"Energy reconstruction and particle identification in a high granularity semi-digital hadronic calorimeter"

Mannai, Sameh

ABSTRACT

This thesis presents studies on various energy reconstruction methods of hadronic showers and Particle identification within a Semi-Digital Hadronic CALorimeter (SDHCAL) proposed for the future electron-positron collider ILC. The SDHCAL technological prototype is the first of a series of new-generation detectors fulfilling almost all the ILC requirements. Beside its semi-digital readout, the main feature of SDHCAL is its high granularity allowing imaging capabilities required for the application of the particle flow algorithm (PFA) in order to improve the jet energy resolution. The SDHCAL technological prototype has been tested successfully in beam tests at CERN and shows good performance. An exhaustive study on the optimization of the energy reconstruction of hadronic showers using GEANT4 simulation is presented and confirms the important impact on energy resolution of a semi-digital readout. Different analytic methods have been developed for the energy reconstruction within the multi-threshold mode of SDHCAL. A further approach based on NeuralNetwork has been also studied. An analysis investigating the energy resolution of pion showers recorded in SDHCAL during beam tests at CERN, has been presented. A linear response and a good energy resolution are obtained for a large range of hadronic energies for both the Digital and the Semi-Digital modes of the SDHCAL prototype. The Semi-Digital mode shows however better performance at energies exceeding 30 GeV. Neural network technique provided a significant improvement of the energy resolution and linearity in comparison with t...

CITE THIS VERSION

Université catholique de Louvain
Secteur des Sciences et Technologies
Institut de Recherche en Mathématique et Physique
Center for Cosmology, Particle Physics and Phenomenology

Energy Reconstruction and Particle Identification in a High Granularity Semi-Digital Hadronic Calorimeter

Doctoral dissertation presented by

Sameh Mannai

in fulfilment of the requirements for the degree of Doctor in Sciences

Jury Members

Prof. E. Cortina (Supervisor)	UCL, Belgium
Prof. G. Bruno	UCL, Belgium
Prof. K. Piotrzkowski	UCL, Belgium
Prof. V. Lemaitre	UCL, Belgium
Prof. I. Laktineh	IPNL, France
Dr. M. Tytgat	UGent, Belgium

March, 2017

Table des matières

Introduction

The Standard Model has been one of the major achievements in particle physics since it was proposed by Glashow, Salam and Weinberg in the 1970s [1–3]. It is the theory describing the elementary particles constituting the matter and their interactions via three fundamental forces : the electromagnetic, the weak and the strong force.

According to the Standard Model, the matter is composed of fermions which are divided into leptons and quarks. The interaction between fermions is mediated by the gauge bosons specific to each force : photons for the electromagnetic force, $W^{\pm}/Z^{0}$ bosons for the weak force and gluons for the strong force. It also predicted the existence of the Higgs boson which explains the origin of mass for $W^{\pm}/Z^{0}$ as well as the masses of all elementary particles. All the predicted particles of the Standard Model have been observed experimentally with the discovery of : Z and $W^{\pm}$ bosons at CERN in 1983 [4, 5], the top quark in 1995 at the Tevatron by CDF and $D\emptyset$ [6]experiments, and the Higgs boson in 2012 announced by the ATLAS and CMS experiments at the CERN Large Hadron Collider (LHC) [7–9].

However, behind these extremely successful achievements, there are still some shortcomings.
Among them we can invoke its failure to explain the identity of the cosmic dark matter. Indeed, only 5% of the mass of our universe composed of atoms, can be described by the Standard Model. The rest of the universe which is made of 23% of Dark Matter in addition to 72% of Dark Energy, could not be explained with the Standard Model. The Standard Model cannot explain why the universe contains more matter than antimatter, neither why the basic particles of matter are the quarks and leptons. Also, although the Standard Model postulates the Higgs field, it does not explain the properties of this field.

However, new models [10, 11] introducing new physics beyond the Standard Model provide explanations for these unresolved questions. The most known extension to the Standard Model is Supersymmetry (SUSY) [12], which can provide answers for many questions such as the one related to Dark Matter. Thus, the Standard Model and the possible Standard Model extensions should be tested and explored experimentally at high precision particle accelerators in order to reveal the suitable theories and to discover new phenomena. A particular emphasis should be accorded to the exploration of the Higgs sector which plays an important role in all of the unanswered questions.

In this context, the discovery of the Higgs boson in the LHC gives us a point of entry into a deep and precise exploration of its properties allowing us to obtain a comprehensive understanding of its coupling to particles. However, even if it is ideal for discovering new particles, LHC may present a limited precision in measurments since the colliding particles are protons. Therefore, a lepton-lepton collider will be the right tool to provide a precise measurment due to the usage of elementary particles in collision with well defined initial states.

Currently, one of the most advanced projects for a lepton collider is the International Linear Collider (ILC), which is discussed in this thesis. ILC is expected to advance our knowledge of particle physics beyond the information provided by the LHC. Indeed, ILC will be able to deliver a precise description of the properties of the Higgs boson and measure its mass, quantum numbers, and couplings with high precision.

In order to perform high precision measurements at ILC, a number of special requirements are imposed on the detector designs. Indeed, the detector concepts are designed to achieve unprecedented advances in collider detector performance by using new technologies and reconstruction techniques based on the Particle Flow Algorithm (PFA). Two detectors (ILD and SiD) have been conceived to meet these challenges in ILC. In particular, the International Large Detector (ILD) is designed in the aim to reach an excellent precision by using PFA. The PFA allows to reconstruct separately each individual particle in the jet in the appropriate sub-detector. This has a direct consequence on the calorimeters design, which must have excellent imaging capabilities. Thus, for an optimal application of PFA, the calorimeter must be highly granular . Such imaging calorimeters are investigated within the CALICE collaboration (CAlorimetry for LInear Collider Experiments) [13].

This thesis focuses on the Semi-Digital Hadronic CALorimeter (SDHCAL) option within the ILD concept. Beside its 2-bit readout with 3 levels of charge deposits in a cell, the main feature of SDHCAL is its high granularity required for an efficient application of the Particle Flow Algorithm. The prototype is the first of a series of new-generation detectors equipped with a power-pulsing mode intended to reduce the power consumption of these high granularity detectors. A dedicated acquisition system was developed to deal with the output of more than 70 millions of electronics

channels in the final object. After its completion in 2011, the prototype was commissioned using cosmic rays and particles beams at CERN. Since then, several beam test campains at CERN have been engaged and allowed to validate the prototype and test its performance.

The major goals of this thesis are first, the study of the energy reconstruction of single pions based on simulation and then the analysis of data recorded with the SDHCAL in beam tests, in order to estimate the energy of pions with the real data using the different techniques developed in simulation. In addition, given the good performance of SDHCAL and the encouraging results we obtained and in order to further improve them, two additional purposes appeared. The first one consists in applying the Neural Network technique for energy reconstruction to a better energy estimation. The second aims to use the imaging capability of SDHCAL in particle identification by using Multivariate methods.

This thesis is organized as following. The reason for building ILC, the physics goals, the design of the accelerator and their detector concepts are discussed in chapter 1. The particle flow algorithm is also introduced in this chapter. An overview on the interactions of particles with matter and the basics of calorimetry for high energy physics are given in chapter 2. Chapter 3 exposes a detailed description of the SDHCAL technological prototype. Chapter 4 details the simulation studies done to evaluate the detector design as well its response to radiation when operating in a similar condition as in beam test, focusing on the energy reconstruction of single pions within SDHCAL. The study is especially dedicated to test, based on simulation, the validity of several energy reconstruction methods and estimate the hadronic energy resolution of the detector. The analysis presented in chapter 5 investigates the energy resolution of pion showers recorded in SDHCAL during 2012 beam tests at CERN. Details about the beam test setup, event building and event selections for this analysis using pion data are provided. This is followed by the presentation of the results of the energy resolution of the hadronic showers for beam test. Finally, chapter 6 presents an analysis which describes a particle classifier based on Multivariate methods using information provided by our high granularity calorimeter. The final goal was to determine if by using multivariate techniques analysis instead of classic cuts, the particles can be better identified and the response of SDHCAL to pions showers is improved.

Chapitre 1

The International Linear Collider

1.1 Physics motivation

The discovery of the Higgs boson in the LHC [7–9] gives us a point of entry into a deep and precise exploration of its properties, in order to obtain a comprehensive understanding of its coupling to particles. The International Linear Collider (ILC), is one of the next high energy particle accelerators, expected to advance our knowledge of particle physics beyond the information provided by the LHC.
The physics program of ILC will focus mainly on the investigation of the Higgs sector. The Higgs boson, discovered at the ATLAS and CMS experiments, has a mass of 125 GeV and, so far, the properties as postulated in the Standard Model. Thus ILC, with its center-of-mass energy at 500 GeV(extendable to 1 TeV), is the right tool to study this Higgs particle at 125 GeV.
Fig. 1.1 shows the Higgs production cross section as function of center-of-mass energy for the Higgs boson, corresponding to the dominant Higgs production channels. The ILC program on the Higgs boson includes experiments at 250 GeV, the peak of the cross section for the channel $e^+e^- \rightarrow ZH$, and at higher energies to access the process $e^+e^- \rightarrow WW \rightarrow \nu\nu H$ with WW fusion production of the Higgs boson. The other dominant channel is the ZZ fusion ($e^+e^- \rightarrow ZZ \rightarrow e^+e^-H$). Feynman diagrams for the three major Higgs production processes at the ILC are sketched in Fig. 1.2.

Also, a measurement with a crucial precision of the top quark properties can be achieved via $t\bar{t}$ pair-production by the process $e^+e^- \rightarrow t\bar{t}$ at a center-of-mass energy $\sqrt{s} \approx 2m_t$ (350 GeV) [17]. Indeed, thanks to the accurately known and readily

variable beam energy, ILC offers the possibility to realize cross section measurements at particle production thresholds. With LHC the top quark mass has been measured with an accuracy of 0.76 GeV [14], while in ILC a statistical precision of the order of 30 MeV can be reached [15].

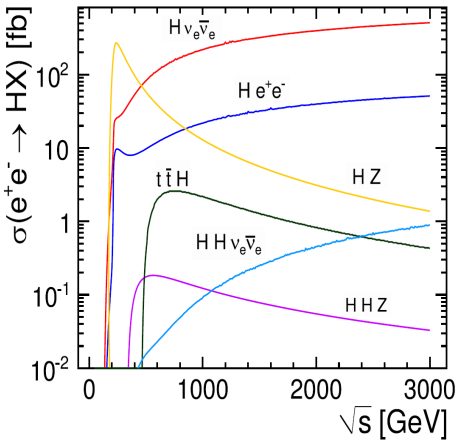

FIGURE 1.1: Higgs production cross section as function of center-of-mass energy $\sqrt{s}$ for M_{Higgs} = 125 GeV [16].

For a higher center-of-mass energy, above 500 GeV, the Higgs mass can be measured via associated production with top quarks $e^+e^- \rightarrow t\bar{t}H$ and via the double Higgs production channels Higgsstrahlung $e^+e^- \rightarrow ZHH$ which are two very important processes studied for the first time at $\sqrt{s} = 500GeV$ and fusion $e^+e^- \rightarrow \nu\bar{\nu}HH$.

Since it is the the heaviest matter fermion in the Standard Model, the top quark is essential to understand the fermion mass generation mechanism. The $e^+e^- \rightarrow t\bar{t}H$ reaction will allow the extraction of the top Yukawa coupling.
The measurement of the triple Higgs boson coupling is indispensable for understanding the electroweak symmetry breaking and can be realized through the processes $e^+e^- \rightarrow ZHH$ and $e^+e^- \rightarrow \nu\bar{\nu}HH$.

In addition, precision studies at ILC will allow to discover new interactions at higher mass scales through pair-production of quarks, leptons, Z and W bosons. Precison studies will also make important contributions to the investigation of new physics models, like supersymmetry [12], proposed to solve Standard Model deficiencies or provide missing elements such as dark matter. This program will allow to search for new particles such as color-singlet supersymmetry particles that are difficult to detect at the LHC because of their expected low energy release and low production rate,

while ILC experiments, can explore these particles in a direct way. Indeed, if new particles are discovered, ILC will be able to measure their quantum numbers unambiguously and determine their couplings to the percent level [17]. Also, these studies include the search for a new neutral weakly interacting particle and which might be the constituent of the dark matter of the universe. Table 1.1 lists the major reactions that will be studied at the ILC in the various stages of its program.

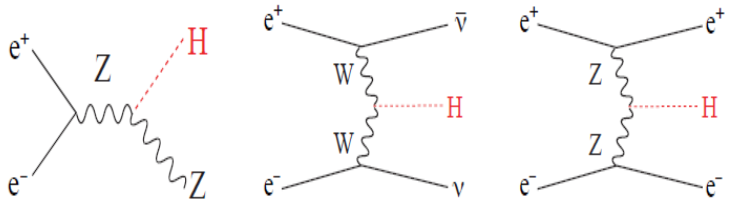

FIGURE 1.2: Feynman diagrams for the three major Higgs production processes at the ILC : $e^+e^- \to ZH$ (left), $e^+e^- \to \nu\bar{\nu}h$ (center), $e^+e^- \to e^+e^-H$ (right) [17].

E(GeV)	Reaction	Physics Goal
91	$e^+e^- \to Z$	ultra-precision electroweak
160	$e^+e^- \to WW$	ultra-precision W mass
250	$e^+e^- \to Zh$	precision Higgs couplings
$350-400$	$e^+e^- \to t\bar{t}$	top quark mass and couplings
	$e^+e^- \to WW$	precision W couplings
	$e^+e^- \to \nu\bar{\nu}h$	precision Higgs couplings
500	$e^+e^- \to f\bar{f}$	precision search for Z'
	$e^+e^- \to t\bar{t}h$	Higgs coupling to top
	$e^+e^- \to Zhh$	Higgs self-coupling
	$e^+e^- \to \tilde{\chi}\tilde{\chi}$	search for supersymmetry
	$e^+e^- \to AH, H^+H^-$	search for extended Higgs states
$700-1000$	$e^+e^- \to \nu\bar{\nu}hh$	Higgs self-coupling
	$e^+e^- \to \nu\bar{\nu}VV$	composite Higgs sector
	$e^+e^- \to \nu\bar{\nu}t\bar{t}$	composite Higgs and top
	$e^+e^- \to \tilde{t}\tilde{t}^*$	search for supersymmetry

TABLE 1.1: Major physics processes to be studied by the ILC at various energies. The table indicates the various Standard Model reactions that will be accessed at increasing collider energies, and the major physics goals of the study of these reactions [17].

The physics program of ILC presents a wide range of physics scenarios which imply several requirements on the detector design. The main requirements for ILC detector are excellent track momentum resolution and jet energy resolution.

The largest production cross section for a center-of-mass energy of 250 GeV corresponds to the Higgsstrahlung process $e^+e^- \rightarrow ZH$ and represents the cleanest way to study the Higgs. Unfortunately, at LHC, many Higgs decays channels, like Higgs produced by gluon-gluon fusion, are considered as nearly impossible to exploit due to the presence of high QCD background because of the use of sub-structured protons. However, thanks to the well defined center-of-mass energy of collision at ILC, the initial state is very well known and the search for a Higgs signal is performed by tagging the Z boson through the detection of its decay products, independently from the Higgs boson decay modes. In this case, when the Higgs boson is identified, it is possible to measure the rates for all its decays, including decays to invisible or unusual final states, with high precision. As explained before, such decays are very difficult to separate from Standard Model background events at the LHC and the use of fully hadronic final states is possible only in a very clean environment of an e^+e^- collider. The precision measurement of the rates of decay of the Higgs boson to the different types of quarks, leptons, and bosons is of main interest to provide a confirmation of the hypothesis that the Higgs field operates alone to create the masses of these particles, or if it has partners that are additional new particles.

The Higgs production for this process can be measured most precisely in case of the Z decaying via leptonic channels $Z \rightarrow \mu^+\mu^-$ and $Z \rightarrow e^+e^-$. Therefore, the tracking detectors proposed for ILC must perform an excellent momentum resolution, $\sigma_{p_T}/p_T^2 = \sigma_{1/p_T}$ of 3.10^{-5}GeV^{-1}, a factor ten better than the one achieved at LEP [25, 39]. The requirements on momentum resolution are needed for reconstruction of the mass of the Higgs boson with high precision from the mass recoiling to the associated Z boson decaying to a lepton pair in the Higgs-strahlung process ($Z \rightarrow \mu^+\mu^-$) [18–20]. This measurement is completely model independent without any assumption on the decay mode of the Higgs particle. Previous studies showed that a precision of about $2-3\%$ on the ZH cross section can be obtained at ILC [21]. However, the decay branching ratios for the leptonic channels is small ($\sim 3.3\%$) which limits the statistical precision and thus the hadronic decay mode of the Z boson ($Z \rightarrow q\overline{q}$) with the largest branching ratio ($\sim 70\%$), should be considered for a Higgs study.

Like many interesting physics processes at the ILC characterised by multi-jet final states, the measurment of the Higgs production process $e^+e^- \rightarrow ZH$ in events with four jets, in the presence of background from WW pairs requires a good jet energy resolution, to distinguish between W and Z bosons. The requirement demands a jet energy resolution that is of order $3-4\%$ above 100 GeV.
Indeed, to achieve the physics program at ILC, we need a clean separation of the

hadronic decays of W and Z bosons with an explicit goal to reconstruct the di-jet invariant masse with a resolution comparable to the gauge boson widths, $\sigma_m/m = 2.7\% \approx \Gamma_W/m_W \approx \Gamma_Z/m_Z$.

Fig. 1.3 shows idealised reconstructed W and Z mass distributions for different assumed mass resolutions. Good separation can be obtained for a mass resolution of 2.5%, which corresponds to a jet energy resolution of 3.5% [39].
As a consequence, to reach the performance requirements of the physics goals at ILC, the detector design have been optimized in order to use the particle flow approach for event reconstruction. In this approach the use of highly segmented calorimeters with imaging capabilities as the one studied in this work is a necessity.

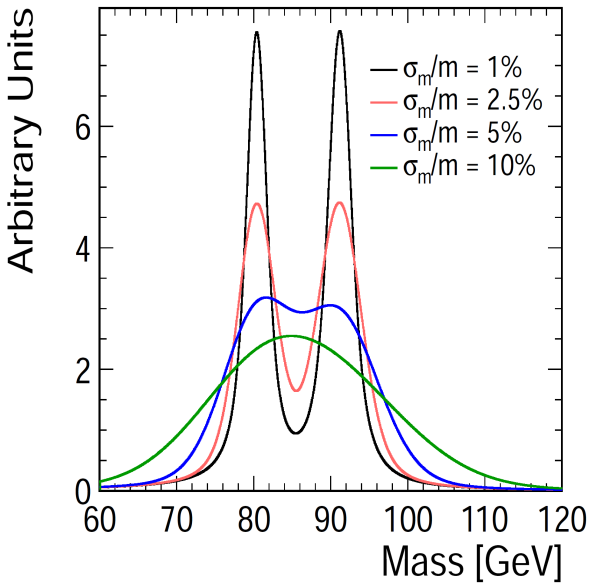

FIGURE 1.3: Ideal W/Z separation vs. jet mass resolution obtained using a Gaussian smearing of Breit-Wigner distributions [16].

1.2 ILC Accelerator

The International Linear Collider (ILC) is the most advanced project of a linear electron-positron collider with a continuous center-of-mass energy range between 200 GeV and 500 GeV. It is a high luminosity collider, based on 1.3 GHz superconducting radio-frequency (SCRF) accelerating technology with a peak of luminosity reaching $\approx 2 \times 10^{34}$ cm^{-2} s^{-1} at 500 GeV center-of-mass. The total length of the ILC accelerator is $\approx$ 31 km long. ILC can be upgraded to a center-of-mass energy up to and beyond 1 TeV by extending the length from 31 to 50 km.

The layout of the accelerator complex, indicating all the major subsystems, is shown in Fig. 1.4, as defined in the technical design report (ILC TDR) [22].

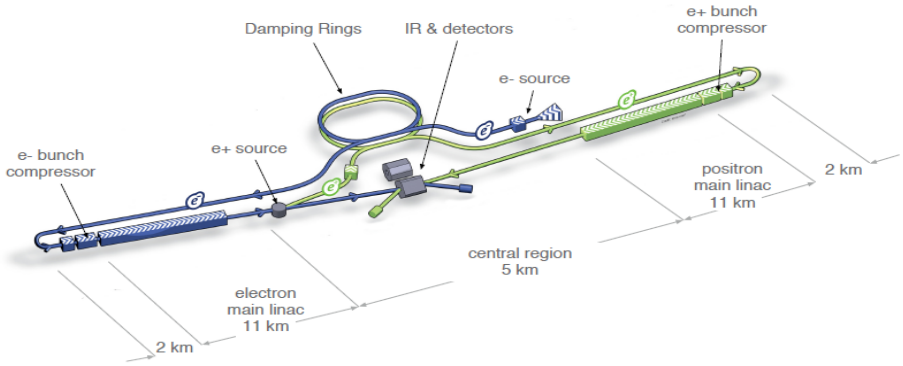

FIGURE 1.4: Schematic layout of the International Linear Collider (ILC) [22].

The main subsystems of the collider are :

– A polarized electron source : generates polarized electrons by illuminating a Gallium-Arsenic (GaAs) photo-cathode using laser light in a DC gun. Electrons are then accelerated, through different stages to 76 MeV using normal-conducting structures, after which the electron bunch train is accelerated to 5 GeV in a superconducting linac. The electron beam have a polarization higher than 80% and the positron beam have a polarization of up to 50%. The use of polarized beams could greatly increase the production rate of a signal and also effectively suppress background noise.
– 5 GeV electron and positron damping rings (DR) : The electrons from the source are injected in an electron damping ring of 3.2 km of circumference, used mainly to reduce the beam emittance.

– The ILC Ring to Main Linac (RTML) : is responsible for beam transport from damping ring to the entrance of the main accelerator. During this step, the electron bunch is compressed and accelerated to 15 GeV.
– The main linacs : using SCRF cavities (Fig. 1.5) operating at 1.3 GHz with an average gradient of 31.5 MV/m [23,24]. In the main linac, electrons are accelerated to an energy of 250 GeV, over a distance of 11 km.
– A polarized positron source : is responsible for the positron beam generation using the electrons accelerated in the main linac. At an energy of 150 GeV, the electrons pass through an undulator to generate polarized photons of an energy of at least 10 MeV. These photons are sent on a thin titanium foil where they generate polarized positrons through pair production. Afterwards, the positrons are captured, separated from electrons and unused photons, accelerated to 5 GeV and transported to the positron Damping Ring. Positron acceleration similar to the electron acceleration occurs in a second main linac.
– Two beam-delivery systems (BDS), each 2.2 km long : transport the positron and electron beams from the linac exits and bring them into collision with a 14 mrad crossing angle, at a single interaction point which can be shared by two detectors in a so-called push-pull approach.

FIGURE 1.5: A 1.3 GHz superconducting radio-frequency cavity in niobium used for e^+/e^- beam acceleration in ILC accelerator [23].

1.3 ILC Detectors

In contrast to the LHC, the ILC accelerator has been designed to have only one interaction region. However, in order to avoid physics bias associated to the geometry or components of the detector while providing cross-checking and confirmation of results, two independent experiments have been planned enabling two detectors running, alternatively, in a common interaction region.

This may occur using a push-pull configuration as shown in Fig. 1.6. In this scheme, when one detector is taking data, the other is out of the beam in a close-by maintenance

position. The detectors can be exchanged at regular intervals, by moving the data-taking detector out of the interaction region, while the other detector is being pulled in. The detector motion is ensured with platforms that preserve the detector alignment and the accuracy of detector positioning($\pm 1mm$).

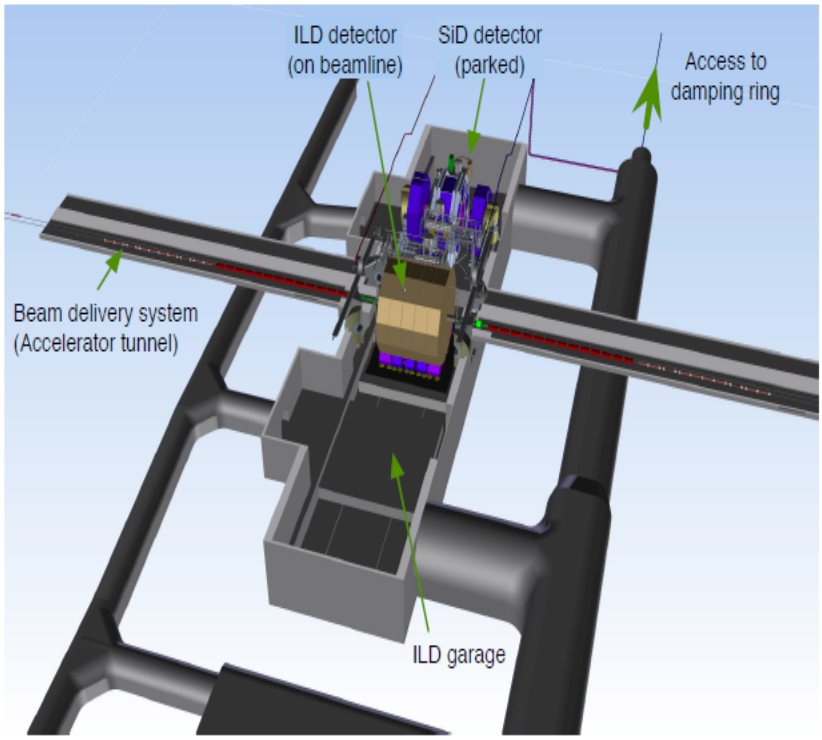

FIGURE 1.6: Example layout of the detector, showing the location of the two detectors in a push-pull configuration [25].

