense Combined Test Beam program [64] was carried out at H8 beam line at CERN for the central region of the ATLAS experiment. For the very first time all the ATLAS detectors took data together. Many components were the final or almost-final ones. The data acquisition software used in the CTB was the prototype of the final system [65]. A complete slice of the trigger system from the first hardware level of trigger (*LVL1*) [66] to the *High Level Triggers (Level 2 - LVL2 and Event Filter - EF)* [67] [68] were tested and used. Data were reconstructed with the ATLAS offline analysis framework software (*Athena*) [69].

The CTB had as main goals the detailed study of combined performances of the detectors with different particles at different energies and the technical test and performance study of the infrastructure hardware and software components (data acquisition software, trigger chain, reconstruction framework). It had also the goal to validate the ATLAS simulation software based on the Geant4 toolkit [58].

More than 90 millions of events were taken in different configurations of the detectors, magnetic field, quantity of passive material along the beam line and beam conditions. Electrons, muons and pions were impinging in different geometric configuration, with energies ranging from 1 to 350 GeV. Geant4 has been chosen as the ATLAS simulation software. It is completely embedded in the ATLAS analysis framework and this adapted to simulate the CTB setup with a detailed description of materials and geometry. The two main aspects to be validated in the simulation are the correct description of the detectors geometries and materials [70] and the correct simulation of the physics processes simulated by Geant4.

The CTB setup is implemented in Geant4 simulation *as-built*: the geometry and budget materials have been measured at the beam line and implemented, as precisely as possible, in the simulation code. Many studies have been performed to validate the geometry descriptions and showed that it is possible to validate the geometry description at a millimeter level comparing real data and simulations. We will not discuss these studies here and we will concentrate in the ability of Geant4 in reproducing the calorimeter response, comparing pions impinging at $\eta = 0.35$ and with energies from 20 to 350 GeV.

In particular we will present here the total energy deposit in the calorimetric setup, the measured resolution and the lateral and longitudinal partial shower profiles and compare the measured quantities and Geant4 predictions.

All these quantities are important to validate the description of the physics processes simulated in Geant4, as this software will be used in the final ATLAS detector simulation and in many calibration schemes of the ATLAS calorimetry (for example the ones based on the *Calibration Hit* tool [71], or the *Weighting methods* for jet calibrations), we will discuss later in this work these recalibration methods. The CTB data are ideal to perform a detailed comparison with the simulations as many components (both hardware and software) are almost the final ones and data were taken in a geometric configuration similar to the final ATLAS one.

4.1.2 Detectors setup

The ATLAS barrel CTB has been performed at the SPS H8 beam-line at Prevessin CERN site. The ATLAS detectors were positioned along the line with a geometry close to the final one. The beam line was also instrumented with many different ancillary detectors [72]: Cherenkov counters, scintillators, beam chambers. Two different lines were used, a secondary one used for energies over 10 GeV and a tertiary line used for energies from 1 to 9 GeV. In the following we will discuss only the data collected using energies above 10 GeV. The tertiary line data require special attention to control the beam contamination and and efficiently suppress the noise that, especially in the LAr calorimeter, can play an important role. The tools to reconstruct and analyze this data are still under development and analysis is ongoing.

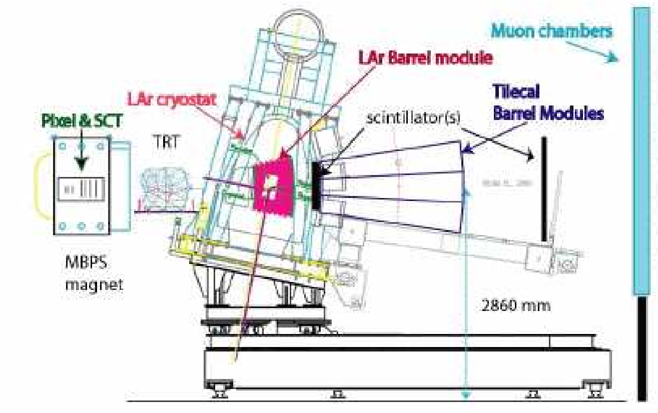


Figure 4.2: Schematic view of the calorimetric setup. A LAr module is equipped inside the cryostat visible in the figure. Three Barrel TileCal module (Module 0 and 2 production modules) follow the cryostat. Between the two calorimeters a scintillator is placed. All the system is positioned on a rotating table that allows to select the impinging direction of the particles.

General setup

In the figure 4.1 a schematic view of the CTB beam-line setup is presented. The beam enters from the left, the triangles represent the bending magnets that extract the beam to the tertiary line (Very Low Energy line) and the ovals represent the focusing quadrupole magnets. *CHRV1* and *CHRV2* are two Cherenkov threshold counters used to separate particles up to energies of about 50 GeV. The detectors marked with *BC* are multi-wire proportional chambers. These are used to monitor the beam-spot size and position. *S1, S2/S3* are scintillators that are used, in coincidence, to form a trigger signal, when the LVL1 system is not used. The beam-line is equipped with additional scintillators (*SMV, SMH, SC*), used to identify, respectively, high energy muons coming from the secondary line (when using the Very Low Energy beam line), out of beam-axis particles, and pions starting the shower only in TileCal. The *SMT* scintillator, positioned just after the beam dump, can be used to tag muons.
The detectors setup is shown in the bottom scheme of figure 4.1. The silicon inner detectors (Pixel and SCT) are placed inside the *Morpurgo* magnet, followed by the TRT system. The calorimetric system consists of a module of the liquid Argon detector (*LAr*) [51] and modules of the Tile hadronic calorimeter (*TileCal*) [54]. The calorimetric setup, that will be discussed in some more details in the next session, is followed by the muons chambers (only the first one is shown in the diagram).

Calorimetry setup

In figure 4.2 a detailed view of the calorimetry setup is shown. The LAr module is placed inside the cryostat. The LAr module is sub-divided into four longitudinal layers: the Pre-Shower (*PS*) used to recover the energy lost, mainly by electrons and photons, in the up-stream material, the *Strips*, highly segmented in eta, the thick *Middle* layer and the *Back* layer. Downstream of the cryostat a scintillator (SC) is placed to select pions starting the shower in TileCal.

Three central Barrel modules of TileCal are exposed to beam: the *Module 0* and two production modules.

A set of twelve scintillators (*Muon Wall*) follows TileCal to measure the punch-through and the muon contamination.

The calorimetric system is placed on a rotating table to expose the setup to different impinging directions of the beam. In the following analysis the data at $\eta = 0.35$ are considered. This value of the pseudorapidity has been chosen since in this geometry hadronic showers are well contained [1].

For TileCal crack studies ($0.7 < \eta < 1.2$) three *Extended Barrel* (*EB*) modules where placed near the rotating table to mimic the final ATLAS setup. In this configuration an electromagnetic iron-scintillators calorimeter (*Phantom*) was placed in front of the EB modules to simulate the presence, in the ATLAS configuration, of the electromagnetic End-Cap LAr calorimeter.

4.1.3 ATLAS Geant4 Simulation

Geant4 is an Object Oriented set of libraries and utilities to develop Monte Carlo applications and simulate the interaction, transport and digitization of particles in matter. ATLAS has chosen this environment to develop its Monte Carlo simulation. Also the Combined Test Beam setup was implemented in Geant4. All the detectors and the ancillary detectors are described in Geant4, geometry has been measured in the real setup and included in the simulation code.

The output of the simulation program can be reconstructed in a way very similar to the real data.

In this section we will describe in some details the general procedure of the simulation.

The detector geometry is described specifying the volumes of which the detector is composed, the materials they are composed of and the sensitive parts (for example the scintillating tiles for the central hadronic calorimeter and the LAr for the EM calorimeter).

Geant4 is provided with a single particle generator and is interfaced to the

---

[1] The negative η side of the LAr calorimeter was not equipped with read out electronics while TileCal barrel coverage arrives up to $\eta = 1$, $\eta = 0.35$ is therefor a good compromise.

physics generators (PYTHIA, MC@NLO, GENSER). In the test beam simulation the single particle generator is used.
After providing to Geant4 engine the description of the setup the simulation steps are the following:

1. *Generation*: the impinging particle is generated monochromatic in energy but with a gaussian smearing in the transverse component of the momentum. A gaussian smearing is also applied for the starting position in the transverse plane. This smearing allows for a detailed simulation of the real beam profile. The tracking of the particles in the different materials, the production of secondaries and the energy loss simulation is done in the next step.

2. *Simulation*: Geant4 is provided with a detailed description of the physics processes that occur when a particle transverses the matter. The interacting particle can undergo several different processes that lead to energy loss and eventually can start a shower. For the electromagnetic physics the undergoing processes are well known and precise parametrization of cross sections and energy losses for the different physics processes exist. The interactions involving hadrons are more difficult to treat and different models and parametrization exist in Geant4. In the next session more details on this aspect will be given.
   The main innovation introduced by Geant4 with respect to other simulation softwares is the introduction of the *production cut* [73]. This technique ensures the energy conservation with reasonable processing time and allows for physics process to produce secondaries even below threshold (an example is the gamma conversion: the positron is always produced, even with zero kinetic energy, for further annihilation).
   The energy deposits in the active material are called *hits*. Before further analyzing this information some processing is performed to simulate the response of the sensitive parts of the detector. For example the Birk's law effect [59] is applied to take into account the saturation of TileCal scintillators. The hits are then collected and made available for the next step.

3. *Digitization*: the Geant4 energy deposits in the active materials are collected according to the division in cells [2]. In this step the simulation of the electronic chain is performed: the signal is shaped and sampled according to the timing samples of the readout. Electronic noise is added to the samples. The results of this step are called *digits* and are very similar to the real output of the detectors data acquisition system.

---

[2] Also for other detectors, not naturally divided in cells, sensitive detector units are defined: a Si strip in SCT or a MDT tube in muon spectrometer.

4. *Reconstruction*: the reconstruction step, starting from the digits is performed to produce the response of the single read-out channel and apply calibration constants and reconstruction algorithm. The algorithms used in this step are very similar to the one used for real data and produce the same output. The idea behind the division of simulated data in hits/digits/reconstructed quantities should allow to apply the same algorithms (and, ideally, even the same implementation of this algorithms) to both real data and simulation. This is desirable for re-usability of the offline code and hence to minimize the risk of mistakes and not to introduce different systematics.
   Since we want to validate Geant4 physics simulation, it is particularly important to analyse simulation and data with ideally the same codes or at least with similar procedures.

For TileCal data the energy reconstruction method is a fit of the time slices that extracts the energy deposit in each PMT.
For LAr a optimal filtering (OF) algorithm is applied to the timing slices for each electronic channel. However the OF coefficients, have been made available only recently for real data while they are available since long in the simulation. In the following the LAr data are analyzed using the so called *parabola fit* method. This is the main difference between energy reconstruction in data and Monte Carlo simulation, the implications of this will be discussed in paragraph 4.3.1.

The role of QGSP physics lists

Geant4 comes with different models for the simulation of interacting hadrons with matter. This models can be grouped in the so called *physics lists* in which the involved physics processes are described. Following the suggestion of the Geant4 validation group and the results obtained in the stand alone analysis presented in the previous chapter, the *QGSP_GN* (version 2.6) physics list has been chosen for the simulation of the CTB setup. The QGSP uses theory-driven modeling for the reactions of energetic pions, kaons and nucleons, employing the quark-gluon string model. In this version of the QGSP physics list the gamma-nucleus interactions are present as well.
A physics list is a collection of particles and models for the physics processes. The QGSP physics list is well suited for the simulation of high energy interacting kaons, pions and nucleons. Where the processes in the list cannot be applied the standard Geant4 parametrization is used.
Quark gluon string model [62] is capable to predict final states (produced hadrons which belong to the scalar and vector meson nonets and the baryon-antibaryon octet and decoupled) of reactions on nucleon and nuclear targets with nucleon, pion and kaon projectiles.
For the nuclear interactions, after the initial parametrization of the target nucleus in terms of the constituent nucleons and the nucleon density, the

Sample	Data RMS (MeV)	MC RMS (MeV)
PS	60 (104)	46 (84)
Strips	20 (100)	11 (70)
Middle	50 (173)	42 (110)
Back	35 (99)	23 (75)
TileCal	25	25

Table 4.1: Values of the RMS of the cell energy distribution for pedestals events for the different calorimetric samples. For LAr the values between parentheses are the one obtained after the re-clustering in bigger cells (see 4.2.2).

inelastic interaction is treated in the Regge-Gribov approach[74], in this method the inelastic interactions are mediated by the exchange of one or more *Pomerons* each one consisting of 2 *strings*.
After the interaction the constituents of the projectile and of the target nucleon are divided in "string constituents" (for mesons the valence quark and the valence antiquark, for (anti-)barions a (anti-)quark and a (anti-)diquark). These constituents are rearranged in string: a string has at each end one of the initial constituents.
A string is "stretched" between flying away constituents and it is split into a hadron and a new string, i.e. a quark-antiquark (or diquark-antidiquark) pair is created and placed between leading constituents.
The new string is then again split until the energy in the strings gets too low for further splitting.
The formed hadron is built at one of the end-points of the original which quark content determines its species and charge.
After the initial interaction a pre-equilibrium decay model with an extensive evaporation phase [75] is used to model the nuclear fragmentation and to describe the evolution of the damaged nucleus. Gamma-nuclear and electro-nuclear reactions are modeled using chiral invariant phase-space decay for energies below 3 GeV.
This physics engine gives better quality of predictions compared to other physics list (i.e. the ones based on parametrization like the LHEP [61] physics list based on the GEISHA package), with the prize of slower simulations.

4.2 Data Analysis

Aim of this analysis is the measurement of response of the calorimetric setup to single pions. In real data pion beams are contaminated by muons and electrons, therefore we will first select a pure sample of pions using the ancillary detectors and the calorimetric deposits (paragraph 4.2.1). Even

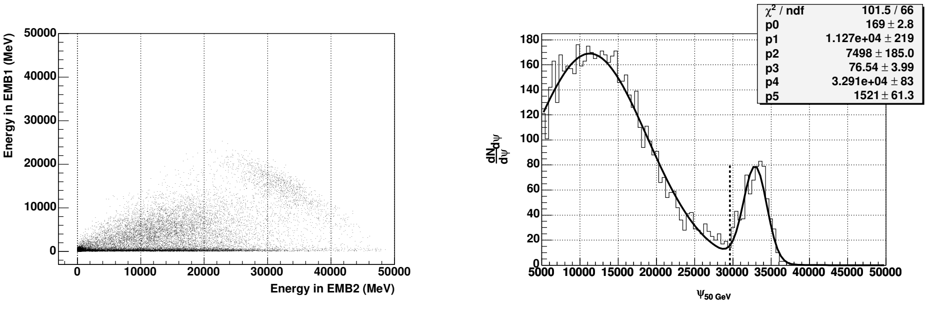


Figure 4.3: On the left: correlation between the energy deposit in the Strip and Middle samples of the LAr calorimeter. On the right: distribution of the parameter Ψ for the same events. The parameter Ψ is defined in the text.

if the simulated data contain a pure sample of pions the calorimetric cuts used in real data are applied also to simulated ones to minimize different systematics.

Hence both data and simulation have been treated with exactly the same analysis procedure to reconstruct energy deposits.

The pion energy is calculated summing the deposit from all the calorimetric cells with $0 < \eta < 0.7$. The energy of the cell is considered only if it is over a threshold of 2 times the RMS of the noise distribution.

We measure the noise RMS from data themselves. For real data the noise is extracted from the out-of beam events, while for Geant4 data the noise is measured from the cells away from the beam axis since no pedestal events are available. In table 4.1 the measured values of the RMS of the distributions of pedestal events are reported. We can notice that the level of noise in LAr in the simulation is smaller with respect to what is measured in the data. This is due to the different reconstruction algorithm used between data and simulation: the optimal filtering method, available for Geant4 simulation, substantially reduces the electronic noise, as compared to the parabola fit method used in the data.

4.2.1 Events Selection

In real data we have to take into account the beam contamination from electrons and muons in pion beams.

Since we have pure pion beams we need to identify electrons and muons with high efficiency. In some cases we can use the ancillary detectors but we have also to exploit the calorimeters themselves which perform differently for π, e and μ.

Muons are easily rejected using the SMT scintillator requiring a signal compatible with the pedestal. To take into account inefficiencies of the scintillator also a cut based on TileCal energy deposit is used. This cut depends on the beam energy. For high energy ($E_{beam} \geq 250$ GeV), where the muon contamination is low and the muon signal is well distinguished from pion signal, a simple cut based on total energy deposit in TileCal (typically: $E_{tile} > 50$ GeV) is used to distinguish pions from muons. For lower energy ($E_{beam} < 250$ GeV) a cut on the fraction E_D/E_{tot} of deposited energy in the third sample of TileCal is used. The values (beam energy dependent) of this cut were obtained first selecting muons with SMT and then looking at the distributions, for each energy, of the fraction of energy in the third sample of TileCal.

Electrons are removed with the information from the first two samples of LAr calorimeter. In the left plot of figure 4.3 the correlation between the Strip (EMB1) and Middle (EMB2) LAr samples is shown for a beam energy of 50 GeV. The electrons are well distinguished from the pions and are visible in the up-right area of the plot. To exploit better this correlation, we rotate the axis so that the x' axis is perpendicular to the anti-correlation band of the electrons and we project the histogram onto the x' axis.

Thus we define the parameter:

$$\Psi = E_{EM2} \cos\theta + E_{EM1} \sin\theta \tag{4.1}$$

where E_{EM2} and E_{EM1} are the energies measured in the second and first electromagnetic samples, and θ is the angle between x' and x axis.

The distribution of the parameter Ψ is shown on the right plot of figure 4.3 a cut on Ψ separates electrons (right peak) from pions. The distribution has been fitted with the sum of two gaussians. To remove $\sim 95\%$ of the electrons, particles yielding to a signal with Ψ values greater than $\mu_e - 2\sigma_e$ (the mean value and the σ of the electron gaussian) have been rejected.

Some additional cuts have been introduced to increase the quality of the events. A cut on the energy deposited in the PS longitudinal sample of LAr has been applied to remove events which start the shower in the material before the EM calorimeter. Less then 5% of energy in the PS is required for beam energies obove 50 GeV. For lower energies we require $E_{PS} < 10\%\ E_{beam}$.

Additional cuts on the beam chambers have been applied to select well collimated particles.

No cut on the signal from the trigger scintillators has been used since it does not further improve the quality of the data.

4.2.2 Corrections

Energy lost in the cryostat

Some corrections to the measured energy are applied to take into account detector effects as well as biases introduced by the analysis.

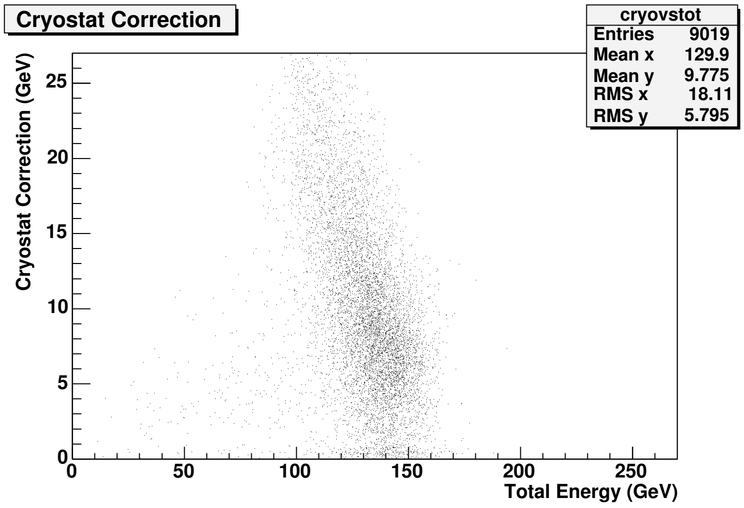


Figure 4.4: Correction for energy loss in the cryostat between LAr and TileCal as a function of the total reconstructed energy for 180 GeV pions.

We take into account the energy loss in the cryostat wall between LAr and TileCal with the parametrization [76]:

$$E_{cryo} = \sqrt{E_{Back} \times E_A} \tag{4.2}$$

where E_{Back} is the energy loss in the last longitudinal sample of LAr and E_A is the energy loss in the first longitudinal sample of TileCal. In figure 4.4 the correction for 180 GeV pions is shown in function of the total energy deposit in the calorimetric system. A mean correction of about 10 GeV (5.5%) is applied to data, the correction is more important for small total energy deposit, indicating that for these events more energy is lost in the cryostat wall.

The effect of noise cut

As already mentioned the total energy deposit is calculated using a sum on all cells with an asymmetric noise cut. This introduces a bias in the calculation of the energy since only positive fluctuations are counted. This effect depends on the number of cells and on the noise level.

To reduce this bias the following procedure has been applied: LAr cells has been grouped in larger *pseudo-cells* to obtain a granularity similar to the one of TileCal [3].

Using the simple approximations: number of cells in a cluster (N_c) constant,

---

[3] A reduced granularity of $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ can be considered enough for hadronic showers.

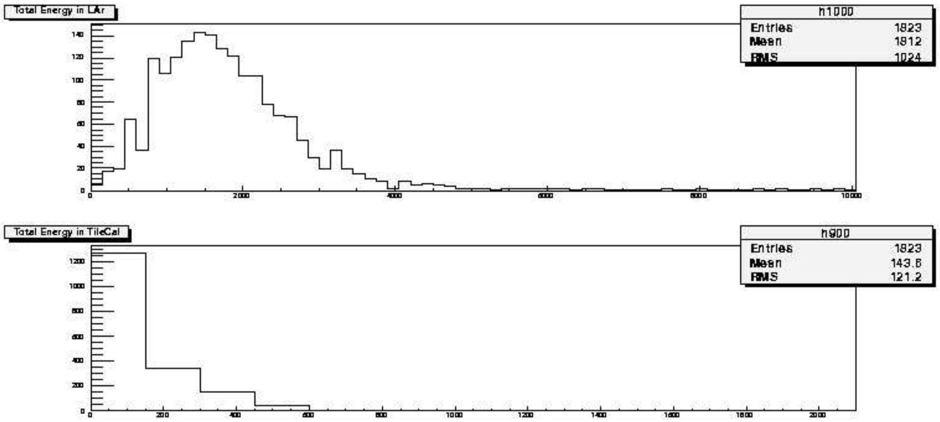


Figure 4.5: Energy deposit in LAr (top) and TileCal (bottom) for pedestal-like events, summing over all the cells with the $E > 2\sigma_{noise}$ cut.

noise in each cell distributed as a gaussian around zero and variance σ^2_{cell} and no correlation between noise of different cells; the bias introduced by the asymmetric noise cut can be written as:

$$E^{bias}_{cell} \approx P(E_{cell} > \delta_{cell}) \times N_{cell} \times \bar{E}_{>\delta_{cell}} \tag{4.3}$$

where $P(E_{cell} > \delta_{cell}) \times N_{cell}$ is the number of cells that fluctuate of a quantity more than δ_{cell}. In our case $\delta_{cell} = 2\sigma_{cell}$ and the bias introduced by a single cell $\bar{E}_{>\delta_{cell}}$ is $k\sigma_{cell}$[4].

Grouping the LAr cells in clusters, composed of N_c cells, we can write $N_{cluster} = N_{cell}/N_c$, $\sigma_{cluster} = \sqrt{N_c}\sigma_{cell}$ and thus:

$$\begin{aligned} E^{bias}_{cluster} &\approx P(E_{cluster} > \delta_{cluster}) \times N_{cluster} \times k\sigma_{cluster} && (4.4) \\ &= P \times \frac{N_{cell}}{N_c} \times k\sqrt{N_c}\sigma_{cell} = \frac{E^{bias}_{cell}}{\sqrt{N_c}} && (4.5) \end{aligned}$$

The effect on energy reconstruction of this method has been studied applying the described procedure to pedestal-like events. The mean value obtained from the distributions of energy deposited in each sample for pedestal events

---

[4] The bias $\bar{E}_{>\delta_{cell}}$ can be simply obtained:

$$\begin{aligned} \bar{E}_{>\delta_{cell}} &= \int_{\delta}^{\infty} \frac{E \times e^{-E/2\sigma^2}}{\sqrt{2\pi}\sigma} dE \\ &= \frac{e^{-\delta^2/2\sigma^2}}{\sqrt{2\pi}}\sigma \\ &= \frac{e^{-2}}{\sqrt{2\pi}}\sigma \qquad \delta = 2\sigma \end{aligned}$$

has then been removed to the measured energy in data. In figure 4.5 the energy distribution in LAr (top) and TileCal (bottom) for pedestal-like events is shown.

The energy deposit in LAr for pedestal events gives a mean value of $E^{bias}_{cluster} =$ 1800 MeV much more than for TileCal where we obtain only 143 MeV. To understand this difference we have to remember that, even if the LAr cell noise is compatible with the one obtained from TileCal (see table 4.1) the number of cells involved is much larger (three TileCal cells forming a $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ tower is formed by 61 LAr cells).

The RMS of distributions for pedestal-like events for both data and Monte Carlo and the different calorimetric samples are presented in table 4.1. The values between parentheses corresponds to the RMS obtained after clustering together LAr cells.

4.3 Results

4.3.1 Total energy deposit

In figure 4.6 the distribution of the energy deposit in LAr (top plot), in TileCal (middle) and sum of the two (bottom) is shown and a direct comparison between data (black points) and MC (red histogram) is made for 180 GeV pions. All the described corrections have been applied.

We can notice the following:

1. The total calorimeter response TileCal + LAr (bottom plot) is in good agreement with Geant4 predictions all over the spectrum of deposited energy.

2. The energy deposits in TileCal (middle plot) at low energy ($E <$ 50 GeV) are over estimated by Geant4 with respect to data. The same happens in the upper end of the spectrum ($E >$ 150 GeV)

3. In LAr (upper plot), the very low energy deposit ($E <$ 5 GeV) has a marked peak due to pions starting the shower in TileCal. The peak is broader in the data than in Geant4. This is an instrumental effect: the energy reconstruction method in the data (the *parabola fit*) is more noisy than the one used in the Monte Carlo analysis (the optimal filtering coefficients for the data were not yet ready when this analysis was performed). The effects of the parabola fit are confined to small energies and is negligible above 10 GeV. Figure 4.7 shows the effect of the parabola fit on the LAr energy deposit for beam energy of 20 GeV. Both plots show the total energy deposit in LAr for data (black dots) and simulation (red histogram). The difference between the two is the different energy reconstructed algorithm used in the Geant4 simulation.

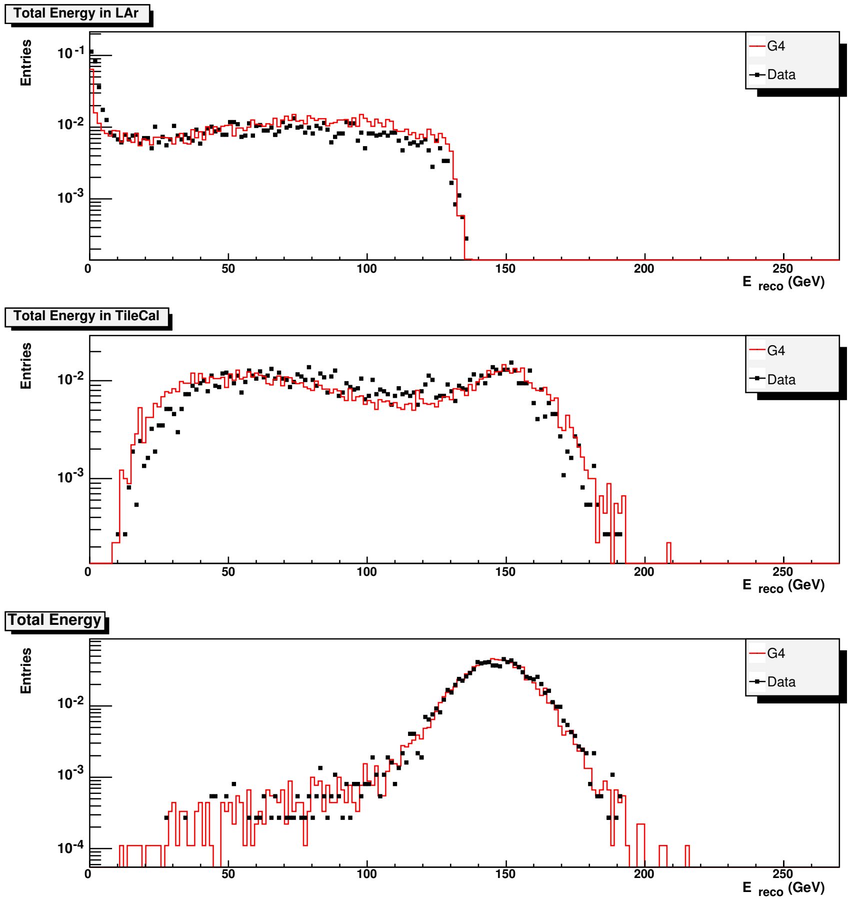


Figure 4.6: Comparison between data (black points) and Monte Carlo simulation (red histogram) of reconstructed energy only in LAr (top), TileCal (middle) and in both sub-systems (bottom). The histograms are normalized to the unit. The beam energy is 180 GeV.

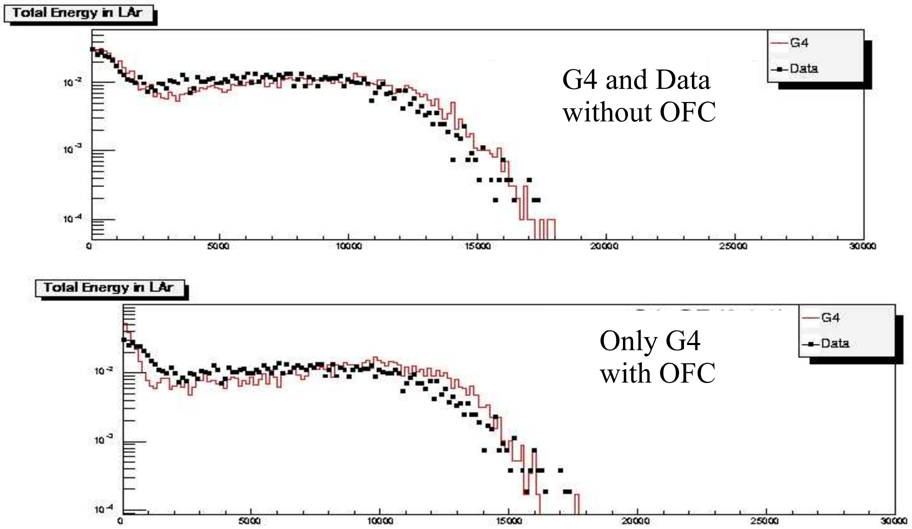


Figure 4.7: Comparison between data (black dots) and Geant4 simulation (red histogram) for the total energy deposit in LAr obtained with two different algorithms in energy reconstruction. In the top figure the Monte Carlo data are reconstructed with the parabola fit method, as in the data, while in the bottom plot the standard optimal coefficients algorithm has been used.

In the top plot the Monte Carlo data are reconstructed using the parabola fit method instead of the optimal filtering coefficients.
It is possible to see that the agreement improves, in particular for the peak at very low energy deposit.

4. In LAr, the large energy deposits ($E > 10$ GeV)are a bit over estimate in Geant4.

As a function of beam energy, the ratio of the calorimeter system (LAr + TileCal) response in data and Monte Carlo is shown in figure 4.8.
The calorimeter response is computed from a gaussian fit to plots similar to the one of figure 4.6 lower, excluding the non gaussian tails by iterating the fit in a $\pm 2\sigma$ interval around the maximum.

4.3.2 Resolution

The measured resolution σ/E_{reco} as a function of the beam energy is shown in figure 4.9.
The upper plot shows the resolution, as a function of beam energy, for data

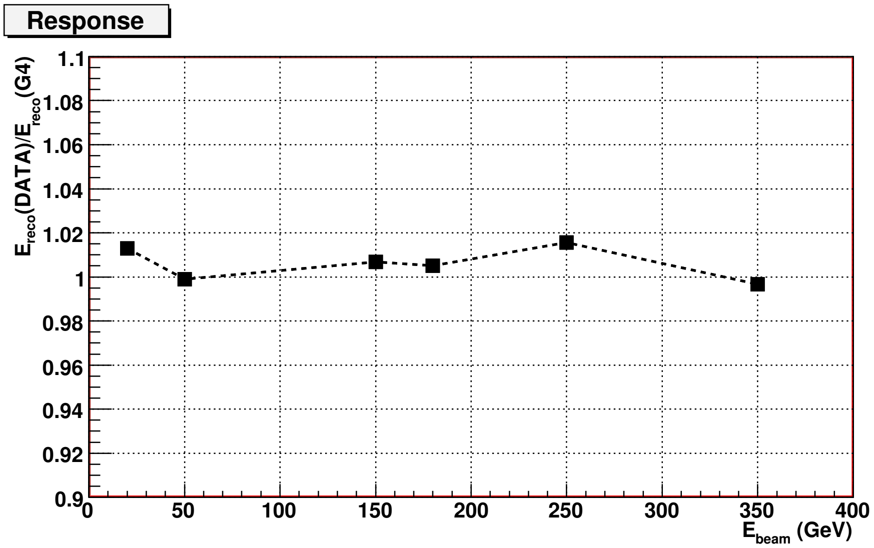


Figure 4.8: The ratio between reconstructed total energy in data over the reconstructed energy in Geant4 simulation as a function of the beam nominal energy. The agreement over the energy range 20 GeV $\leq E_{beam} \leq$ 350 GeV is well inside the ±2% level.

(in black) and for Geant4 (in red). The ratio of the measured resolutions between data and Geant4 as a function of beam energy is shown in the bottom plot.

We can notice a general good agreement, above 100 GeV, between simulation and data. The level of agreement is much better than 10% in this region of energies.

The resolution, for this high energy points ($E > 100$ GeV) has been fitted with $\sigma/E = a/\sqrt{(E)} + b$ obtaining from data $a = (0.65 \pm 0.20)$ GeV$^{1/2}$ and $b = 0.03 \pm 0.01$ to be compared with $a = (0.59 \pm 0.16)$ GeV$^{1/2}$ and $b = 0.04 \pm 0.01$ obtained from Geant4 Monte Carlo, the values obtained from the fit are compatible with the design resolution (see section 3.1.1).

For energies belov 100 GeV the resolution obtained in data sample is 20-30% larger than the one obtained from simulation. This is due to the LAr energy recostruction method, different between data and simulation.

A special run, using parabola fit reconstruction, has been simulated with Geant4 (see figure 4.7) for the 20 GeV energy point. The measured resolution from this special dataset agrees at the level of 9% with real data ($\sigma/E(data) = 0.25$ and $\sigma/E(MC) = 0.23$).

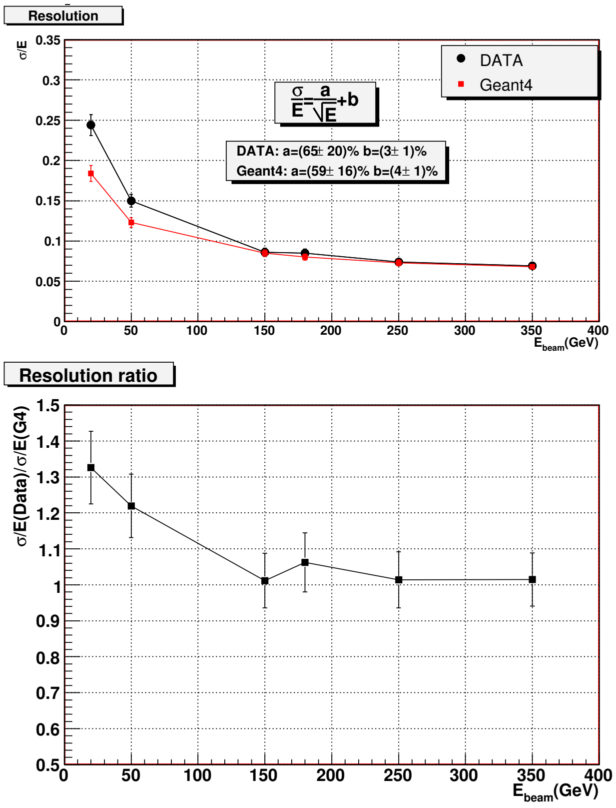


Figure 4.9: Top: Resolution obtained from data (black points) and simulation (red points) as a function of beam energy. Bottom: The ratio between data and Geant4 as a function of beam energy.

4.3.3 Lateral and longitudinal partial shower profiles

The study the shower profile both lateral and longitudinal can be done exploting the segmentation of the calorimetric setup.

The lateral shower profile has been studied comparing the energy deposit in each calorimetric tower defined by TileCal granularity ($\Delta\eta \times \Delta\phi = 0.1 \times 0.1$). In figure 4.10 the energy deposit in each tower is shown as a function of the mean pseudorapidity of the tower.

Data from three different energies are shown: from top to bottom the beam energy is 20, 180, 350 GeV. In all three cases the beam is centered at at $\eta = 0.35$. Geant4 simulation predicts larger energy deposit in the central tower that what is measured with the data, and less energy in the towers adjacent to the central one.

Geant4 shower predicts too compact (laterally) showers than what is measured from data, this is especially visible for the low energies.

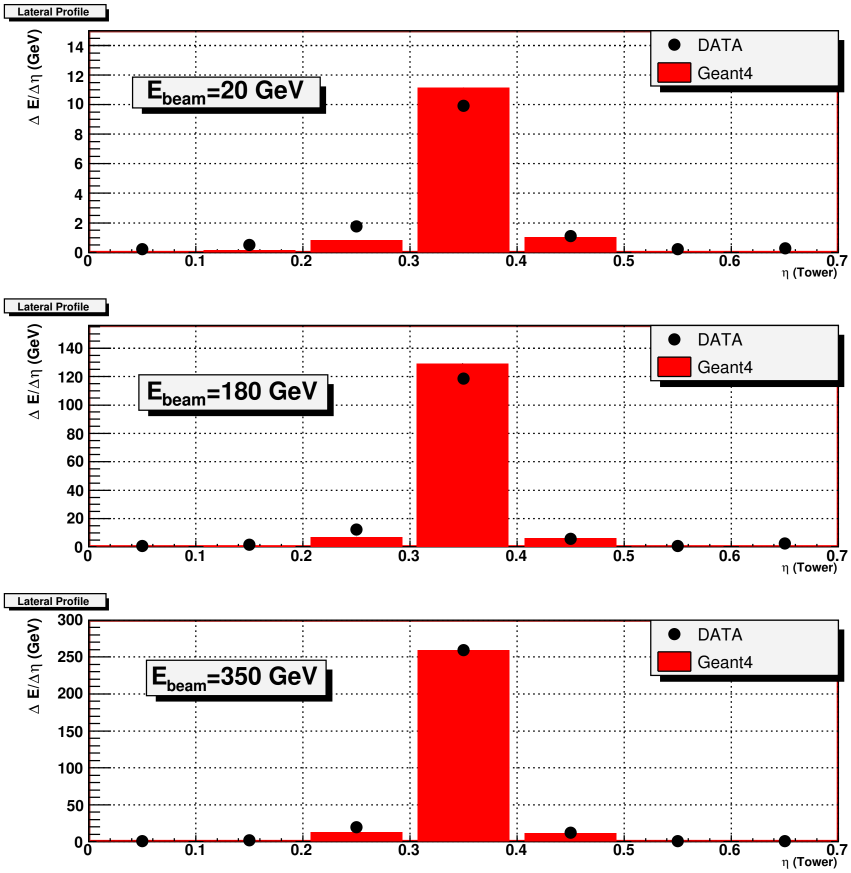


Figure 4.10: Lateral shower profile for pions at 20, 180 and 350 GeV. Energies deposit in the calorimetric towers in function of the η central value of the towers are shown. Black points are data, while red histograms are Geant4 predictions.

This effect is symmetric around the hit tower hence cannot be produced by an off-centred beam. For the analysis of the longitudinal shower profile we normalized the energy deposit in each calorimetric sample to the thickness (in interaction length λ) of the samples.

The first two samples of LAr (PS and Strips) have been considered together in the following analysis.

To avoid fluctuations in the starting point of the hadronic shower it has been requested that more than 10% of total measured energy is deposited in the first three samples of LAr (corresponding to $\sim 1.1\ \lambda$).

From a Geant4 geantino scan it is possible to determine the thickness of the LAr samples in X_0 units. From the Calorimetry TDR [77] the total thickness for LAr is 1.2 λ_I, from these numbers it is possible to compute the thickness of the calorimetric longitudinal samples. The thickness that have

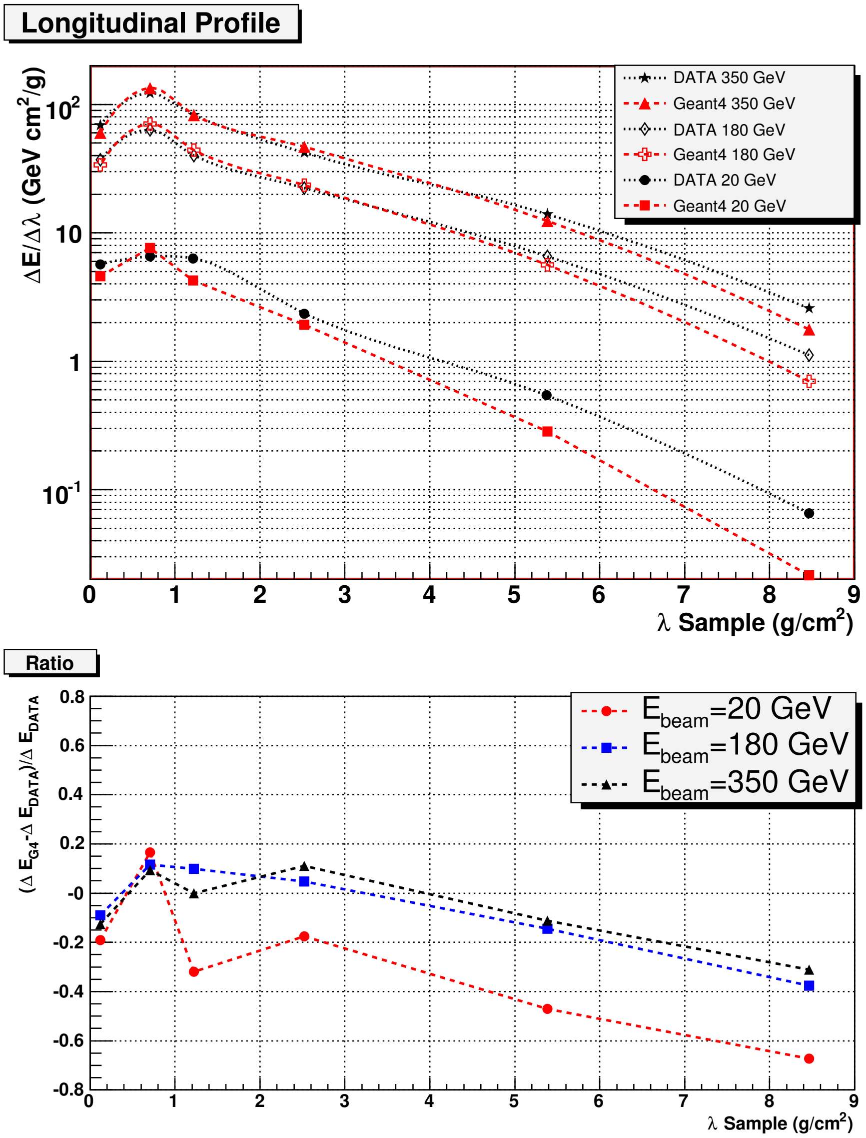


Figure 4.11: Top: Longitudinal shower profiles for 20, 180 and 350 GeV pions for Geant4 (red) and for data (black) as a function of the depth λ in the calorimeter. Bottom: The fractional difference $(\Delta E_{MC} - \Delta E_{Data})/\Delta E_{Data}$ as a function of the depth for the considered energies.

Sample	Thickness λ_I (g/cm^2)
PS+Strips	0.239
Middle	0.933
Back	0.102
Cryostat	0.500
A Cells	1.487
BC Cells	4.248
D Cells	1.912

Table 4.2: Thickness of the longitudinal segments of the calorimetric setup in units of interaction length.

been used for TileCal calorimeter are the one quoted in TileCal TDR [54]. In table 4.2 the thickness for each sample is quoted once corrected for the impinging direction $\eta = 0.35$.

For the cryostat wall between LAr and TileCal a thickness of 0.5 λ_I has been considered [77].

Figure 4.11 (top) shows the energy deposits in each longitudinal sample, normalized to its thickness, as a function of the depth in the calorimeter λ. Red points represent the Geant4 predictions for the three beam energies 20, 180, 350 GeV while measurements from data are presented with the black points.