The two detectors (Fig. 1.7) named the International Large Detector (ILD) and the Silicon Detector (SiD), have been conceived, developed and optimised to fulfill the ILC physics program over the full range of center-of-mass energies from 200 GeV up to 1 TeV. The detector concepts are designed to achieve unprecedented advances in collider detector performance by using new technologies and reconstruction techniques based on the Particle Flow Algorithm (PFA)(see section 1.4).
In particular, this will allow to meet one of the most important chalenges in ILC experiments which is the achievement of high-resolution jet energy reconstruction. For this

purpose, highly granular electromagnetic and hadron calorimeters as well as highly efficient tracking systems are required.

SiD is a compact detector with a 5 Tesla magnetic field and a tracking detector made of silicon, with silicon-tungsten electromagnetic calorimetry (ECAL) and highly segmented hadronic calorimetry (HCAL).
ILD is a larger detector using a tracking system based on a time-projection chamber (TPC) combined with silicon tracking. Its granular calorimeter system is contained inside a 3.5T magnetic field and provides very good particle-flow reconstruction.

Both detector concepts offer robust and stable performance over a large energy range. Their technical details were described in a letter of intent, validated and signed by the International Advisory Group in 2009 [26]. Since then, major progress has been made and a relatively mature version of ILC detectors was presented in the report (Detailed Baseline Detector) published in 2013 in ILC TDR (Technical Desin Report) [22].

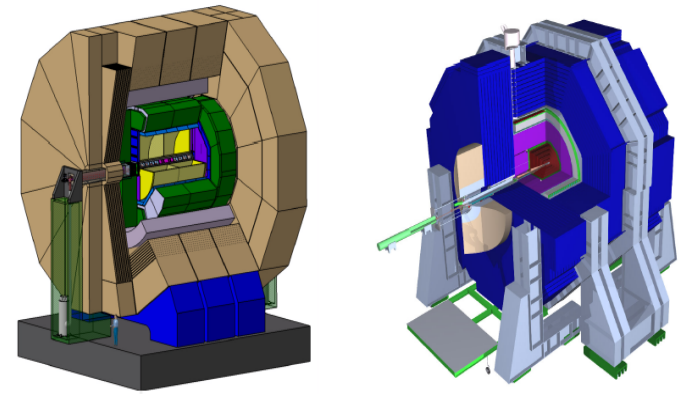

FIGURE 1.7: (a) : The International Large Detector ILD. (b) : The Silicon Detector SiD.

1.3.1 ILD

The ILD design results from the combination of the earlier GLD and LDC detector concepts [27–29]. An artistic view of the detector is shown in Fig. 1.8(a), a cross section of the detector is seen in Fig. 1.8(b).

ILD has been conceived and optimised in the aim to reach an excellent precision in jet energy reconstruction by using PFA. This has an important impact on the design of the main sub-detectors of ILD. In particular, the choice of PFA imposes a number of special requirements on the tracking and calorimeter systems.

1.3.1.1 Tracking system

An important feature of ILD is its powerful tracking system allowing superb tracking capabilities and excellent detection of secondary vertices. It consists of Silicon based vertexing and tracking system combined with a high precision large volume time projection chamber. All sub-detectors, together with the calorimeter system, are placed in a 3.5 T strong magnetic field.

– **Vertex Detector - VTX** : The vertex detector allows the measurments of the primary vertices as well as the secondary decay vertices of short lived particles, such as D or B mesons, by reconstructing the trajectory of their decay products. This requires a thin detector, to minimize interactions before the calorimeters, and precise vertex reconstruction. The goal of the ILD vertexing system is to reach a resolution on the track impact parameter [1] of $\sigma_b < 5 \oplus 10/p \sin^{3/2} \theta \mu m$ [22].
The required performance of the ILD vertex detectors resumes in a spatial resolution near the IP of 2.8 μm besides a low material budget below $0.15\% X_0/layer$.
In practice, the ILD vertex detector consists of three, nearly cylindrical, concentric layers each comprising two layers, with radii ranging from 16 mm to 60 mm. Each layer is equipped with pixel sensors on both sides, $\geq$ 2 mm apart, resulting in six measured impact positions for each charged particle traversing the detector.
– **Silicon Tracking** : The ILD tracking system includes four silicon microstrip detectors : two barrel components, the Silicon Inner Tracker (SIT) between the VTX and the TPC and the Silicon External Tracker (SET) between the TPC and the electromagnetic calorimeter, one end cap component, the End Tracking Detector (ETD) behind the endplate of the TPC, and the forward tracker (FTD) between the beam pipe and the TPC. The layout of the system is shown in Fig. 1.9.
The ETD, SIT and SET detectors improve the momentum resolution, allow to link the VTX detector with the TPC, and help in extrapolating from the TPC to the calorimeter by providing precise space points before and after the TPC. The FTD covers the very forward region and ensures efficient and precise tracking down to very small angles.
– **Time Projection Chamber** : The TPC is the central tracker of the ILD tracking system, where tracks can be measured continuously with a large number of three-

1. the shortest distance from the interaction point to the track.

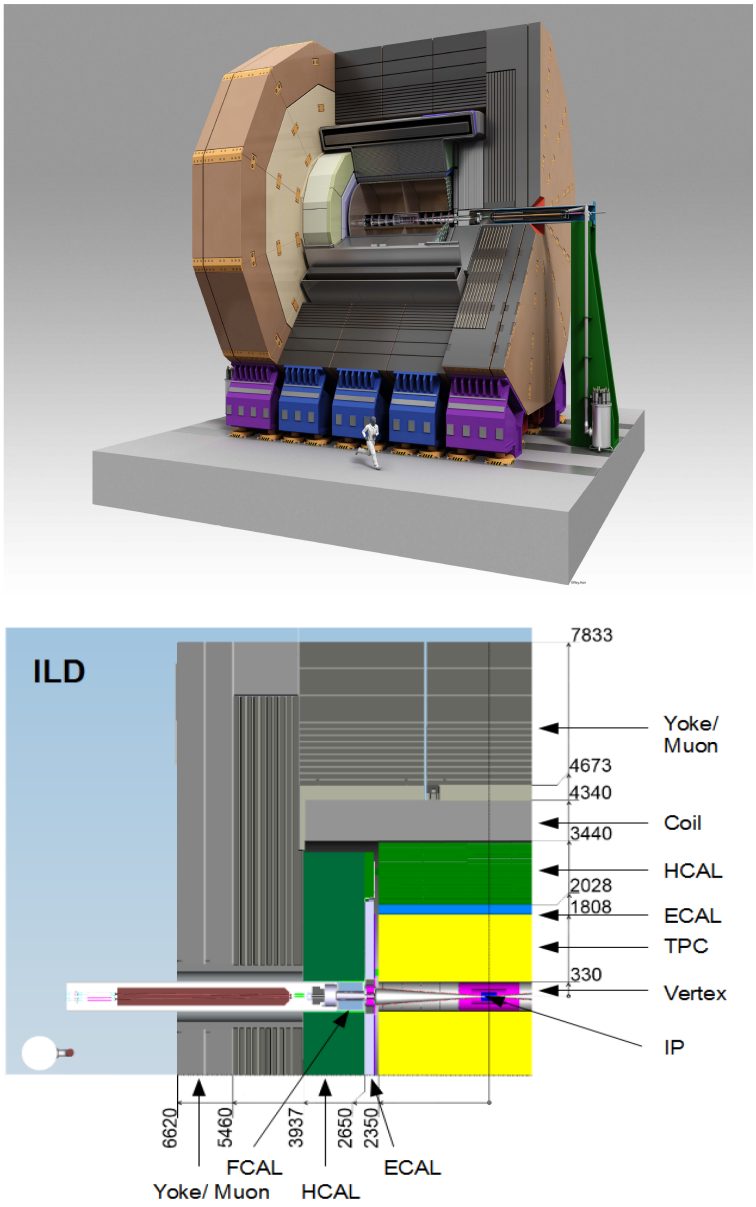

FIGURE 1.8: Views of the ILD detector concept. Dimensions are in mm [22].

dimensional (r, ϕ, z) space points. It offers several advantages such as a good spatial resolution compared to silicon detectors, while maintaining a low budget material as

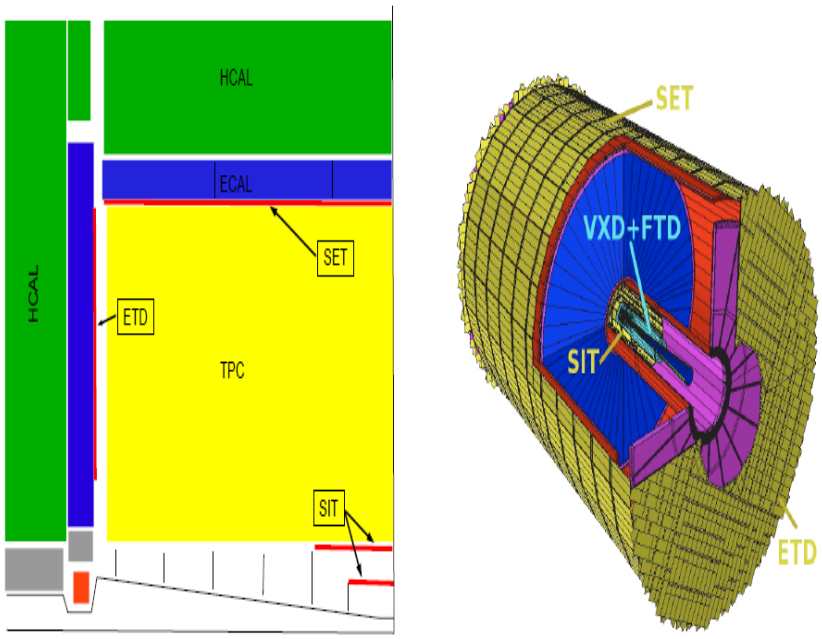

FIGURE 1.9: Left : a cross view of the ILD silicon tracking system. Right : a 3D GEANT 4 simulation description of the silicon system as sketched in the cross view on the left [22].

needed for the best calorimeter and PFA performance. Charged particles traversing the TPC volume will ionise the gas. In the presence of an electric field parallel to the beam, ionisation electrons drift towards the endcaps where they undergo a charge multiplication process, using either Gas Electron Multiplier (GEM) or Micromegas technology [30, 31]. Both options would use pads of size $1 \times 6mm^2$ which has been found to provide good resolution and to guarantee a low occupancy [43]. An alternative technology based on a pixelated readout with much smaller pads is being investigated [32]. The determination of the three-dimensional coordinates is thus provided by reading the position of the fired pad. In addition, the measure of the energy loss per unit of length dE/dx along the track trajectory in the gas, can provide valuable particle-identification discrimination.

1.3.1.2 Calorimeter system

The ILD design is optimised in order to implement the PFA approach as a basis for event reconstruction. This has direct consequences on the design of the calorimeters by imposing a number of basic requirements.
Indeed, particle flow allows to reconstruct individually each particle in the appropriate detector. Thus, the electromagnetic and hadronic calorimeters must have excellent imaging capabilities that allow correct assignment of energy clusters deposits to charged or neutral particles. This implies that calorimeters needs to be highly granular transversely and finely segmented longitudinally. In addition, in order to avoid significant energy leakage, calorimeters should be sufficiently thick and hermetic. The calorimeter depth should be optimised with respect to the highest-energy collisions envisaged, to ensure a full containment of hadronic showers.

Electromagnetic Calorimeter The electromagnetic calorimeter (ECAL) is composed of a central barrel part and two endcaps as shown in Fig. 1.10.

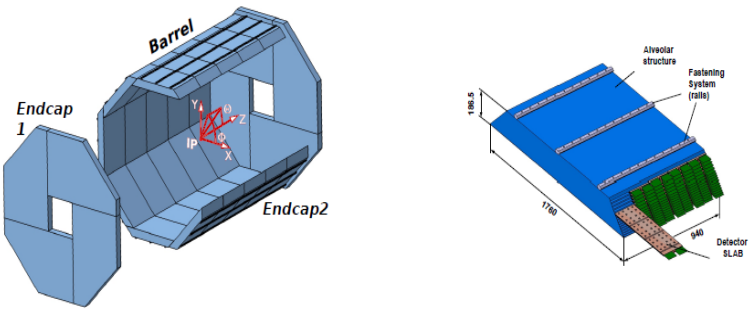

FIGURE 1.10: ECAL barrel and end-cap (left) and one ECAL module (right) [22].

It is a highly segmented electromagnetic calorimeter (ECAL) providing 30 alternating layers of active and absorber materials with varying absorber thicknesses. In order to reach the most compact design and a better separation of nearby electromagnetic showers, tungsten has been chosen as absorber material (radiation length X_0 =3.5 mm, Moliere radius R_M=9 mm, hadronic interaction length λ_I =99 mm), thanks to its small Molière radius and its large ratio of interaction length to radiation length ; this with a depth of roughly 24 X_0, compared to lead, will allow a good separation of electromagnetic and hadronic showers while minimizing the leakage. For the active

layer, silicon diodes (SiECAL) or scintillator strips (ScECAL) are considered. In both options, the sensitive layers are segmented into cells with a lateral size of 5×5 mm^2. In the baseline model, the SiECAL consists of 30 silicon (Si) and 29 tungsten (W) layers. However in order to optimize it, other configurations of different number of silicon layers have been studied. In this context, the jet energy resolution, using $Z \rightarrow q\bar{q}$ events generated at center-of-mass energies of 91, 200, 360, and 500 GeV was estimated for each of these configurations. The results are presented in Fig. 1.11(left) and show a degradation of 10% in jet energy resolution when reducing from 30 to 20 the number of silicon layers at 91 GeV, while a small degradation is observed for higher energies in this case. However, when the number of silicon layers decreases below 20, a significant deterioration of the resolution is observed.

A SiECAL prototype has been conceived and validated in beam test within the CALICE collaboration [33]. It consists of 30 silicon and 29 tungsten layers with a transversal segmentation of 10×10 mm^2. The obtained energy resolution for electrons in this case is $\sigma_E/E = 16.6\%/\sqrt{E} \oplus 1.1\%$
In addition, a ScECAL prototype has been also validated into beam test [34]. Nevertheless, scintillators individual tiles of size 5×5mm^2 are difficult to realise. However, this can be resolved by using strips of 5×45 mm^2 arranged in alternative directions so that an effective granularity approaching 5×5 mm^2 can be achieved. In the case of ScECAL CALICE prototype the active layer consists of scintillators strips of 10×45 mm^2 arranged alternately in order to realize an effective granularity of 10×10 mm^2. This option has the advantage to reduce the number of electronic channels but the event reconstruction becomes more difficult, in particular in dense jets. The obtained energy resolution in this case is $\sigma_E/E = 13.5\%/\sqrt{E} \oplus 3.5\%$.

An alternative option consists in an hybrid ECAL option. It is a mixture of silicon layers and scintillator-strip layers with 27 layers of tungsten absorber. Different configurations of the active layers were studied and the jet energy resolution in $e^+e^- \rightarrow q\bar{q}$ (q = uds) events generated at center-of-mass energies of 91, 200, 360, and 500 GeV was estimated for each of these configurations. Fig. 1.11(right) shows the jet energy resolution as a function of the fraction of scintillator layers in the ECAL for each of the jet samples. At low energies, the increasing of the number of scintillator layers has no effect on the jet energy resolution. At higher energy, the resolution is deteriorated when increasing the number of scintillator layers, especially for a fraction above 50%.

Hadronic Calorimeter The Hadronic calorimeter (HCAL) allows the separation of energy deposits belonging to charged and neutral particles, for the reasons explained in section 1.4, in order to guarantee a precise measurement of the energy of neutral

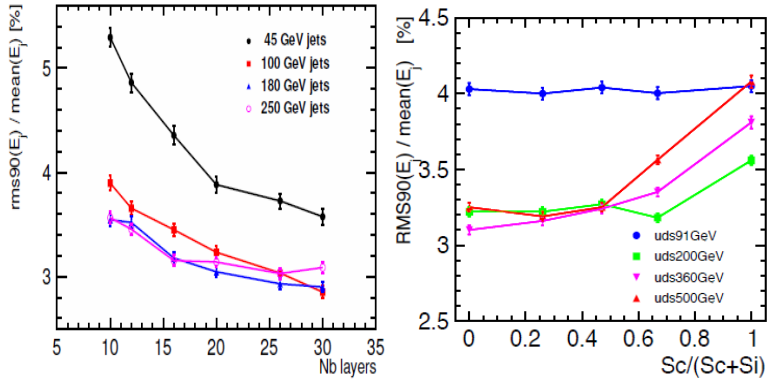

FIGURE 1.11: Left : Jet energy resolution as a function of the number of silicon layers for the SiECAL option (Silicon only ECAL). The resolutions are shown for $e^+e^- \rightarrow q\overline{q}$ events at different center-of-mass energies. Right : Jet energy resolution in $q\overline{q}$ events at different center-of-mass energies, using a hybrid ECAL with silicon (Si) and scintillator (Sc) layers. The jet energy resolution is shown as a function of the fraction of scintillator layers in the ECAL, (Sc/(Sc+Si)). The total number of ECAL layers (Sc+Si) is 28 [22].

hadronic particles. Therefore, in order to achieve this goal with a high accuracy of measurment, the calorimeter must be highly granular for an optimal application of PFA.

It is a sampling calorimeter consisting in alternating active layers and stainless steel absorber layers. Two technologies options have been conceived, the scintillator based AHCAL and the Glass Resistive Plate Chamber (GRPC) based SDHCAL. The active medium of AHCAL is made of scintillating tiles providing a transverse segmentation of 3×3 cm^2. However, the cell size parameter has been optimized for AHCAL using full detector simulation with particle flow reconstruction. The performance of AHCAL with respect to the cell size for different energies is shown in Fig. 1.12(left). The energy resolution is not improved for cell sizes below 3×3 cm^2. However, the resolution is deteriorated for larger cell sizes. As a consequence, the choice of 3×3 cm^2 has been adopted for AHCAL.

A physics prototype has been built and validated within the CALICE collaboration in 2007 [35]. The hadronic energy resolution measured with this prototype is $\sigma_E/E = 58\%/\sqrt{E} \oplus 1.6\%$ [36].

For the second option, SDHCAL, the GRPC is used as active layer with a fine granularity of 1×1 cm^2 and three-threshold electronics (2 bit) readout. The fine granularity is motivated by PFA application but also required for a precise energy measurment

since this latter is obtained by counting the particles generated in a hadronic shower. The cell size was selected for SDHCAL based on a study using a full ILD detector model with 48 active layers indicating that a 1 cm size cell realizes better performance than a one of 3 cm size (Fig. 1.12 right). The SDHCAL prototype as well as its optimization will be described and discussed in this thesis.

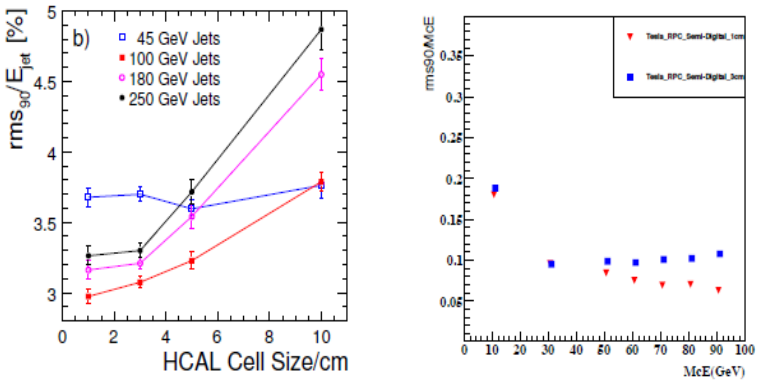

FIGURE 1.12: Optimization of the hadron calorimeter cell sizes. Left : Particle flow jet energy resolution as a function of the AHCAL cell size. Right : Single K_0 energy resolution for particles showering in the SDHCAL for two different cell sizes [22].

In the baseline model, two different geometries are being proposed for the barrel calorimeter. The first version is separated longitudinally into 2 rings and azimuthally into 16 modules, and is conceived to host the AHCAL. The second version is presented in Fig. 1.13 and is conceived to host SDHCAL. In addition, a system of special calorimeters are intented for the very forward regions of the detector, close to the beam pipes : LumiCal and BeamCal. They have the advantage to be extremely resistant to radiation, up to one MGy per year. The role of LumiCal is to measure the luminosity with an accuracy of better than 10^{-3} at 500 GeV center-of-mass energy using Bhabha scattering. BeamCal will perform a fast estimation of the luminosity based on beamstrahlung radiation.

1.3.1.3 Magnet System and Muon system

The magnet system is composed of three main parts : the superconducting solenoid coil surrounding the calorimetric and tracking systems, will create a magnetic field up

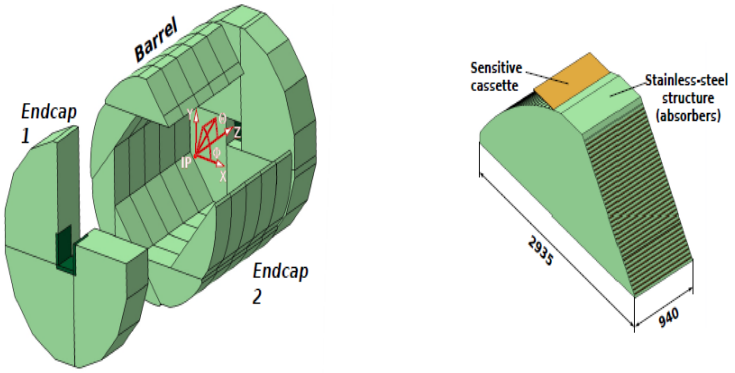

FIGURE 1.13: HCAL barrel and end-cap (left) and one HCAL module (right) [22].

to 3.5 T ; an anti-DID [2](Detector Integrated Dipole) reduces the beam background in the vertex and tracking volume ; the iron yoke, consisting of the barrel yoke and the two end-cap yokes, will be used to return the magnetic flux.

Muon system is used to identify the muons and to measure their momentum. It is also used as a tail catcher, to recover the energy of the hadronic showers which leak through the external side of the calorimeter system. The muon detectors are hosted in the iron yoke which consists of the barrel yoke and two end-cap yokes, of dodecagonal shape. The barrel yoke is composed of one sensitive layer in front of the iron yoke, 10 layers spaced 14 cm apart, followed by three sensitive layers spaced by 60 cm apart. The end-caps of the yoke are equipped each with 10 layers spaced by 14 cm, followed by two sensitive layers spaced by 60 cm. The layout of the muon system is shown in Fig. 1.14. Two options are proposed for the sensitive layers, scintillator strips or resistive plate chambers (RPC).

2. The Detector Integrated Dipole is provided by integrating a dipole coil with the detector solenoid. The anti-DID configuration can be achieved when the field is aligned with the outgoing beam

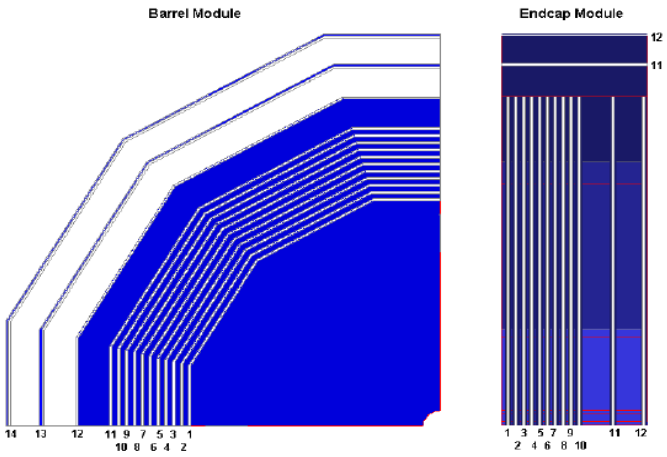

FIGURE 1.14: Sensitive Layers of ILD Muon System Catcher [22].

1.4 Particle Flow Algorithm

The particle flow event reconstruction algorithm (PFA) [37–39] has been of prime importance and a guide for the detectors designs at ILC. An explicit goal of its application is the capability to reach a jet energy resolution that is of order 3-4% at 100 GeV [17]. As explained, this is required to separate W and Z di-jet final states.

In the traditional reconstruction approach the so called energy flow, the jet energy is measured in the calorimeter system : $E_{Jet} = E_{ECAL} + E_{HCAL}$ with a typical jet energy resolution of $55\%/\sqrt{E}$. However, the average energy content of a jet is about 60% charged particles (mainly hadrons), around 27% photons, about 10% neutral hadrons and the rest is carried by neutrinos.
Unlike in the traditional reconstruction approach, the PFA allows to measure separately the energy of each individual particle in the jet in the appropriate sub-detector. Therefore, the energy of the charged particles will be measured only by the tracker, while photons and neutral hadrons are measured with the calorimeters, thus, only 10% of the jet energy is measured with the hadronic calorimeter. Hence, the jet energy is estimated as $E_{JET} = E_{Track} + E_\gamma + E_n$, where E_{Track}, E_γ and E_n are respectively the sum of the energies of the charged particles, of photons and of neutral hadrons.
The principle of the Particle flow approach is illustrated in Fig. 1.15.