The fractional difference $(\Delta E_{MC} - \Delta E_{Data})/\Delta E_{Data}$ is shown, as a function of the depth λ, in the bottom plot of figure 4.11.

The simulation predicts less energy at high λ values, and in the first calorimetric sample (PS and Strips), the simulated shower starts deeper than the measured one and is more compact. There are two possible explanation for this discrepancy: the material distribution is not precisely simulated or the physics list used to describe the shower development does not precisely describe the real data.

Similar effects to the one discussed above have been already obtained in the past analysis of TileCal test beam in stand alone more (see [78] and [79]). In particular (see figure 3.12) less energy was simulated in the first part of the shower with respect to the data.

However should be noticed that the effect now is smaller compared to the one obtained in the past. The new version of Geant4 has improved on the hadronic shower development.

4.3.4 Conclusions

The validation of the Geant4 simulation can be performed only with a detailed study of the calorimeter response to test beam data and comparing data and Monte Carlo using the same procedure. This was not yet possible because reconstruction methods of LAr are not yet ready for the Test Beam data. The results described in this chapter are still preliminary. However from the preliminary analysis presented here some preliminary conclusions can be stated.

The available data and simulation do not use exactly the same energy reconstruction algorithms and this introduces discrepancies, especially on the noise level, difficult to control, in particular at low energies.

However we can conclude that Geant4 predicts a total energy deposit and resolution in good agreement with the data ($\pm 2\%$).

The shower profile comparisons is less satisfactory and indicate that more tuning of Monte Carlo shower development simulation is still needed: simulated showers are more compact in both dimensions with respect to the measured ones.

A detailed study of shower shape can be performed only with a precise and reliable energy reconstruction algorithm in real data. Optimal filtering coefficients in LAr are mandatory, especially at low energies, followed by a noise-suppressing clustering algorithm.

The preliminary comparison between Geant4 simulation and combined test beam presented in this chapter shows an improvement, respect to the past, in the level of agreement between data and Monte Carlo, however the needed precision in simulation claims for further improvements.

In the next chapter Geant4 simulations will be used to develop a weighting technique to improve jet energy measurements. A reliable Monte Carlo description is needed to be able to cope with the precision required to perform physics analysis involving jet measurements.

A jet is composed of a large number of hadrons impinging the calorimetric system over a wide angular region ($\Delta R = 0.7$). Other experimental effects can play also a role and can complicate the task of jet reconstruction. Cracks between the calorimeters, leakage, out of cone energy are some of the effects that should be understood in detail when measuring jets.

Chapter 5

Jet Energy Calibration

Jets are the manifestation of scattered sub-nuclear partons. The jet reconstruction algorithms aim to the jet identification from the energy deposit in the calorimeter cells. Once the jet is reconstructed its energy must be corrected both for detector effects (e/h, cracks, dead material, leakage) and for physics effects (jet clustering losses, initial and final state radiation, underlying event).
In this chapter a jet calibration scheme will be presented and discussed. After an introduction on jet reconstruction at a hadron collider we will present a method designed to correct jet energy measurement for detector effects while improving jet energy resolution. The method uses Monte Carlo simulations to extract weighting functions to re-calibrate jets. The algorithm to extract the weights and its limits of validity will be discussed. Results on linearity and resolution obtained on QCD di-jet events and $t\bar{t}$ *events will be presented.*

5.1 Introduction

According to the current theory, jets are the manifestation of scattered subnuclear partons.

Because of Quantum Chromo Dynamics (QCD) confinement [80] scattered partons loose energy pulling out from the vacuum more partons (fragmentation) which have then to rearrange themselves in colour singlet (hadronization). The final result is a set of collimated hadrons (mainly π^0 and $\pi^\pm$) that are absorbed and measured by the calorimetric system.

A jet is thus measured as a set of calorimetric *clusters*. Clusters are collections of geometrically closed calorimetric cells.

The calorimeter is mechanically divided in cells which allow to define "natural" clusters: pseudo-projective towers of fixed dimension in the $\Delta\eta \times \Delta\phi$ space. More sophisticated clusters, as it will be discussed in the following, can be used to minimize the electronic noise contribution.

Jet measurements are performed in different steps: first clusters have to be merged together (*jet clustering*) and the kinematic variables of the jet have to be calculated from the cluster response (*recombination scheme*).

Corrections to the measured energy must be applied to take into account different effects, such as:

- Detector effects: the ATLAS calorimetric system is non-compensating, since a jet is a mixture of hadrons and photons (π^0) the response of the calorimeters will be smaller than expected (the enery scale of a jet is different from the electromagnetic scale at which the calorimeters have been calibrated). Moreover, the presence of dead material and cracks, and the use of different technologies in different pseudo-rapidity regions makes more difficult to achieve a good uniformity of response.

- Physics effects: the energy associated to the jet from the calorimetric clusters is in general different from the energy of the parent parton because of final state radiation, fluctuations in the underlying event, clustering effect (un-clustered energy).

ATLAS strategy is to apply the needed corrections separately. First detector corrections are applied to reconstruct, from calorimetric response, the energy of the particles composing the jet (see figure 5.1). To obtain the correction factors Monte Carlo simulations can be used once the simulations have been validated [1]. This corrections typically depend on the energy response of the clusters.

The physics effects, mainly depending on the $\mathrm{P_T}$ of the jet, must be corrected taking into both Monte Carlo simulations and real data. Some simulation

---

[1] For the validation, test beam data are used, as it has been shown in the previous chapter, however the simpler test beam geometry and the use of single particles must be carefully taken into account when extrapolating to the performances to the final detector.

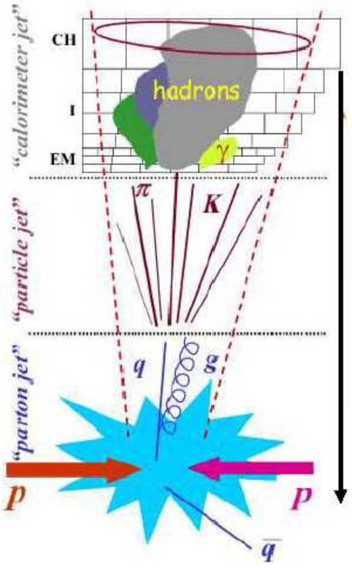


Figure 5.1: Schematic picture of the jet formation: the parton jet fragments hadronize in hadrons (particle jet) that are detected by the calorimeter (calorimeter jet). The jet reconstruction algorithm defines jets from calorimetric jets. Constants have to be applied to get the particle and parton energies.

parameters must be tuned and cross checked with real data (PDFs, fragmentation functions, modelling of ISR and FSR underlying event, pile-up). The aim of the work presented in this chapter is to develop a calibration algorithm for jets to correct for detector effects only. The algorithms relies on Geant4 simulations of the detector.

5.2 Jet reconstruction

The jet reconstruction can be divided into three steps: the jet clustering algorithm which associates the calorimetric clusters, the recombination scheme aiming to define the kinematic variables for the jet and the calibration, which provides the final evaluation of the jet kinematics [81].

5.2.1 Jet clustering

In order to compare the jet measurements with the theoretical perturbative QCD predictions, a jet reconstruction algorithm has to fulfil a number of criteria.

We must be able to apply the same jet algorithm to different sets of data: calorimetric clusters, Monte Carlo particles, partons. This is required to

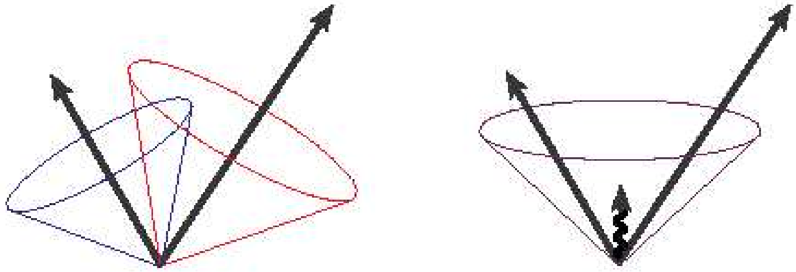


Figure 5.2: An illustration of infrared sensitivity. Jet clustering begins around "seed" particles shown here as arrows with length proportional to energy. The presence of a soft radiation between two jets can cause the merging of two jets.

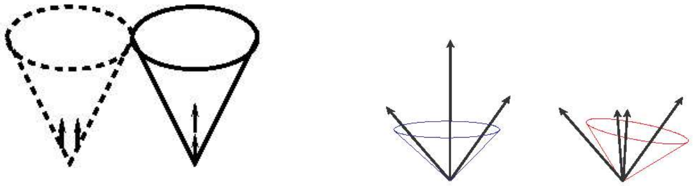


Figure 5.3: An illustration of collinear sensitivity. In left figure the energy produced by two particles can fail to produce a seed because the energy is split among several detector towers. The right figure shows an example of sensitivity to ordering of particles that act as seeds.

compare experimental data with theoretical calculations and simulation results.

The jet clustering can be applied easily to Monte Carlo particles: we replace the cell energy deposits with the energy and momentum of the simulated particles and then form the jet.

Jet calibration is done in two steps: first we correct for detector effects (cracks, leakage where energy is lost in specific geometrical regions) and non-compansation. Then we take into account physics effects (ISR, FSR, unclustered energy). We should finally be able to measure from experimental data the kinematic properties of the partons originating the jet.

To perform the second step, often data themselves are used, constraining jj invariant mass in known processes (hadronic W boson decay in two jets) or using $Z + jet$ events with the following $Z \to ll$ decay.

In the following of this section we will refer to the different possible data sets (clusters, particles, partons) with the name jet *components*.

From the theoretical point of view we have to leave unmodified the strategy to construct a jet independently from the jet *components*. The design of a jet clustering algorithm is driven by several considerations, the following features are desirable and even necessary:

- In order to compare the experimental quantities with theoretical calculations the properties of the jets should not depend on the presence of arbitrary soft *components* (theoretically this property is known as *infrared safety*). Consider the case shown in figure 5.2 the presence of a soft radiation between two *components* can change the jet properties. In the absence of soft radiation two jets are reconstructed around the two *components*, while with the presence of the soft radiation they are merged to form one single reconstructed jet.

- It should be collinear safe: the jet properties should not depend on any collinear radiation in the event. Left diagram of figure 5.3 shows a case in which a jet is found only if the two *jet components* lie in the same calorimetric tower.

- It should be collinear factorisable: the jet properties should not change replacing one *component* with a set of collinear *components* with the same total momentum. The right diagram of figure 5.3 shows a case in which, due to the energy ordering of jet *components* acting as seeds, the properties of the reconstructed jet are not collinear factorisable.

Different jet clustering algorithm have been implemented in the ATLAS software and are used for physics analysis.

The most used at hadron colliders, historically developed first, is the cone algorithm, which associates *jet components* in a cone of radius R around a point in the $\eta - \phi$ space. If the jet clustering is started around *components* carrying at least a minimum energy E_{seed} the algorithm is called seeded, however to avoid infrared and collinear problems, a *seedless* cone algorithm should be considered.

In a seedless algorithm trial cones are positioned anywhere in the $\eta - \phi$ space: cones of radius R are built around all the *jet components* (regardless of their position, their E_T ordering and, most important, of their energy content). For each *component k*, with coordinates (η^k, ϕ^k), the center of the cone $\vec{C}^k = (\eta^{Ck}, \phi^{Ck}) = (\eta^k, \phi^k)$ is defined. A *component i* is included in the cone if $\sqrt{(\eta^i - \eta^{Ck})^2 + (\phi^i - \phi^{Ck})^2} \leq R$. Then the E_T-weighted centroid [2] $\overline{\vec{C}}^k$ is evaluated:

$$E_T^{Ck} = \sum_{i \subset C^k} E_{Ti} \qquad \overline{\eta}^{Ck} = \frac{\sum_{i \subset C^k} E_{Ti}\eta_i}{E_T^{Ck}} \qquad \overline{\phi}^{Ck} = \frac{\sum_{i \subset C^k} E_{Ti}\phi_i}{E_T^{Ck}}$$

In general the centroid $\overline{\vec{C}}^k$ is not identical to the center of the cone $\vec{C}^k$ and the cone is not stable. Therefor, an iterating procedure is applied until the cone becomes stable and the coordinates of the jet centroid do not change more than some percent.

The described procedure can lead to a final jet list where some of the jets overlap. A *split and merge* procedure has to be used to merge or separate jets with overlaps in order to avoid the assignment of *jet components* to two jets. The usual way to do this is to merge two jets if the overlapping energy percentage is over some threshold [3].

In ATLAS, due to the large number of calorimetric regions (with the granularity of $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ we have approximately 200000 calorimetric towers), the described algorithm is so expensive in term of execution time that a seedless algorithm is not practically usable. The problem can be partially solved considering as seeds (thus in a seeded algorithm) only those clusters above some (low) $\mathrm{E_T}$ threshold ($\mathrm{E_T}$ is typically chosen to be 1-2 GeV). This seeded cone algorithm is safe only for values of the transverse energy of the jet greater than the threshold value.

An alternative jet clustering algorithm designed to be both infrared and collinear safe is the K_T algorithm [82]. In this case, *components* are clustered on the base of closeness both in (η, ϕ) space and in transverse momentum. The algorithm flows as follows:

1. For each *component i* the quantity $d_i = P_{Ti}^2$ is defined

---

[2] Different choices in the calculation of the centroid position and in the jet energy can be used leading to the choice of a different recombination scheme, the role of which will be discussed.

[3] It should be however noticed that the split and merge procedure in general breaks the collinear and infrared safety of a jet algorithm. The procedure can be avoided using different jet algorithms (K_T) where the *jet components* are assigned uniquely to one jet.

2. For each pair of *jet components* (i, j) (with $i \neq j$) the quantity d_{ij} is defined as follows:

   $$d_{ij} = min(P_{Ti}^2, P_{Tj}^2)\frac{(\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2}{D^2}$$

   where D is a parameter of the jet algorithm.

3. Find the minimum (d_{min}) among of all the d_i and d_{ij}.

4. If d_{min} is a d_{ij} merge *jet components* i and j to have a new *particle* k with: $p_k^\mu = p_i^\mu + p_j^\mu$

5. If d_{min} is a d_i a jet has been found.

6. Iterate until all the *jet components* are either merged or labelled as jets

Since the K_T algorithm can be very expensive in term of execution time time, usually pre-clustering techniques are applied to reduce the number of input *jet components*.

The calibration algorithms in ATLAS are usually studied using the cone algorithm, thus in the following all the results will refer only to the cone algorithm. The choice of the ATLAS collaboration at present is to consider a seeded cone algorithm with $E_{Tth} = 2\ GeV$, $R = 0.7$ or $R = 0.4$, two overlapping jets are merged if they share more than 50% of the energy of the least energetic one.

5.2.2 Recombination

After the jets are reconstructed the raw kinematic variables are calculated. Historically, variables are defined by the Snowmass convention [83]. Transverse energy $\mathrm{E_T}$, the pseudorapidity η and the azimuthal angle ϕ of a jet are:

$$\begin{aligned} E_T &= \sum_i E_{Ti} \\ \eta &= \frac{\sum_i E_{Ti}\eta_i}{E_T} \\ \phi &= \frac{\sum_i E_{Ti}\phi_i}{E_T} \end{aligned}$$

where i is the i-th *jet component*.

The problem with this approach is that E_T, η and ϕ variables are only approximately equal to the "real" quantities, becoming exact in the limit of small jet mass ($M^J \ll E_T$). Moreover the Lorentz invariance is lost. For these reasons, the approach preferred in ATLAS (as well as that of CDF and D0 at Tevatron [81]) is the E-scheme (or 4-vector recombination) where one simply adds the 4-momentum (p^μ) of the *jet components* (considered massless): $p^\mu_{jet} = \sum_i p^\mu_i = (E, p_x, p_y, p_x)$ with the definition of the Lorentz invariant variables:

$$p_T = \sqrt{p_x^2 + p_y^2} \qquad \phi = \tan^{-1}\frac{p_y}{p_x} \qquad y = \frac{1}{2}\ln\frac{E+p_z}{E-p_z}$$

5.2.3 Calibration and weighting techniques

The ATLAS calorimeters are calibrated at the electromagnetic scale (EM). This means that the calibration constants have been calculated from test beam data, so that the sum of the energy deposits in the cells gives the correct energy scale for electrons and γ. The measured energy of a jet must be corrected for the detector and physics effects to report the measured energy to the jet “true” jet energy.

The definition of the “true” jet energy is not trivial. The most natural definition is to consider as the “true” jet energy the energy of the parton that originated the jet itself, however this will experimentally not correspond to the sum of the energies of the particles, mainly pions, composing the jet due to unclustered energy. In the rest of this work we will define the jet “true” energy as the sum of the energies of the particles produced in the fragmentation process that fall in the jet volume. By construction this corresponds to the *particle jet energy.*

Aim of this work is to present an algorithm to correct for non-compensation, cracks, non uniformity, leakage. This definition of truth is therefore the most appropriate to disentangle the physics effects (ISR/FSR, out of cone, ...), to be treated with the *in-situ* calibration, from the detector effects that we want to correct for.

To achieve this one of the most used techniques is to apply a set of weights to the calorimetric cells belonging to the jet. The method has been developed and widely applied by the H1 experiment [84].

We propose a cell based weighting scheme a-la-H1 for ATLAS [85]. A similar method has been already used for the ATLAS combined test beam of 1996 and 1998 [86, 87].

In the following a detailed study of the algorithm characteristics, of its implementation and performances are presented.

5.3 Jet calibration

The basic idea of the method presented here is to apply a weight to the cell energy (E_i) to correct for the detector effects. The weights depend on the energy deposit in the cell i, on the cell position and on the energy of the jet

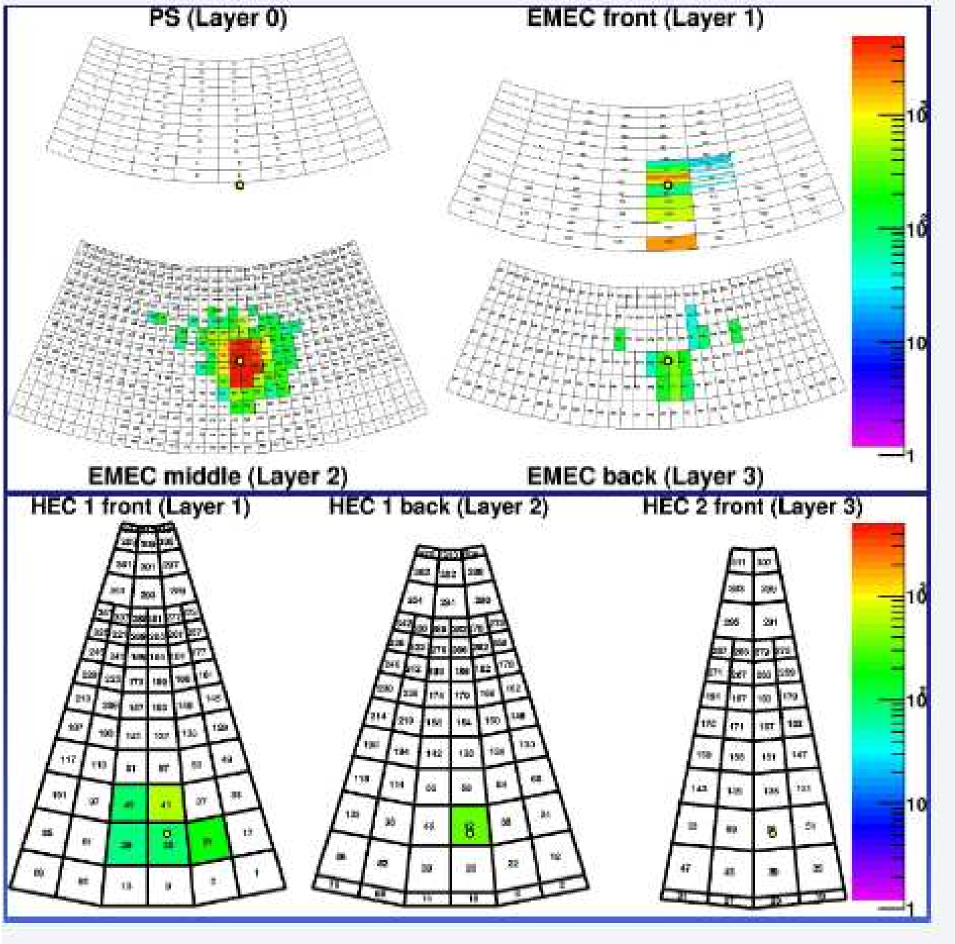


Figure 5.4: Topological cluster for a 120 GeV pion in the electromagnetic and hadronic end-cap calorimeters (2002 test beam data). The 7 calorimetric layers are visible in the $\eta - \phi$ plane, the small circle represents the impact point of the pions, the coloured cells belong to the reconstructed topo-cluster (see section 5.3.1), different colours represent different energy deposits (scale on the right in MeV). In EMEC Middle (Layer 2), the seed cells are well visible (red) surrounded by the other cells belonging to the cluster. Note that the topo-cluster extends, starting from the seed, in all three dimensions: inside the same electromagnetic layer and in the front and back layers.

itself (E_J^{par}). The calibrated energy of the jet J (E_J^{rec}) is defined as the sum of the cell energy deposits multiplied by the weight:

$$E_J^{rec} = \sum_{i \subset J} w_i(E_J^{par}, E_i)E_i \tag{5.1}$$

E_J^{par} is the particle jet energy.

With this procedure we want to correct the cell energy to recover the correct calibration, hence the linearity of the calorimeter response. Moreover this method leaves space to improve the calorimeter resolution. These considerations suggest to determine the weights by minimizing the dispersion of the recosntructed jet energies with the constraint of the linearity. The procedure is the following: many two jets (QCD) Monte Carlo events are analyzed using equation 5.1 to reconstruct the jet energy, a χ^2 is then build (equation 5.2) and minimized with respect to the weights w_i. The minimization is constrained by the linearity (equation 5.3). In order to obtain the dependence of w_i of the true jet energy, this process is done in narrow bins of E_J^{par}.

The minimized function is:

$$\chi^2 = \sum_e \frac{(E_{J,e}^{rec} - E_{J,e}^{par})^2}{(E_{J,e}^{par})^2} \tag{5.2}$$

where e runs over all the events and the linearity constraint is:

$$\sum_e (E_{J,e}^{rec} - E_{J,e}^{par}) = 0 \tag{5.3}$$

Since $e/h > 1$ we expect that the $w_i > 1$ to compensate for the hadronic components of the jet.