The efficiency of a PFA application relies on the capacity of the calorimeter system to discriminate the energy deposits of neutrals from those of charged particles. Calorimeters are thus, required to perform a clean separation of tracks belonging to charged and neutral particles as well as excellent two-track separation capability for nearby showers. Therefore, the primary emphasis, in the calorimeter design, is shifted towards a fulfilment of excellent imaging capabilities than achieving the best intrinsic single particle energy resolution.

However, even if ILC calorimeters are non-compensating, the high granularity offers the possibility for an efficient software compensation to equalize the response to electromagnetic and hadronic showers. It is conceivable that after applying a software compensation technique, the energy resolution is improved. The compensation algorithm makes use of the fact that electromagnetic sub-showers have different spatial characteristics compared to purely hadronic energy deposits. Using the high granularity of the calorimeter, fluctuations between the electromagnetic and hadronic component of the shower can be corrected on an event by event basis. The AHCAL has an e/π ratio of approximately 1.2. The application of a software compensation improves the resolution by close to 20%, reaching a stochastic term of 45$\%$ [36]. Ongoing studies focuses on using a software compensation in order to improve the energy resolution in SDHCAL.

In a realistic calorimeter, an overlapping of nearby showers produced by charged and neutral particles may occur, and thus generates errors in matching particle clusters to tracks. This leads to a confusion within the calorimeter between the individual particles contributions and thus deteriorate the jet energy resolution. In this case, using the particle flow, the jet energy resolution is defined by :

$$\sigma_{Jet} = f_{Track} \cdot \sigma_{Track} \oplus f_{\gamma} \cdot \sigma_{ECAL} \oplus f_{n} \cdot \sigma_{HCAL} \oplus \sigma_{conf} \tag{1.1}$$

where f is the energy fraction of the particles, the indexes $Track$, γ and n indicate the charged particles, photons, and neutral hadrons respectively. σ_{conf} is the confusion term which limits the particle flow performance as shown in Fig. 1.16.

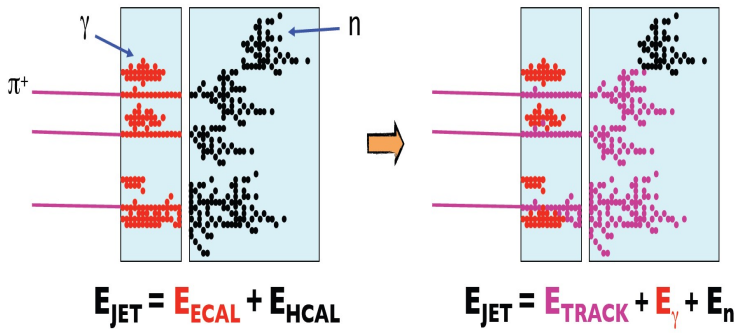

FIGURE 1.15: Schematic representation of the Particle flow approach.

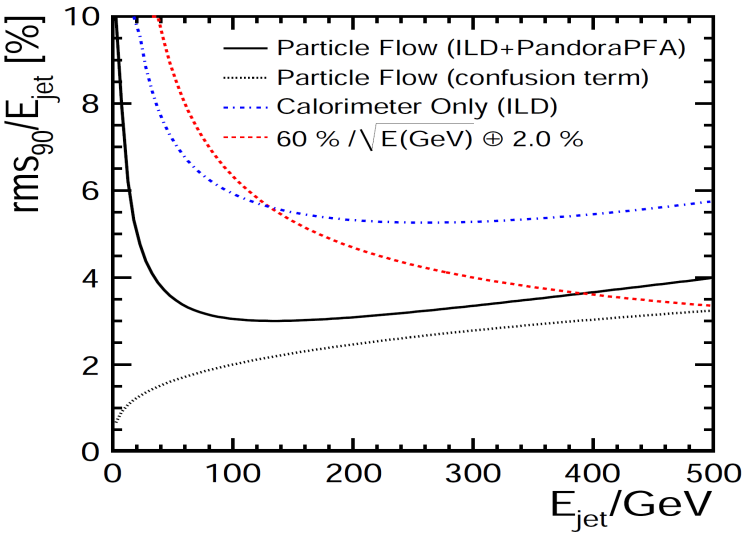

FIGURE 1.16: The jet energy resolution as a function of the jet energy obtained from PFlow calorimetry (PandoraPFA and the ILD concept). The blue dot-dashed line corresponds to the resolution obtained using only the calorimeter information, the black dotted line is the PFA contribution of the confusion term, the solid black line is the PandoraPFA jet energy resolution, compared to the jet energy resolution expected from the single track shower reconstruction approach (red dashed line). Figure taken from [39].

1.5 Conclusion

The ILC accelerator project described in this chapter is technically ready to be constructed and exploited. Several countries have shown an interest to accommodate the experience, including Japan where a site for ILC has been already shosen in mountainous regions of Japan : Kitakami in the Tohoku district in northern Japan. The ILC construction should start when the Japanese government gives its green light for the project.

In addition to the ILC accelerator studies, other projects of electron-positron colliders are proposed. The first project is the Compact Linear Collider CLIC with a center-of-mass energy between 1 and 3 TeV [40]. The detector concepts for CLIC use the ILC detector designs as a baseline. Two circular electron-positron collider projects are also under study : FCC (Future Circular Collider) [41] and the CEPC (Circular Electron Positron Collider) [42].
The conclusions and applications of the techniques described in this work can be adopted with minor changes to any of these projects.

Chapitre 2

Calorimetry

The knowledge of the basic physics of the interactions of elementary particles at colliders and the effects produced by these processes is fundamental in particle physics experiments. Indeed, the understanding of these processes is the basis of the particle detection and its measurement and thus determines the performance of a detector. The detectors at colliders, such as calorimeter devices, are used to measure the properties of the final state particles of physics processes of interest. These measurements result from the interaction of the particles with the detector matter which manifests in different kinds of processes (ionization, Cherenkov radiation, transition radiation), depending on the particle type, its energy, and the type of matter. Thus, each subdetector is designed to measure separately some particular properties of certain types of particles by focusing on specific interaction types. Finally, the measurements provided by all the subdetectors are combined together to determine the properties of most of the resulting particles.

This chapter is intended as an overview of calorimetric particle detection in high energy physics. First, the physics processes involved when high energy particles interact with matter are described. Then, a description of the electromagnetic and hadronic shower development in calorimeters is presented. Finaly, some characteristics of calorimeters are discussed.

2.1 Passage of Particle Through Matter

When traversing matter, radiation interact with the basic constituents of the material resulting in a loss of its energy as well as a deflection of its incident direction.

Indeed, during its interaction with the detector, the incident particle deposits its energy progressively in the active material of the calorimeter untill it is fully absorbed. This absorption leads to the initiation of a shower of secondary particles. The deposited energy by these secondary particles is then detected and used to measure the primary energy of the incident radiation.

Passing through material, the interaction of the radiation with the atoms may occur through different kinds of processes, depending on the radiation type, its incident energy and the type of the material.Processes related to the charged particles and photons involve mainly the electromagnetic interactions concerning in particular inelastic collisions with the atomic electrons of the material. However, charged particles can be divided into two categories : electrons and positrons (1), and heavy particles (2) such as muons, pions and protons.
Although, the theory behind their energy loss processes is well developed and detailed in many studies [44–48], a brief idea on these processes is presented in the following.

Energy loss of heavy charged particles For heavy charged particles, the inelastic collisions with atomic electrons are the major reactions which can occur, resulting in an ionization or excitation of the atoms of the material. Indeed, during each collision, the particle transfers a small fraction of its energy to the atoms of the material. With enough transferred energy, an electron can be extracted of an atom which is thus ionized. Otherwise, the atom is left in an excited state. However, this reaction occur many times with a large number of collisions per unit path lengh. Thus, during its travel through matter, the particle can lose all its energy by cumulating all the transferred ammounts of energy in collisions.

Using relativistic quantum mechanics, the average energy loss per unit path lengh in a uniform medium for a heavy charged particle $< -\frac{dE}{dx} >$ can be described by Bethe-Bloch equation in the region $\beta\gamma \in [0.1, 1000]$:

$$\langle -\frac{dE}{dx} \rangle = Kz^2 \frac{Z}{A} \varrho \frac{1}{\beta^2} \left[\frac{1}{2} \ln \frac{2m_e c^2 \beta^2 \gamma^2 T_{max}}{I^2} - \beta^2 - \frac{\delta}{2} \right]$$

$$\frac{K}{A} = \frac{4\pi N_A r_e^2 m_e c^2}{A}$$

With classical electron radius : $r_e = \frac{e^2}{4\pi\epsilon_0 m_e c^2}$.

$T_{max} \approx 2m_e c^2 \beta^2 \gamma^2$ is the maximum kinetic energy transferred to a free electron in a single collision, for $M \gg m_e$.

where I is the mean excitation energy, δ the density effect correction, m_e the electron mass, N_A the Avogadro constant, z the charge of the incident particle, Z and A the atomic number and atomic mass of the material. $\beta = v/c$ is the speed of the particle relative to the speed of the light in vacuum, and γ is the particle Lorentz factor $\frac{1}{\sqrt{(1-\beta^2)}}$.

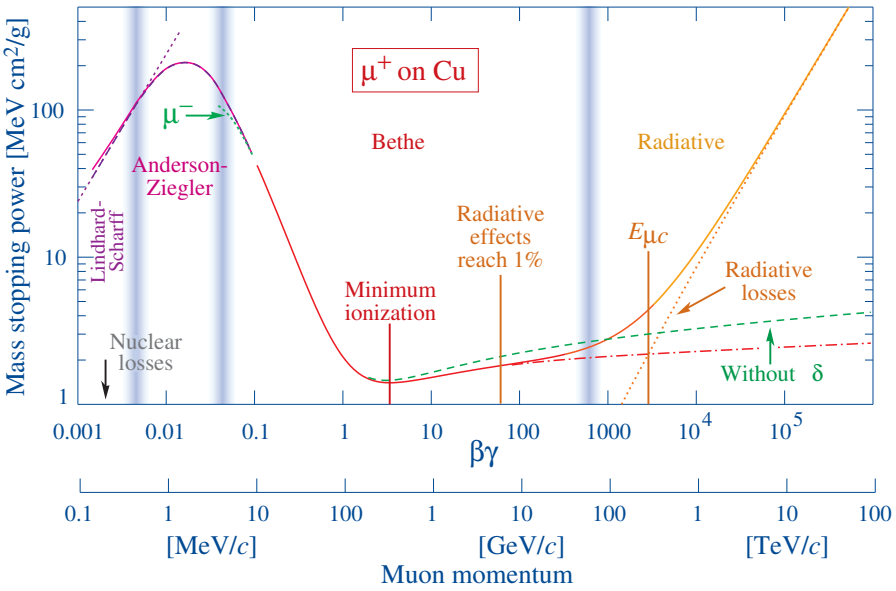

FIGURE 2.1: Stopping power $< -\frac{dE}{dx} >$ for positive muons in copper. The stopping power is plotted as a function of the particle velocity $\beta\gamma$. Figure taken from [46]

Fig. 2.1 shows the stopping power for positive muons in copper. The Bethe-Bloch range is shown in the region $0.1 \leq \beta\gamma \leq 1000$.
For very low energy region, where the particle velocity is comparable to the velocity of the electrons of the atoms, a number of complicated processes occur and $\frac{dE}{dx}$ reaches a maximum followed by a decrease. In the Bethe-Bloch regime, at non-relativistic energies, the $\frac{dE}{dx}$ decreases with increasing velocity untill a minimum is reached. Particles at this point are called Minimum ionizing Particles (MIPs). According to Fig. 2.1, muons having energy between 100 MeV and 100 GeV are considered as MIPs. Beyond this point, $\frac{dE}{dx}$ rises again while above the critical energy $E_{\mu c}$, radiative effects dominate.
However, one can notice the weak energy dependence of the energy loss between its minimum and the critical energy. The energy loss of muons stays very low in this re-

gion with a high value of $E_{\mu c}$ and thus one needs a very large calorimeter to stops muons. Therefore calorimeters are not appropriate for a muon energy measure. Nevertheless, generally in this range, muons are used in detector calibration.

Energy loss of electrons and positrons Two principal processes characterize the energy loss of electrons and positrons through matter. Firstly, like heavy charged particles, their energy loss is due to ionization which dominates at low energies. Besides, because of their small mass, a loss of energy by radiation due to bremsstrahlung come into play and domninates at high energies above a critical energy. This process leads to the emission of a photon when an electron (or a positron) is scattered in the electric field of the nucleus. In addition, further rare processes may occur in the low energies range.
Fig. 2.2 shows the energy lost by electrons and positrons in lead. At energies above 10 MeV, bremsstrahlung is the dominant process which yields a high energy loss for electrons compared to muons. This makes easy the total absorption of electrons in thin calorimeters. At energies up to 10 MeV, ionization dominates. However, at energies of a few MeV, other processes such as $Møller$ scattering ($e^-e^- \rightarrow e^-e^-$), Bhabha scattering ($e^+e^- \rightarrow e^+e^-$) and electron positron annihilation ($e^+e^- \rightarrow \gamma\gamma$), are contributing as well at a small factor.

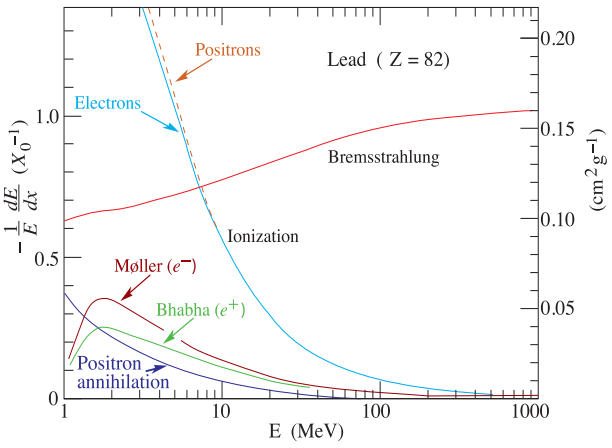

FIGURE 2.2: Fractional energy loss per radiation length in lead as function of electron or positron energy. Figure taken from [46]

Energy loss of photons The main processes involved in photon interaction in matter are the photoelectric effect, Compton scattering (including Thomson and Rayleigh scattering) and pair production. Fig. 2.3 shows the cross sections of these processes versus the incident photon energy. At low energies, although the photoelectric effect dominates, Compton and Rayleigh scatterings contribute as well. With increasing energy, the cross section of both photoelectric effect and Rayleigh scattering decrease and Compton scattering becomes dominant between 1 to 10MeV. It should be noted that for Thomson and Rayleigh scattering processes, photons retain their original energy and no energy is transferred to the medium but only the photon direction is changed. Thus, these two pocesses are not considered responsible for the energy loss of photons in matter. At higher energies, these processes can be neglected and only the process of pair production becomes considerable. Note that this last process may occur only when the energy of the photon reach the threshold of twice the electron mass.

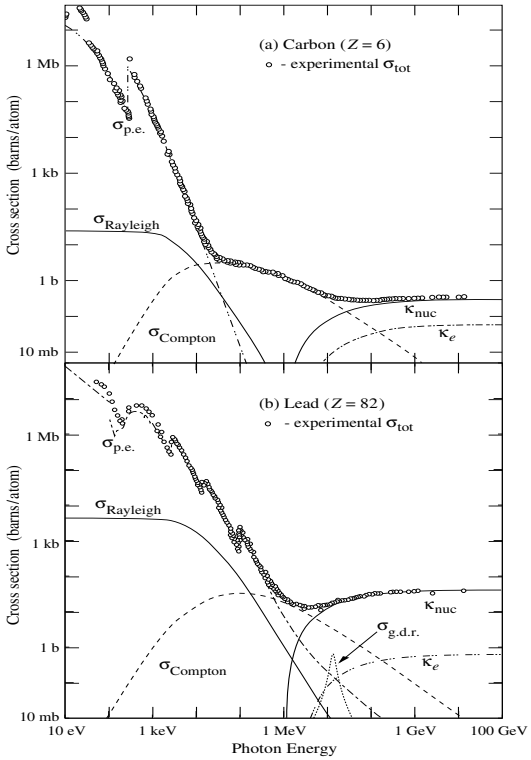

FIGURE 2.3: Cross section of photon interaction processes as a function of the photon energy in carbon (top) and lead (botom). Figure taken from [46]

2.1.1 Electromagnetic Showers

The interaction of high energy electrons, positrons and photons via bremsstrahlung emission and pair production gives rise to electromagnetic showers. High energetic photons emited by electrons or positrons, through bremsstrahlung, leads to the creation of an e^+e^- pair which, in turn, emit additional energetic bremsstrahlung photons converted then into further e^+e^- pairs, and so on. This process continues until the energy of the produced pair falls below the critical energy. The result of these repeated interactions is a cascade of photons, electrons and positrons as shown in Fig. 2.4. The longitudinal shower development represents the number of particles produced in the shower as a function of the penetration depth in the material and can be described using the notion of radiation lentgh X_0 which depends on the characteristics of the material :

$$X_0(g/cm^2) = \frac{716gcm^{-2}A}{Z(Z+1)\ln(287/\sqrt{Z})}$$

where Z and A are the atomic number and mass of the absorber material, respectively. The radiation lentgh represents the distance traveled by an electron in a material, after which it loses all but $\frac{1}{e}$ of its original energy by bremsstrahlung.

The transverse shower development is due to the multiple scattering of electrons and positrons arround the shower axis. It can be described using the Moliere Radius, R_M which is the radius of a cylinder containing on average 90% of the shower energy and defined by :

$$R_M(g/cm^2) = 21MeV\frac{X_0}{E_c(MeV)}$$

The Moliere Radius represents the average lateral deflection of electrons at the critical energy E_c after traversing one radiation length X_0 and is almost independent of the energy of the incident particle.

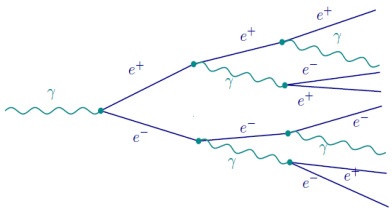

FIGURE 2.4: An example of an electromagnetic shower initiated by a high energy photon.

2.1.2 Hadronic Showers

Like electrons and positrons, charged hadrons interact with matter via electromagnetic processes. In addition, since it is a composite particle made of quarks, a hadron traversing a medium undergoes strong interactions with the nuclei of the absorber material. This makes hadronic interactions more complex than electromagnetic ones.

The result of these hadronic interactions is the production of secondary particles which form an hadronic cascade. Typical processes in an hadronic shower are shown in Fig. 2.5.

The hadronic shower is mainly composed by charged and neutral pions, but also other types of hadrons, nucleons (protons and neutrons) and neutrinos. Neutral pions, which represent one third of the pions, decay quickly into 2 photons which will interact further and develop electromagnetic subcascades in the hadronic shower. Note that the energy of a neutral pion, is totally transferred to the electromagnetic component and thus, cannot contribute anymore to the hadronic component.

The occurrence of strong interactions produces spallation protons and neutrons as well as evaporation neutrons and the nucleus is recoiled into highly excited states, which will be then de-excited by the emission of photons.
Charged particles produced in the hadronic shower deposit their energy via ionisation, which can be then detected and measured.
In contrast to this, the energy deposited by a part of the secondary particles, which are subject to strong interactions, does not contribute to the calorimeter signal, since it is undetectable for many reasons. This may occur either when a part of the hadron energy is used to break up the nucleus, or the decay products deposit a very small amount of their energy, such as muons generated by pions and kaons decays, or also are absorbed in an extremely short time before reaching the detection layer. Furthermore, some particles like neutrinos can escape from the calorimeter. This fraction of energy is considered as invisible. Therefore, for the same particle energy a hadron signal is smaller than an electron signal and the electromagnetic energy resolution is better than the hadronic one. The signal amplitudes of electron and hadron showers can be equalised by compensation techniques,.

To sumarize, the hadronic cascade contains three components : a hadronic component (1), produced by mesons, nuclear fragments and nucleons (protons and neutrons), an electromagnetic component (2) and a neutral component (3) formed by neutrons and neutrinos.

The longitudinal development of the hadronic shower is characterised by the average nuclear interaction length λ_I, which is the distance traveled by a hadron before it

undergoes a strong interaction in the material. It can be estimated as [49] :

$$\lambda_I = 35 g/cm^2 A^{1/3}$$

The interaction length λ_I is much larger than the radiation length X_0, which characterize the electromagnetic shower development. This leads to a larger longitudinal development of hadronic cascades compared to electromagnetic ones. Accordingly, the hadron calorimeter should be much larger than the electromagnetic calorimeter.

Similarly, the lateral development of hadronic showers is also increased compared to electromagnetic cascades, and can be described by a radial width. Indeed, the lateral distribution of a hadron cascade is initially very narrow but becomes wider with increasing calorimeter depth. Frequently, we define a radius to require a 95% lateral containment of the hadronic shower in calorimeter.

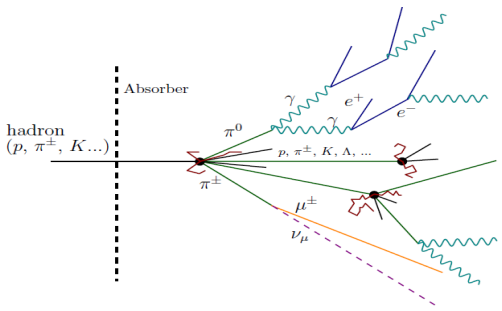

FIGURE 2.5: An example of a hadronic shower.

2.2 Calorimeters and Energy resolution

Calorimeters are instruments used in high energy physics to measure the energy of particles generating electromagnetic and hadronic showers. Accordingly, they are classified into electromagnetic calorimeters used to measure photons and electrons, and hadronic calorimeters, used to measure hadrons.

Calorimeter Types Depending on the design of the detector, calorimeters are divided into two types : Homogeneous and sampling calorimeters.

– **Homogeneous calorimeters :** Homogeneous calorimeters are constructed from a single medium playing the role of absorber and active detector at the same time.

Thus it has the advantage that all the deposited energy in the calorimeter can be detected and measured.

– **Sampling calorimeters :** are made of alternating layers of passive dense absorber and active layers. The absorber is used to degrade the energy of the incident particle, whereas the active layer is sensitive to the deposited energy which is measured either as charge signal or as light signal.

Calorimeter energy resolution The main parameter which characterises electromagnetic and hadronic calorimeters is the energy resolution.
Since the developemnet of the shower is a statistical process, we may describe the energy resolution, according to Poisson statistics, by :

$$\frac{\sigma}{E} \propto \frac{a}{\sqrt{N}} \propto \frac{a}{\sqrt{E}}$$

Where N is the number of secondary particles contributing to the measured signal. The coefficient a is the stochastic term arising from fluctuations in the shower processes.

However, this expression match with an ideal calorimeter, which is only subjet to statistical fluctuations, while for realistic calorimeters, additional sources which influence the behaviour of the intrinsic energy resolution should be considered.
One contribution to the energy resolution comes from the electronic noise of the readout chain, which increases with decreasing energy of the incident particles. Besides, one should add other contributions which do not depend on the energy of the particle, such as instrumental effects that cause variations and nonuniformities of the calorimeter response.
Thus, the energy dependence of the energy resolution may be parameterized as the quadratic sum of three terms :

$$\frac{\sigma}{E} = \frac{a}{\sqrt{E}} \oplus b \oplus \frac{c}{E}$$

The stochastic term $\ll$ a $\gg$ is related to statistical fluctuations of hadronic or electromagnetic showers. The asymptotic term $\ll$ b $\gg$ comes from the saturation of the calorimeter, when the deposited energy is too high and from the leakage. Problems of uniformity and calibration of the detector also increase the value of b. The term $\ll$ c $\gg$ is related to the behavior of the detector. The detector noise can influence the value of this term.

2.3 CALICE : Calorimeter for a future Linear Collider Experiment

Calorimeters play an important role in high energy physics experiments. They should offer the possibility to provide measurements with a high level of accuracy. To achieve such a measurement accuracy and for an optimum performance, several requirements on the detectors design have to be made, by employing innovative technologies and advanced analysis techniques. In particular, calorimeters form a key element in the future Linear Collider (ILC) experiments, where it is expected that unprecedented jet energy resolution will be achieved, by using the particle flow approach which is being adopted as a guide in the detectors design. A large common international effort has been devoted, within the CALICE collaboration, to develop different technologies for the calorimeters proposed for the International Linear Collider(ILC). Fig. 2.6 illustrates all these technologies for both electromagnetic and hadronic calorimeters.

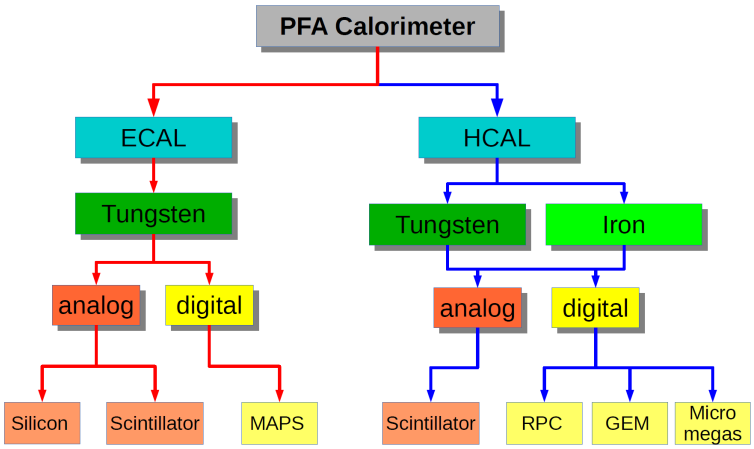

FIGURE 2.6: Tree of electromagnetic calorimeter technologies ECAL (left) and hadron calorimeter technologies HCAL (right) [22].

CALICE has been able to organise several test beam experiments at different beam facilities such as DESY, CERN and FNAL, where all major technologies have been exposed to particle beams, and large amounts of data has been obtained with the large prototypes, capable of providing full shower physics data. These prototypes are recapitulated in the following [50] :

SiW ECAL : The CALICE SiECAL group has designed and constructed a physics prototype [51], equiped with 30 sampling layers, each with an active area of 18×18 cm^2 and a granularity of 1 × 1 cm^2. The active layer is subdivided into a central part with 3 × 2 array of silicon wafers and a bottom part consisting of a 3× 1 array of silicon wafers as illustrated in Fig. 2.7. The absorber material used in SiECAL is tungsten with a various thickness. The total depth of the prototype is 20 cm corresponding to $24X_0$ obtained with 10 layers of $0.4X_0$ tungsten absorber plates, followed by 10 layers of $0.8X_0$, and another 10 layers of $1.2X_0$ thick plates.

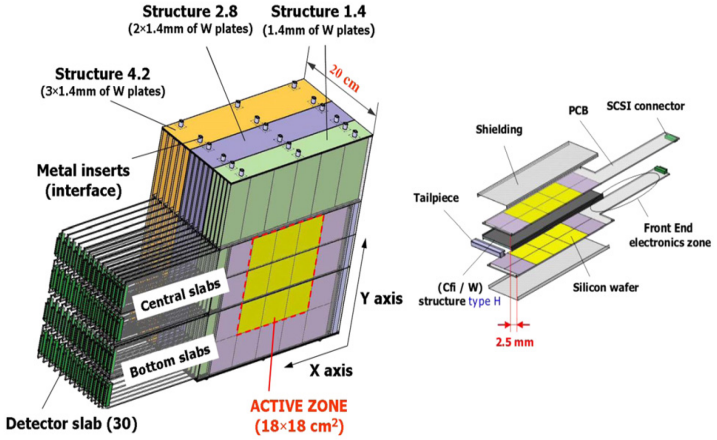

FIGURE 2.7: Schematic view of the CALICE silicon/tungsten ECAL prototype (left) and detailed view of a slab (right) [52].

SciW ECAL : The CALICE scintillator strip-based ECAL (ScECAL) consists of 30 active layers interleaved with absorber plates. Each active layer includes 72 scintillator strips (Fig. 2.8), readout each by a photosensor : Multi Pixel Photon Counter (MPPC, a silicon photo-multiplier). The first ScECAL prototype, was tested at DESY in 2013 [53].

Sci Fe AHCAL : The scintillator-based analogue HCAL (AHCAL) prototype is a steel scintillator sandwich calorimeter made of 38 active layers of scintillator tiles (5 mm thick) and steel absorber plates of 17 mm thickness [54]. The transverse dimensions of the active layer are 90×90 cm^2. The tile size is 3×3 cm^2 in the central region ; surrounded by three rows of 6× 6 cm^2 tiles and finally 12× 12 cm^2 tiles are used in the outer region (Fig. 2.9). In total the Sci AHCAL prototype counts 7608

scintillator tiles, with an analog readout via a wavelength-shifting fibre by a silicon photo-multiplier (SiPM).

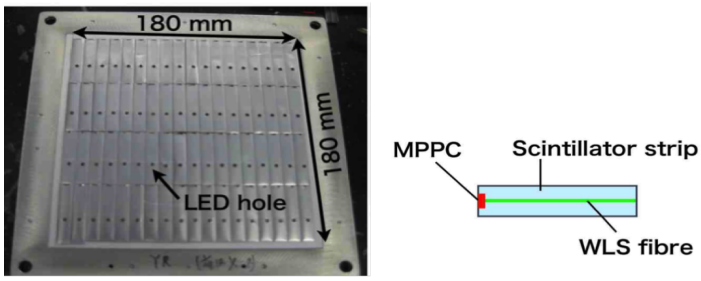

FIGURE 2.8: Left : An active layer of the SciECAL prototype, consisting of 72 scintillator strips. Right : Schematic view of a scintillator strip with a WLS fibre and a MPPC [50].

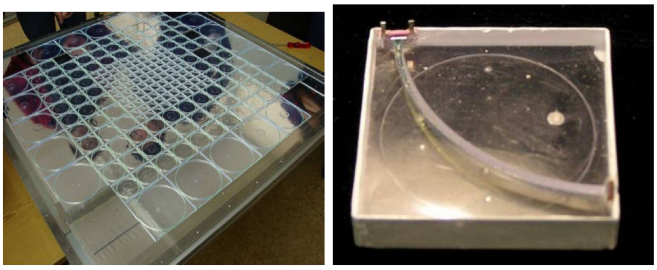

FIGURE 2.9: Active layer of the AHCAL prototype, scintillator tile with SiPM [50].

RPC Fe DHCAL : The Digital Hadron Calorimeter (DHCAL) uses Resistive Plate Chambers (RPCs) as active layers with a granularity of 1×1 cm^2 and a 1-bit readout electronics. The large physics prototype with up to 54 layers and close to 500,000 readout channels was built in 2008-2011. The transverse dimensions of the active layer are 96×96 cm^2 [55]. The DHCAL prototype has been tested and validated in several beam test campaigns in the years 2010 - 2012 and allowed to collect a large sample of data. A photograph of the setup in the CERN test beam is shown in Fig. 2.10.

RPC Stainless Steel SDHCAL : The SDHCAL technological prototype (Fig. 2.11) is made of 48 sampling layers using Glass Resistive Plate Chambers (GRPCs) as active layers alternated with steel absorber plates. The GRPCs are 1 m^2 area each, with a fine granularity of 1×1 cm^2 pad size, associated with power-pulsed embedded electronics. In contrast to the RPC DHCAL using a simple binary readout, the SDHCAL uses three-threshold electronics (2 bit) readout. The role of the different thresholds is to help identifying among few, many or too many particles crossing the same pad [88]. The detailed description of the SDHCAL technological prototype is the subject of the next chapter.

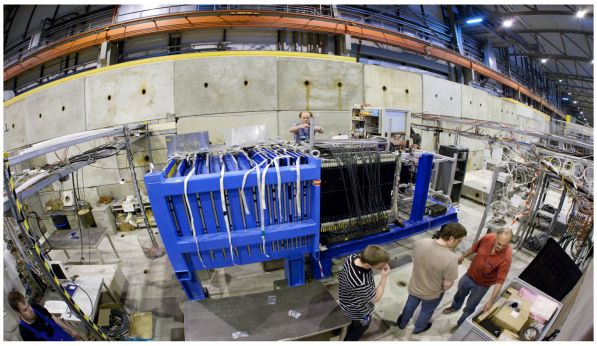

FIGURE 2.10: Photograph of the DHCAL setup in the PS test beam at CERN [50]

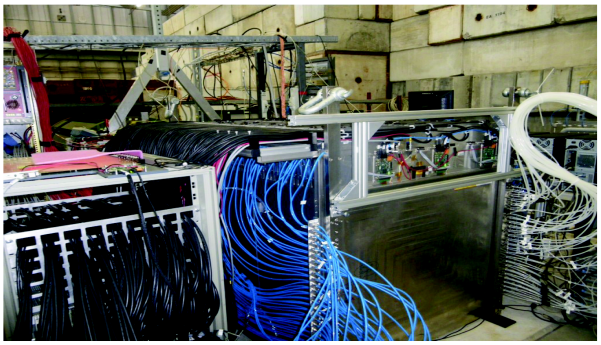

FIGURE 2.11: The Semi-Digital Hadronic Calorimeter prototype during the April 2012 test beam at CERN [50].

2.4 Conclusion

The CALICE collaboration has been developing calorimeters for experiments at a future linear collider. New materials and technologies are used to construct highly granular calorimeters designed for the particle flow concept. Different calorimeter prototypes has been built and tested in several test beams at DESY, CERN and Fermilab. In particular, the SDHCAL prototype has been conceived and built to meet the ILC requirements. The next chapter focuses on the description of the design of the SDHCAL technological prototype.

Chapitre 3

Construction and commissioning of a technological prototype of a high-granularity semi-digital hadronic calorimeter

The CALICE collaboration has built several prototypes of electromagnetic and hadronic calorimeters of different technologies to be used in the future Linear Collider [35, 57–61]. The semi-digital hadronic calorimeter (SDHCAL) is one of those prototypes and the first technological one built. Beside its semi-digital readout, the main feature of SDHCAL is its high granularity required for the application of the Particle Flow Algorithm (PFA) in order to improve the jet energy resolution in the ILC experiments. The prototype is the first of a series of new-generation detectors equipped with a power-pulsing mode intended to reduce the power consumption of this high granularity. A dedicated acquisition system was developed to deal with the output of more than 440000/m^3 electronics channels. After its completion in 2011, the prototype was commissioned using cosmic rays and particle beams at CERN. The data taken during these beam tests is the one it has been analyzed in this work. The topics of this chapter are the description of the SDHCAL prototype and its working principles. A detailed study on the construction and commissioning of this prototype has been published [62].

After a general description of the SDHCAL prototype in the first section, the GRPC

detector is presented in the second section where the geometry, material specification and electronics are described. In addition, the GRPC operation mode as well as the gas mixture and the thresholds choice have been detailed. The power pulsing mode is discussed in the third section. Finally, the fourth section is dedicated to the acquisition system developed for SDHCAL.

3.1 General Description

The SDHCAL technological prototype (Fig. 3.1) of $[1 \times 1 \times 1.3]m^3$, equivalent to $6\lambda_I$, was conceived and built fulfilling almost all the ILC requirements by associating different technological options [62, 63]. It is a sampling calorimeter made of 48 Glass Resistive Plate Chambers (GRPC), of 1 m^2 area each, associated with power-pulsed embedded electronics. Each GRPC is placed inside a *cassette* hosted by a self-supporting mechanical structure made of 1.5 cm thick plates of stainless steel. The cassette walls, made also in stainless steel, as well as the structure plates, form the calorimeter absorber. The cassette is 11 mm thick, of which 5 mm are the thickness of the walls and 6 mm are occupied by the GRPC and its readout electronics. In total 2 cm of stainless steel are separating two consecutive GRPCs.

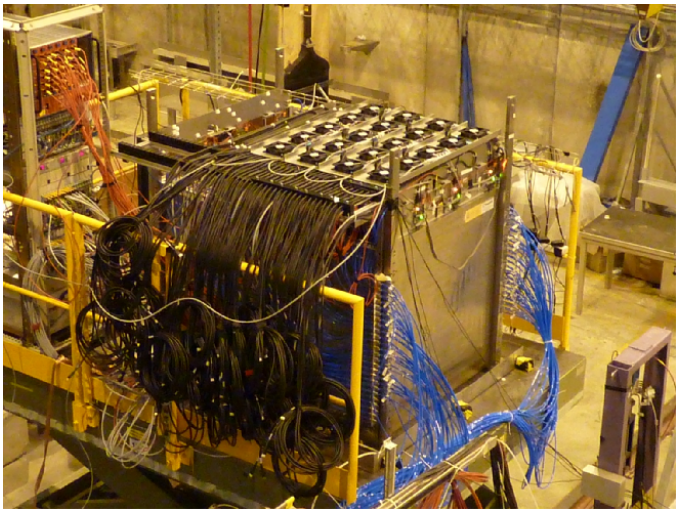

FIGURE 3.1: The SDHCAL prototype during 2012 test beam at SPS/CERN H2 line.

SDHCAL provides a high granularity in both the transverse and the longitudinal directions with the challenge to cope with millions of electronic channels, needed to read out such a high-granularity, and yet still compact and hermetic detector.
The homogeneity is achieved by the use of large Glass Resisitive Plate Chambers (GRPC) and by having all services (gas inflows and outflows, high voltage, data readout, etc.) to be on only one side of the GRPC and outside the HCAL active volume.

The fine segmentation is one of the most important singularities of SDHCAL which has been considered in order to allow the application of particle flow algorithm. It is achieved longitudinally by planes, separated by 2 cm as described before, and transversally by 1 cm^2 pads segmentation, used to readout the signal created by the passage of charged particles in the GRPC detector. The second important singularity of SDHCAL is its semi-digital readout electronic (2-bit) using three-thresholds instead of one threshold used in the binary case. The three thresholds will provide further informations usefull to tag pads fired by few, little or many crossing particles and to improve the hadronic shower energy measurment.

The technological prototype was tested successfully in beam tests at CERN several times and shows nice performance. Fig. 3.2 shows an event display of a hadronic shower and another one of an electromagnetic one recorded in the SDHCAL during 2012 test beam at the CERN-SPS.

3.2 The GRPC Detector

Glass Resistive Plate Chambers (GRPC) are gaseous detectors consisting in two parallel plates separated by a gas gap of a few millimeters. They are known as robust detectors that exhibit an excellent time resolution due to the small gas gap thickness as well as a good spatial resolution. An additionnal advantage is that they can provide good energy resolution and high efficiency if operated in adequate operational conditions [64–70].

Resistive Plate Chambers (RPC) made with Bakelite electrodes are used as muon detectors in the LHC experiments. The resistivity of bakelite ($\rho \approx 10^{10}\Omega$cm) allows it to be very efficient for particles rates up to a few hundred Hz/cm^2. However, the use of bakelite entails some problems in constructing very thin ($<$1 mm) plates with a uniform surface, which is important for the ILD calorimeter that will be installed inside the magnet, and which must be as compact as possible. Besides, the maintenance of bakelite requires the use of linseed oil to smooth surfaces in order to reduce the intrinsic noise level and allow the use of RPC for particles rates of a few kHz /cm^2. However, the use of this oil causes several problems leading to a significant reduction

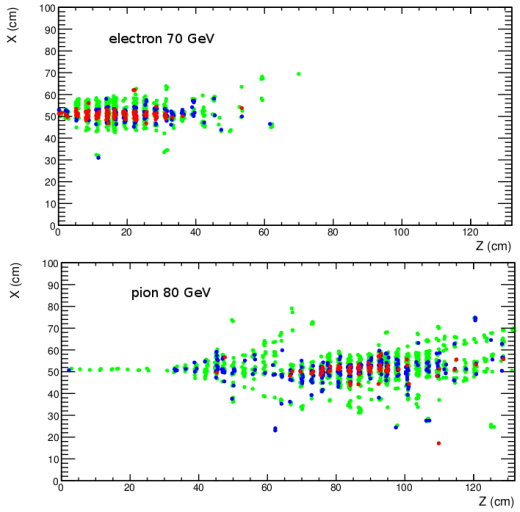

FIGURE 3.2: The shower profile of a 70 GeV electron and an 80 GeV pion. The hits in green color correspond to the fired pads in the detector when only the 1rst threshold is crossed, the blue ones when the 2nd threshold is crossed but not the third and the red ones when the 3rd threshold is crossed [90].

in the detection efficiency with age [76, 77]. In addition, such acquisition frequency is not required in the case of ILC collider thanks to the leptonic context of the experiment, as well as the beam structure, with a maximum rate of 100 Hz/cm^2 in the areas closest to the beam and 1.5 Hz/cm^2 in the central part. The glass allows such an acquisition frequency and makes it possible to obtain large homogeneous surfaces without having to use an additional treatment (linseed oil). This choice also reduces the manufacturing cost of the detector.

3.2.1 GRPC Description

The 1 m^2 chambers proposed for SDHCAL consist of the elements shown in Fig. 3.3. The GRPC has two glass plates with a thickness of 0.7 mm and 1.1 mm for the anode and cathode respectively. The glasses have a resistivity of 10^9 to 10^{12} Ωcm, and are separated by a 1.2 mm thick gap filled with a gas mixture constituted predominately of tetrafluoroethane (TFE). The gas gap thickness is maintained constant thanks to small ceramic ball spacers glued between the electrodes. The outer faces of the electrodes are covered with a conductive painting allowing the application of a high voltage to

create an electric field inside the gas gap. Readout pads of 1cm×1cm area are isolated

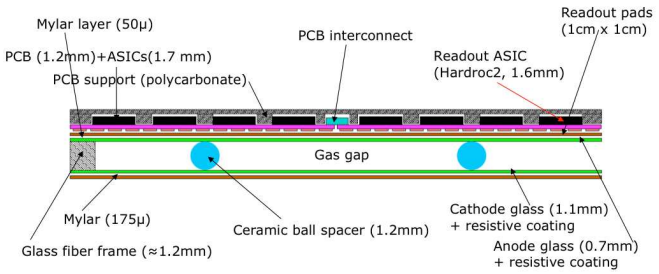

FIGURE 3.3: Schematic cross section view of a GRPC.

from the anode glass by a thin Mylar foil of 50 μm. To reduce cross-talk, adjacent pads are separated by a space of 406 microns [1]. The electronic channel cross-talk between two adjacent pads is less than 2% [63]. The pads are located on one side of an electronic board : PCB (Printed Circuit Board). A 1 m^2 GRPC is tiled with 6 PCB of size of $\frac{1}{3} \times \frac{1}{2}$m^2. Each PCB is composed of 6 ASUs (Active Sensor Unit) and each ASU hosts 24 readout chips. A specific ASIC (Application Specific Integrated Circuit) has been designed to read out the GRPC detectors of the SDHCAL. This chip is named HARDROC2 (HAdronic Rpc Detector ReadOut Chip) [71,72] and is sketched in Fig. 3.4. Each ASIC is connected to 8×8 pads through the PCB. The assembly of the ASUs used to read out 1m^2 is schematically shown in Fig. 3.5. Fig. 3.6 shows the electronics of 1m^2 GRPC. There are 144 ASICs and thus 9216 readout channels per one GRPC plane. The whole SDHCAL prototype has over 440000 readout channels.
The set of 48 ASICs (the pair of two ASU boards called a slab) is connected to a Detector InterFace (DIF) board (Fig. 3.6 left). A square meter of front-end electronics requires 3 DIFs. The DIF is connected through USB and HDMI connectors to the DAQ. Beside the ASIC configuration, the DIF central role is the transmission of data from the active layer to the external DAQ devices. The general architecture of the DIF is shown in Fig. 3.7.
Finally, a polycarbonate spacer ('PCB support' in Fig. 3.3) is used to 'fill the gaps' between the readout chips and to improve the overall rigidity of the cassette (detector/electronics 'sandwich') as well as the lateral homogeneity of the active layer.

1. This space is not a dead one since the passage of a charged particle in this could be detected thanks to the charge induction in the adjacent pads.

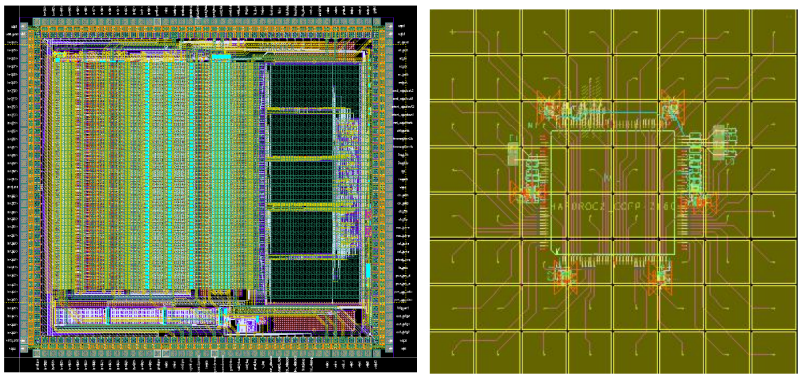

FIGURE 3.4: Left : HARDROC layout. Right : Semi transparent view of the ASIC from the copper pads side. The image shows how a matrix of eight by eight 1cm2 pads is connected to the ASIC's pins. Two neighboring pads are separated by 406 microns [62].

FIGURE 3.5: A scheme of the ASU used to read out 1m2 detector by assembling six small ASUs. Inside the blue and red shapes are boards ensuring connection between 2 ASUs (blue) and between ASU and DIF (red) [62].

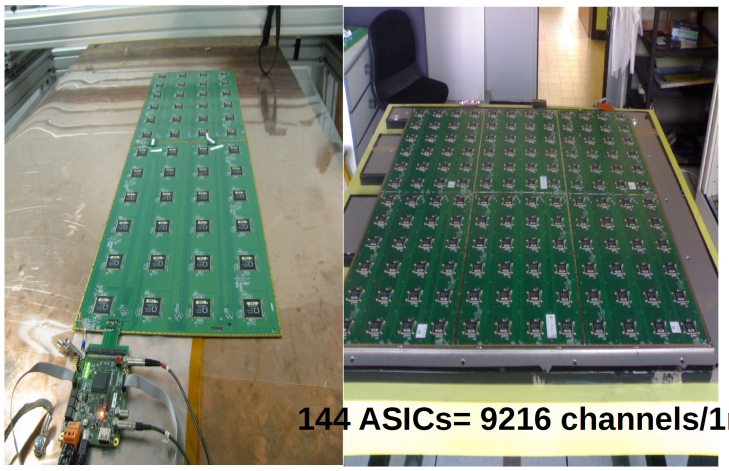

FIGURE 3.6: The electronic of 1m2 GRPC composed by 6 ASUs (left) and one slab of two ASUs connected to the DIF (right).

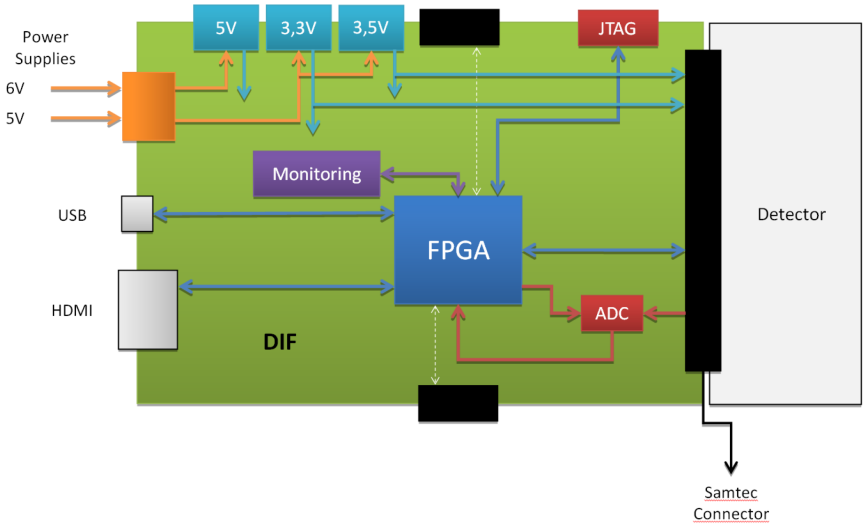

FIGURE 3.7: The DIF architecture [62].

3.2.2 GRPC operational mode

Depending on the intensity of the electric field inside the gas gap, the GRPC can be operated in three modes : avalanche, saturated avalanche and streamer. In order to fulfill the performance requirements of SDHCAL, the GRPC has been operated in avalanche mode which proceeds in this way :

– The process is initiated by the passage of a charged particle through the gas where a high voltage is applied to the electrodes. The high voltage applied to the GRPC was 6.9 kV and was motivated by efficiency studies [63].
– An ionization of the gas occurs and generates electron/ion pairs which are then accelerated by the strong electric field.
– The ions reach the cathode while each individual electron sufficiently accelerated will start a new ionization, repeated, resulting in a charge multiplication untill an avalanche is formed in the direction of the anode.
– Since the mobile charge carriers drift towards their electrode, they induce charge signals on neighbouring pickup copper pads which will be detected by charge amplifier of the DAQ. The charges are accumulated locally on the glass electrodes and then neutralized rapidely which allow high detection frequencies.

In streamer mode, the electric field inside the gap being higher, the multiplication of the charges is faster. At high gains photons contribute to the avalanche development and cause a spread of the avalanches in all directions. When the number of electrons in the avalanche reaches a critical value, a plasma is then formed between the two electrodes. This mode of operation generates large variations in charge and its relaxation time of the resistive electrode is relatively longer than in the avalanche mode. This mode is adequate for low-rate experiments, especially for cosmic-ray experiments.

3.2.3 The gas mixture

The gas filling of the GRPC is crucial for exploiting the best detector performance and to overcome some indesirable effects which helps to ensure a reliable and stable functionality of the detector.
Indeed, an electric field of high magnitude provokes a non-negligible dissipation of energy in the gas. However, this effect can be limited by an appropriate choice of the gas mixture.

In the case of SDHCAL, the gas mixture consists of 93% of tetrafluoroethane (TFE), 5% of CO_2 and 2% of SF_6 [73]. The first gas provides the primary electrons-ions when ionized by the crossing charged particle while the second and the third are UV photons and electrons quenchers respectively. Their role is to limit the size of the

avalanche that follows the creation of primary electrons-ions. It also reduces the probability of producing additional avalanches away from the charged particles impact in the GRPC.

3.2.4 The thresholds

The SDHCAL uses 3 thresholds for 2-bit readout provided by 3 variable fast shapers (15-25 ns), associated to each HARDROC, each one followed by a low-offset discriminator. Each fast shaper covers a different dynamic range leading to a total channel dynamic range from 10 fC to 15 pC. The first threshold is set to a low value, that of the two others to medium and high values in order to auto-trigger down to 10 fC and up to 15 pC. The 3 thresholds were set to 114 fC, 5 pC and 15 pC respectively. This choice is motivated by simulation studies that will be discussed in chapter 4.

3.3 Power Pulsing

One of the features of SDHCAL is its high granularity required for the application of the particle flow algorithm (PFA) and high precision energy measurment. However, the challenge is to cope with the millions of electronic channels needed to readout such a high-granularity while remaining a compact, hermetic and low consuming detector.