5.3.1 Cell clustering

The typical calorimetric clusters used as starting point for jet reconstruction algorithms are usually towers, obtained by grouping cells belonging to the same $\Delta\eta \times \Delta\phi$ regions.

To minimize the electronic noise term, only the cells with an appreciable signal ($E > \delta_{cut}$) can be used in the jet energy calculation algorithm. This however introduces a positive bias in the jet energy calculation since cells with a positive noise fluctuation are included, while negative noise fluctuation does not count.

To overcome this problem a clustering algorithm has been developed that groups together neighbouring cells if $|E| > \delta_{cut}$. The use of absolute value reduces the positive bias since also the negative over-fluctuating noise cells will be included.

The clustering starts from a seed, i.e. a cell with a ratio $|E|/\sigma_{noise}$ greater

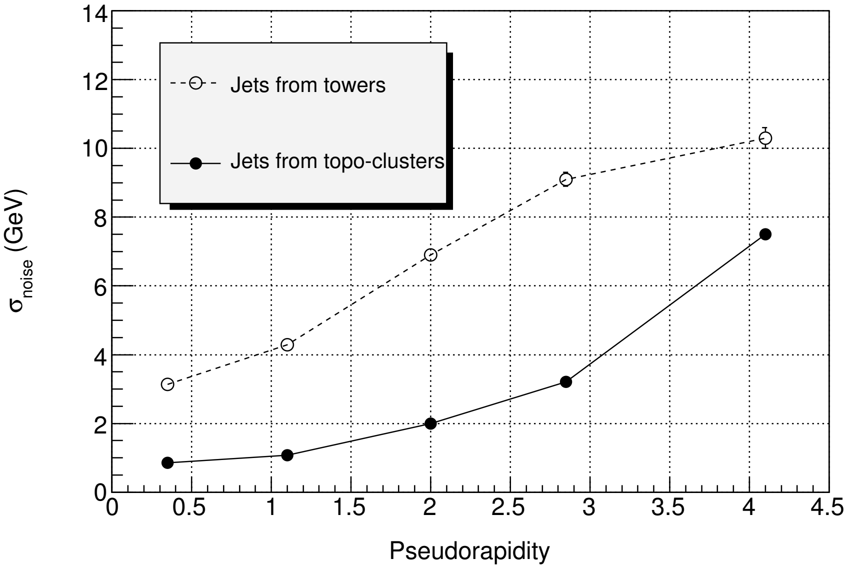


Figure 5.5: Contribution of noise (measured in GeV) to the jets reconstructed from calorimetric towers (open circles) and from topological clusters (full circles) as a function of pseudorapidity.

than N_{seed}, all the neighbouring cells are included in the cluster if $|E|/\sigma_{noise}$ is greater than N_{neig}.

The procedure is then repeated until the cluster is "stable", that is no now cells are added. A further step is performed: the condition $|E|/\sigma_{noise} > N_{out}$ is checked on all the cells surrounding the cluster and, if satisfied, these cells are included. This algorithms is known as "topological clustering" and the clusters are called "topo-clusters".

The neighbours are searched for not only in the $\eta - \phi$ plane but also in different calorimeter sampling, leading to clusters that span over the three dimensions (η, ϕ, r).

It can happen that two or more clusters overlap. In this case a split-and-merge procedure is used. The energy deposit in a cell belonging to two different cluster is assigned to both clusters dividing it proportionally to the distance of the cluster centroid.

This algorithm produces a list of clusters that are used as input for the jet reconstruction algorithm.

The topological clustering algorithm is fully defined by the triplet of parameters $(N_{seed}, N_{neig}, N_{out})$ that are tuned on single particles to minimize the contribution of noise on resolution, ATLAS choice for the three parameters is (4,2,0).

An example of a topo-cluster from a 120 GeV single pion obtained at the combined electromagnetic-hadronic end cap test beam is shown in figure 5.4 where the different layers of the calorimetric system are presented in the $r - \phi$ plane (corresponding to $\eta - \phi$ coordinates). The coloured cells belong to the topo-clusters and the seeding cells are well visible in the EMEC Middle layer.

Figure 5.5 shows the contribution of noise to the jet energy as a function of the jet pseudorapidity. The noise contribution has been evaluated by summing for each cell belonging to the jet only the electronic noise contribution. A gaussian fit has been performed on the resulting energy for jets constructed both from calorimetric towers and topo-clusters. Figure 5.5 shows that topological clusters are able to reduce noise by a factor 3. The increase of noise at large η is due to the fact that the technologies used in the calorimeters at small angles have higher noise.

The number of cells that makes a cluster is reduced by topological clustering: figure 5.6 (top-left plot) shows the number of cells composing the most energetic jet in QCD di-jet events reconstructed from topological clusters versus the number of cells composing the same jet reconstructed from calorimetric towers. The topological clustering algorithm efficiently suppress those calorimetric regions containing only electronic noise, thus jets reconstructed from topological clusters contain about 5 times less cells and noise is efficiently suppressed.

The average number of clusters composing a jet is about 20 (top-right plot) with the leading cluster carrying on average about 30% of the total jet energy (bottom plot).

5.3.2 Definition of η regions

The ATLAS calorimetric system is composed of calorimeters which use different technologies and have different segmentation and cover different pseudorapidity regions. This has the benefit of choosing radiation-hard material where it is mostly needed, at the price of a more complex mechanical system and more difficult calibration method.

Due to the different materials, thus different e/h values, in the various regions and the presence of cracks and dead material, we expect that the calorimeter response (defined as the ratio betweeen the measured energy in the calorimeter and the true energy) to hadrons depends on the pseudorapidity. This is clearly visible in figure 5.7.

Calorimeter response, at the EM scale (before any correction is applied), to jets of energy between 550 and 750 GeV as a function of pseudorapidity is shown. As expected the response is smaller than one due to non-compensation and cracks. In the figure, the boxes schematically identify the different calorimetric systems. Each box shows the coverage of the cor-

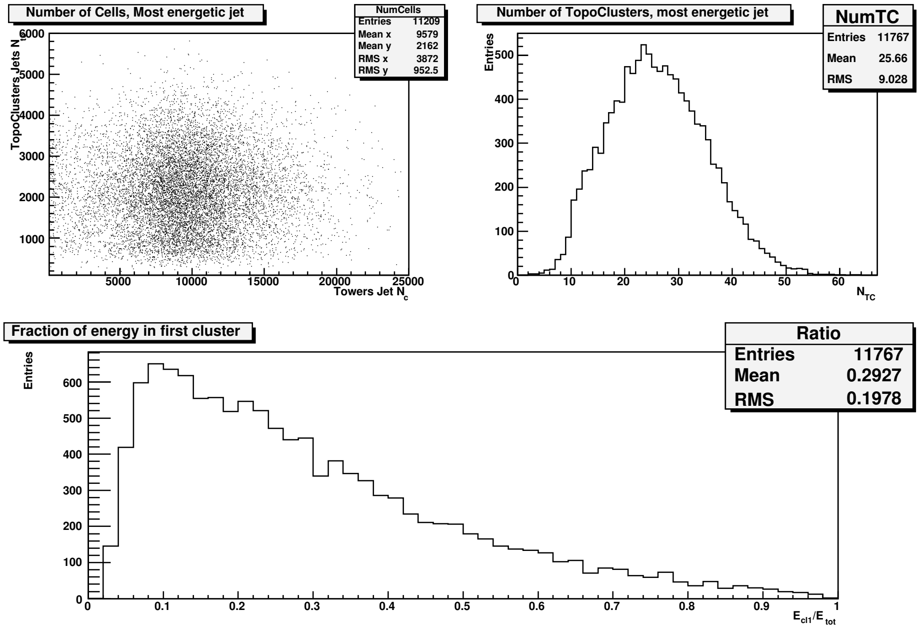


Figure 5.6: Number of cells composing the jet built from topo-clusters versus number of cells for the same jet built from calorimetric towers (upper left). Number of clusters composing the jet (upper right). Fraction of the energy carried by the most energetic cluster (bottom). All plots refer to the most energetic jet in QCD di-jet events with 280 GeV/c< P_T <560 GeV/c.

responding calorimeter, for example TileCal Barrel modules extend from $\eta = 0$ to $\eta = 1.0$. The position of the cracks between calorimeters is well visible in this plot. The most important cracks are between barrel and end-cap regions at about $\eta = 1.5$ and between end-cap and forward regions at $\eta = 3.4$.

To improve the performances of the calibration algorithm we will follow the geometrical partitions of the calorimetric system, dividing the entire η range in sub-regions. A set of weighting function will be extracted separately for each region. The red vertical lines of figure 5.7 represent these partitions that are also summarised in table 5.1.

5.3.3 Weight calculation

The large number of cells (about 200000) of the ATLAS calorimetric system and the limited statistic available impose to reduce the number of parameters w_i of equation 5.1. In principle, in fact, we can have a different w_i for each specific cell i.

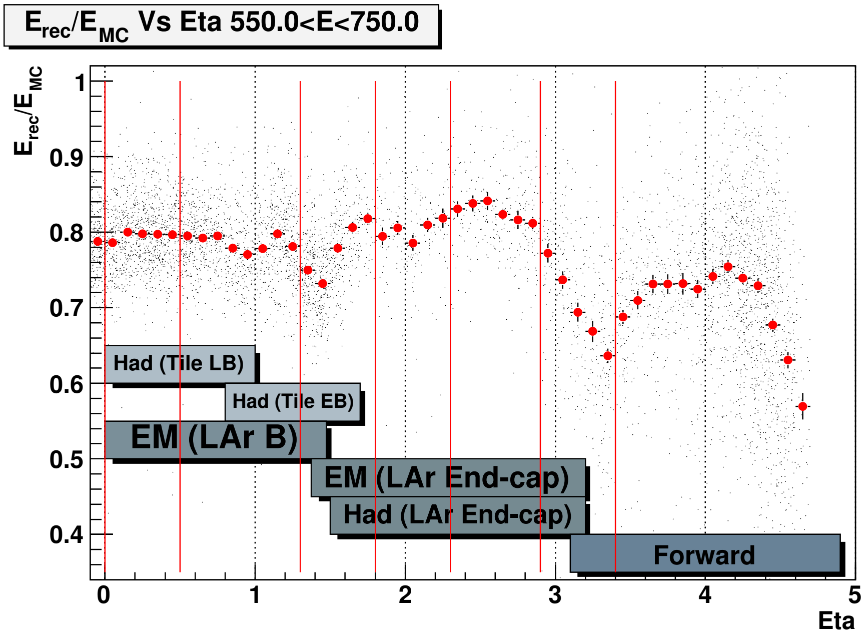


Figure 5.7: Response E_{rec}/E_{MC} as a function of pseudorapidity. E_{MC} is the true energy as defined in 5.3 and E_{rec} is the response to the jet at the electromagnetic scale. Points represent profile of the distribution. Boxes indicate the coverage of the different calorimeters.

The first reduction on the number of parameters assumes that all the cells belonging to the same longitudinal sample have the same weight. Some longitudinal samples are also grouped together to obtain a more uniform radial thickness.

Moreover the weight applied to cells belonging to the presampler and to the first sample of electromagnetic calorimeters is set to 1 as most of the electromagnetic component of a jet is absorbed here and the calorimetric system is calibrated at the electromagnetic scale.

For each pseudorapidity region table 5.2 lists the samples (or groups of them) that are used to estimate the weights. The weights of all other samples are set to 1.

The weights depend on the cell energy E_i. We have divided the energy spectrum of the cells in bins. The bins are chosen to be almost equally populated. Cells in the same bin b have the same weight $w_{s,b}(E_J^{par}, \mathbb{E}_b)$.

The center values of the 17 bins $\mathbb{E}_b$ are represented in figure 5.8. The first bin $\mathbb{E}_0$ goes from 0 to 200 MeV. The last bin is used for cells with $E_i > 200$ GeV.

Region		Calorimeters
I	$\lvert\eta\rvert \leq 0.5$	LAr Barrel, TileCal Barrel
II	$0.5 < \lvert\eta\rvert \leq 1.3$	LAr Barrel, TileCal Barrel and TileCal Extended Barrel
III	$1.3 < \lvert\eta\rvert \leq 1.8$	LAr Barrel, LAr end-cap, TileCal Extended Barrel
IV	$1.8 < \lvert\eta\rvert \leq 2.3$	Electromagnetic and Hadronic end-cap
V	$2.3 < \lvert\eta\rvert \leq 2.9$	Electromagnetic and Hadronic end-cap
VI	$2.9 < \lvert\eta\rvert \leq 3.4$	Electromagnetic and Hadronic end-cap, Forward calorimeter
VII	$3.4 < \lvert\eta\rvert$	Forward calorimeter

Table 5.1: Calorimetric regions: each region is calibrated separately. The calorimeters technologies used in the different regions are showed in the table.

With these simplifications equation 5.1 becomes:

$$E_J^{rec}(EM) = \sum_s \sum_{b=0}^{17} E_{s,b} \tag{5.4}$$

$$E_J^{rec} = \sum_s \sum_{b=0}^{17} w_{s,b}(E_J^{par}, \mathbb{E}_b) E_{s,b} \tag{5.5}$$

with:

$$E_{s,b} = \sum_{\mathbb{E}_{b-1} < E_i < \mathbb{E}_b, i \in s} E_i \tag{5.6}$$

where the index s runs on all the calorimeter samples and the index b runs on all the cell energy bins.

The minimization function 5.2 and the linear constraint 5.3 becomes:

$$\chi^2 = \sum_e \frac{(\sum_s \sum_{b=0}^{17} w_{s,b}(E_{J,e}^{par}, \mathbb{E}_b) E_{s,b,e} - E_{J,e}^{par})^2}{(E_{J,e}^{par})^2} \tag{5.7}$$

$$0 = \sum_e (\sum_s \sum_{b=0}^{17} w_{s,b}(E_{J,e}^{par}, \mathbb{E}_b) E_{s,b,e} - E_{J,e}^{par}) \tag{5.8}$$

The number of weights is still too large (for the pseudorapidity region I we have 3 weighted samples for a total of $17 \times 3 = 51$ weights function of the jet energy). The weights $w_{s,b}$ are parametrized as a function of the cell energy bin $\mathbb{E}_b$ and E_J^{par}. We adopted the following parametrisation that uses only two independent functions of the jet energy to describe the weight dependence on $\mathbb{E}_b$ [85]:

$$w_{s,b}(E_J^{par}, \mathbb{E}_b) = A_s(E_J^{par}) + \frac{B_s(E_J^{par})}{\mathbb{E}_b} \tag{5.9}$$

Region	Eta	Weighted samples
I	$\|\eta\| \leq 0.5$	EMB2+EMB3, Tile1+Tile2, Tile3
II	$0.5 < \|\eta\| \leq 1.3$	EMB2+EMB3, Tile1+Tile2, Tile3
III	$1.3 < \|\eta\| \leq 1.8$	EMB2+EMB3, EME2+EME3, Tile1+Tile2+Tile3, HEC1+HEC2+HEC3+HEC4
IV	$1.8 < \|\eta\| \leq 2.3$	EME2+EME3, HEC1, HEC2+HEC3+HEC4
V	$2.3 < \|\eta\| \leq 2.9$	EME2+EME3, HEC1, HEC2+HEC3+HEC4+ +FCAL1+FCAL2+FCAL3
VI	$2.9 < \|\eta\| \leq 3.4$	EME2+EME3, HEC1, HEC2+HEC3+HEC4+ +FCAL1+FCAL2+FCAL3
VII	$3.4 < \|\eta\|$	EME2+EME3+HEC1+ +HEC2+HEC3+HEC4, FCAL1, FCAL2+FCAL3

Table 5.2: Calorimetric samples that have been weighted in each eta region. The symbol “+” indicates that the two samples have been associated in one sample.

The dependence of the parameters $A_s(E_J^{par})$ and $B_s(E_J^{par})$ from the jet energy E_J^{par} is computed minimizing 5.7 with the constraint 5.8 in bins of E_J^{par}. Finally a set of discrete values A_s^k and B_s^k (index k is the true energy bin) are obtained and are fitted to determine the final parametrisation for the functions A_s and B_s as described in the next section.
The “true” jet energy E_J^{par} is known in a Monte Carlo run, but of course is not known in the real data.
When the calibration procedure is applied to data, an iterative numerical method is used. We start with the uncalibrated jet energy (E_J is computed at the electromagnetic scale) to compute the weights and a first approximation of the calibrated jet energy $E_J^{'rec}$. This value is then used to calculate a new set of weight and a new jet energy $E_J^{''rec}$. The procedure is repeated until a stable solution is found ($E_J^{(n+1)rec} - E_J^{(n)rec} < 0.1\% E_J^{(n)rec}$). The iteration typically stops after 4-5 iterations. In section 5.3.4 the implications of using an iterative method will be discussed.

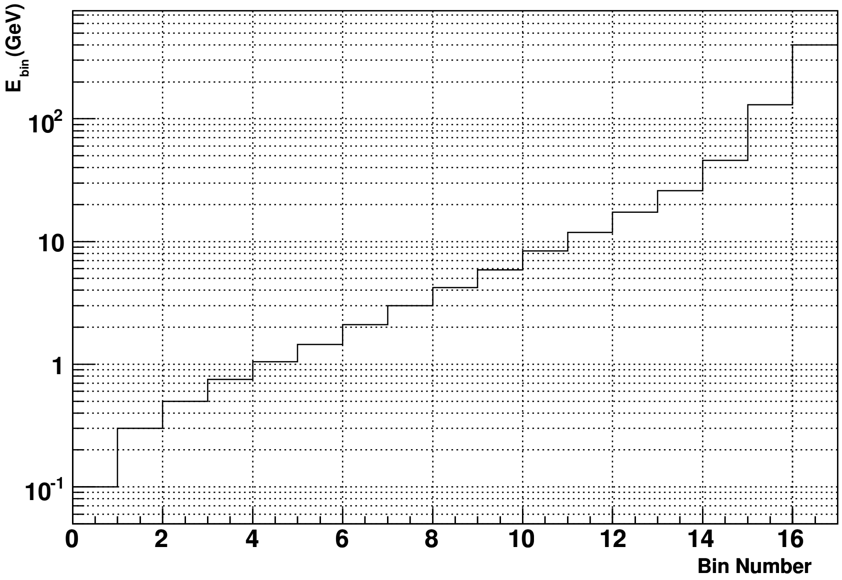


Figure 5.8: Central energy values of the bins used for the cell energy deposits.

Weights

Weights are computed from di-jet QCD samples reconstructed with the cone algorithm with $\Delta R = 0.7$. For each reconstructed jet, the true energy E_J^{par}, to be used in 5.2, is computed as the sum of the energies of the Monte Carlo particles entering in the cone used to reconstruct the jet.

Figure 5.9 shows the dependence of the parameters A_s and B_s on the jet true energy for the middle and back samples of the electromagnetic barrel calorimeter (jets in region II of table 5.2: $0.5 < |\eta| < 1.3$). Figure 5.10 shows, for the same samples, the value of the weight as a function of jet energy and cell deposits. The dependence of A_s and B_s on the "true" energy is rather mild and coefficients tend to a constant at large jet energies.

Both A and B, fitted with a function of the form $\sum_i (a_i/E^i)$, saturates, for large jet energies, to a value greater than 1. In fact the weight corrects for the calorimetric e/h value ($e/h > 1$) and for dead material losses.

5.3.4 Limitations

Low E_T regime

The functional form chosen to describe the A_s and B_s dependence on the jet energy fits well the region at large jet energy but it is not adequate at small energies. This is visible in figure 5.9 where the values of A_s and B_s obtained

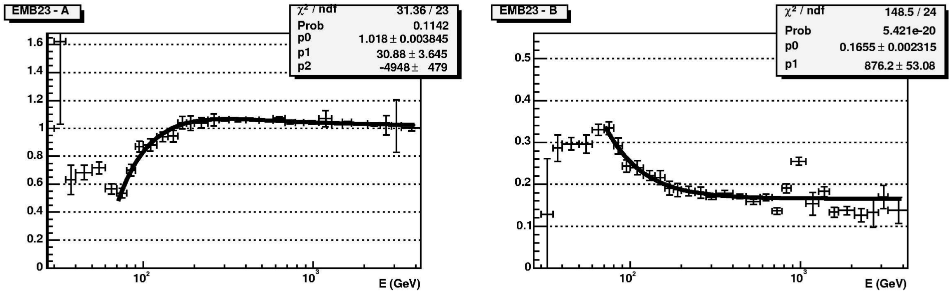


Figure 5.9: The parameters $A_{EMB23}(E_J)$ (left) and $B_{EMB23}(E_J)$ (right) as a function of jet true energy for the calorimetric region II ($0.5 < |\eta| < 1.3$).

for jets with $E < 100$ GeV do not agree with the fit. The behaviour can be observed in all the η regions, the lower value for which the fit reproduces the the computed values of A_s and B_s depends on η and always roughly corresponds to $\mathrm{E_T} \simeq 20 - 30$ GeV.
Jets with $E_T < 20$ GeV are thus calibrated using the weights corresponding at $E_T = 20$ GeV, the effect of this is a less precise determination of the calibrated jet energy causing worse linearity and resolution at low energies, however the results obtained with this simplified approach are still satisfactory as it will be shown in 5.3.5.

Iterative solution limitation

The parameters A_s and B_s are obtained as functions of the true jet energy. However since the true jet energy is not known in real data, an iterative procedure is used, starting from the jet energy at the electromagnetic scale as described in 5.3.3. For each jet J, this procedure correspond to searching the solution E_J^{par} of equation 5.1.
For sake of simplicity and let us write the equation in the compact form:

$$E_J = f_J(E_J) \tag{5.10}$$

This equation is solved by iteration, starting from the jet energy at the electromagnetic scale. The sequence of approximate solutions E_i converges to the exact (and unique) solution E_J if [88]$f_J(E_J)$ has a derivative less than one:

$$|f'_J(E_J)| < 1 \tag{5.11}$$

in the interval $[E_{min}, E_{max}]$ and if the starting value (in our case the jet energy at electromagnetic scale) lies inside this interval. This condition is sufficient.

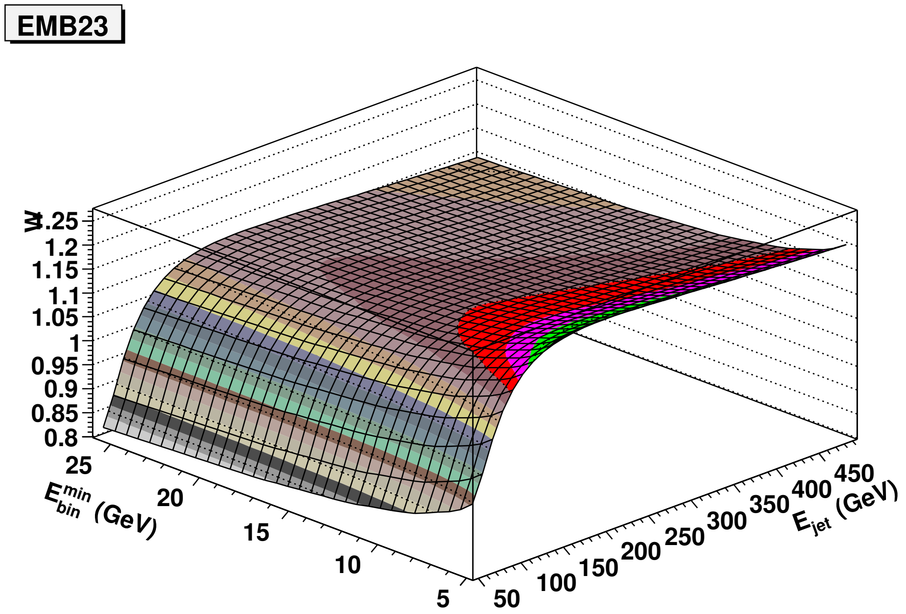


Figure 5.10: Weight of sample EMB2+EMB3 for the calorimetric region II ($0.5 < |\eta| < 1.3$) as a function of jet energy and cell energy.

The functional form of A and B (polynomials in $1/E$) guarantees the derivability of f_J for $E_J > 0$ and E_{max} can be chosen as large as needed. However a lower limit E_{min} exists, below which, the condition 5.11 is not any longer true.

Figure 5.11 shows the quantity f'_J as a function of jet $\log(\mathrm{E_T})$ for the calorimetric region II. For each jet the calculation of f'_J at the solution E_J is drawn, red points represent the profile. For jets transverse energy below $\log(\mathrm{E_T}) < 1.4$ ($\mathrm{E_T} \simeq 20 - 25$ GeV) the condition 5.11 is no more satisfied and we can not assume that the iterative procedure converges to the exact solution. In general this will cause a worsening of the linearity and resolution for energies below this limit since some jets will not be calibrated to the correct energy.

This limit has been checked for every region and we have found that, in all cases, it is below (or very close to) $\mathrm{E_T} \simeq 20 - 30$ GeV, that is already the validity limit of the parameterisations for A and B.

5.3.5 Results on QCD samples

In order to evaluate the performances of the calibration algorithm we have applied the procedure to different data samples.

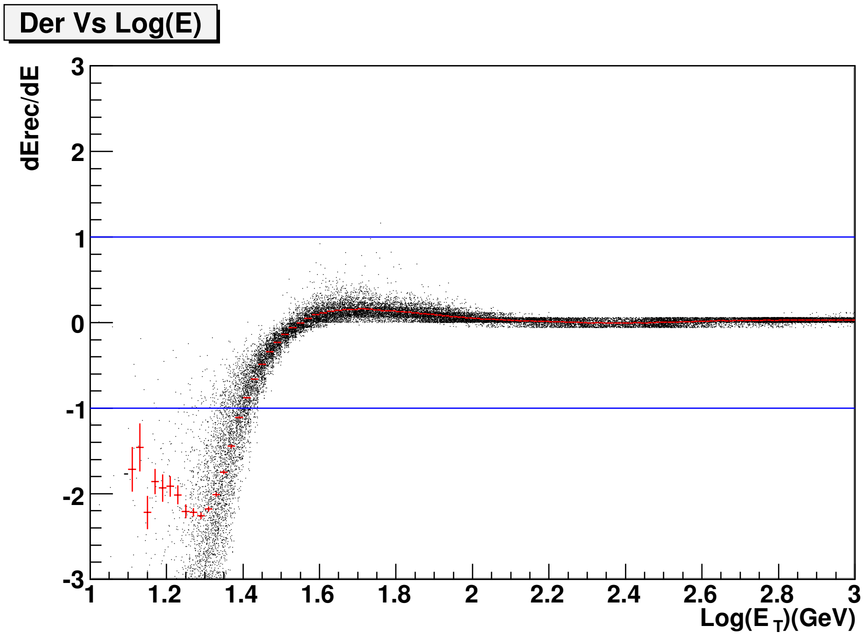


Figure 5.11: The quantity (dE_J^{rec}/dE_J) for region II as a function of jet log($\mathrm{E_T}$).

Figure 5.12 shows, for calorimetric region I ($|\eta| < 0.5$) and for four different jet energy bins, the effect of the calibration procedure on jets from QCD di-jet events reconstructed with $\Delta R = 0.7$.

The ratio E_T^{rec}/E_T^{par} between the reconstructed jet transverse energy and the true $\mathrm{E_T}$ (called linearity) is shown before (blue) and after (red) the calibration has been applied. Linearity and resolution are then computed, for all the jet energy bins. This is done with a gaussian fit limited to a $\pm 2\sigma$ range around the mean.

Figure 5.13 shows the linearity and resolution for jets in region I as a function of $\mathrm{E_T}$.

Linearity is recovered at the level of 2% on a wide energy range (40 GeV – 3 TeV) and we observe an improvement in $\mathrm{E_T}$ resolution for all jet transverse energies above 30-40 GeV.

The $\mathrm{E_T}$ resolution has been evaluated fitting the results with the equation:

$$\frac{\sigma(E_T)}{E_T} = \frac{a}{\sqrt{E_T}} \oplus b \oplus \frac{c}{E_T} \tag{5.12}$$

The results from the fit for both the electromagnetic and the calibrated scales are summarised in table 5.3 for all the seven calorimetric regions. The linearity is recovered at the level of 2% in all the regions.

An improvement in the sampling term of about 20-30% is achieved on all the calorimetric regions. For region VII only a and b are used in the fit. The

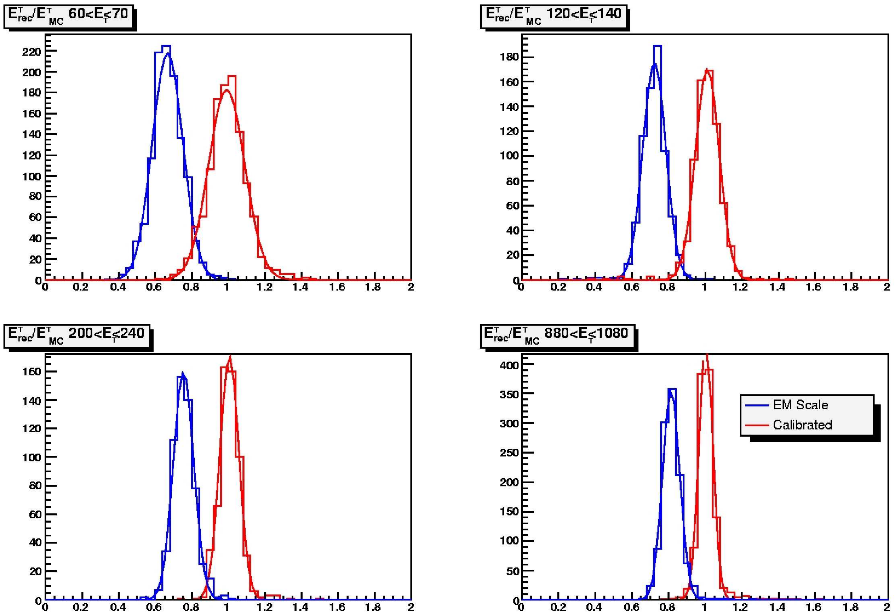


Figure 5.12: E_T^{rec}/E_T^{par} ratio for jets in region I. Blue histograms is obtained for jets at the electromagnetic scale, while red histograms are obtained after the calibration has been applied. Four different transverse energy bins are shown: 60 GeV $< E_T^{par} <$ 70 GeV (top left), 120 GeV $< E_T^{par} <$ 140 GeV (top right), 200 GeV $< E_T^{par} <$ 240 GeV (bottom left), 880 GeV $< E_T^{par} <$ 1080 GeV (bottom right). In all cases the improvement in linearity and $\mathrm{E_T}$ resolution is clearly visible.

contribution of the noise term c to the $\mathrm{E_T}$ resolution is negligeable at high values of the pseudo-rapidity η.

As previously discussed, jets at low $\mathrm{E_T}$ are calibrated with a simplified parametrisation: weights are calculated for $\mathrm{E_T}$ = 20 GeV and used for all jets with $\mathrm{E_T}$ < 20 GeV. We expect that in this region the performances will be poorer than at larger energies. In fact figure 5.14 shows in detail the linearity for the low transverse energy regime in the region I. The linearity is recovered at the level of ±5% for $E_T < 35$ GeV.

So far we presented results obtained on jets reconstructed with a cone algorithm and $\Delta R = 0.7$ however, in many ATLAS physics studies such as the one that will be discussed in the next chapter (top quark hadronic decay), jets are reconstructed using a cone size of $\Delta R = 0.4$ because in such events the jet multiplicity is larger and a narrower cone is needed to distinguish the different jets.

We have reconstructed the di-jet QCD sample with the cone algorithm using $\Delta R = 0.4$. For the energy calibration we have used the weighting param-

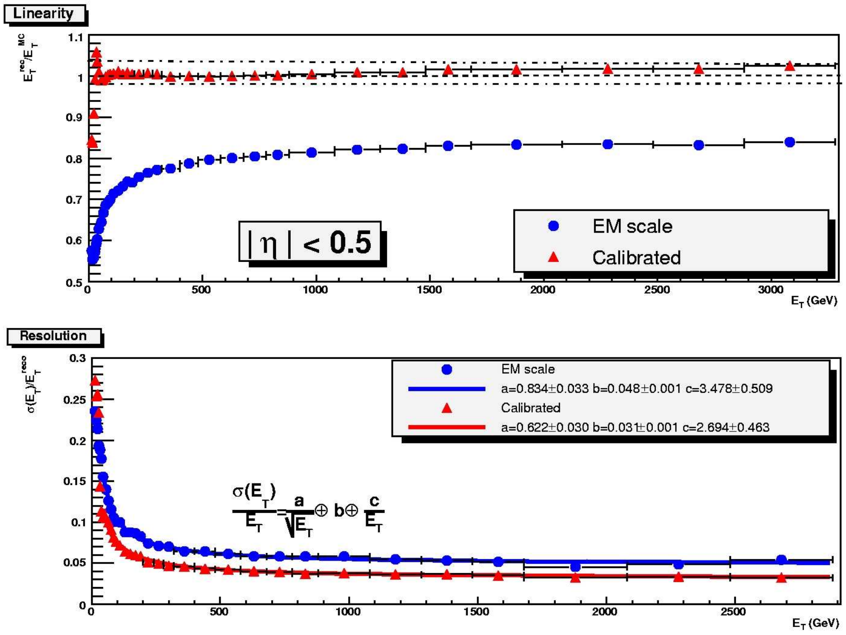


Figure 5.13: Linearity (E_T^{rec}/E_T^{par}) and transverse energy resolution ($\sigma(E_T^{rec})/E_T^{rec}$ obtained for region I. Blue points refer to the electromagnetic scale, red points to the calibrated scale. Results from fit with eq. 5.12 of resolution are also reported. Jets are reconstructed with $\Delta R = 0.7$.

eters found with $\Delta R = 0.7$ cone jets. Figure 5.15 shows the linearity and $\mathrm{E_T}$ resolution plot for the first pseudo-rapidity region.
The weights still work well. The reason is that most of the correction is done locally, at the cluster level and the procedure is not sensitive to global features of the analysis. The linearity is recovered at the level of ±2% again from E_T =30 GeV up to E_T=3 TeV. In table 5.4 the results of the fit to the transverse energy resolution, both at the electromagnetic and calibrated scales are reported. The transverse energy resolution is a bit worse than that obtained with $\Delta R = 0.7$. However the calibration procedure improves the measured $\mathrm{E_T}$ resolution (improvement of 30%) with respect to the electromagnetic scale. Linearity is restored to the level of 2% in all regions.

5.3.6 Results on $t\bar{t}$ sample

In next chapter we will discuss the application of this calibration method to a physic study: the measurement of top quark mass with the data collected in the first period of data taking. We sill analyze semi-leptonic $t\bar{t}$ events ($t\bar{t} \to \mathrm{W^+W^-b\bar{b}} \to \ell^{\pm}\mathrm{jjb\bar{b}}$) with four jets (two of which produced in the

Region	EM scale			Calibrated scale		
	a (GeV^1/2)	b	c (GeV)	a (GeV^1/2)	b	c (GeV)
I	83%	4.8%	3.5	62%	3.2%	2.7
II	79%	5.3%	3.5	65%	3.5%	3.6
III	80%	6.0%	2.4	42%	4.6%	3.6
IV	52%	3.8%	3.2	36%	2.7%	2.9
V	43%	3.8%	2.3	27%	3.5%	3.1
VI	73%	7.7%	1.0	59%	4.8%	2.7
VII	67%	1.4%	-	50%	4.0%	-

Table 5.3: Values of parameters of 5.12 (electromagnetic and calibrated scale) for the seven different eta regions obtained on jets reconstructed with $\Delta R = 0.7$. For region VII only the first two terms where used for the fit.

Region	EM scale			Calibrated scale		
	a (GeV^1/2)	b	c (GeV)	a (GeV^1/2)	b	c (GeV)
I	88%	5.0%	3.1	70%	3.1%	1.7
II	87%	5.3%	3.2	66%	3.5%	4.4
III	85%	6.3%	3.1	55%	4.6%	3.9
IV	59%	3.9%	2.4	41%	2.7%	2.6
V	47%	4.4%	3.4	21%	4.1%	4.2

Table 5.4: Values of parameters of 5.12 (electromagnetic and calibrated scale) for the first five different eta regions obtained on jets reconstructed with $\Delta R = 0.4$. The results correspond to a local production of the first five calorimetric zones.

fragmentation of the b-quarks) in the final state. We will not discuss the special calibrations that should be applied to b-jets and we will apply the same calibration functions to all jets.

The analysis is done using a cone algorithm with $\Delta R = 0.4$ (see below) and using our scheme of jet energy calibration. The weights used are the same discussed above and computed on a QCD di-jet sample with $\Delta R = 0.7$.

To improve jet reconstruction performances a cone size $\Delta R = 0.4$ is used in $t\bar{t}$ events. These events present high jet multiplicity and a reduced cone size, with respect to standard $\Delta R = 0.7$, reduces the overlap between jets improving the correct jet number reconstruction and thus improving the selection efficiency, since, as we will discuss in section 6.3.1, semi-leptonic $t\bar{t}$ events are selected requiring exactly four reconstructed jets in the final state.

While the details of the analysis will be discussed in the next chapter, here we insist on the performances of our calibration method.

Figure 5.16, obtained from all jets with $|\eta| < 2.5$ for an integrated luminosity

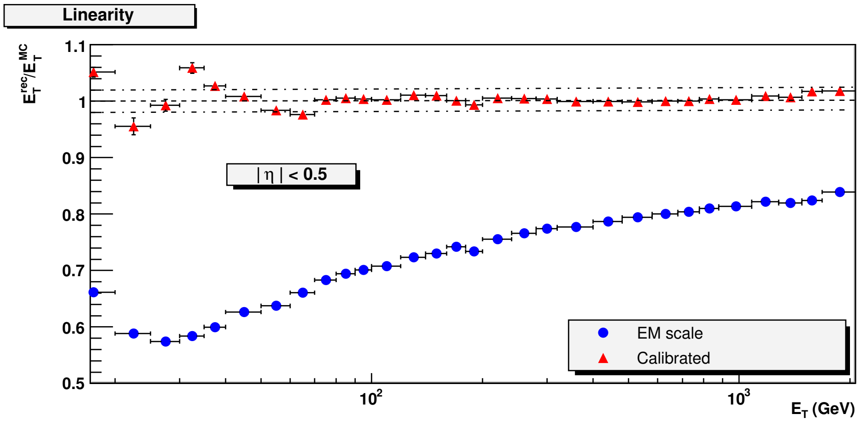


Figure 5.14: Linearity for jets in region I showing the low energy regime. The effect of simplified weighting scheme on linearity is visible with a degradation at the level of ±5%.

of 100 pb^{-1} , shows that linearity is restored at the level of 2-3% from E_T = 35 GeV up to 600 GeV (this limit is set by the available statistic). The ratio is systematically higher than 1: about 1-2% higher than the linearity obtained on QCD jets, we can notice that also at the electromagnetic scale linearity is higher of about the same amount, compared to QCD events. This indicates that the effect is not due to calibration procedure and it is probably due to the different event topology.
The transverse energy resolution has been fitted with the function:

$$\frac{\sigma(E_T)}{E_T} = \frac{a}{\sqrt{E_T}} \oplus b$$

and we obtained, after calibration, an improvement for the term a from 107% at the electromagnetic scale to 75% at the hadronic scale and an improvement for term b from 4% to 2.4%.
Thus the calibration algorithm developed shows good performances also on an independent physics sample and behaves robustly since calibration performances do not show strong dependence on the parton type which the the jet. In QCD events jets are primarily originating from gluons, while in $t\bar{t}$ events quark (both light and b) is the parent parton type for jets.

5.3.7 Conclusions

In this chapter we presented a jet calibration algorithm that improves the transverse energy resolution and restore the calorimeter linearity. The method uses Monte Carlo simulation to extract weights functions of the jet energy

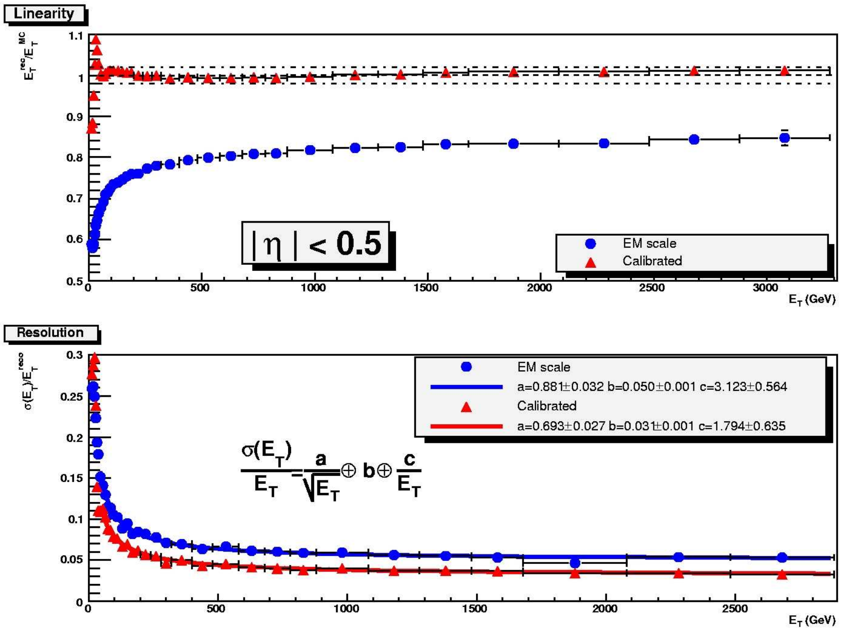


Figure 5.15: Linearity (E_T^{rec}/E_T^J) and resolution ($\sigma(E_T^{rec})/E_T^{rec}$) obtained for region I. Blue points refer to the electromagnetic scale. Results from fit with eq. 5.12 of resolution are reported. Jets are reconstructed with $\Delta R = 0.4$.

and cell energy deposit.

The results obtained on a QCD di-jet event samples and $t\bar{t}$ events from ATLAS Geant4 simulation has been presented and discussed.

It shows good performances on a wide $\mathrm{E_T}$ range improving, at the level of 2%, the jet transverse energy linearity and improving the resolution of about 20-30% in the central regions (10-20% for high pseudo-rapidity regions).

The low transverse energy regime ($\mathrm{E_T} < 30$ GeV) presents some peculiarities and more work is needed to fully understand the algorithm behaviour. However we have shown that a simplified algorithm can be used with adeguate results for this events. The $\mathrm{E_T}$ linearity is partially recovered at the level of 5%.

To further verify the performances of the jet calibration algorithm more detailed studies on physics channels are required. One of this study will be presented in the following chapter in which jets from the events $t\bar{t} \rightarrow \mathrm{W^+W^-b\bar{b}} \rightarrow \ell^{\pm}\mathrm{jjb\bar{b}}$ are calibrated with the proposed method and the top quark mass will be measured.

The calibration method performances can be further checked using the combined test beam data. A set of weights can be extracted from the Monte Carlo single pion simulations and then applied to the real data. Before this

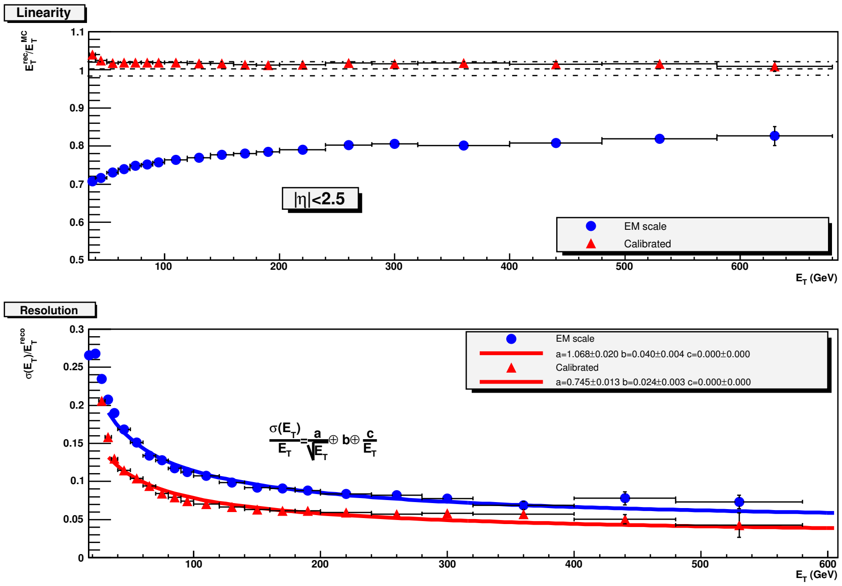


Figure 5.16: Linearity and resolution for jets coming from $t\bar{t}$ events in which one top decays totally hadronically.

study can be performed the understanding of the low energy data, acquired during the combined test beam, is mandatory since a jet is mainly composed of relatively low energy particles.