For this purpose, the HARDROC2 has been designed to be power-pulsed by switching on/off the power consuming parts of the readout chip by sending external control signal and therefore takes advantage of ILC operating cycle [74]. Otherwise, the use of a cooling system on each detector layer become inevitable.
Indeed, in the case of ILC, the beam bunch crossing in which collisions occur is expected to last only 0.95 ms every 200 ms. According to the beam time structure of ILC shown in Fig. 3.8, it is possible to power on the ASICs during the BC and to power them off the rest of the inactive period of 199.05 ms. This intends to reduce the power consumption per channel from 1.425 mW, when everything is on during the 200 ms, to less than 10 μW when the ASICs are switched on only during the bunch crossing (0.5% of the time).

Although one expects that running the SDHCAL prototype with the power-pulsing mode using the future ILC duty cycle will produce much less heating, the need of an appropriate cooling system should be assessed taking into account the geometry of the future SDHCAL and the constraint imposed by the global detector structure. Indeed, a thermal study based on the simulation of the ASICs power dissipation in the prototype was performed without a cooling system. A relative increase of 4.6 $^{\circ}$C was found in

the central part of the prototype with respect to the ambient temperature. This increase of temperature would lead to more noise in the central part of the detector and thus a degradation of the prototype data quality [62].

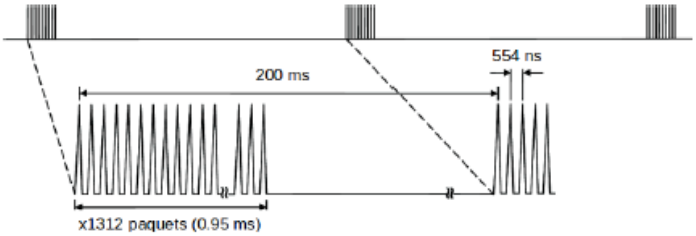

FIGURE 3.8: The ILC beam Spill.

3.4 Data acquisition

The communication between the detector and the acquisition computers occurs in two ways. First, HDMI is used as transmission protocol that maintains the system synchronized. Second, USB transmission protocol is used for data transmission. The DAQ chain of SDHCAL prototype is sketched in Fig. 3.9.

The DAQ chain consists of a Synchronous Data Concentrator Card (SDCC) followed by additional boards called Data Concentrator Cards (DCCs) which enable the transmittion of the commands sent, synchronously using private protocol on HDMI cables [2], from the computer network to the different Detector InterFace boards (DIFs), which are connected to the detector layers. The use of intermediate DCC is necessary since one SDCC has only a number of 9 HDMI outputs which does not cover the connection of the 243 DIFs of SDHCAL. With one DCC(having 9 HDMI outputs) we can connect 81 DIFs to the DAQ, which still not sufficient, while with two DCC blocks we can connect up to 729 DIFs which is largely sufficient to operate the whole prototype. The acquisition computers consist of one PC used to send the commands to the SDCC and several other PCs used to host data.

2. The SDCC is connected to the computer network through USB.

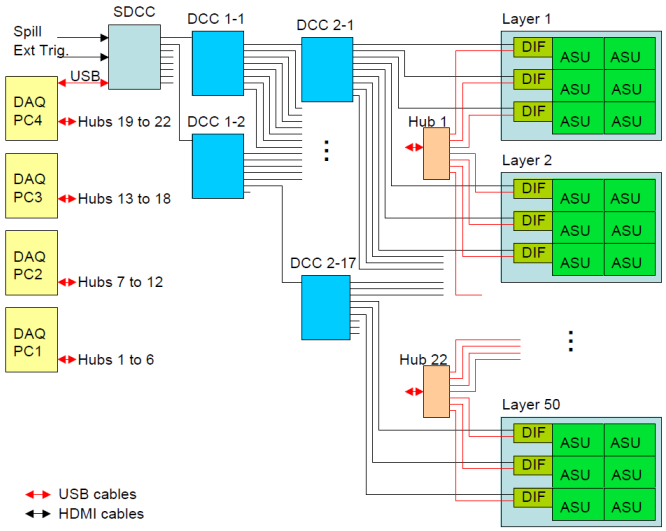

FIGURE 3.9: SDHCAL DAQ Architecture [62].

3.4.1 The readout modes

The data acquisition operates with two main readout modes.

– **Triggerless mode** :The acquisition is started upon the arrival of the Spill signal. During this process, the data is stored in the internal ASICs memory once the lowest threshold is crossed. The acquisition stops either at the end of the beam Spill or when a RAMfull signal is sent by the ASIC itself to the SDCC through the DIF. This latter case occurs when the memory of one of the detector ASICs is full, while each ASIC can store up to 127 events. Consequently, the acquisition is stopped for all the ASICs followed by a readout process performed by the DIFs. The data is then recorded in the DIF's FPGA memory and sent to the acquisition computers through USB and HDMI. Simultaneously, a Busy signal is sent to the SDCC to not allow starting new acquisition during the readout process. The acquisition starts again when the last ASIC sends all its data. The stored RAMfull events contain relevant informations helpful for data reconstruction (that will be discussed in section 5.2 of chapter 5) which are included in the data stream according to a specified format. First, a timestamp called Bunch Crossing IDentifier (BCID) is recorded by the DIF and corresponds to the time at which the acquisition process was stopped and the RAMfull event was readout. Whose counter is incremented by a step of 200 ns and is synchronized with the ASIC's one. In addition, we find the RAMfull event number and the time of each fired pad with respect to the start of the acquisition

cycle recorded by the BCID counter as well as the thresholds information for each fired pad. Further informations on the position of the fired pad are stored and gives its DIF (from 0 to 255), ASIC (from 0 to 47) and the channel numbers (from 0 to 63).

– **Trigger mode** :The data acquistion starts after the arrival of an external signal used to trigger the readout, such as a coincidence signal from scintillator Photomultipliers tagging the passage of particle. As for the Triggerless mode, when the acquisition stops, a StopAcquisition signal is sent to the ASICs followed by a StartReadout allowing to transmit the data to the DIFs. A RAMfull signal can occur when the memory of one ASIC gets full before the arrival of the Trigger signal. Trigger mode is useful to discriminate the collected data due to beam particles from those due to the noise when using a small number of chambers. It can also be helpful to improve the purity of data when a Cherenkov or muon veto signal enter the coincidence.

3.4.2 Raw Data Format

For every readout operation, the raw data is stored in the DIF's FPGA memory then sent to the computer network through USB with a specific data format using the LCIO (Linear Collider I/O) [75] event data model, where the data is organized into collections. Indeed, when the acquisition starts, a preliminary LCIO file is created. Each acquisition window is stored as an LCEvent and consists of :

– A collection of RawCalorimeterHits where for each hit corresponds a time information(BCID), the status of the crossed threshold as well as a unique key to determine the hit position by identifying the DIF, the ASIC and the Channel associated to the hit.

– A collection of generic objects containing specific informations associated to the DIF. These informations are the following :

1. DIF Trigger Counter (DTC) : It counts the number of readouts. It is reset at the first acquisition of each run.
2. Information Counter (IC) : It is used to count the dead time when ASICs are not acquiring. It is reset every acquisition. It is also used for BCID counter overflow and it is reset at the first acquisition of each run.
3. Global Trigger Counter (GTC) : It counts the number of triggers received by the DIF when the Trigger mode is runing and counts the number of readouts in the Triggerless mode. It is reset at the first acquisition of each run.
4. Absolute BCID : It is incremented with the 5 MHz clock received from the SDCC. It is reset at the first acquisition of each run.
5. BCID DIF : This is incremented with the 5 MHz clock coming from the SDCC and is synchronized with the ASIC BCID.

3.5 Conclusion

The semi-digital imaging hadronic calorimeter prototype has been conceived and built for the future linear collider experiments. It has been proven that a high granularity, compact and homogenous calorimeter with PFA capability can be acheived using cheap and simple but robust and efficient GRPC detectors. Further efforts are deployed to improve the SDHCAL concept focusing on developing new and more robust mechanical structure as well as the construction of larger GRPCs up to 3 m^2 and revisit their gas distribution. An upgrade of the ASICs daisy chaining is also considered.

Chapitre 4

Energy Reconstruction in a High Granularity Semi-Digital Hadronic Calorimeter using Monte Carlo Simulation

The study of the performance of the detector goes through different stages consisting of the simulation, reconstruction and data analysis. The simulation is an important key to evaluate the detector design as well its response to radiation when operating in real conditions. In the case of calorimeters, the most important factor for which detectors are designed is the energy measurment of the incident particle with an excellent accuracy and thus the energy resolution. Simulation studies allow to test the validity of different energy reconstruction methods and estimate the energy resolution of the detector.

The data analysis involves the reconstruction of the raw detector data supplied from the acquisition system. This will produce meaningful physics events associated with the hits left in the sensitive material by the passage of the particles through it. Indeed, these latter will induce signals in the detector electronics and thus provide information allowing the reconstruction of the hit coordinates in the detector. This is then followed by the application of selection criteria to identify a specific signal process while rejecting the background processes. At this stage, the energy reconstruction methods elaborated with the simulation studies can be applied to selected data to extract some physical results of interest such the energy resolution.

This chapter presents first, a description of the simulation of SDHCAL prototype. The second part, focuses on the energy reconstruction of single pion events within SDHCAL. A detailed study about different reconstruction methods didcated to energy resolution optimization is presented.

4.1 The simulation of SDHCAL

4.1.1 The GEANT4 simulation

In order to study the response of SDHCAL to radiation and for energy reconstruction optimization studies, a Monte Carlo simulation study was conducted using GEANT4 [78],[79]. However, inspite of its importance in particle physics, it should be noted that a Monte Carlo simulation of hadronic showers, especially using the GEANT4 package, is complicated and needs a deep understanding of the hadronic shower developpment and the underlying physics processes to obtain a reliable Monte Carlo code. This can be achieved by comparison of the experimental results with its MC equivalent. In this context, a deep study was conducted in our group in order to validate the simulation of hadronic showers in SDHCAL proptotype [80].

The simulation was carried out with version 4.9.3 of GEANT4. The geometry and the calorimeter structure as well as its material composition implemented in GEANT4 are almost identical to the physical prototype and detailed with high accuracy.
First, primary particles, with chosen position, energy and charge, are generated with the GEANT4 event generator. The incoming particle is directed towards the detector where it penetrates and starts interacting with the matter of the detector. As a result, hits are created in a selected volume, called the sensitive detector. This step needs a precise modeling to describe the passage of a particle through the detector and to determine the trajectory of the secondaries inside it, and thus the shower developpment. GEANT4 provides different models describing the interactions of particles with matter and it's recommended to use the FTFP_BERT physics list [81] for high energy physics calorimetry.

The sensitive detector is assigned to the gas gap, so that only energy deposits located inside the gas generate recorded hits. The SDHCAL simulation output contains the following information : the list of steps [1] inside the gas gaps ; the position of the step relative to where the primary particle was generated ; the deposited energy in these steps ; the entrance and the exit point positions of each step in gas gaps ; and the occurrence time of each step in the gap.

1. A step in GEANT4 is a segment of a particle path. In addition, each time the particle meets a material boundary or has an interaction a new step is created.

However, during the GEANT4 simulation, the GRPC's response is not simulated, neither the avalanche growth and only the interaction of the charged particles with the gas is taken into account. Thus, the readout of the simulation providing the energy deposited in pads is not compatible with the information provided by SDHCAL giving the induced charge accumulated in the copper pads.
Indeed, the aim of SDHCAL is not to measure the induced charge of the pads but an attempt to distinguish between pads crossed by one, few or many charged particles using the information of three thresholds applied to the induced charge.
Hence, a digitization procedure is necessary to simulate the GRPC's response and thus to achieve a realistic simulation of SDHCAL.

4.1.2 The Digitization

The digitizer module uses the information of hits provided by GEANT4 to modelize the avalanche growth in the GRPC and estimate the charge collected in each pad. The threshold is then applied to this charge and the pad is considered fired when the charge is above the threshold value. It should be noted that, although the digitization uses the information associated with the energy deposits in GEANT4 Cells, it is not implanted in GEANT4 and is realized independently with a dedicated algorithm [82]. The digitization procedure can be summarized in the following :

The induced Charge spectrum The digitization requires the determination of the induced charge spectrum distribution created by the passage of a charged particle in the GRPC gas gap. The avalanche growth in RPCs operating in avalanche mode has been studied in many papers [83–86]. According to these studies, the induced charge spectrum of RPC in avalanche mode can be described by a Polya function. Stimulated by this and because SDHCAL is operating in avalanche mode, this widely used description has been adopted in SDHCAL. In this way, the induced charge distribution can be described with :

$$P(q;\beta,\bar{q}) = \frac{1}{\Gamma(1+\beta)}\Big(\frac{1+\beta}{\bar{q}}\Big)^{1+\beta} q^{\beta} e^{[-\frac{q}{\bar{q}}(1+\beta)]} \tag{4.1}$$

Where $\bar{q}$ is the average value of the deposited charge in the gas gap, β is related to the width of the charge distribution and Γ is the Gamma function.
These parameters have been determined in SDHCAL in two ways.
The first one [87], uses the cosmic rays to measure the charge spectrum. The obtained result is then fitted with a polya function and leads to a good agreement. The problem with this method is that the measure was realized with a small GRPC operating in different conditions (high voltage, gas mixture...) than in SDHCAL prototype which

is using much more larger GRPCs. However, the validation of the description of the charge spectrum with a Polya function is confirmed.

A second method allows the determination of the polya parameters consisting in measuring the efficiency of muon beams while scaning the charge threshold. Then a fit is performed with the following function :

$$\varepsilon(q) = \varepsilon_0 - c \int_0^q \frac{1}{\Gamma(1+\beta)} \left(\frac{1+\beta}{\bar{q}}\right)^{1+\beta} q'^{\beta} e^{[-\frac{q'}{\bar{q}}(1+\beta)]} dq' \tag{4.2}$$

where ε_0 is the efficiency at $q = 0$ and c a normalization constant in pC^{-1} .

Fig. 4.1 (left) shows the result of the measurments of the efficiency as a function of the threshold for data performed with a few layers of the detector.

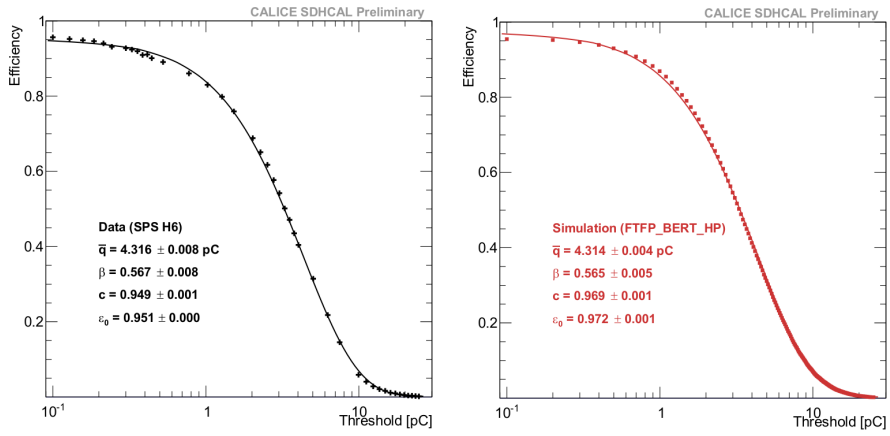

FIGURE 4.1: Efficiency versus threshold for data (left) and simulation (right)[82].

In this spirit, a threshold scan study of the efficiency is realized with the GEANT4 simulation samples, where for each hit the induced charge is drawn randomly following the Polya distribution. This step aims to reproduce the same results obtained with the muon test beam data by comparing the parameters, β and $\bar{q}$, extracted from the fit performed with equation 4.2. The Polya parameters are tuned until an agreement between simulation and data is found. At this point, when validated, the parameters are then fixed for the rest of the digitization study.

Charge sharing Because we are dealing with the charge created by a local ionisation inside the gas gap which could be shared between different neighbouring pick

up pads in the readout plane, the charge contribution of each pad should be estimated separately. This can adequately be determined by invoking the surface charge density $\sigma(x, y)$, created by the induced charge in the pads.
First a modelisation of the charge spliting between neighbouring pads is necessary to take into account the multiplicity effect and to compute the fraction of charge of each pad.
Indeed, the pad (P_0) where at least one charged particle crossed the gas gap is selected. The neighbouring pads, in a same sensitive layer, are situated at a distance smaller than a given distance r_{max}, set to 30 mm, from the pad P_0. GEANT4 hits positions are then localized in the considered pads. Each hit represents an avalanche and is considered surrounded by an area where the induced charge is distributed.

The integral of the surface charge density is then computed for each pad with respect to its boundaries. This is then normalized to the total integral of the density charge in the area delimited by r_{max}. The procedure of charge spliting is recapitulated in Fig. 4.2. The charge ratios R_i between P_0 and its neighbouring pads are defined with an overlap of a sum of Gaussian functions :

$$R_i = \frac{\int_{a_i}^{b_i}\int_{c_i}^{d_i}\sum_{j=0}^{n-1}\alpha_j e^{\frac{(x_0-x)^2+(y_0-y)^2}{2\sigma_j^2}}dxdy}{N} \tag{4.3}$$

where $a_i,\ b_i,\ c_i,\ d_i$ represent the border positions of the pad i, x_0, y_0 are the hit centre coordinates and N is the normalisation factor defined as :

$$N = \int_{-r_{max}}^{+r_{max}}\int_{-r_{max}}^{+r_{max}}\sum_{j=0}^{n-1}\alpha_j e^{\frac{(x_0-x)^2+(y_0-y)^2}{2\sigma_j^2}}dxdy \tag{4.4}$$

Note that the charge surface density is considered as a 2D Gaussian function. The integer n (number of Gaussian functions), set to 2, and the parameters α_j and σ_j are free parameters tuned using muon data to reproduce multiplicity. Finally, the charge of each pad Q_{pad} is derived and when it is above a certain threshold, the pad is considered fired :

$$Q_{pad} = R_i Q_{induced} \tag{4.5}$$

For the absorber material, in this study, we used either Steel and Tungsten in order to compare between them. In addition, in order to assess the impact of the pad size on the energy resolution, a simulation of the calorimeter with different pad sizes was

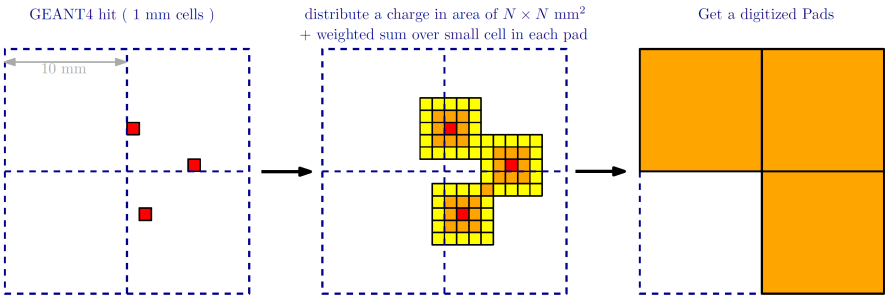

FIGURE 4.2: Schematic description of the charge spliting procedure. A matrix of 5× mm of cells is created around the simulated GEANT4 hit. A weighted sum is calculated for each pad (bounded by the blue lines). The digitized pads are the ones which reach the threshold. [101].

performed. We deduced that it has a crucial effect on the linearity and the energy resolution. The larger the pad size is, the poorer the linearity and the energy resolution. In the same context, we have also investigated the effect of varying the GEANT4 physics list in the simulations. The results of these works are shown in the Appendix A.2.

4.2 Energy Reconstruction

In this work the simulation is used to study the energy reconstruction and to estimate the hadronic energy resolution with SDHCAL. This study was performed with our preliminary GEANT4 simulation and does not include digitization.
In the following, after a brief introduction of the energy reconstruction with the binary mode in the first section, the multi-thresholds mode will be presented in the second section where different methods elaborated to study the hadronic energy reconstruction are discussed.

4.2.1 Binary mode

The response of SDHCAL to negative pions at different energies is studied in this section. This will allow to study the linearity and the energy resolution, in the digital case, according to different thresholds.
A first method for reconstructing the energy of the incident particles uses only the number of hits of each event by assuming that the incident energy is proportional to

the hit counting number.
Under this assumption, an estimate of the energy in the digital mode can be made by counting and eventually weighting the total number of the fired pads(N_{hit}). A pad is considered fired only when the deposited energy is above a given threshold : $E_{rec} = C \cdot N_{hit}$ with C a constant determined from a linear fit.
This can be observed in Fig. 4.3(top), where the correlation between the average number of hits and the incident energy at different thresholds values shows that one can find the energy of the hadronic shower up to 50 GeV for almost all the thresholds [88]. A deviation of N_{hit} from its expected value at high energies is noticed for some thresholds. To overcome this problem, one can replace the constant weight C by a linear function $C_0 + C_1 \cdot N_{hit}$. The linearity can thus be restored at high energies and the energy is constructed using the quadratic formula : $E = (C_0 + C_1 \cdot N_{hit}) \cdot N_{hit}$.
Indeed, at high energy the hadronic shower core is very dense and we can have several avalanches in a same pad. Thus a binary readout will suffer from saturation effect and leads to a deterioration of the detector response.
Such a behaviour was not only observed in linearity but also in hadronic energy resolution, at high energy, as it can be depicted in Fig. 4.3(bottom). The energy resolution defined by $\frac{\sigma_{rec}}{\langle E_{rec} \rangle}$ in the binary mode is studied for different thresholds and fitted with respect to the formula :

$$\frac{\sigma_{rec}}{\langle E_{rec} \rangle} = \frac{a}{\sqrt{E_{beam}[GeV]}} \oplus b \oplus \frac{c}{E_{beam}[GeV]} \tag{4.6}$$

As explained in Chapter 2 'a' is the stochastic term, 'b' is the constant term taking into account the calibration uncertainty and the energy leakage and 'c' is a noise term including the electronic noise and digitization effects.

A Semi-digital readout (2-bit) provided by SDHCAL reduces this effect and improves the energy estimation by using the information related to three thresholds useful to tag pads fired by few, little or many crossing particles inside the shower.
The choice of the thresholds values as well as the energy reconstruction in the multi-threshold mode will be discussed in the next paragraph.

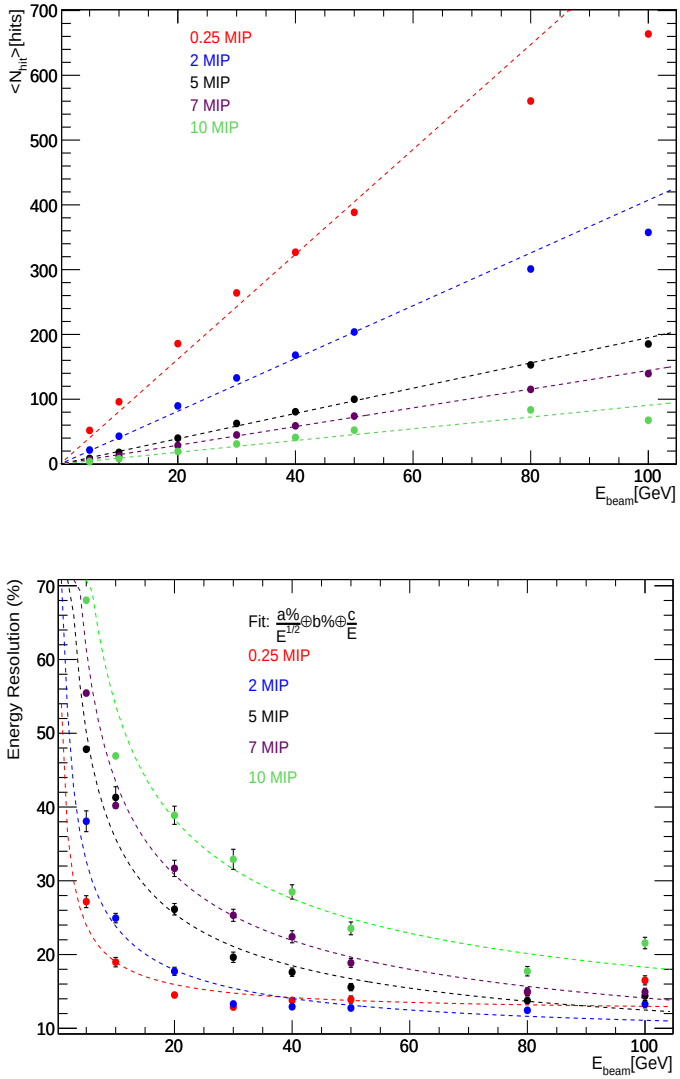

FIGURE 4.3: DHCAL response to Pions for different thresholds.Top : Mean number of hits as a function of the simulated beam energy of pions. The dashed lines indicate the result of a fit that $\langle N_{hit} \rangle = slope \cdot E_{beam}$ made for different threshold values. Bottom : $\frac{\sigma(E)}{E}$ the pion energy resolution as a function of the simulated beam energy. The resolution is computed using the Gausian fit parameters of the total number of hits distribution (binary mode with $E = C \cdot N_{hit}$ function).

4.2.2 Multi-threshold mode

To measure the deposited energy in the calorimeter with the semi-digital approach, we count the number of hits above each of the three thresholds affected by three weights (α, β, γ). The reconstructed energy is expressed as :

$$E_{\mathrm{rec}} = \alpha \cdot N_1 + \beta \cdot N_2 + \gamma \cdot N_3 \tag{4.7}$$

where N_1,N_2,N_3 represent respectively the number of hits above threshold 1(S_1) but below threshold 2(S_2), above threshold 2 but below threshold 3(S_3) and above threshold 3.