Chapter 6

Early top physics studies

The top quark is a fundamental parameter of the standard model and ATLAS has the goal to increase the precision in the top mass measurement with respect to available data from CDF and D0. In addition it is also the ideal signal for commissioning studies thanks to its clear signature and high production cross-section. Even with reduced detector performances its mass can be measured with a statistical error of 1 GeV already with and integrated luminosity $L = 275\ \mathrm{pb}^{-1}$*, achievable in some days of data taking. The absence of b-tagging makes the* $W + Njets$ *background important.*
In this chapter we will present the top quark and W *boson masses measurements using early data (the commissioning data). We will first describe the signal selection and reconstruction, the description of background processes and finally results without and with background will be presented. Exploiting the reconstructed* W *boson mass, a strategy to clean the sample, will be presented.*

6.1 Introduction

The top quark mass is a fundamental parameter of the standard model. Together with a small set of other quantities, its mass is used as an input parameter for the theoretical predictions of the electroweak observables. Even if the top mass is measured with a better precision than the other quarks, the current precision is the dominant effect on the theoretical uncertainties.

The first data available at ATLAS will be used for the commissioning and the calibration of the detector. The top production is an ideal laboratory for initial studies: in addition to the high production rate, the semi–leptonic channel is easy to trigger, thanks to the isolated lepton and the high missing energy coming from the neutrino. Even at low luminosity ($L = 10^{33}\mathrm{cm}^{-2}\mathrm{s}^{-1}$) more than 50 events per day after trigger and analysis selection can be recorded. The additional presence of four jets, two of which being b-quarks, makes this channel interesting for many detector performance studies related to lepton identification, jet reconstruction and calibration, missing energy and b-tagging. In conclusion top events will be very useful to give a prompt feedback on detector performances and as a calibration tool.

The events are selected requiring an isolated high $\mathrm{P_T}$ lepton, missing energy and at least four jets. The background, mainly coming from $W + jets$, $Z + jets$, $WW/ZZ/WZ$ events, is reduced to a negligible level by the selection cuts and by the request of one or more b-jet.

The W boson and top quark reconstructions are simple, for the hadronically decaying top, through the combination of two non-b (light) jets (to reconstruct the W boson) and the subsequent association with one b tagged jet (for the top quark reconstruction). The main background is due to wrong combinations in assigning the jets to the W boson or to the top quark.

One of the main uncertainty in the top mass measurement originates from the jet energy scale (see table 1.4): with the design detector performances, an error in the light jet energy scale of 1% causes a systematic top mass error of 0.3 GeV/c^2 . The error in the mass measurement increases to 0.7 GeV/c^2 for the b-jet energy scale error of 1%.

Light jets can be calibrated in top events with the required precision employing a W boson mass constraint. $Z + jet$ with the subsequent decay $Z \rightarrow ll$ and $\gamma + jet$ can be also used to perform *in-situ calibration*. However all this techniques will be available after detailed studies of the backgrounds and of elctromagnetic absolute scale and thus will require some time to be performed.

We have shown that the jet energy scale, with the use of ATLAS Monte Carlo simulations, can be set, in $t\bar{t}$ events, to better than 2-3% (see figure 5.16) over a wide energy range.

This result is supported by the Monte Carlo simulations (with all the caveats that we have discussed in the previous chapter) and from Test Beam data.

	Expected performance at “day 1”	Possible data sample to improve performances
E.M. calor. uniformity	1%	Minimum Bias, $Z \rightarrow ee$
Electron energy scale	1-2%	$Z \rightarrow ee$
Had. calor. uniformity	2-3%	Single pions, QCD jets
Jet energy scale	$\leq 10\%$	$Z + j \rightarrow ll + j$, $W \rightarrow jj$ (from $t\bar{t}$ events)
Tracker alignment	20-200 μm in Rϕ	isolated muons, $Z \rightarrow \mu\mu$

Table 6.1: Examples of expected detector performances for ATLAS detector at LHC startup and possible physics samples to be used to improve the performances.

We will start from this result and, in the rest of the chapter, we will investigate the possibility of a top quark mass measurement with the first collected data; we will show that the top quark can be clearly identified as a resonance in the mass spectrum of three jets and its mass can be measured even with limited detector performances, in particular we will extract the top mass without information from the b-tagging and without in-situ calibration. The precision on the determination of the top mass is not limited by statistics; its main contribution originates from the energy scale of the jets; corrections to precisely fix the energy scale will be done with the so called *in-situ* calibration and will be possible only after the detector performances are carefully understood.

6.2 ATLAS physics potential with the first data

The ATLAS commissioning period is well advanced and will continue until after data taking with proton-proton collisions has started. It may be divided in several phases, including sub-detector level calibrations and noise studies, recording of cosmic ray events, studies with beam-halo muons during the single beam runs and finally with the very first collisions to commission the trigger.

After this first debugging stage of the detectors and the trigger physics analysis can be performed in order to give a fast feed-back to the detector communities to assess the data quality.

6.2.1 LHC status and first data taking

According to present LHC schedule [89], the machine will be cooled down in September 2007, and will then be commissioned for a few months start-

Channel	recorded events for 1 fb^{-1}
$W \rightarrow \nu\mu$	7×10^6
$Z \rightarrow \mu\mu$	7×10^6
$t\bar{t} \rightarrow \mu + X$	8×10^4
QCD jets $\mathrm{P_T}$ >150 GeV/c	$\sim 10^6$ (10% trigger bandwidth)
minimum bias	$\sim 10^6$ (10% trigger bandwidth)
$\tilde{g}\tilde{g},\ m_{\tilde{g}} \sim 1$ TeV	$10^2 - 10^3$

Table 6.2: For some physics processes, the number of events expected to be recorded for an integrated luminosity of 1 fb^{-1}

ing with single-beam runs. The machine will be commissioned with reduced beam energy of 450 GeV and first colliding beams are expected by the end of 2007. These will be engineering runs and physics will be possible only in parasitic mode.

After a shut-down of few months, physics run will take place for about 7 months in 2008 rumping up luminosity to the nominal initial value of 10^{33} $\mathrm{cm}^{-2}\mathrm{s}^{-1}$ (low-luminosity regime).

There are several uncertainty on this plan, therefore it is assumed [90] that the integrated luminosity collected by the end of 2008 will range between a very modest 100 pb^{-1} and a very ambitious 10 fb^{-1}.

6.2.2 ATLAS status during first runs

ATLAS will start with two pixel layers (instead of three) and without Transition Radiation Tracker in the region $2 < |\eta| < 2.4$. In addition part of the high-level trigger and data acquisition processors will be deferred, with the consequence that the output rate of the LVL1 trigger will be limited to 35 kHz.

The impact of this staging on physics will be significant but not dramatic. The main loss is a descoped B-physics programme because, due to the reduced level-1 bandwidth, the single muons triggers will have to be raised.

Table 6.1 summarizes the expected performances of ATLAS at “day 1”, that is, at the moment when data taking starts. This is an estimation of what we will know of the detector before data taking, based on construction quality checks, on the known precision of the hardware calibration and alignment systems, on test-beam measurements and on simulation studies.

The performances should be significantly improved as soon as the first data will be available, and thanks to the large event rates expected at the LHC, the ultimate statistical precision should be achieved already after a few days/weeks of data taking [1].

The detailed study of the known SM processes will be possible already in the first data taking periods at LHC. In many cases, these processes offer themselves the potential for important measurements (for example the top quark mass measurement), and them will become background to non standard model physics.

Table 6.2 shows the data samples expected to be recorded in some physics processes for an integrated luminosity of 1 fb^{-1} . The trigger selection efficiencies have been included. The following goals can be addressed with such a data sample:

- Calibrating the detector and commissioning the trigger. $Z \rightarrow ll$ is a gold-plated process for a large number of studies to set the absolute electron and muon scales in the electromagnetic calorimeter and in the trackers, whereas $t\bar{t}$ will be used to establish the jet energy scale and to understand the b-tagging performance.
- Perform extensive measurements of the main SM physics processes, e.g. cross sections and event features for minimum-bias, QCD di-jet, W/Z bosons and $t\bar{t}$ production.
- Prepare the road to discoveries measuring the backgrounds to possible new physics channels. Processes like $W/Z + jets$, QCD multi-jet and $t\bar{t}$ are omnipresent backgrounds for a large number of searches and need to be understood in all details.

6.3 Top quark mass measurement with commissioning data

In the following we will concentrate in analyzing the $t\bar{t}$ events in the semi-leptonic channel ($t\bar{t} \rightarrow \mathrm{W^+W^-b\bar{b}} \rightarrow \ell^{\pm}\mathrm{jjb\bar{b}}$) to determine the top quark mass with the first collected data using the jet calibration algorithm that we have developed. Two main assumptions are made in the following: the jet energy scale is set by the weighting technique that we have described (i.e. no in-situ calibration) and no b-tagging information is used (*commissioning approach* [91]).

This conditions are very likely to be the case at the ATLAS startup.

Events used in the analysis come from the official ATLAS Monte Carlo production known as the “Rome” production. Data are reconstructed with the Athena (version 11.0.41) software and the analysis is performed on AOD

---

[1] Assuming a combined selection and trigger efficiency of 1.5% in $t\bar{t}$ semi-leptonic events a similar statistic to current CDF and D0 one can be collected in ATLAS within some days of data taking at 10% of low-luminosity ($L = 10^{32}\mathrm{cm}^{-2}\mathrm{s}^{-1}$).

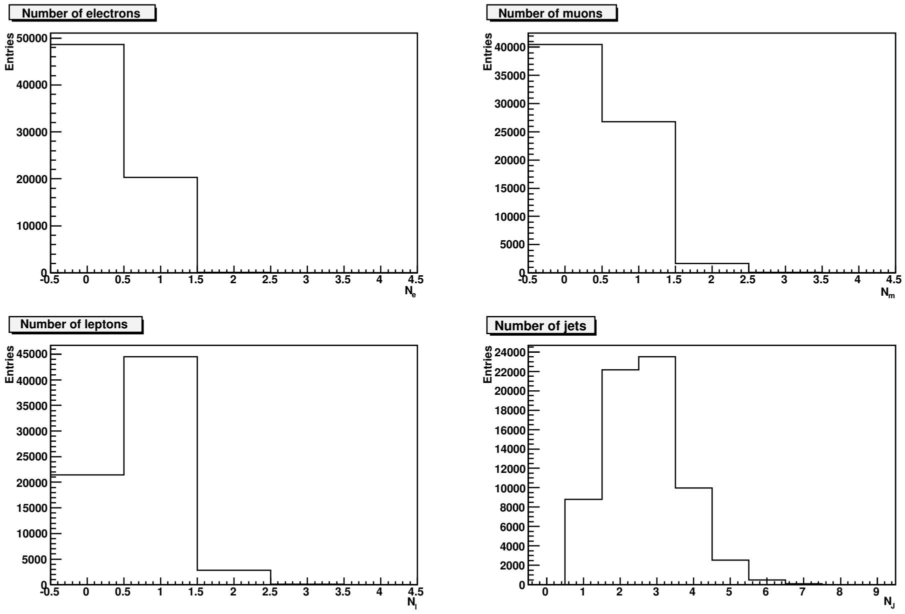


Figure 6.1: Multiplicities for pre-selected events: electrons (top left), muons (top right), generic lepton (bottom left) and jet multiplicity (bottom right).

data-sets using the EventView package.

Events are generated at the nominal top mass of 175 GeV/c^2 with the TopRex [92] Monte Carlo, interfaced to the PYTHIA shower and hadronization package. TopRex includes the polarization of the top quarks, in order to describe spin correlations.

Interaction of particles with the ATLAS detector is simulated with Geant4 simulation engine. To simulate the initial detector conditions the b-tagging has been switched off and all the jets are reconstructed as light jets and calibrated with the same scheme.

6.3.1 Events selection

The events are pre-selected requiring an isolated lepton and high missing transverse energy. To reduce the combinatorial background the events are required also to have exactly 4 jets.

Multiplicity distributions for $t\bar{t}$ events are shown in figure 6.1: top-left plot shows the number of electrons in the event, while the top-right plot shows corresponding distribution for muons. The multiplicity of leptons (muon and electron with $\eta < 2.5$ and $\mathrm{P_T} > 20$ GeV/c) is shown in the bottom-left plot, no tau lepton selection is performed. The number of jets in the event is

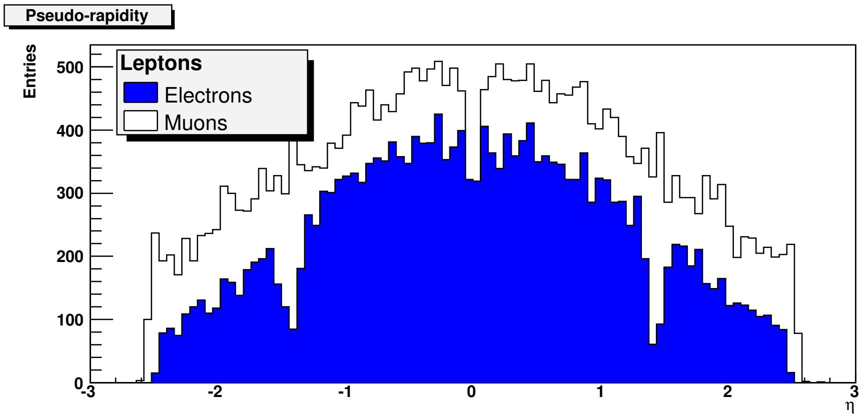


Figure 6.2: Pseudo-rapidity distribution of selected leptons.

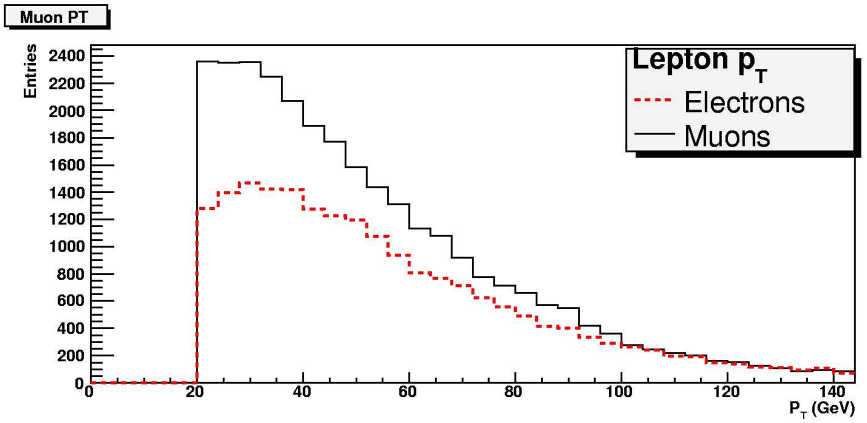


Figure 6.3: $\mathrm{P_T}$ distributions for selected electrons and muons.

shown in the bottom right distribution.

Lepton selection

Events are selected requiring one isolated lepton with $\mathrm{P_T} > 20$ GeV/c and $|\eta| < 2.5$.

The isolation criteria requires $E_T^{iso} < 15$ GeV in a $\Delta R = 0.45$ cone around a selected electron, or $E_T^{iso} < \mathrm{E_T}(\mu)$ in the cone around a muon candidate. Lepton selection is refined requiring shower shape cuts (for electrons) in the electromagnetic calorimeter and matching direction with inner-detector information (for both electrons and muons).

Figure 6.2 shows the pseudo-rapidity distribution for muons and electrons in $t\bar{t}$ signal events. The absolute number of selected muons is larger than

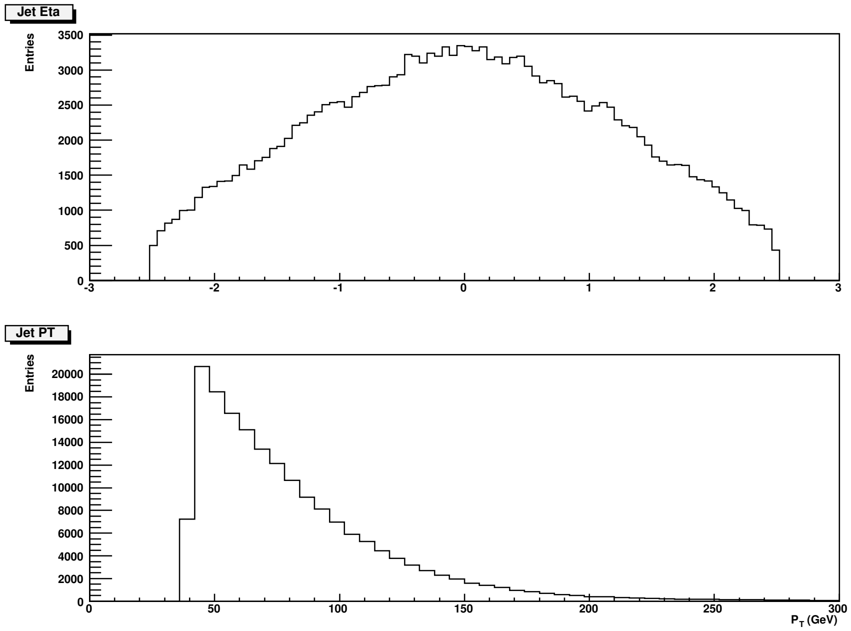


Figure 6.4: Jets η (top) and $\mathrm{P_T}$ (bottom) distributions for pre-selected $t\bar{t}$ events.

the corresponding number of electrons, indicating higher efficiency for muon reconstruction and selection. The acceptance cracks are also visible: around $\eta = 0$ (especially for the muon system) and the two barrel–end-cap electromagnetic calorimetric cracks around $\eta = 1.5$.

The $\mathrm{P_T}$ distributions for selected muons and electrons are shown in figure 6.3.

Jet selection

Jets are reconstructed from topo-clusters (topological clustering is defined in 5.3.1) using a cone algorithm with $\Delta R = 0.4$. After the split and merge procedure the jets are calibrated using the calibration scheme that we have discussed in chapter 5.

Jets are accepted if $\mathrm{E_T} > 40$ GeV and $|\eta| < 2.5$. Overlap with e/γ is removed requiring no reconstructed electrons or photons around jet-axis in a cone of $\Delta R = 0.5$.

Distributions of η and $\mathrm{P_T}$ for selected jets are shown in figure 6.4. The $\mathrm{P_T}$ cut at 40 GeV/c has visible effect in the transverse momentum distribution.

Events with exactly four jets are accepted and used for reconstruction.

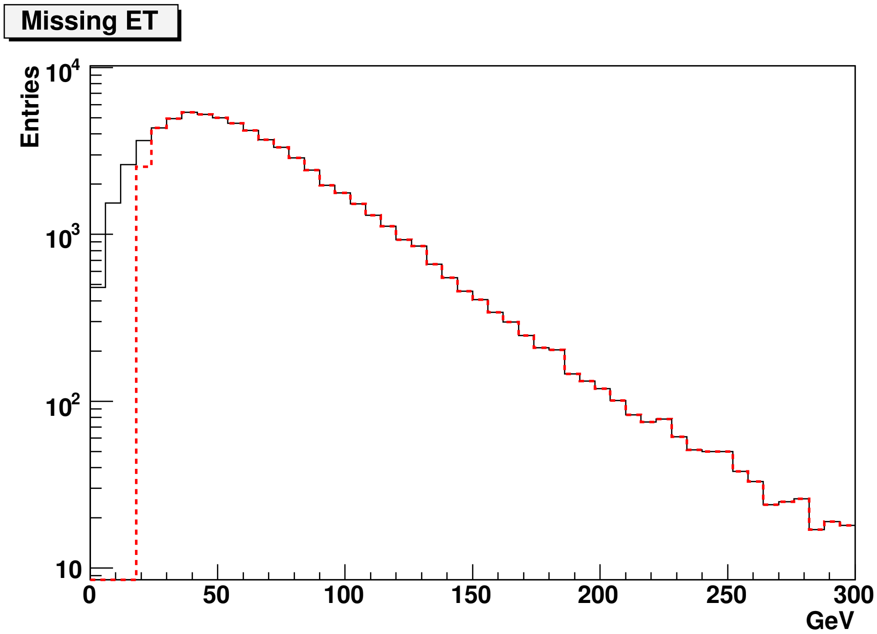


Figure 6.5: Missing energy for all $t\bar{t}$ events, $\mathrm{E_T^{\mathit{miss}}} > 20$ GeV cut is also shown.

Missing transverse energy cut

In ATLAS missing energy is calculated from the calorimetric cells applying an *object-base* calibration. Single calorimetric cells are calibrated according to the physics quantity they belong to, and the most appropriate calibration scheme is applied. Additional corrections to muons in the event and to unclustered energy are also applied.
In this analysis the events are selected requiring $\mathrm{E_T^{\mathit{miss}}} > 20$ GeV. Distribution of reconstructed $\mathrm{E_T^{\mathit{miss}}}$ for $t\bar{t}$ events is shown in figure 6.5. The $\mathrm{E_T^{\mathit{miss}}}$ selection at 20 GeV is also shown.

Selection efficiency

The efficiencies for the described selection cuts and the number of events expected for an integrated luminosity of L=100 pb^{-1} ($\sigma_{t\bar{t}} \times BR = 252$ pb) are reported in table 6.3 . The relative (and absolute) efficiencies are reported for each one of the described selections on the lepton, jets, $\mathrm{E_T^{\mathit{miss}}}$. About 2000 events remain after all cuts corresponding to $\epsilon = 7.9\%$.

Selection	Events	Efficiency
semi-leptonic $t\bar{t}$ (100 pb^{-1})	26068	
Num. Lepton = 1, $\mathrm{P_T} > 20$ GeV/c and $\lvert\eta\rvert < 2.5$	16815	64.5%
Num. Jets = 4, $\mathrm{P_T} > 40$ GeV/c and $\lvert\eta\rvert < 2.5$	2281	13.5% (8.7%)
$\mathrm{E_T}^{miss} > 20$ GeV	2057	90.2% (7.9%)
$W + 4jets$ (100 pb^{-1})	123900	
Num. Lepton = 1 $\mathrm{P_T} > 20$ GeV/c and $\lvert\eta\rvert < 2.5$	66514	53.7%
Num. Jets = 4 $\mathrm{P_T} > 40$ GeV/c and $\lvert\eta\rvert < 2.5$	6902	10.4% (5.6%)
$\mathrm{E_T}^{miss} > 20$ GeV	6414	92.9% (5.2%)

Table 6.3: Selection efficiencies for $t\bar{t}$ events and the main background $W + 4jets$. The number of events expected for L=100 pb^{-1} and the relative and absolute efficiencies are reported

6.3.2 Top quark mass measurement from hadronic top decay

The hadronic top quark decay is reconstructed combining three out of the four jets in the event. Among the four possible combination the one that gives the highest $\mathrm{P_T}$ is selected as the reconstructed top quark. The distribution of the invariant mass of such three - jet system is presented in the left plot of figure 6.6. The number of events corresponds to the statistics accumulated in an integrated luminosity of 275 pb^{-1} . This integrated luminosity has no deep reason: we have used all the simulated events. These come very close to the statistics that we expect to have after few days of LHC running at low luminosity.

The top quark mass peak is well visible over the combinatorial background. The distribution has been fitted with a gaussian distribution for the signal together with a fourth order Chebyshev polynomial for the background. In order to test the stability of the fit, different orders of polynomial have been tried and it has been checked that the fitted top quark mass does not change significantly varying the background parametrization. Varying the order of the Chebyshev polynomial from third to ninth the fit results are always consistent within the error bars. The fit result, for 175 GeV/c^2 generated mass, is $m_t = 170.4 \pm 0.7$ GeV/c^2 with a width of $\sigma = 14 \pm 1$ GeV/c^2.

The measured mass is smaller compared to the generated mass due to the out-of-cone energy. Part of the energy falls outside the reconstruction cone and is not associated with the jet. An in-situ calibration should be performed to recover for this energy (see section 1.2.4). We have not considered, in this work, this calibration. In fact it will be possible only after the detector has

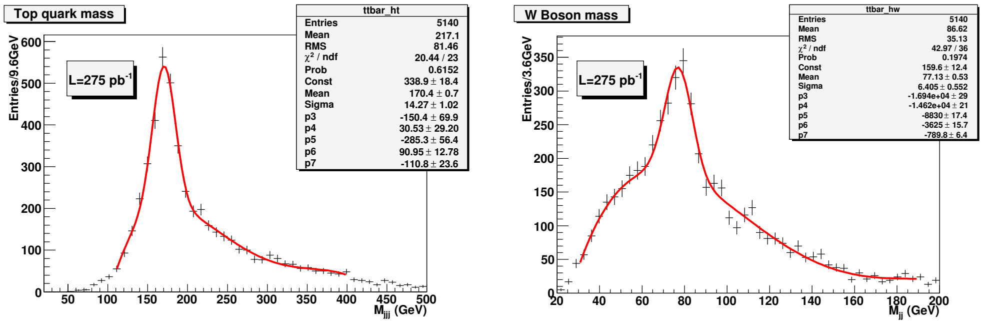


Figure 6.6: Reconstructed Top (left) and W (right) masses for and integrated luminosity of 275 pb^{-1} . Only combinatorial background included.

been fully understood.

The distribution of left in figure 6.7 shows the invariant mass for the three-jet combinations that have not been considered because give a smaller $\mathrm{P_T}$ of the reconstructed top. The shape of the distribution confirms the correct choice for most of the reconstructed top events: the top quark peak is not visible.

Using the events with the reconstructed top mass it is also possible to reconstruct the hadronic W boson from a two-jet combination out of the three selected as the top decay product.

In a similar way as for the top quark reconstruction, the two-jet combination giving the highest reconstructed $\mathrm{P_T}$ of the pair in the top center of mass frame is considered as the correct combination. The invariant mass, obtained with an integrated luminosity of 275 pb^{-1} , of such combination is visible in right distribution of figure 6.6. A gaussian fit for the signal together with a fourth order Chebyshev polynomial for the background has been performed obtaining a W boson mass $m_W = 77.1 \pm 0.5$ GeV/c^2 and a width $\sigma = 6.4 \pm 0.5$ GeV/c^2.

Right distribution of 6.7 shows the two-jet combinations not giving the highest $\mathrm{P_T}$ in the reconstruction of the W boson and are not considered.

Though some events are consistent with W bosons mass, most give a rather flat distribution confirming the correctness of the reconstruction algorithm.

The identification of the quark originating from the b-quark can drastically reduce the combinatorial background (see distributions 1.7 obtained with ultimate detector performances).

6.3.3 Backgrounds

As it has been discussed in sec. 1.3.1 the major sources of background to the $t\bar{t}$ signal are the QCD multi-jet production and the $W + jets$ production

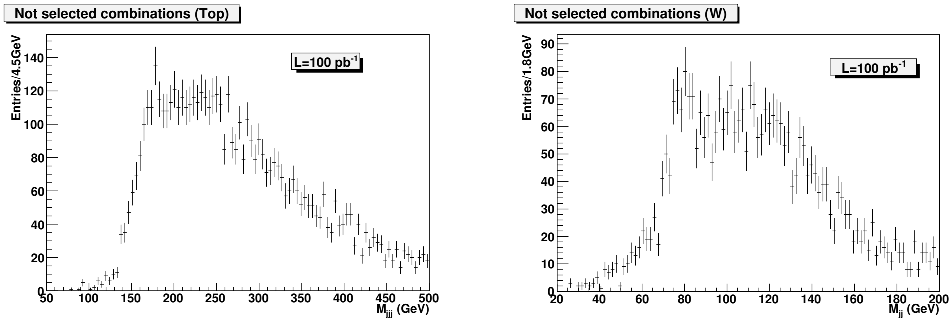


Figure 6.7: Left: the three-jet “wrong” combinations of reconstructed top mass. Right: the “wrong” two-jets combinations of reconstructed W mass obtained from the “correct” three-jet combination of reconstructed top. See the text for details.

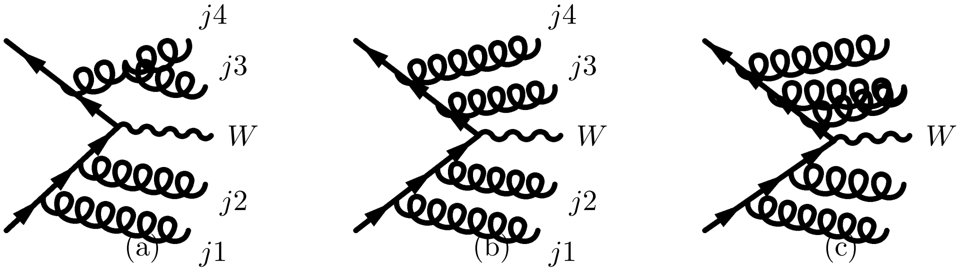


Figure 6.8: Feynman Diagram for $W + multi - jet$ process, main background to top signal. All three processes have in the final state a W boson and four jets, but in the first case (a) one of the jet comes from a soft radiating gluon, in the third case (c) two jets are merged into a single jet by the reconstruction. Details can be found in the text.

followed by the leptonic decay of the W boson.

The requirements of the isolated lepton and of the missing energy suppress strongly the QCD multi-jet background. To reject $W+jets$ events, b-tagging would be necessary: only 0.002% of the $W + jets$ would pass the final request of at least 2 b-jets (see table 1.3). The $Z + jets \rightarrow ll + jets$ reducible background with one misidentified lepton, is not considered in the following, it is estimated to be, however, efficiently suppressed by the $\mathrm{E}_{\mathrm{T}}^{miss}$ request and it is only of the order of 14% of the $W + jets$ background (table 1.3).

It is thus expected an important background from $W + jets$ events, and in particular of the $W + 4jets$, due to the absence of b-tagging.

Figure 6.8 (b) shows the Feynman diagrams for the W plus 4 partons (W + 4p) process. This is not the only process that can lead to a final state with four jets. Jet counting is a bit problematic, as we will discuss now and we need to deal with it with some care.

The diagram represented in (a) represents a “hard” production of $W + 3p$

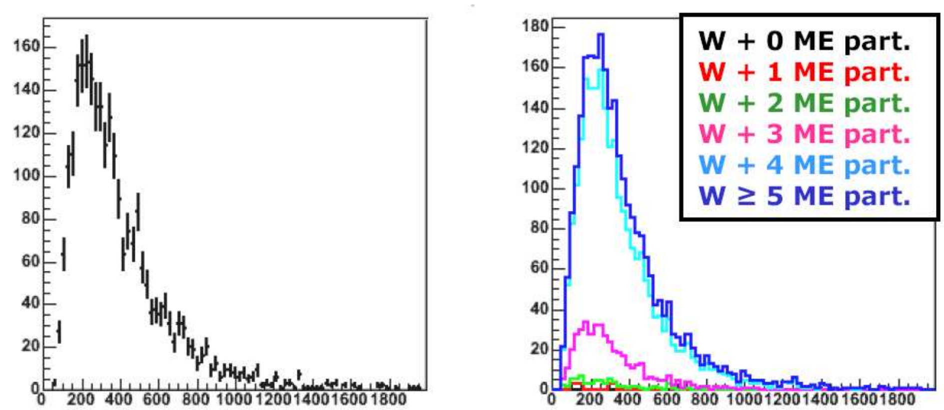


Figure 6.9: jjj invariant mass to the top signal for $W + 4jets$ (left) and for $W + Njets$ (right) with $0 \leq N \leq 5$. The different cumulative distributions are shown. The increase in the background is about 10% when considering also $W + Njets$ with $N \neq 4$. The data refer to 127 pb^{-1} [93].

followed by the emission of an additional "soft" gluon leading to a final reconstructed state $W + 4jets$. This must be considered a four-jet event.
The opposite happens in diagram (c), in which a $W + 5p$ events is reconstructed as a four jets due to the overlap of two final-state jets. The same happens when the fifth jet in the event is outside acceptance or it is below the selection threshold.
In conclusion all the $W + Np$ with $N > 0$ at generation level will contribute to the $t\bar{t}$ background.

Background generation

The traditional approach of parton shower (PS) generators should not be used for W boson plus N partons ($W + Np$) since they do not reproduce correctly the hard gluon emission. The best approach is thus to use matrix element (ME) Monte Carlo for the high $\mathrm{P_T}$ radiation (partons are generated if resulting $\mathrm{P_T}$ is greater than 40 GeV) followed by a PS Monte Carlo for the simulation of the showering. The ME generator used for this study is ALPGEN interfaced to the HERWIG PS Monte Carlo. However the simulation of this kind of background events is not completely straightforward due to the double-counting problem.
This issue can be understood with a simple example: consider a $W + 3p$ event generated by the ME generator, this event is then further processed by the PS Monte Carlo that can add an additional soft gluon in the final state (diagram (a) of figure 6.8).
We have to discard this event if it is already included in the phase space

accessible to the $W+4p$ ME generation, otherwise we would count the event twice when adding the two samples ($W+3p$ ME with a fourth parton from PS and $W+4p$ ME). The practical solution to this problem is to generate the $W+Np$ events at ME level separately, combine them with PS Monte Carlo, check “a posteriori” for events that could be possibly generated twice and, in case, remove them.

Figure 6.9 shows the contribution to the jjj invariant mass (with jets selected as in the analysis) from the different $W+Njets$ backgrounds [93]. On the left the background corresponding to 127 pb^{-1} for only the $W+4jets$ is shown, while on the right the total background from $W+Njets$ with $0 \leq N \leq 5$ is shown. The cumulative distributions for different N values are shown. The background is increased of 10% mainly due to events with 3 and 5 partons in the final state.

In the following only the $W+4p$ background will be included in the analysis [2]. To take into account the absence of $W+Np$ events (with N different from four) other backgrounds and the uncertainty in the $W+4p$ cross-section the background is increased of a factor of 50%.

At LHC the detailed measurement of the $W+Np$ cross section will require the use of the data themselves. A possible strategy is the measurement of the cross-section of $Z+Njets$ channel (with $Z \rightarrow ll$) from the data and the use of the Monte Carlo to extract the ratio $\sigma(Z+Njets)/\sigma(W+Njets)$ (with bosons decaying leptonically) that presents smaller theoretical uncertainties.

Background treatment

The ALPGEN generation cross section for $W+4p$ with the following decay of the W boson in lepton plus neutrino is $\sigma \times BR = 1200$ pb.

The background has been reconstructed using our jet calibration scheme and the $t\bar{t}$ commissioning selection and reconstruction has been applied on 120000 events corresponding to an integrated luminosity of $L = 100\ \mathrm{pb}^{-1}$. This is all the available fully simulated events for this type of background.

The resulting distribution for the jjj invariant mass is shown plot at the left of figure 6.10. The distribution peaks at an invariant mass of about 200 $\mathrm{GeV/c}^2$ and extends to large mass with a long tail (the same scale as for the signal in figure 6.6 has been used).

Figure 6.10 (right plot) shows the invariant mass of the two-jet system selected as described in 6.3.2. A total of about 6000 events pass the selection leading to an efficiency of 5.2% (see table 6.3).

---

[2] At the time of writing this work, the solution at the “double-counting” problem has been only recently implemented, correctly generated $W+Np$ events will be available in the near future. However, as it has been discussed, the contribution to the background of events with N different from 4 is expected to be small (of the order of 10%).

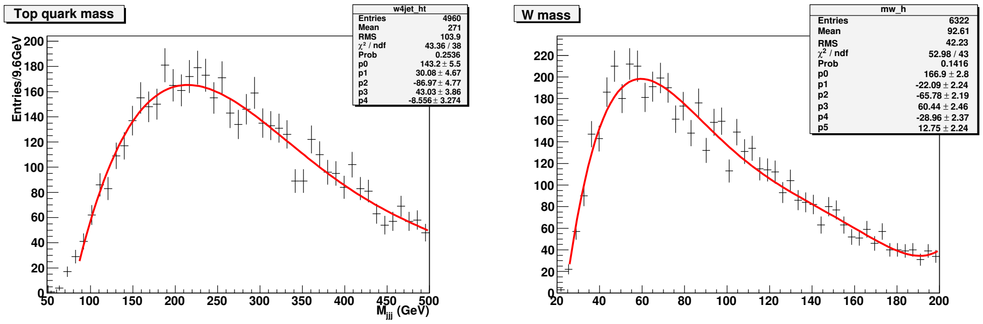


Figure 6.10: Left: distribution for jjj invariant mass in $W+4p$ events. Right: the jj invariant mass in the same events. Distributions are obtained as described in the text and correspond to an integrated luminosity of $L = 100$ pb^{-1}. The fit with Chebyshev polynomials are also displayed.

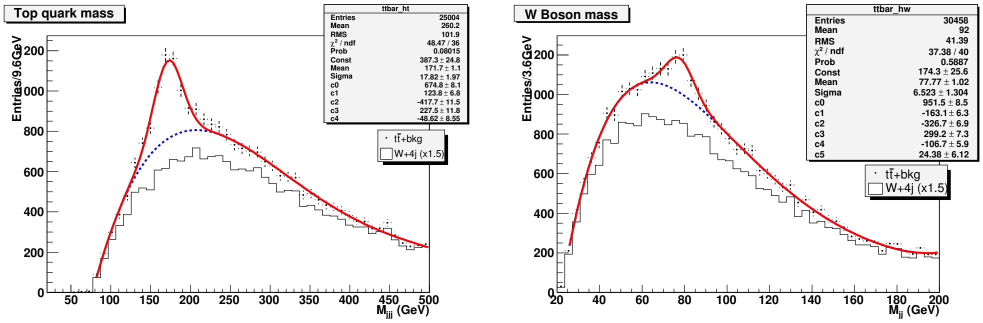


Figure 6.11: Reconstructed Top (left) and W (right) masses for and integrated luminosity of 275 pb^{-1} . $W + 4jets$ background included. The background has been scaled by a factor 1.5.

To add the background to the signal we have to rescale the background to an integrated luminosity $L = 275$ pb^{-1}. We performed a parametrization of the background shapes with a Chebyshev polynomial (a fourth and fifth order polynomial for top and W boson backgrounds respectively), the fit results are shown, together with the distributions, in figure 6.10.
The χ^2/ndf for the two fits are 43.36/38 for the jjj invariant mass and 52.98/43 for jj invariant mass.
The background shape for an integrated luminosity of $L = 275$ pb^{-1} has then been obtained from this parameterizations. A number of events corresponding to $L = 275$ pb^{-1} has been generated with jj and jjj invariant masses distributed according to the obtained parameterizations. In this way, with the correct shape of the background (figures 6.10), we have the correct statistics for the final analysis.

Results with background

The reconstructed top mass for the signal plus the background (black dots) is presented in figure 6.11 left plot. The top quark mass peak is well visible over the background. The distribution has been fitted with a gaussian function plus a Chebyshev polynomial of the fourth order. The result of the fit is represented by the red curve in the figure and the fit results in $\chi^2/ndf = 48.47/36$.

From the fit we obtain a top mass of $m_t = 172 \pm 1$ GeV/c^2 for a generated top mass of 175 GeV/c^2 . The result is in agreement with the top quark mass obtained without the background ($m_t = 170.4 \pm 0.7$ GeV/c^2). The width of the top quark peak is obtained from the standard deviation of the gaussian function and is $\sigma_t = 18 \pm 2$ GeV/c^2. The presence of the background increases a bit the top quark mass width from the value obtained with only the $t\bar{t}$ combinatorial background $\sigma_t = 14 \pm 1$ GeV/c^2 (see figure 6.6) to $\sigma_t = 18 \pm 2$ GeV/c^2. Though the two values are not incompatible, it is also reasonable that the noise due to background events has deteriorated a bit this parameter.

The figure also shows (histogram), the $W + 4jets$ background. The available number of background events was corresponding to an integrated luminosity of only 100 pb^{-1} . It has been scaled to the expected integrated luminosity of the first data taking ($L = 275$ pb^{-1}) and furthermore increased of a factor 50%.

The blue line represents the $W + 4jets$ and the $t\bar{t}$ combinatorial background has it has been obtained from the fit. The difference between this line and the histogram is due to combinatorial from $t\bar{t}$ events with the wrong jjj assignment. The signal over background ratio is $S/B = 0.3$ with 1764 events entering the gaussian fit.

In figure 6.11 the right distribution shows the jj invariant mass in $t\bar{t}$ events with the background $W + 4jets$ events. The combinatorial background, already important in absence of $W + 4jets$ events (figure 6.6 right) has been worsened. The background peaks at an invariant mass of 60 GeV/c^2 and extends to large values of m_{jj}.

A fit to the distribution with a gaussian function plus a fifth order polynomial has been performed resulting in $\chi^2/ndf = 27.38/40$. The results for the W boson mass are $m_W = 78 \pm 1$ GeV/c^2 and $\sigma_W = 7 \pm 1$ GeV/c^2. Both are in agreement with what has been obtained considering only $t\bar{t}$ signal events ($m_W = 77.1 \pm 0.5$ GeV/c^2 and $\sigma = 6.4 \pm 0.5$ GeV/c^2, see figure 6.6). Only 793 events enter the gaussian fit leading to a modest $S/B = 0.1$.

This number of event is smaller than the one obtained from the top quark mass peak since the W boson reconstruction is performed only after the hadronic top quark decay products have been associated and it is less efficient than top quark reconstruction. For each reconstructed top quark three possible jj associations are possible, this corresponds to three W boson

Sample	m_t (GeV/c^2)	σ_t (GeV/c^2)	S/B
Only signal ($L = 275$ pb^{-1})	170.4 ± 0.7	14 ± 1	0.93
$W + 4jets$ (x1.5) included	172 ± 1	18 ± 2	0.3
as before plus W mass cut	172 ± 1	21 ± 2	1.2
Sample	m_W (GeV/c^2)	σ_W (GeV/c^2)	S/B
Only signal ($L = 275$ pb^{-1})	77.1 ± 0.5	6.4 ± 0.5	0.51
$W + 4jets$ (x1.5) included	78 ± 1	7 ± 1	0.1

Table 6.4: Top quark mass measurement and associated W mass measurements results without and with background.

candidate for each reconstructed top event. The combination giving highest P_T of the W boson candidate is kept and the other two are discarded. However, as it is visible in figure 6.7, some of this discarded combinations give an m_jj peaking at the m_W nominal value.

Even if it is difficult to use this analysis for a precise W boson mass measurement, the jj invariant mass distribution suggests a possible method to reduce the background in the m_{jjj} distribution by selecting events with a reconstructed W mass in a window around the nominal W boson mass.

Events with 70 GeV/c^2 $< m_{jj} <$ 90 GeV/c^2 are selected and the reconstructed top mass distribution is shown in figure 6.12.

A much cleaner top mass peak is visible compared to figure 6.11. The combinatorial background, represented by the difference between the blue line and the histogram is drastically suppressed and the S/B ratio improves of a factor 4 up to $S/B = 1.2$. This however has been obtained by reducing of about one half the number of events entering the gaussian peak.

The gaussian fit gives top quark mass of $m_t = 172{\pm}1$ GeV/c^2 consistent with the previous values. The width of the peak is increased to $\sigma_t = 21{\pm}2$ GeV/c^2, is anyway compatible with the previous value. The fit gives a value for the χ^2/ndf of 73/35.

6.3.4 Conclusions

In this chapter we have discussed ATLAS potential for top physics with the very first collected data. In particular we concentrated our attention on the top mass measurement.

During the first period of data taking we can consider that the detector will not have the design performances. In particular we concentrated in the scenario where no b-tagging is available and the jet energy scale is set by the jet calibration methods without the use of the *in-situ* calibration.

We presented a strategy to measure top quark mass with some hundreds of pb^{-1} of collected data, corresponding to some days of ATLAS operation

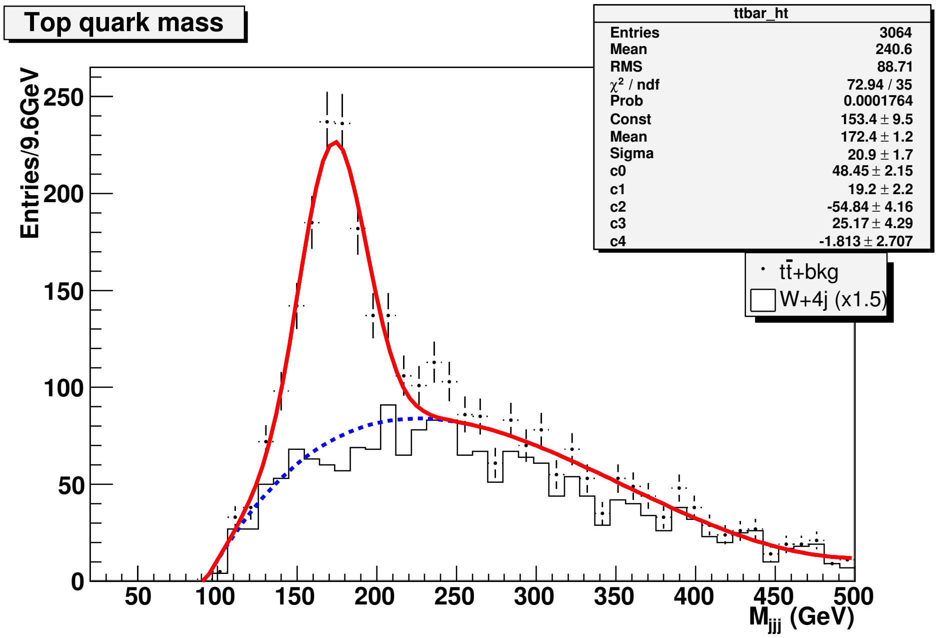


Figure 6.12: Reconstructed top mass obtained with a cut on associated W boson mass: events outside the window $70 < m_{jj} < 90$ GeV/c^2 are discarded. The distribution is fitted with a gaussian peak plus a Chebyshev polynomial. The background from $W + 4jets$ has been included.

at low luminosity regime. The $t\bar{t}$ events in the semi-leptonic decay channel have been considered. The hadronically decaying top can be entirely reconstructed from three jets while the isolated lepton and the neutrino, arising from the leptonically decaying top, are used to select the event.