In a first attempt, our strategy of optimization of the energy resolution in the semi-digital method, consits of two steps.
The first one, is the determination of the best thresholds values giving good linearity and energy resolution.
Once the thresholds are chosen, the second step consists in the determination of the best set of the calibration weights (α, β, γ) to give a reconstructed energy closest to the beam energy and allowing a better energy resolution.
To realize the second task we used a minimisation procedure based on a χ^2 expression defined by :

$$\chi^2 = \sum_{i=1}^{N} \frac{(E^i_{\mathrm{beam}} - \left(\alpha \cdot N_1 + \beta \cdot N_2 + \gamma \cdot N_3)^i\right)^2}{E^i_{\mathrm{beam}}} \tag{4.8}$$

Where $'i'$ is the event number and N is the total number of events used in the minimization.
In this case, in the χ^2 formula the conventional σ^2 is approximated to the Beam energy E_{beam}. Indeed, according to equation 4.6, in a first approximation, the energy resolution is proportional to the stochastic term $a/\sqrt{E}$ with respect to the beam energy, resulting in $\sigma^2 \approx E_{beam}$. Another parametrization for a better estimation of σ^2 $\left(\sigma^2 = p_0 \cdot E_{\mathrm{beam}} + p_1 \cdot E^2_{\mathrm{beam}} + p_2\right)$ is recently used and allows a slight improvement in the results.

The choice of the Thresholds During this work, the choice of the thresholds was made qualitatively using the results obtained with the binary mode and based on the criterion of having a good linearity and energy resolution.

According to Fig. 4.3, the thresholds (S_2) and (S_3) that give a good linearity can be set to 5 and 10 average charge of a MIP. As the first threshold (S_1) must be lower,

we used the one giving the best energy resolution which is 0.25 of an equivalent MIP average charge.

To confirm this choice, several values of the three thresholds were used to minimize Equation 4.8. S_1, S_2 and S_3 were allowed to vary respectively from 0.1 to 3 MIPs, 4 to 8 MIPs and 9 to 15 MIPs. The minimization is performed for a range of energy from 10 to 100 GeV by a step of 10 GeV. The results of this study confirm the validity of the set of values previously chosen.
Nevertheless, a recent ongoing study, using a new simulation code taking into account the digitization procedure, revealed a new set of thresholds having achieved a slight improvement in the detector performance [80].
Basically, the energy resolution and the deviation to the linearity defined with $\frac{E_{\rm rec}-E_{\rm beam}}{E_{\rm beam}}$, are ploted for the thresholds S_2 and S_3 while varying their values.
According to the results illustrated in Fig. 4.4, the best linearity and energy resolution are achieved with the lower(higher) values of threshold $S_2(S_3)$ and thus it is recommended to use $(0.114, 1, 15pC)$ as thresholds, the average MIP-induced charge being around 1.2 pC with a rather large spread [87].

It is now essential to properly determine the best parametrization of the calibration weights (α, β, γ). We studied different methods of energy reconstruction. These methods depend on the nature of the calibration weights, each offer a modellisation depending or not on the energy of the beam. In addition, a further method based on Neural Network has been used for energy reconstruction. In the following, we will discuss the results of each method.

4.2.3 Analytic methods for energy reconstruction

4.2.3.1 Energy Dependent weights

Because of the complexity of hadronic showers and its evolution with energy, it is not possible to consider the calibration weights constants. Nevertheless, in order to understand the behaviour of these weights with respect to the energy and thus the total number of hits, we assumed first the possibility of a set of calibration weights contants and depending on energy.

Energy Dependent weights :First method : This method was our first attempt to get an idea on the energy resolution and the linearity with the multi-thresholds mode of SDHCAL [88]. In this method, the calibration weights are supposed to be constants and are optimized for each energy point.

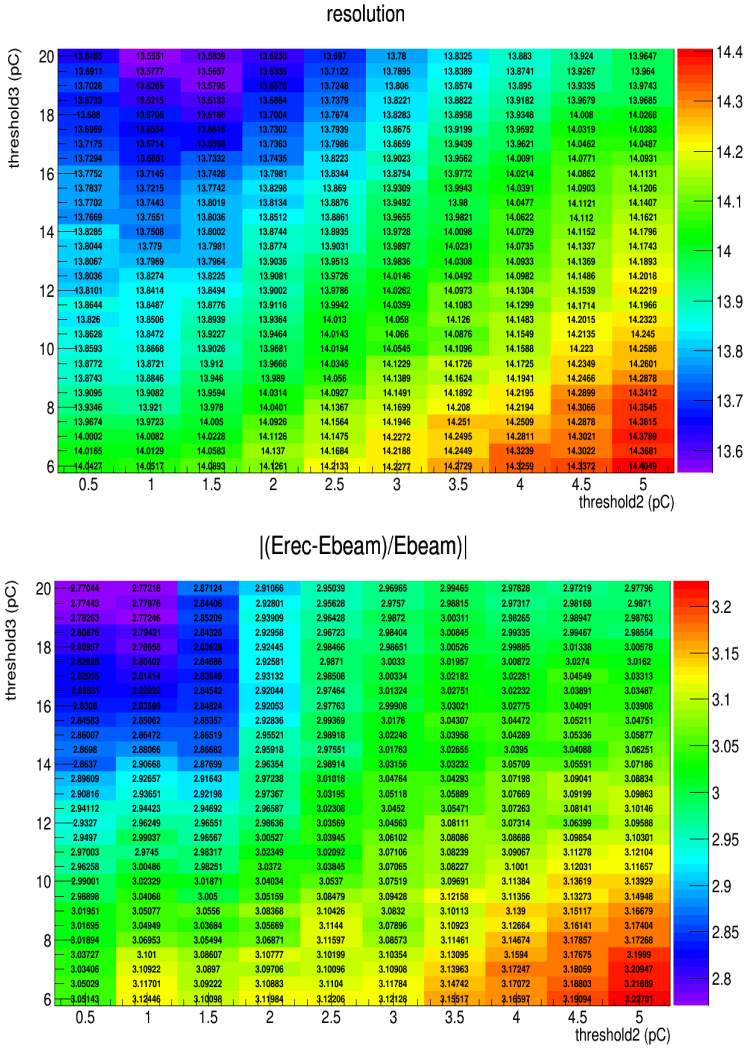

FIGURE 4.4: Linearity and energy resolution behaviour with respect to the thresholdsS_2 and S_3 [80].

The calibration constants were determined by minimizing equation 4.8. The thresholds were fixed to $(0.25, 5, 10MIP)$ and only α, β and γ are allowed to vary in specific intervals going from 0.01 to 1.5 by a step of 0.01. The evolution of these parameters with respect to N_{hit} is presented in Fig. 4.6.
Once the calibration constants are determined, they are used to compute the reconstructed energy according to equation 4.7 by replacing α, β and γ by their corresponding values available for each energy. Fig. 4.5 shows the reconstructed energy distributions of two studied energies. Each reconstructed energy distribution is plotted and then fitted using a Gaussian. Then from the fit parameters we can deduce the standard deviation σ and the average reconstructed energy to calculate the energy resolution and the linearity.
As it can be noticed in Fig. 4.7, thanks to this method the linearity is restored. In addition, a better energy resolution compared to the binary mode is achieved with the multi-threshold mode.The difference in energy resolution between the two modes is more important above 40 GeV. A direct comparison of the energy resolution results of both modes is shown in Fig. 4.8. The energy resolution shown in Figure 4.8 is fitted with equation 4.6. The fit parameters are summarized in Table4.1.

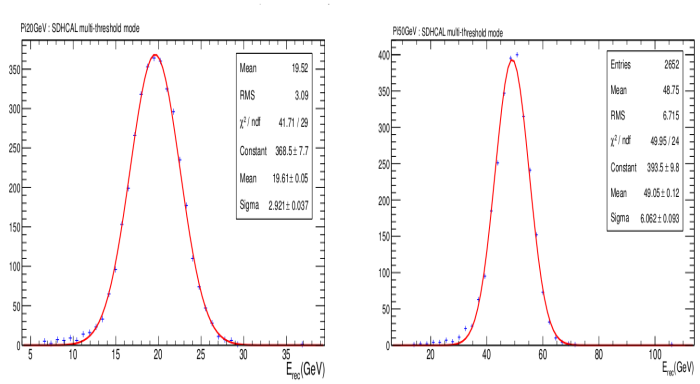

FIGURE 4.5: Reconstructed energy for pion showers using information from the three thresholds (multithreshold mode with energy dependent weights described in section 4.2.3.1). Left is for beam energy of 20 GeV, and right for beam energy of 50 GeV

TABLE 4.1: Energy resolution fit parameters computed with equation 4.6.

	a [%]	b [%]	c [GeV]	χ^2/ndf
Gaussian fit	49.49 ± 0.014	9.025 ± 0.014	0.0006 ± 0.014	3.91

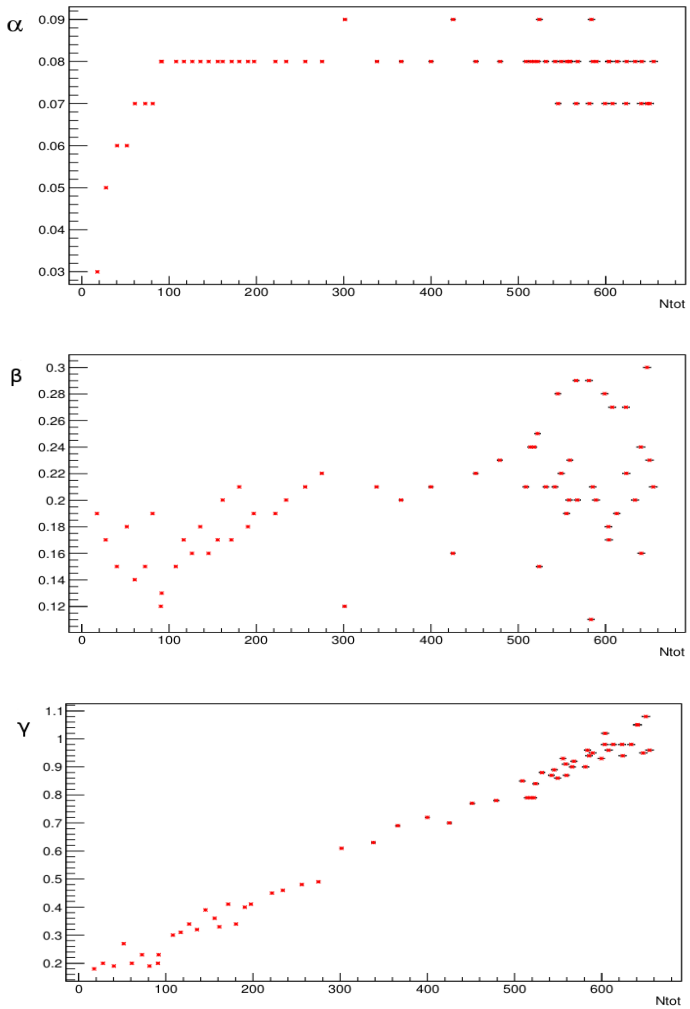

FIGURE 4.6: Dependence of the calibration weights of the total Number of Hits.

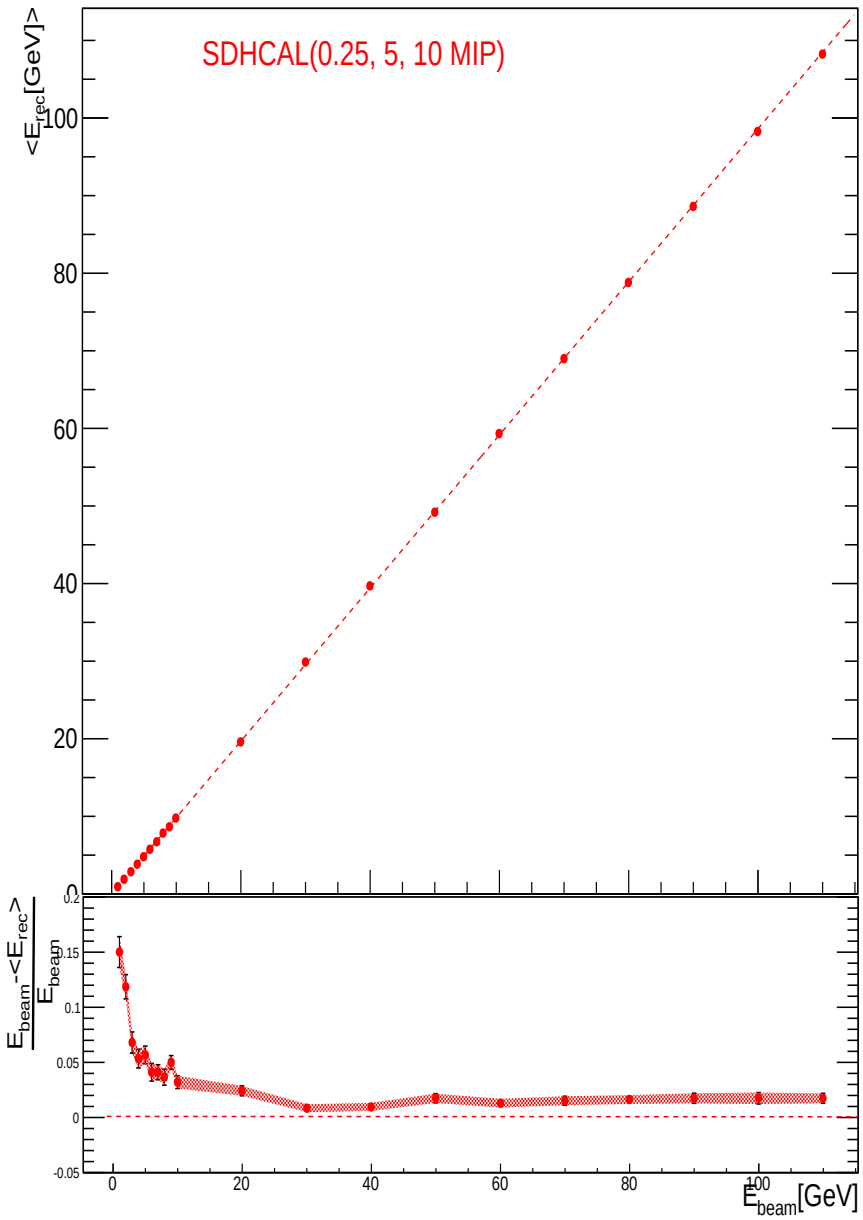

FIGURE 4.7: SDHCAL :(Top) Mean reconstructed energy for pions as a function of the beam energy computed using the three thresholds information (multi-threshold mode). The line corresponds to x=y. (Bottom) Relative deviation of the mean reconstructed energy with respect to the beam energy as a function of the beam energy for pions

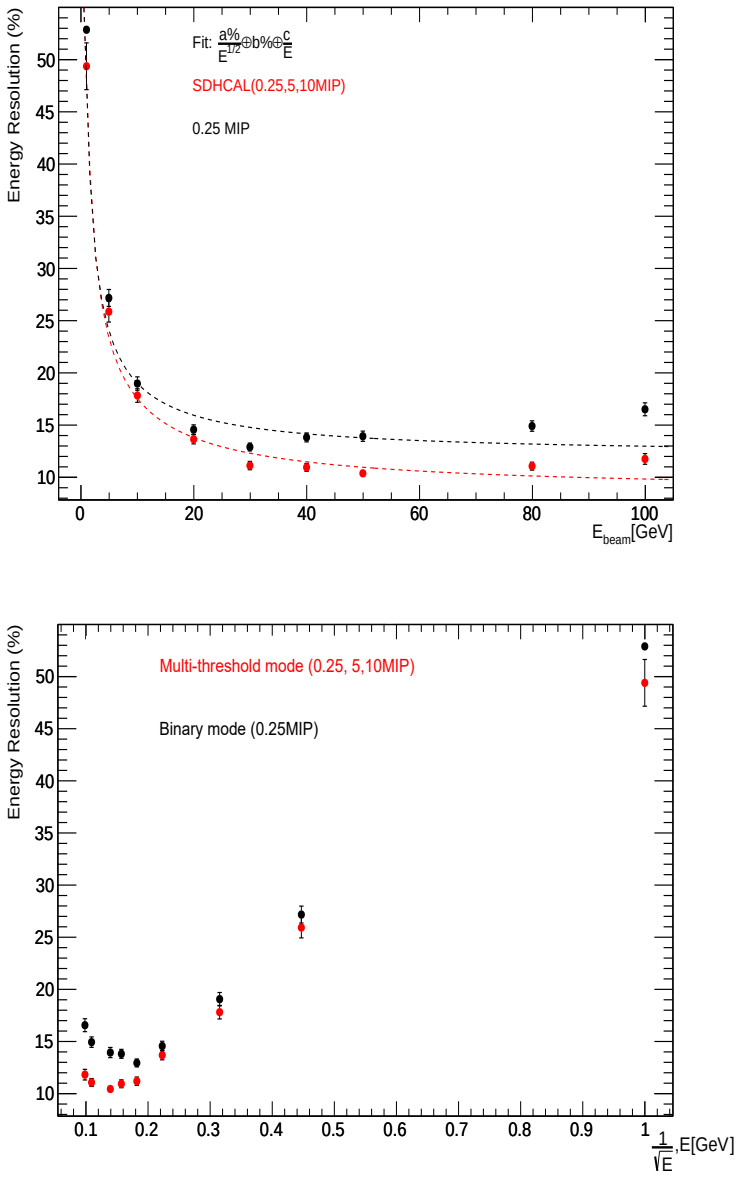

FIGURE 4.8: The relative resolution of the reconstructed pion energy as a function of the beam energy. For the black circles graph, the reconstructed energy is computed using only the total number of hits (binary mode). For the red circles graph, the reconstructed energy is computed using the three thresholds information (multi-threshold mode).

Energy dependent weights :Second method The same analysis was reproduced with the root class "TMinuit" [89] which also optimize for this part the calibration weights for each energy. For this task I used the same weights parametrization. The difference between this method and the previous one is only in the way of computing the weights. In the previous method the weights are computed using a simple C++ script and they are varrying in specific intervals, while in this method I used the class TMinuit where the weights are generated randomly. The purpose from this study is to check the results obtained with the previous method.
The results obtained, using TMinuit minimization, are very similar to those of the previous analysis. As can be noticed in Fig. 4.9, the calibration weights α and γ keep the same behaviour as previously. The β parameter has not exactly the same behaviour as previously since it is varying from 0 to 0.3 while the range from 0 to 0.1 was not exploited in the previous method of computaion of the weights. The behaviour of the weights confirms again their dependence on the total number of hits and the beam energy.
Table 4.2 summarize its values besides their corresponding energy resolution and mean reconstructed energy obtained for some energies.
The reconstructed energies distributions obtained in this way are fitted as before and are shown in Appendix B.1.

Fig. 4.10 shows the mean reconstructed energy (top) and the energy resolution (bottom) with respect to the beam energy. First, as expected, the optimization using TMinuit gives results in a good agreement with the previous analysis for both linearity and energy resolution. In addition, the energy resolution is again improved compared to the binary mode as was predicted from our preliminary minimization study.
This is an encouraging result which highlits the very good impact of the use of the informations related to the three thresholds on the energy resolution.

Although this minimization method shows promizing results, it is not reliable because it has a drawback which is the dependence of the calibration weights on the beam energy supposed to be unkown during the experiment. This led us to try a further method based on constants and free energy dependent weights. However, thanks to this method, the option of using a multi-threshold readout rather than a binary one was decided. In addition, this method allowed to study the behaviour of the weights with respect to the beam energy and the total number of hit which helped to find the best modelization of these parameters.

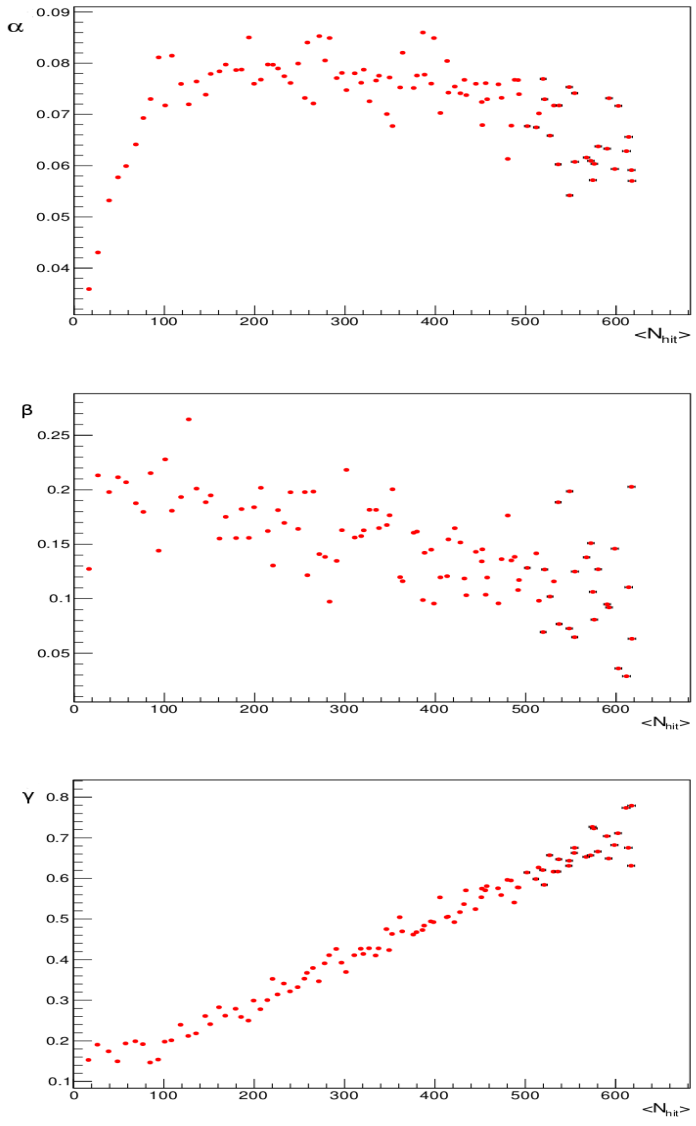

FIGURE 4.9: Dependence of The calibrations weights of the total Number of Hits estimated with TMinuit minimization.

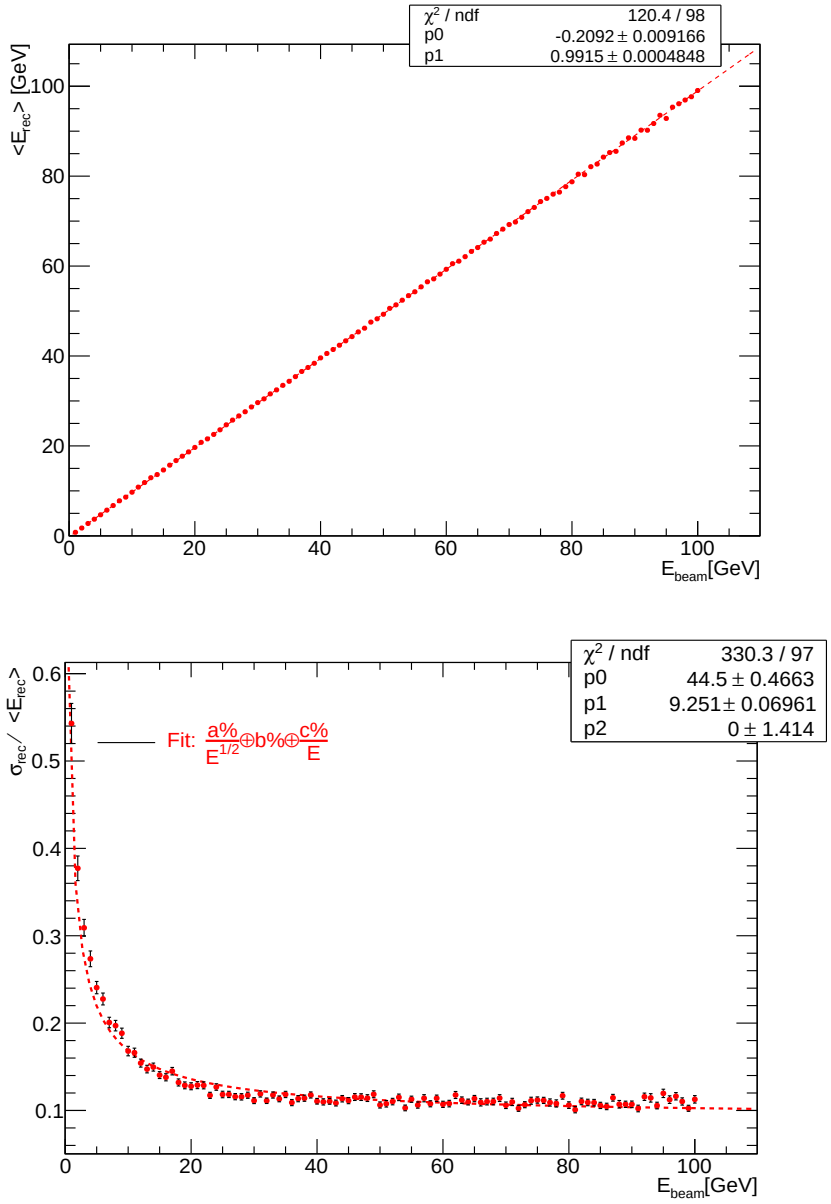

FIGURE 4.10: Top : Mean reconstruced energy as a function of the beam energy with energy dependents weights method using TMinuit minimization. Bottom : Energy resolution as a function of the beam energy with energy dependent weights method using TMinuit minimization.