The top quark mass can be measured as the invariant mass of the three-jet system leading to the highest P_T reconstructed top. The W boson daughter can be identified out of the three jets originated from the top decay as the two-jet system giving the highest P_T reconstructed W boson in the top center of mass frame. In both cases a peak is well visible over the combinatorial background.

The most important background ($W+Njets$) has been discussed. The background with $N = 4$ accounts for 90% of the total background and has been included in the analysis with a “safety factor” of 1.5 that takes into account the absence of the other backgrounds and of the uncertainity on W plus jets cross-section.

Including the background the reconstructed top quark peak is still visible over the background. The W boson peak is weakly visible but its presence suggests a way to clean up the sample. Events with a reconstructed W mass between 70 and 90 GeV/c^2 have been selected leading to an improvement of the S/B ratio of a factor 4 for the top quark mass peak.

The reconstructed top and W candidates invariant mass distributions have been fitted with a gaussian distribution, representing the signal, plus a Chebyshev polynomial describing the background (combinatorial and $W + 4jets$).

Results for measurements with and without background and for the events passing the W mass cut are summarised in table 6.4. The measured value of the top quark mass can be used from the first days of the data taking, to assess the performances of jet reconstruction algorithms and provides a fast feedback to detector communities on the status of the apparatus and calibrations.

The measured mass do not peak at the nominal value due to the unclustered energy that is not assigned to the reconstructed jets. A further final *in-situ* calibration must be performed to report the measured jet energy to the parton energy.

The $t\bar{t}$ events themselves used in this analysis can be used to set the jet energy scale at the level of 2% (respect to the quarks originating the jets), provided that the hadronic W boson can be reconstructed efficiently constraining the invariant mass of the decay products of the W boson. After applying this calibration the measured top mass is $m_t = 176 \pm 13$ GeV/c^2 [94]. However the W boson signal is weakly visible over the large background (right distribution of figure 6.11), thus the possibility to use the b-tagging is mandatory to reduce the combinatorial background and drastically reduce the background originated from $W + Njets$ events.

This work has shown that even with the very first collected data the ATLAS experiment has the potential to perform physics measurements. Even if the detector will not have final performances, the high collected statistics will allow to achieve, within some days of data taking, the first measurements in the standard model field.

The *commissioning* studies, such as the one presented here, will allow to understand in detail the detector performances and consequently ATLAS will have the potential to extend our knowledge of the standard model performing precise measurements.

The very first year of data taking will be an extremely exciting period for the LHC experiments, the physicists will have the possibility to open the road for discoveries and eventually reach the final goals and in the field of new physics discoveries.

Conclusions

In the past ten years the design, construction and testing of one of the next generation high energy experiments has become closer to its conclusion: the ATLAS detector is in an advanced status of construction and it is on its way to start data taking by the end of 2007 when LHC will deliver the first collisions.

A main component of the ATLAS detector is the calorimetric system, composed of different sub-systems, that has the important responsibility of precisely measure the energy of electrons, photons and jets and to provide the calorimetric first level of trigger. Liquid argon calorimeters for the electromagnetic part and for the hadronic calorimetry in the forward regions are completed by an scintillating tiles-iron sampling hadronic calorimeter in the barrel region.

The response to single particles of both calorimeters has been studied at the test beam and calorimeter performances have been studied in detail by the respective sub-detector communities. The challenging physics studies that will be performed in ATLAS rely also on the results obtained during these test beam periods to assess the performances of the detector in the field of jet measurements. From this point of view, extremely useful data have been collected during the combined test beam that give the possibility of studying combined performances in a setup closer to the final one. In fact even if, in many physics studies, hadrons are usually grouped into jets, the detailed knowledge of single pion response is fundamental to validate the tools that are used to simulate, reconstruct and calibrate the jets.

Aim of this thesis is to study the combined performances of the calorimeters starting from the results of the test beams up to the reconstruction of the top mass peak using jet events. In this work we used combined test beam data and, for the first time, we have studied the combined response of the final ATLAS calorimeters to pions and we have compared the results with Monte Carlo simulations.

The present study addresses three main points relevant to the physics with jet in ATLAS, especially during the first period of data taking: the comparison between test beam data and Monte Carlo predictions to perform the validation of the simulation software, the development and the study of performances of an original jet calibration method, and the measurement of the top quark mass from three jets invariant mass.

Precise measurements of top quark properties, and in particular of its mass, will play a fundamental role in LHC experiments. Thanks to its mass close to the electro-weak scale, top quark is an ideal tool to test the Standard Model and constrain new physics models. This is not the only important role for the top quark: in almost all the Standard Model extensions the top quark production is one of the main backgrounds and hence it is important to study its properties in details. At last the top quark can be extensively used during the commissioning phase to give prompt response to the sub-system communities on the detector performances. The large $t\bar{t}$ production cross-section and the semi-leptonic decay $t\bar{t} \rightarrow \mathrm{W}^{+}\mathrm{W}^{-}\mathrm{b}\bar{\mathrm{b}} \rightarrow \ell^{\pm}\mathrm{jjb}\bar{\mathrm{b}}$ offers the possibility to study in detail all the detector aspects: lepton trigger and reconstruction, jet reconstruction and calibration, $\mathrm{E}_{\mathrm{T}}^{miss}$ studies, b-tagging performances.

In this work we first address the problem of the validation of the Monte Carlo simulations, and in particular of the simulation of high energy interacting pions. This is possible comparing the response of the calorimeters to single pions data acquired during test beam programs and compare the result with the simulations. We have shown that, in general, a good level of agreement is reached. We want to stress here that this work has been used also to give important experimental feed-back to the developers of the simulation programs that could improve the Monte Carlo code also using these results.

Confident of the good level of agreement reached between data and Monte Carlo we have developed a method that uses simulations to calibrate the jet energy. The method uses a set of weights, extracted from di-jet QCD simulated events, to calibrate jet energy and allow to recover linearity to a level of 2% over a wide energy range and, at the same time, improve the transverse energy resolution of about 30%.

After assessing the performances of the method on QCD events we applied the method to a particularly interesting physics channel: the top quark measurement from the $t\bar{t} \rightarrow \mathrm{W}^{+}\mathrm{W}^{-}\mathrm{b}\bar{\mathrm{b}} \rightarrow \ell^{\pm}\mathrm{jjb}\bar{\mathrm{b}}$ channel. We have shown that this measurement can be performed during the very first data taking period with an integrated luminosity corresponding at some days of operation at low luminosity. We obtained this result also considering a detector with reduced performances, in particular without the use of b-tagging information. One of the main sources of systematic error in the top mass measurement is the jet energy scale and we have shown that the precision reached in jet energy measurement, obtained with our method, allows for a preliminary top mass measurement in the very beginning of the data taking period.

We believe that this work, albeit far from complete, shows that, from the very beginning of ATLAS life, we will be able to perform interesting and precise measurements in the Standard Model field, it also shows that the use of test beam data is fundamental to understand the performances of the detector, to assess the quality of our simulation codes and to help in developing new strategies, like jet calibration, to fulfill ATLAS physics goals.

Bibliography

[1] S. L. Glashow. Partial symmetries of weak interactions. *Nucl. Phys.*, 22:579, 1961.

[2] S. Weinberg. A model of leptons. *Phys. Rev. Lett.*, 19:1264, 1967.

[3] A. Salam. Proccedings of the eight nobel symposium on elementary particle theory. 1968.

[4] W. Fetscher and H. J. Geber. Precision tests of the electroweak model. In *World Scientific, Singapore*, 1993.

[5] S. Weinberg. *The Quantum Theory of Fields.* Cambridge University Press, 1995.

[6] P. W. Higgs. Broken symmetries, massless particles and gauge fields. *Phys. Lett.*, 12:132, 1964.

[7] N. Cabibbo. Unitary simmetry and leptonic decays. *Phys. Rev. Lett.*, 10:531, 1963.

[8] M. Kobayashi and T. Maskawa. *Prog. Th. Phys.*, 49:652, 1973.

[9] F. Abe *et al.* *Phys. Rev. Lett.*, 73:225, 1994.

[10] F. Abe *et al.* *Phys. Rev.*, D50:2966, 1994.

[11] S. Abachi *et al.* *Phys. Rev. Lett.*, 74:2632, 1995.

[12] Vogt Kidonakis. *Phys. Rev.*, D68:114014, 2003.

[13] S. Willenbrock *et al.* *Phys. Rev.*, D56:5919.

[14] E. Boos A. Belayev. *Phys. Rev.*, D63:034012.

[15] M. Smith *et al.* *Phys. Rev.*, D54:6696.

[16] S. Eidelman *et al.* (Particle Data Group). The cabibbo-kobayashi-maskawa quark-mixing matrix. *Phys. Rev. Lett.*, B592 1, 2004.

[17] ATLAS Collaboration. Atlas detector and physics performance. Technical design report, CERN, 1999.

[18] J. Schwindling A. I. Etienvre, J. P. Meyer. Top quark mass measurement in the lepton plus jets channel using full simulation. ATL-PHYS-INT-2005-002, July 2005.

[19] C. Peterson *et al.* *Phys. Rev.*, D27:105, 1983.

[20] CDF Collaboration. *Phys. Rev. Lett.*, 80:2767, 1998.

[21] D0 Collaboration. *Phys. Rev.*, D58:052001, 1998.

[22] I. Efthymiopulos. High p_t top mass reconstruction in the single lepton plus jets channel using a large calorimeter cluster. ATL-COM-PHYS-99-050, 1999.

[23] L. La Rotonda A. Lagatta and M. Cobal. Top mass evaluation in the $t\bar{t}$ dilepton channel. ATL-COM-PHYS-99-044.

[24] C. T. Hill. *Phys. Lett.*, B266:419, 1991.

[25] C. T. Hill. *Phys. Lett.*, B345:483, 1995.

[26] E. Richter-Was and M. Sapinski. Search for the sm and mssm higgs boson in the $t\bar{t}h$, $h \to b\bar{b}$ channel. ATL-PHYS-98-132.

[27] K. J. F. Gaemers and G. Hoogeveen. *Phys. Lett.*, B146:347, 1984.

[28] A. Stange D. Dicus and S. Willenbrock. *Phys. Lett.*, B333:126, 1994.

[29] K. Lane and E. Eichten. *Phys. Lett.*, B352:382, 1995.

[30] E. Eichten and K. Lane. *Phys. Lett.*, B327:129, 1994.

[31] C. T. Hill. *Phys. Rev.*, D49:4454, 1994.

[32] R. Casalbuoni *et al.* Z. phys. C69:519, 1996.

[33] N. Cartiglia and J. Parsons. Study of atlas sensitivity to a heavy resonance decaying to $t\bar{t}$. ATL-COM-PHYS-99-038.

[34] G. Maholn and S. Parke. *Phys. Rev.*, D53:4886, 1996.

[35] V. Simak A. Lagatta and J. Smolik. $t\bar{t}$ spin correlations and the potential for observation of cp violation in the production vertex. ATL-COM-PHYS-99-049.

[36] D. Cavalli *et al.* The charged higgs in hadronic decays with the atlas detector. ATL-COM-PHYS-99-030.

[37] L. D. Chikovani and T. D. Djobava. Atlas sensitivity to the flavor-changing neutral current decay $t \rightarrow zq$. ATL-COM-PHYS-99-034.

[38] S. McGrath J. Dodd and J. Parsons. Study of atlas sensitivity to rare top decays. ATL-COM-PHYS-99-039.

[39] Y. P. Gouz and S. R. Slabospitsky. Double top production at hadron colliders. hep-ph/9811330.

[40] L.R. Evans. Cern lhc project report 101. Technical design report, CERN, 1997.

[41] CMS Collaboration. Cmt technical proposal. Technical design report, CERN, 1994.

[42] Alice Collaboration. Alice technical design report. Technical design report, CERN, 2001.

[43] LHCb Collaboration. Lhcb technical proposal. Technical design report, CERN, 1998.

[44] ATLAS Collaboration. Atlas technical proposal. Technical design report, CERN, 1994.

[45] ATLAS Collaboration. Atlas magnet system technical design report. Technical Design Report CERN/LHCC/97-18, CERN, 1997.

[46] ATLAS Collaboration. Atlas inner detector technical design report. Technical Design Report CERN/LHCC/97-16, CERN, 1997.

[47] ATLAS Collaboration. Atlas muon spectrometer technical design report. Technical Design Report CERN/LHCC/97-22, CERN, 1997.

[48] ATLAS Collaboration. Atlas hlt, daq and dcs technical proposal. Technical Design Report CERN/LHCC/00-17, CERN, 2000.

[49] ATLAS Collaboration. Atlas computing technical design report. Technical Design Report CERN/LHCC/05-22, CERN, 2005.

[50] M. Aderholz *et al.* Monarc phase 2 report. Technical Design Report CERN/LCB/00-01, CERN, 2000.

[51] ATLAS Collaboration. Liquid argon calorimeter. Technical Design Report CERN/LHCC 96-41, CERN, 1996.

[52] W. H. Press *et al. Numerical Recipes in C.* Cambridge Univeristy Press, 2002.

[53] B. Dowler *et al. Nucl. Instr. and Meth.*, A482:94, 2002.

[54] ATLAS Collaboration. Tile calorimeter. Technical Design Report CERN/LHCC 96-42, CERN, 1996.

[55] R. Wigmans. *Calorimetry Energy Measurement in Particle Physics.* Oxford Univeristy Press, 2000.

[56] T. A. Gabriel *et al.* Energy dependence of hadronic activity. *Nucl. Instr. and Meth.*, A338:336, 1994.

[57] V. Tsulaia A. A. Solodkov. Tilecal beam test simulation application in the fads/goofy framework (geant4). ATL-TILECAL-2003-002.

[58] M. G. Pia *et al.* The geant4 object oriented simulation toolkit. In *EPS-HEP99 Conference, Tampere*, 1999.

[59] D. L. Smith R. L. Craun. Analysis of response data for several organic scintillators. *Nucl. Instr. and Meth.*, 80:239–244, 1970.

[60] W. R. Nelson *et al.* The egs4 code system. SLAC-Report-265, December 1985.

[61] R. Brun, F. Bruyant, A. C. McPherson, M. Marie, and P. Zanarini. Geant - detector description and simulation tool, cern program library long writeup w5013. CERN Fata Handling Division, DD/EE/84-1, 1987.

[62] A. B. Kaidalov and K. A. Ter-Martirosyan. *Phys. Lett.*, B117:247, 1982. *Sov. J. Nucl. Phys* 39:979, 1984.

[63] TileCal Collaboration. Testbeam summary paper. Not yet published.

[64] M. Cobal and B. Di Girolamo. Atlas barrel combined run in 2004, day by day.

[65] S. Gadomski *et al.* Deployment and use of the atlas daq in the combined test beam. In *14th IEEE - NPSS Real Time Conference 2005 Nuclear Plasma Sciences Society RT 2005 , Stockholm, Sweden , 4 - 10 Jun, 2005.* CERN-ATL-COM-DAQ-2005-023, CERN-ATL-COM-DAQ-2005-023, ATL-COM-DAQ-2005-023.

[66] J. Garvey. Beam test of the atlas level-1 calorimeter trigger system. CERN-ATL-DAQ-CONF-2005-006, ATL-COM-DAQ-2005-010, ATL-DAQ-CONF-2005-006.

[67] A. Anjos *et al.* Deployment of the atlas high-level triggers. In *14th IEEE - NPSS Real Time Conference 2005 Nuclear Plasma Sciences Society RT 2005 , Stockholm, Sweden , 4 - 10 Jun 2005.* CERN-ATL-COM-DAQ-2005-023, CERN-ATL-COM-DAQ-2005-023, ATL-COM-DAQ-2005-023.

[68] M. Bosman, C. Meessen, A. Negri, E. Segura, S. Sushkov, F. Touchard, and S. J. Wheeler. Development and tests of the event filter for the atlas experiment. In *14th IEEE - NPSS Real Time Conference 2005 Nuclear Plasma Sciences Society RT 2005 , Stockholm, Sweden , 4 - 10 Jun 2005.* CERN-ATL-COM-DAQ-2005-025, CERN-ATL-COM-DAQ-2005-025, ATL-COM-DAQ-2005-025.

[69] W. T. L. P. Lavrijsen. The athena startup kit. In *Computing in High Energy and Nuclear Physics, 24-28 March 2003, La Jolla, California*, 2003. arXiv:cs.SE/0306083 v1.

[70] J. Boudreau and V. Tsulaia. The geomodel toolkit for detector description. In *CHEP04, Interlaken, Switzerland, 27 Sp - 1 Oct 2004.*

[71] P. Loch. Suggestions for a general energy reconstruction scheme for the atlas calorimeters. ATL-CAL-97-091, ATL-AC-PN-91.

[72] B. Di Girolamo, A. Dotti, V. Giangiobbe, P. Johansson, L. Pribyl, and M. Volpi. Beamline instrumentation in the 2004 combined atlas testbeam. ATL-COM-TECH-2005-001, CERN-ATL-COM-TECH-2005-001, 2005.

[73] Geant4 group. *Geant4 User's Guide - For Application Developers.*

[74] A. Baker and A. Krzywicki. *Phys. Rep.*, 28C:1, 1976.

[75] H. P. Wellisch and D. Axen. *Phys. Rev.*, C54:1329–1332, 1996.

[76] M. Cavalli-Sforza *et al.* Analysis results of the first combined test of the largon and tilecal barrel calorimeter prototyptes. ATL-TILECAL-95-067, ATL-L-PN-67.

[77] ATLAS Collaboration. Calorimeter performance. Technical design report, CERN, 1996.

[78] F. Gianotti *et al.* Geant4 hadronic physics validation with lhc test-beam data: first conclusions. CERN-LCGAPP-2004-10.

[79] A. Dotti, A. Lupi, and C. Roda. Results from atlas tile calorimeter : a comparison between data and g4 simulation. In *9th Topical Seminar on Innovative Particle and Radiation Detectors 23 - 26 May 2004 Siena, Italy.*

[80] W. J. Stirling R. K. Ellis and B. R. Webber. *QCD and Collider Physics.* Cambridge University Press, 1996.

[81] Gerald C. Blazey *et al.* Run ii jet physics. 2000.

[82] S. D. Ellis and D. E. Soper. Successive combination jet algorithm for hadron collisions. *Phys. Rev.*, D48:3160–3166, 1993.

[83] J. Huth *et al.* In *Proceedings of Research Directions Fot the Decade: Snowmass 1990*, pages 134–. E. Berger, 1990.

[84] W. Braunschweig *et al.* Desy internal report. DESY 89-022, 1989.

[85] I. Vivarelli. *Hadronic Decay of the Supersymmetric Heavy Neutral Higgs Boson*. PhD thesis, Scuola Normale Superiore Perfezionamento in Fisica, 2005.

[86] M. Casado and M. C. Sforza. H1-inspired analysis of the 1994 combined test of the liquid argon and tilecal calorimeter prototypes. ATL-TILECAL-96-075, 1996.

[87] S. Akhmadaliev *et al.* Results from a new combined test of an electromagnetic liquid argon calorimeter with a hadronic scintillating-tile calorimeter. *Nucl. Inst. and Meth.*, A449:461–477, 2000.

[88] A. N. Kolmogorov A. D. Aleksandrov and M. A. Lavrent'ev. *Mathematics - Its Content, Methods and Meaning*. Dover Publications, INC., 1999.

[89] Katy Forax. Lhc construction/installation general coordination schedule. LHC-PM-MS-0005.

[90] M. Mangano F. Gianotti. Lhc physics: the first one-two years. 2005.

[91] S. Bentvelsen and M. Cobal. Top studies for the atlas detector commissioning. COM-PHYS-2005-039, ATL-PUB-2005-024.

[92] Toprex. *Comput. Phys. Commun.*, 148:87, 2002.

[93] I. Van Vulpen and W. Verkerker. Top physics during atlas commissioning. Presentation given at: ATLAS Physics Week (Oct. 2005).

[94] J. Lu and D. M. Gingrich. Investigation of light-jet energy calibration and top-quark mass measurement using $t\bar{t}$ events. ATL-COM-PHYS-2005-047.

Aknowledgments

At the end of this work, I feel the need to thank some people that, directly on indirectly, helped and supported me in these years.

I would like to start with Prof. Tarcisio Del Prete for his patience and for his guide during my PhD studies, without his help all the aspects of my work would have been much more difficult.

I would like also to thank the entire Pisa group: Prof. Vincenzo Cavasinni for his help and fruitful discussions, special thanks go to Chiara Roda particularly important for the work on jet calibration that would not have been possible without her help, Anna Lupi for the work that we did together at the beginning of my activity, Danilo Cimino because he could stand me in the office, the students Michele Cascella and Paolo Francavilla. A special thanks goes to my colleagues Iacopo Vivarelli and Francesca Sarri, we share our working duties but they are first good friends and it is also for their happiness that working in this group is so productive and nice. I would like also to thank Giulio Usai, presently at CERN, that introduced me to my present activities.

I would also like to thank the external referees for reading and commenting this work: Prof. Luciano Mandelli, Prof. Lorenzo Foá, Prof. Guido Tonelli.

Finally, there are some people that, even if did not contribute directly to this work, helped me, with their presence and support, in the difficult moments and shared with me the joy of the success.

Let me start with my family: my mother and my father, they are guides in all the important decisions of my life. I would like to thank also my sister and my brother to whom I wish to realize all their dreams.

There is then a person, Doris, that only relatively recently entered my life, but from that moment on she has become more and more important. I'm really glad of sharing my days with her.

Concluding I would like to mention all my friends in Pisa, Como and at CERN for the good time spent together.