TABLE 4.2: Calibration weights, Energy resolution and Mean reconstructed energy computed with TMinuit minimization

$\mathrm{E_{beam}}[GeV]$	α	β	γ	R(%)	$\langle E_{rec}\rangle[GeV]$
10	0.07	0.23	0.20	16.8	9.7
20	0.08	0.18	0.26	12.76	19.67
30	0.07	0.20	0.35	11.1	29.63
40	0.08	0.16	0.43	11.04	39.58
50	0.08	0.16	0.46	10.61	49.28
60	0.07	0.15	0.52	10.69	59.28
70	0.06	0.18	0.60	10.58	69.23
80	0.07	0.10	0.66	10.6	78.75
90	0.06	0.08	0.72	10.7	88.41
100	0.06	0.06	0.78	11.27	99.05

4.2.3.2 Constant weights method

In contrast to the method presented in section 4.2.3.1, in which the minimization is done energy by energy, this second method consists in minimizing equation 4.8 for all the energies at the same time and thus provides a set of universal calibration constants available for each energy.

To achieve this task we used again the class TMinuit of root, which generates randomly a set of weights several times, then computes the minimization function each time separetly and finally selects the best ones among them who reach a minimum of equation 4.8 as illustrated in Fig. 4.11. Table 4.3 gives the calibrations constants values obtained here.

The energies reconstructed using these weights are ploted and fitted using a Gaussian and are presented in Appendix C.1.

Fig. 4.12 shows the mean reconstruced energy (top) and the relative resolution (bottom) of the reconstructed pions energy as a function of the beam energy using this method.

We can notice then a degradation of the linearity at high energies as well as a slight degradation in the energy resolution.

Actually, this is expected because during this method, inspite of the fact that the calibration parameters depend strongly on the total number of hits and are different for each energy, as demonstrated in the previous paragraphs and illustrated in Figs. 4.6

and 4.9, we imposed to this parameters to be constants, which is improper to the complexity of the hadronic shower structure and its evolution with energy.

For this reason, it is fundamental to correct for this and therefore review the modelization of the weights.

TABLE 4.3: The constant weights computed with TMinuit.

α	β	γ
0.06	0.19	0.59

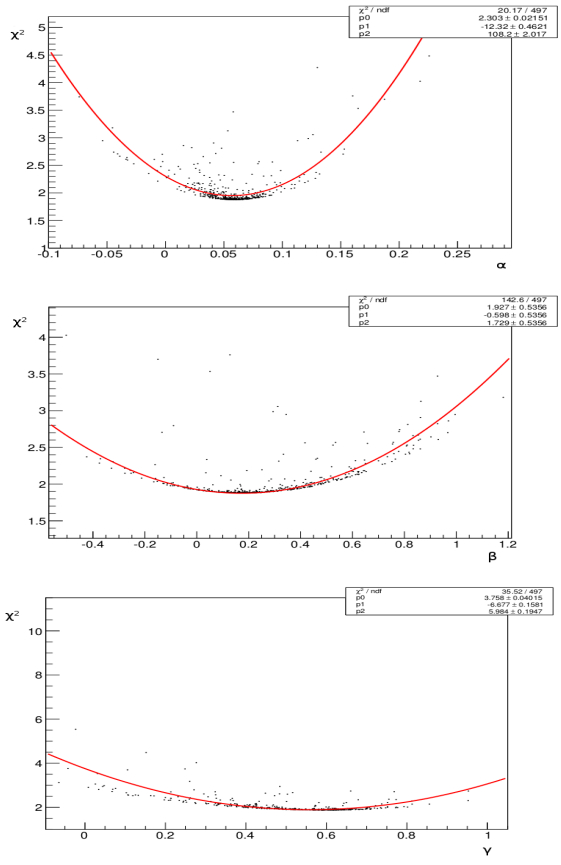

FIGURE 4.11: Minimization of the calibration constants with TMinuit.

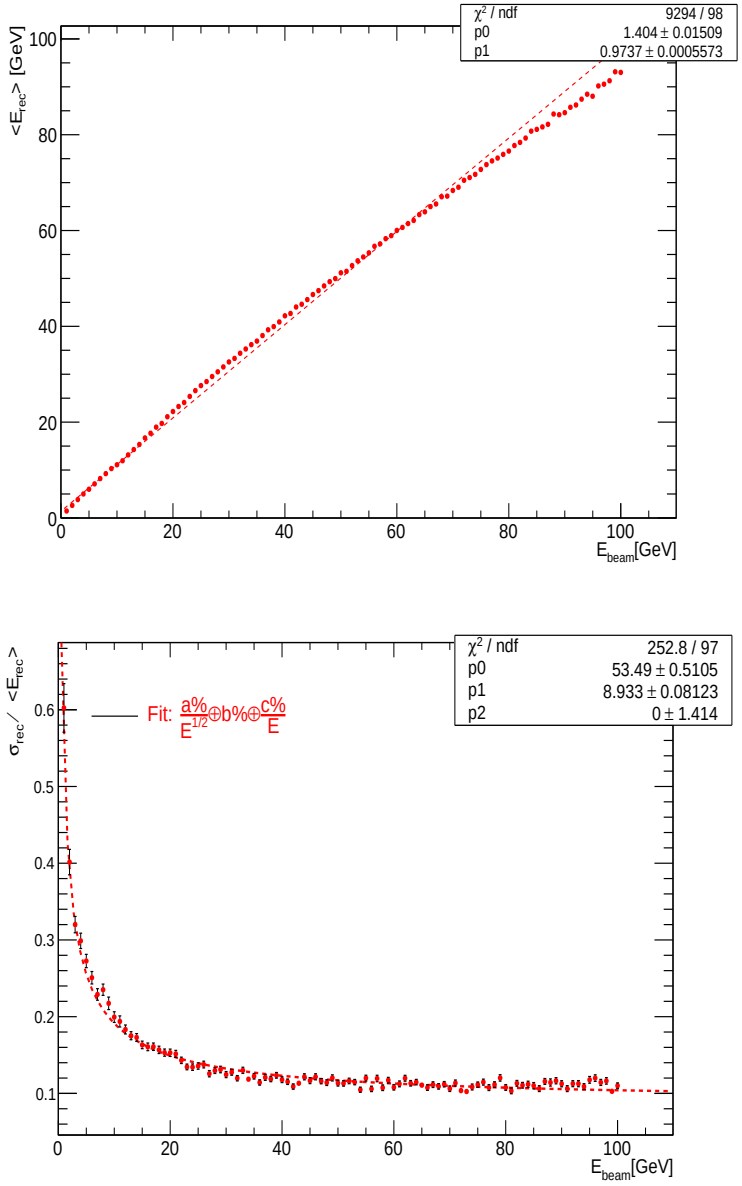

FIGURE 4.12: Mean reconstruced energy (Top) and energy resolution (Bottom) as a function of the beam energy computed with constant weights method using TMinuit minimization described in section 4.2.3.2.

4.2.3.3 Quadratic weights

Taking into account the results of the last method, the weights modelization should be rectified with respect to the total number of hits N_{hit}. For this purpose, we used two types of parametrization.

In the first parametrization, we inspected the behaviour of the weights as a function of N_{hit}. According to Figs 4.6 and 4.9, the parameter α is modelized with an exponential law with respect to $N_{\rm hit}$: $A_0(1 - exp(-A_1 \cdot N_{\rm hit}))$.
The parameters β and γ are modelized with a polynomial function of first degree : $B_0 + B_1 \cdot N_{\rm hit}$ and $C_0 + C_1 \cdot N_{\rm hit}$.

The new reconstructed energy formula is then defined by :

$$E_{\rm rec} = A_0.(1 - exp(-A_1.N_{\rm hit})).N_1 + (B_0 + B_1.N_{\rm hit}).N_2 + (C_0 + C_1.N_{\rm hit}).N_3 \tag{4.9}$$

The results of this parametrization are shown in Figs 4.13 and 4.14. The energy resolution keep the same behaviour compared to the previous results with the multi-threshold mode. But a deviation in linearity at hight energies is noticed. The reconstructed energy distributions obtained from this method are given in Appendix D.1.

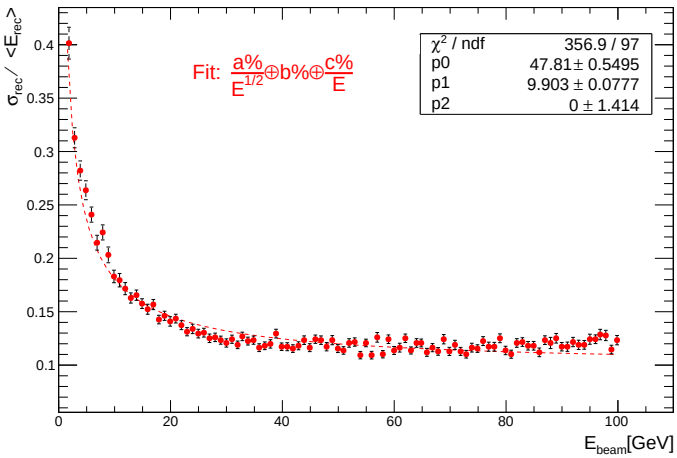

FIGURE 4.13: Energy resolution obtained from exponential parametrization of α, and polynomial parametrisation of β and γ in the reconstructed energy formula.

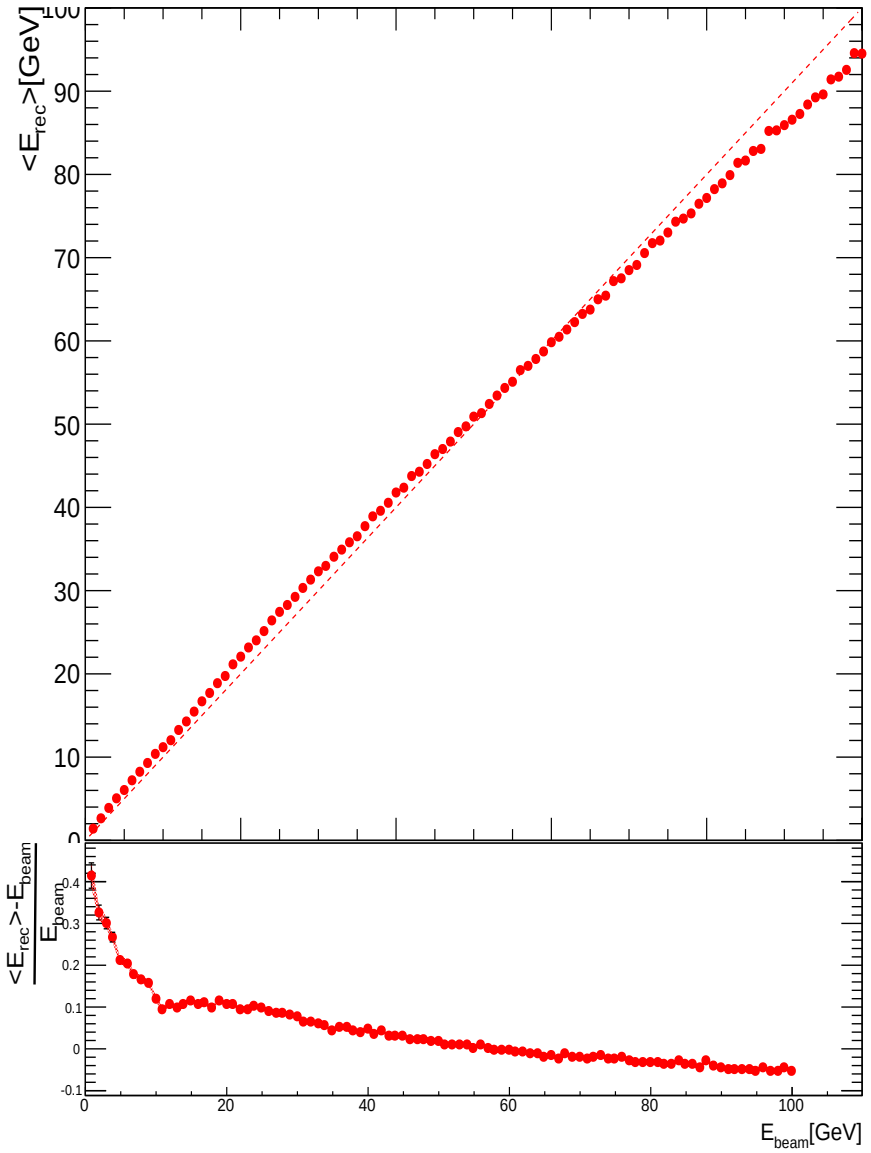

FIGURE 4.14: SDHCAL :(Top) Mean reconstructed energy for pions as a function of the beam energy computed using exponential parametrization of α, and polynomial parametrisation of β and γ. The line corresponds to x=y. (Bottom) Relative deviation of the mean reconstructed energy with respect to the beam energy as a function of the beam energy for pions

In the same spirit, a second parametrization of the weights based on a polynomial function of second degree with respect to N_{hit} was tested. In this way, the energy is reconstructed with equation :

$$E_{\mathrm{rec}} = \alpha(N_{\mathrm{hit}}) \cdot N_1 + \beta(N_{\mathrm{hit}}) \cdot N_2 + \gamma(N_{\mathrm{hit}}) \cdot N_3 \tag{4.10}$$

$$\alpha(N_{\mathrm{hit}}) = A_1 + A_2 \cdot N_{\mathrm{hit}} + A_3 \cdot N_{\mathrm{hit}}^2 \tag{4.11}$$

The same parametrization was applied to β and γ. The distributions of energy estimated with this method are ploted in Appendix E.1. The results obtained for energy resolution and linearity are respectively shown in Figs 4.16, 4.15. Although we notice a degradation of the linearity at high energies, the use of a quadratic parametrization has very good impact on the energy resolution at energies higher than 60 GeV, reaching better resolutions than the previous methods. Note that this method is adopted as the official method used in the energy reconstruction by SDHCAL. It was used in the publication of the last results of SDHCAL [90],[91].
To further improve the energy estimation in SDHCAL, we tested another technique based on artificial neural networks.

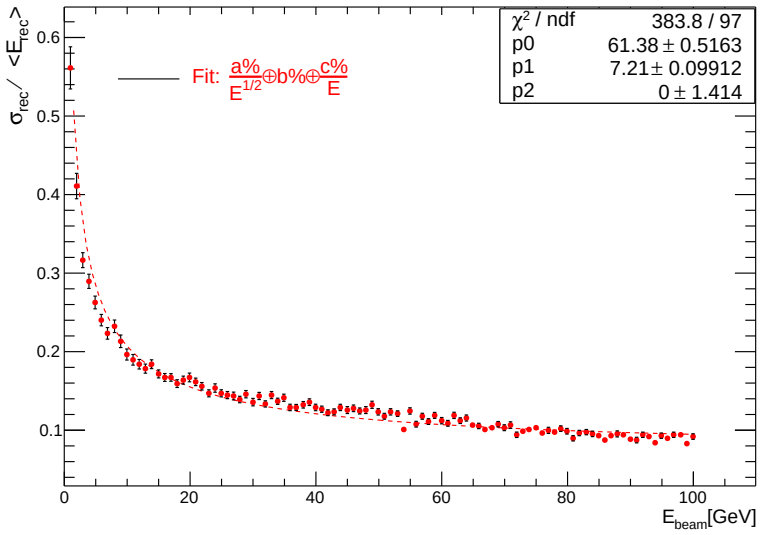

FIGURE 4.15: Energy resolution obtained from quadratic parametrization in the reconstructed energy formula.

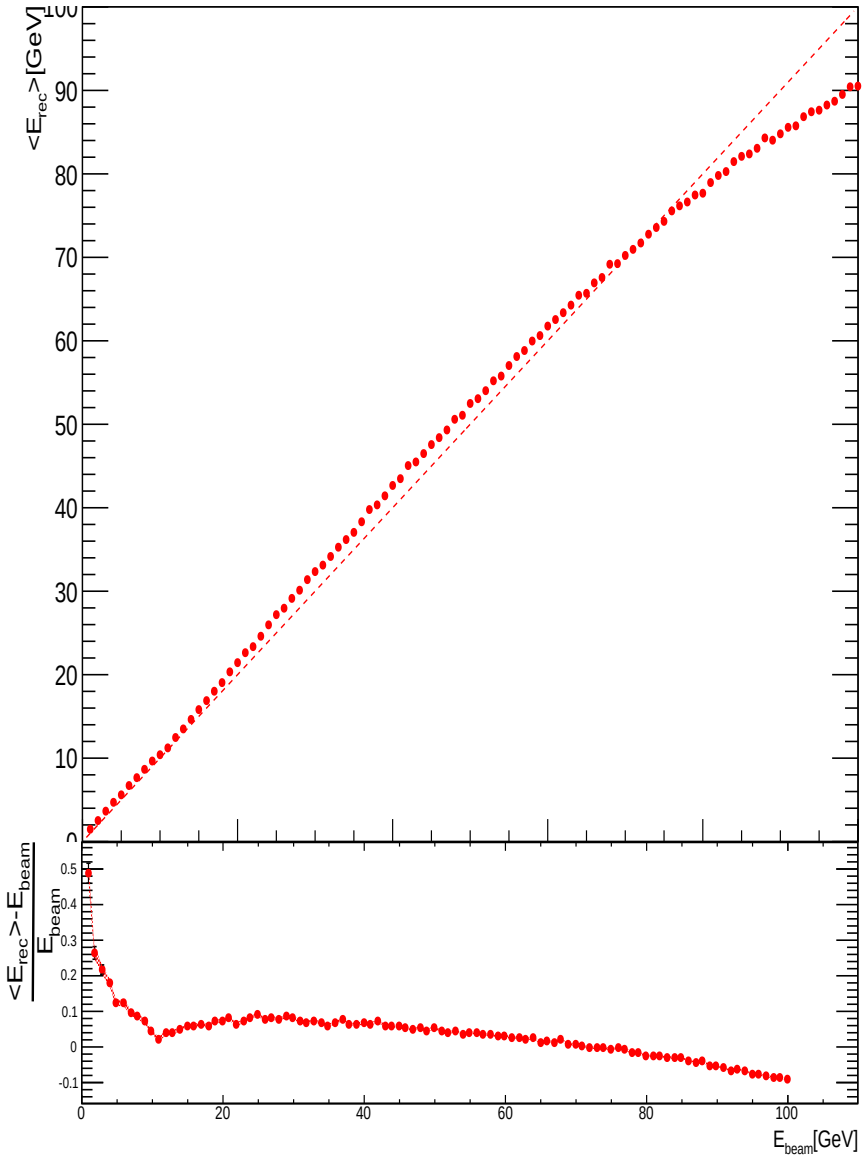

FIGURE 4.16: SDHCAL :(Top) Mean reconstructed energy for pions as a function of the beam energy computed using quadratic parametrization. The line corresponds to x=y. (Bottom) Relative deviation of the mean reconstructed energy with respect to the beam energy as a function of the beam energy for pions

4.2.4 Neural Network technique for energy reconstruction

The studies presented in the previous sections showed that SDHCAL has proved its ability to achieve a good energy resolution using a weighting method of the reconstructed energy based on the information provided by its multi-threshold mode. In order to further improve the energy resolution obtained within SDHCAL, the use of a neural network in energy reconstruction has been studied and applied in Monte Carlo simulation data. The present work will include a description of the neural network approach, the results of its application to simulation as well as the algorithm performance and its comparison with the traditional weighting methods of energy reconstruction

4.2.4.1 Overview

Artificial Neural Networks (ANNs) are non-linear computational algorithms inspired by biological neural systems of the human brain and based on the same principles [92]. Introduced in 1988 [93], [94], it has been successfully tested in a wide range of high energy physics experiments and several physics results have been achieved using this method [95–98]. In this context, ANN was applied for the energy reconstruction of hadronic showers in SDHCAL. The analysis was performed using the ROOT class TMultiLayerPerceptron [99]. The neural network has been built, trained and tested using the simulation samples. An ANN can be modeled by a set of input nodes constituting the input layer and receiving the input signals $(x_1, x_2, ..., x_{\rm n})$ where each input is normalized and weighted by $w_{\rm j}(j = 1, ..., n)$, a number of hidden layers and an output layer of computation nodes. The nodes in each layer are connected to the nodes of the neighbouring layers. Each node presents a non-linear information processing unit called neuron(Fig. 4.17). Each input signal $x_{\rm j}$ connected to a neuron k, is multiplied by a synaptic weight $w_{\rm kj}$. To obtain the desired output from the network, we should adjust the weights. The adjustment algorithm is referred as Training taking into account the comparison between the output of the network and the desired target corresponding to the training sample. One of the methods widely used for its good performance is called "learning by epoch". It consists first in the summation of information for the whole pattern and then updates the weights. Each update minimizes the summed error defined by :

$$Error = \frac{1}{N}.\sum_{j=0}^{N}(d_{\rm j} - o_{\rm j})^2 \tag{4.12}$$

Where $d_{\rm j}$ is the desired target, $o_{\rm j}$ is the output computed by the network and N the number of trained data patterns. The performance of an ANN is given by the accuracy of prediction measured by equation 4.12 and the convergence of the learning process.

At the output of the ANN, we have a plot with two error curves. The first represents the error calculated for the training sample called training error and the second curve corresponds to the error calculated to another independent test sample not used in the training then called the generalization error. The purpose of the learning process is to obtain a low generalization error caracterizing the performance of the ANN to model new data not used in the training.

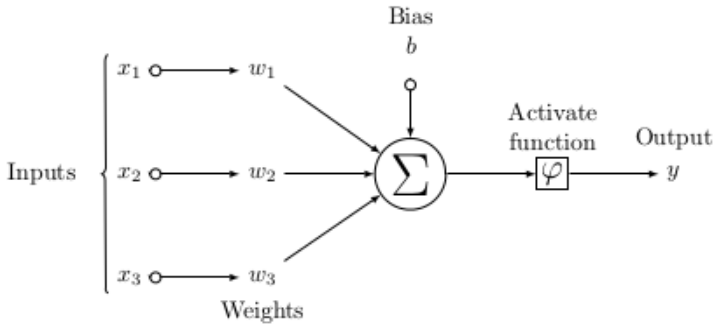

FIGURE 4.17: Example of modelization of an artificial neuron using three inputs

4.2.4.2 Training and Testing of the neural network

As stated previously, in the present analysis, we used the root class TMultiLayerPerceptron to build our neural network [99]. The neural network was trained and tested with pions Monte Carlo samples. The input variables are choosen to be the number of hits related to the 3 thresholds introduced previously : N_1, N_2 and N_3.

First, in order to identify the optimal architecture of the neural network, different architectures of one and two hidden layers with varying numbers of neurons in the hidden layers have been investigated. The neural network architectures performance is evaluated according to the statistic root mean square error, taken as a selection criterion, defined by :

$$\sqrt{\frac{\sum_{i=0}^{N}(E^i_{\text{beam}} - E^i_{\text{rec}})^2}{N}} \tag{4.13}$$

Where E^i_{beam} is the beam energy, E^i_{rec} is a predicted value of the reconstructed energy by the neural network for a given event and N the number of trained data patterns.
When comparing the different architectures, it appeared that the two hidden layer architectures perform better than those with only one hidden layer even when increasing

the number of neurons. It was also found that ANN architecture with two hidden layers with respectively six and two neurons performed the best. The neural network architecture used in this analysis is illustrated in Fig. 4.18. The correlations between the

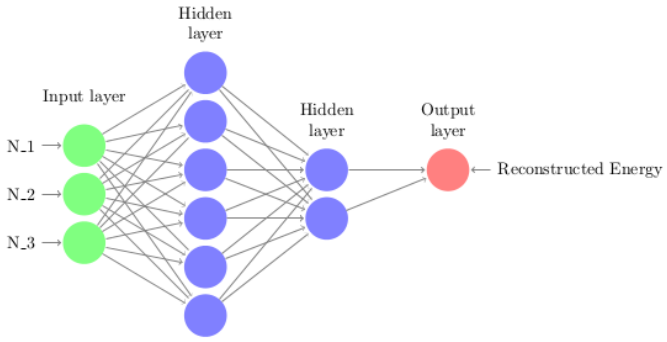

FIGURE 4.18: ANN Architecture used in the analysis.

input variables and the real particle energy are automatically learnt during the training of the neural network. The simulation samples was divided into training and test sets to evaluate the neural network performance. The training process was conducted using 50 samples corresponding to the simulated odd energies from 1 to 99 GeV. The data sets used during the test process correspond to 40 simulation samples associated to the even energies from 10 to 90 GeV. This energy range has been chosen in order to avoid the effect of the boundary otherwise energies would not be reconstructed correctly by the neural network. During the training phase, the network target is fixed to the beam energy. As a result, from the test samples, the neural network should estimate the energy of the incident particle defined as the output variable E_{rec}.

4.2.4.3 Results

Fig. 4.19 shows the behaviour of the error estimated by the ”learning by epoch” method during the training step described by the training curve(red) and the test curve(blue). The stability of the neural network is achieved when the two error curves reach the minimum and the convergence of the learning process.

The energy estimation, for the simulation data sets, performed by the neural network is shown in Fig. 4.20. The energy distributions are fitted with a Gaussian as illustrated in Figs. 4.20 and Appendix F.1. Using the fit parameters, the energy resolution and

linearity are computed and ploted repectively in Figs. 4.21 and 4.22.
The results of this method are promizing. The neural network technique gives a good prediction for the incident energy while achieving a good energy resolution and linearity.

The neural network method has been compared to the traditional methods for energy reconstruction used previously and based on expressing the reconstructed energy by a weighted sum $E_{rec} = \alpha \cdot N_1 + \beta \cdot N_2 + \gamma \cdot N_3$.
Figs. 4.23 and 4.24 show a direct comparison between the results obtained with the different methods of energy reconstruction highlighting the goodness of the results obtained with the neural network.
The neural network is able to achieve a linearity in the range of $\pm 5\%$ well below the one obtained by the quadratic weighting method. In addition, one can notice in Figs. 4.23, F.2 and F.3 that the neural network technique provides similar results to those obtained with the quadratic weighting method of the energy reconstruction with a clear improvement at energies higher than 40 GeV.

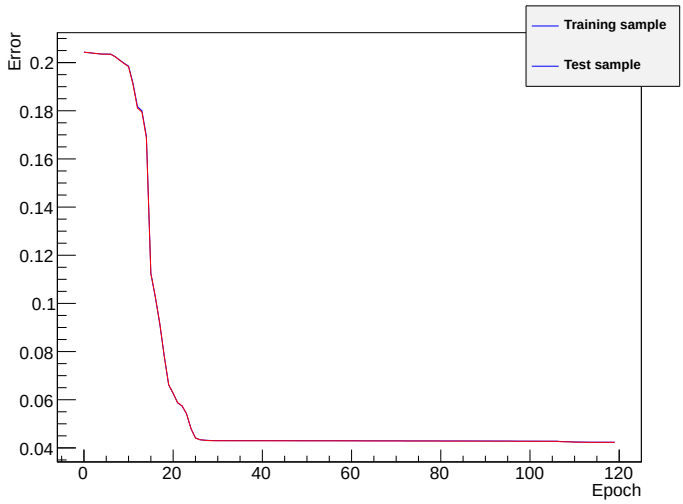

FIGURE 4.19: Accuracy of prediction measured by ANN.

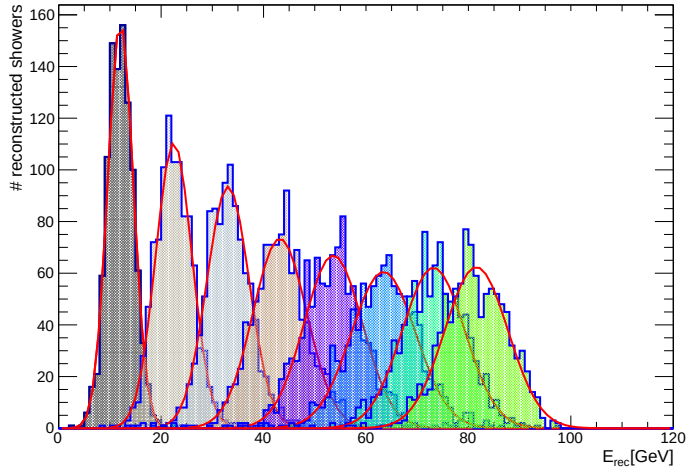

FIGURE 4.20: Pion reconstructed energy distributions obtained with neural network technique from 12 to 82 GeV by a step of 10 GeV.

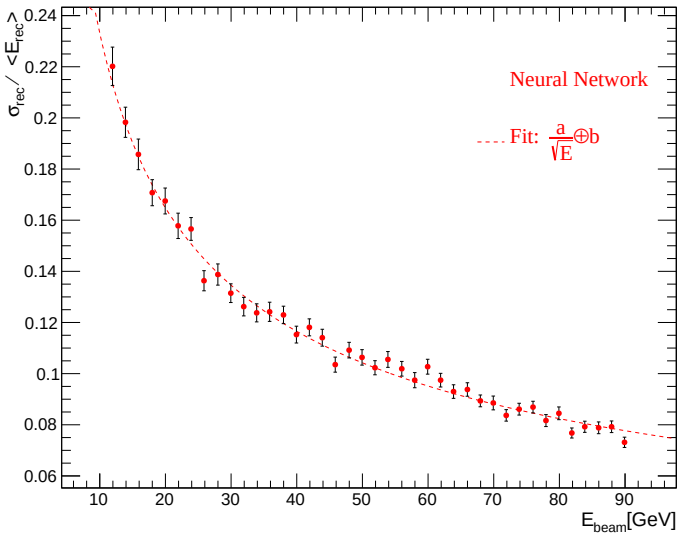

FIGURE 4.21: Pion Energy resolution as a function of the beam energy obtained with the neural network technique.

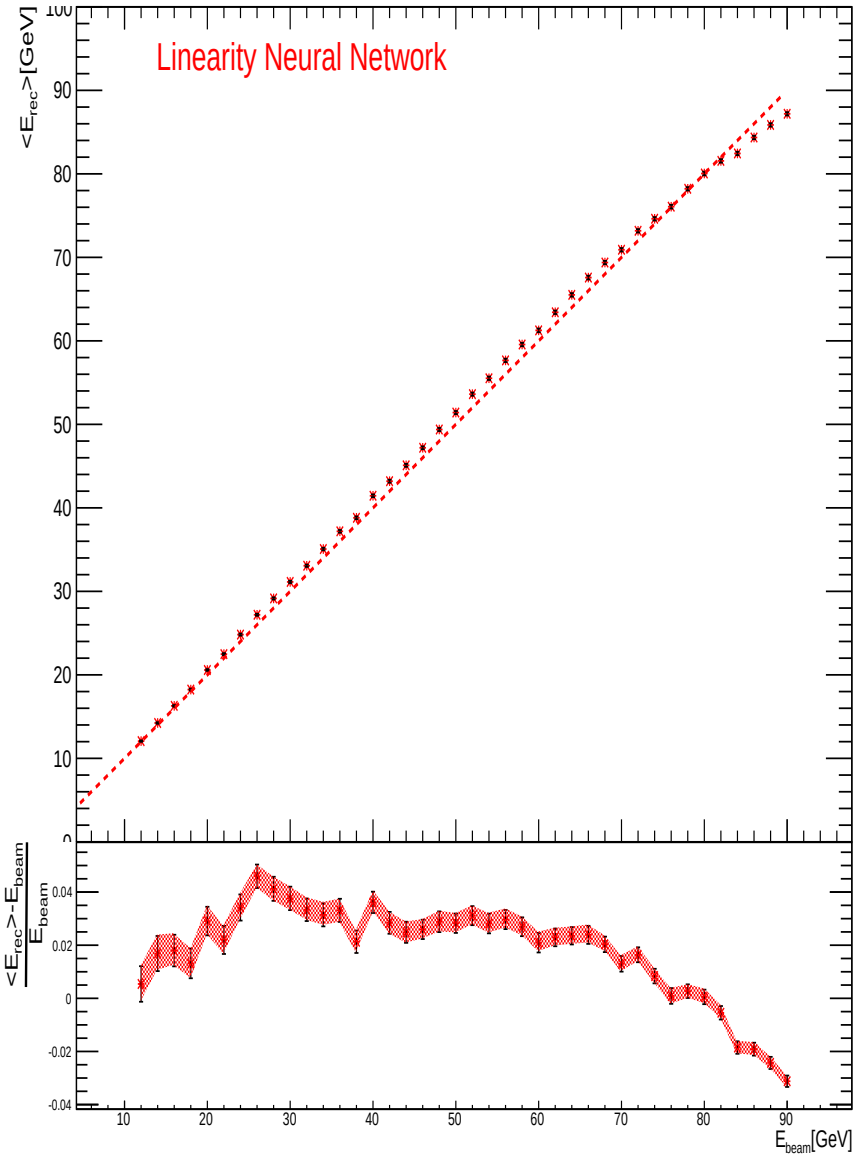

FIGURE 4.22: SDHCAL :(Top) Mean reconstructed energy for pions as a function of the beam energy computed using the neural network technique. The line corresponds to x=y. (Bottom) Relative deviation of the mean reconstructed energy with respect to the beam energy as a function of the beam energy for pions

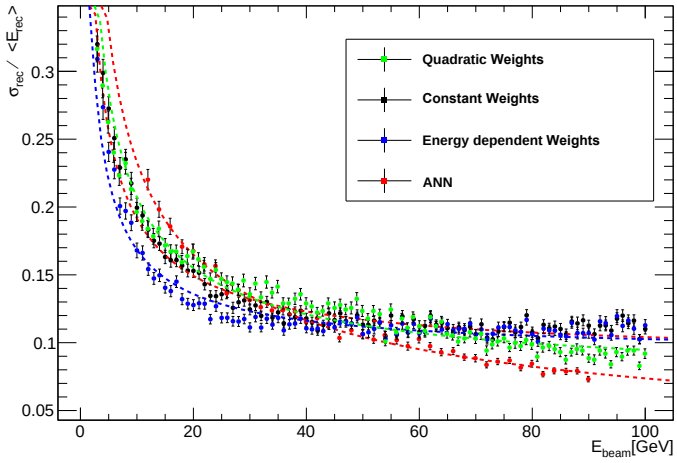

FIGURE 4.23: Comparison of the energy resolution between the different methods of energy reconstruction.

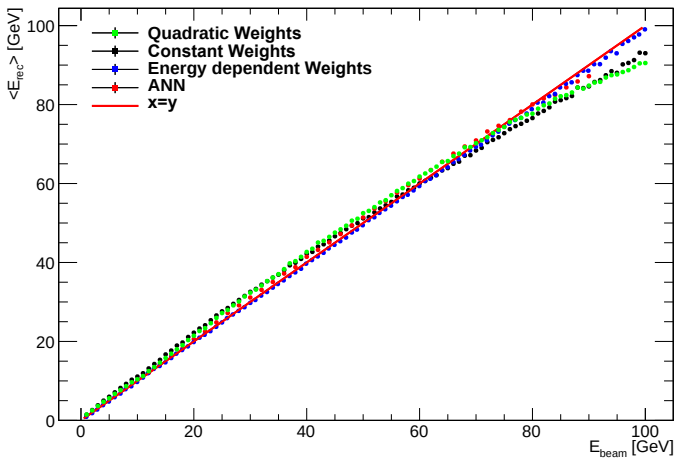

FIGURE 4.24: Comparison of the linearity between the different methods of energy reconstruction.

4.3 Conclusion

An exhaustive study on the optimization of the energy reconstruction using Monte Carlo simulation was presented and detailed in this chapter.

First, a based GEANT4 simulation of SDHCAL prototype as well as a digitization procedure including a modelization of the GRPC response have been described.

Then, a comparison between the binary and multi-threshold modes within SDHCAL have been discussed where the energy reconstruction has been studied for both of them. The multi-threshold capabilities of SDHCAL at high energies have been demonstrated and confirm the important impact on energy resolution of a semi-digital readout. The improvement in energy resolution at high energies is related to a better treatment of the saturation effect thanks to the information provided by the second and third thresholds.

Different analytic methods were developed for the optimization of the energy reconstruction within the multi-threshold mode of SDHCAL. All methods use the fact that the energy of the hadronic shower is expressed as a weighted sum of the number of hits related to the three thresholds N_1, N_2 and N_3. Different weight parametrizations were applied to reconstruct the energy. Results for the linearity, the deviation of reconstructed energy from beam energy, and the energy resolution improvement, given by each parametrization was discussed. The best energy reconstruction technique depends on the energy of the shower.

A further approach of energy reconstruction based on neural network was presented. In this technique, the number of hits N_1, N_2 and N_3 are the variables used by the neural network to estimate the reconstructed energy.

Finally the results obtained with the different techniques were compared. The neural network technique led to the largest energy resolution improvement at high energies reaching 7% at $90GeV$ beside a good accuracy in the energy reconstruction.

The study of energy resolution presented in this work corresponds to the single hadron energy resolution which is different from the jet energy resolution. The jet energy resolution and the single hadron energy resolution are related according to equation 1.1 and thus $\sigma_{Jet} > \sigma_{HCAL}$.

All these results have been summarized in these publications [88], [100] .

In the following, these methods will be applied to the test beam data collected in 2012 at CERN.

Chapitre 5

Energy Reconstruction in SDHCAL using Data

Early in chapter 3 it has been demonstrated that SDHCAL has proved its ability to cope with the requirements of ILC experiments which is its main goal. In addition, one of the basic motivations which drove the construction of a semi-digital hadronic calorimeter optimized for particle flow is the validation of its capability of achieving good resolution of the hadronic energy measurement. Thus, in this chapter we will mainly focus on the energy reconstruction and energy resolution using data taken from test beam. This is one of the most important criteria by which to judge the quality of SDHCAL.

The analysis presented in the following investigates the energy resolution of pion showers recorded in SDHCAL during 2012 beam tests at CERN. After a description of the test beam campains at CERN, the event building procedure is described in section 2 followed by the description of the SDHCAL data quality control in section 3. In section 4 the event selection of the pion data is detailed followed by the results on the linearity and the energy resolution of the hadronic showers, presented in section 5. Different parametrizations of the reconstructed energy are developed and tested in both binary and multi-threshold modes to linearize the response of the calorimeter and improve the energy resolution. Finally, the use of a neural network method in energy reconstruction applied to test beam data will be discussed in section 6. The detailed description of the idea behind the different methods has been discussed previously in chapter 4.

5.1 CERN beam test campains

In order to evaluate the performance of the SDHCAL prototype, several test beams were conducted using the CERN beam facilities. In 2012, the SDHCAL prototype was installed at the CERN Super Proton Synchrotron north experimental area hall EHN1 and exposed to pions, muons and electrons in the H8 beam line in May(2 weeks), the H6 beam line in August(2 weeks) and the H2 beam line in November(2 weeks).

A sketched layout of the T2(T4) wobbling stations for a setting of 150 GeV/c positive secondary hadrons in H2(H6) and 150 GeV/c negative hadrons in H4(H8) is shown in Figs. 5.1 and 5.2.

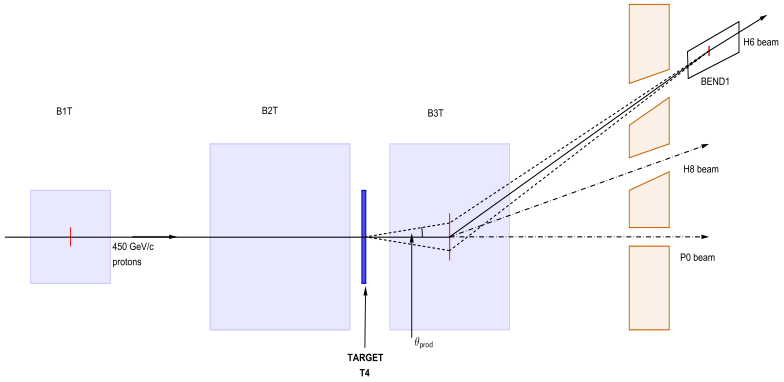

FIGURE 5.1: Schematic view of the T4 wobbling station with positive hadrons in H6 and negatives in H8. B1T, B2T, B3T are the three wobbling magnets.

The SPS provides a high energy 450 GeV proton primary beam, which is then sent on a beryllium target(T2/T4) [102]. An interaction of the protons with the beryllium nuclei occurs and leads to the generation of a secondary beam of moderate high intensity, between 10^4 and 10^7 particles per spill, and consisting mainly in different types of hadrons and a small electron (or positron) component. The composition of the secondary beam depends on the beam momentum, production angle and target length.

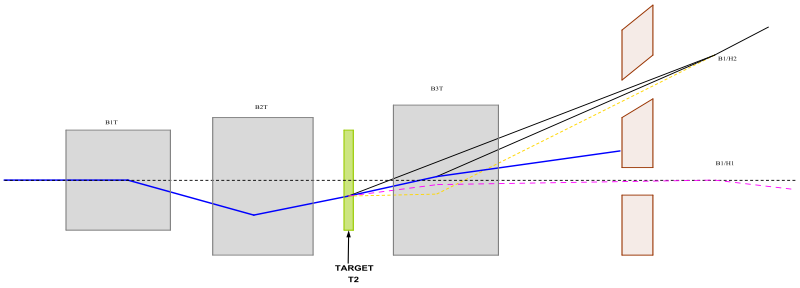

FIGURE 5.2: Schematic view of the T2 wobbling station with positive hadrons in H2 and negatives in H4. B1T, B2T, B3T are the three wobbling magnets.

The T2 target is immediately followed by a wobbling magnet(B3T) monitoring the momentum and the angle of the secondary particles and can be used in two ways : The B3T can be used to sweep +150 and -150 GeV/c particles into the H2 and H4 beam lines after traversing small holes in special thick dump-collimators called TAXes. Otherwise, when applying a strong field in B3T, one can sweep away all the charged particles and keep only the neutrals in the beam line. In this case, negative and positive secondary hadrons beams can be produced from K^0 or Λ^0 decays in the neutral beam.

However, the positive secondary hadron beams are a mixture of pions and protons while the negative hadron beams are almost pure pions beam with some small electron component. The secondary beams can also contain an amount of muons produced by the decay of pions.

Since the beam reachnig the detector cannot be completely pure, a Cherenkov counter can be used to separate electrons from negative pions and protons from positive pions. Alternatively, an absorber transparent to hadrons can be placed along the beam line to remove an electron component from the beam. The absorber is a sheet of lead where the electrons loose most of their energy due to Bremsstrahlung. This option is rather used in SDHCAL test beams to reduce electron contamination in pion beams. However this is not efficient at low beam energies where losses of hadrons are larger.

At this stage, the secondary beam is either distributed simultaneously to different experiments or if needed can be sent to a second target to create a lower energy tertiary beam that can be momentum selected.

In addition, in order to optimize the received beam, several detectors situated in each beam line upstream the calorimeter can be used. These detectors include mainly scin-

tillators to count the beam intensity, multiwire proportional chambers(MWPC) and drift-chambers to monitor the beam profile, the beam coordinates and to record the incoming particle track. A Cherenkov counter can also be used for particle identification purposes [102].

Normally, the settings of the wobbling stations T2 and T4 should be done by the expert beam physicists where the most frenquently changing settings are magnet currents and collimator openings managed using well defined reference settings. However, small changes can also be done by the users via the 'Cesar' application for beam line control [103], according to the experiment requirements depending on the particle type, on the beam momentum and the wanted intensity.

In this context, during beam tests, the beam optics was configured in a way to satisfy the performance requirements of SDHCAL.
First, it was demonstrated that the GRPCs have a low detection rate capability limited to 200Hz/cm^2 [104]. Thus the efficiency will be dramatically impacted at high particle rate. For this reason, it was necessary to enlarge the beam size and reduce its intensity. Consequently, a rate of less than 1000 particles/spill(100 Hz/cm^2) was fixed to provide an optimal efficiency.
The SDHCAL was exposed to different types of particles and enregies : Pions of 10, 20, 30, 40, 50, 60, 70 and 80 GeV and electrons of 10, 20, 30, 40, 50 and 60 GeV were studied in the H2 beam line. More energy points were covered in H6 beam line for pions : 5, 7.5, 10, 15, 20, 25, 30, 40, 50, 60, 70, 80 GeV. Few muons runs were also taken, useful to study and characterize the GRPC.

The pion runs contain an electron contamination. To reduce it, the use of 4mm lead absorber was found efficient at high energy(E$>$20GeV) but not helpfull at low energy where the electron contamination is significant and the losses of pions are important. In addition, pion and electron runs suffer from muon contamination.
As discussed previously, the positive pion runs contain protons as well, especially at energies above 20 GeV where proton contamination can reach 60% at 100 GeV [105]. However, the proton contamination should not be considered as a problem for our study. Indeed, the hadronic showers developed by protons and pions are expected to have similar behaviour at low and medium energies even if the ammount of the electromagnetic component in the showers generated by the two species differs slightly. Since protons have the shortest interaction length compared to pions, larger differences could be observed at high energies. Protons showers are expected to be more contained in the calorimeter with respect to that of pions of the same energy and consequently protons will generate more hits in the calorimeter.

During the test beam, the SDHCAL prototype was operating under a 2-bit readout system using three thresholds with values of 0.114 pC, 5 pC and 15 pC. The threshold one is set to a low value, that of the two others to medium and high values in order to

autotrigger down to 10fC and up to 15pC. The choice of these values is also coherent with the result of the simulation studies on threshold optimization discussed early in chapter 4. The role of the three thresholds is to distinguish among scenarios with few, many or too many particles traversing one pad, expected to provide a better estimation of the number of charged particles in the hadronic shower and thus improve the energy measurment as demonstrated by the simulation study presented previously. A high voltage of 9.6 kV is applied to the GRPC, operated in avalanche mode, while its filling gas is a mixture of 93% of tetrafluoroethane (TFE), 5% of CO_2 and 2% of SF_6.

5.2 Event building

The triggerless acquisition mode, discussed in section 3.4, was performed for the 2012 runs. Firstly, within this acquisition system, the collected data provide the hits information related to the beam particles (pions, electrons, muons) but also cosmics and are assigned to physical events. In addition, the acquisition records also fake events due to the intrinsic noise of the detector. Thus a procedure for physical event reconstruction has to be added.

5.2.1 Time analysis

The potential physical event candidates are reconstructed from hits collection using a time clustering procedure.
Fig. 5.3 shows a time spectrum of an acquisition window where hits are classified according to their time occurence during the acquisition. It includes 40 GeV pions, muons from the beam and cosmic rays as well as noise. One acquisition window is of $\approx$ 30 ms with a granularity of 200 ns (the ASIC internal clock period). The time of each fired pad with respect to the start of the acquisition cycle is recorded.
The time clustering is initiated when a minimum number of hits belonging to a selected clock tick t_0 is reached. In SDHCAL prototype, for each of this selected time slots, the minimum number of hits required to build an event corresponds to a threshold M_{hit} = 7. The choice of M_{hit} value is based on a noise rejection criterion and provides an elimination of a negligible fraction of hadronic showers produced by pions of energy larger than 5 GeV [62].
If in a time slot t_0, the total number of hits $N_{hit} > M_{hit}$, a physical event candidate is considered and built by combining hits belonging to the adjacent time slots in a window of ($t_0 \pm t_{win}$). The width of the time window is estimated from the distribution of the time cluster size built arround t_0 and including adjacent time slots in which at least one hit is recorded (Fig. 5.4). The events are found to be contained mainly in two

time slots of 200 ns.
No hit is allowed to belong to two different events and events with common hits are rejected. The informations about the hits associated to the physical events provide the hit coordinates determined from the position of the Dif in the layer and the detector (DifID), the position of the pad (ChannelID) in the ASIC (AsicID) and their corresponding thresholds. The distribution of number of hits of the physical events for 40 GeV pion run is shown in Fig. 5.5.

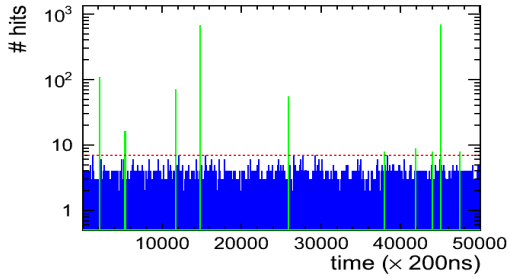

FIGURE 5.3: Hit time spectrum for a 40 GeV pion beam run. The physical event candidates are indicated (in green) over the background noise (in blue). The red line represents the threshold over which the events are considered.

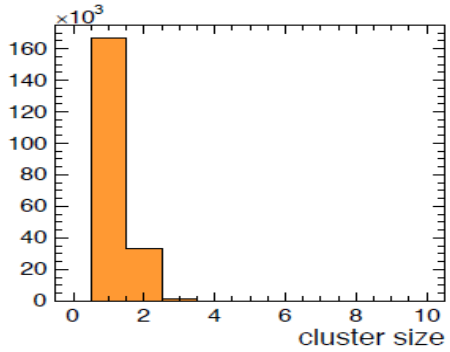

FIGURE 5.4: Time cluster size of hadron interaction events measured in number of time slots of 200 ns.

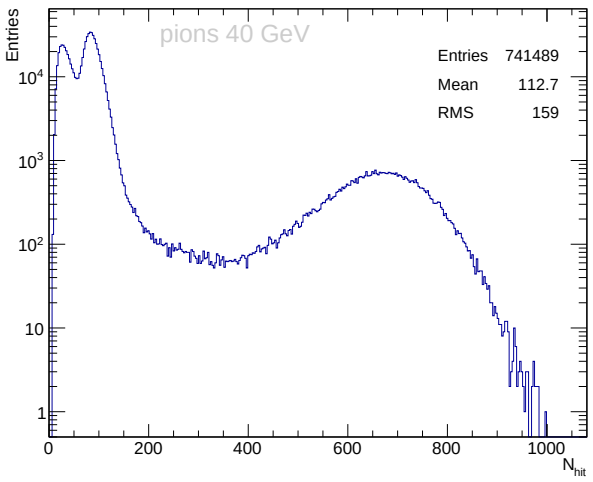

FIGURE 5.5: The distribution of number of hits of the reconstructed events for 40 GeV pions. The first two peaks corresponds to the cosmic and beam muons respectively. The contribution around Nhit = 700 corresponds to pions.

5.2.2 Noise estimation

The remaining hits not related to a particle-interaction event and rejected by the time clustering procedure are used to estimate the intrinsic noise of the detector. Fig. 5.6 shows the distribution of number of hits per clock tick due to intrinsic noise with an average of ≈ 0.35 hit per 200 ns which leads to a mean noise estimation of 1.75 hits per reconstructed physical event which demonstrates that the GRPCs are almost noise-free.

An additional kind of very rare noise, referred to as coherent noise, is related to the prototype's grounding quality. The frequency of this coherent noise was indeed reduced by improving the grounding of the whole detector. This was achieved by connecting the cassettes to each other using metallic connectors adapted to the cassette geometry. In addition, for each active layer of the prototype the grounding of the electronic board was connected to that of the acquisition board and both of them were then connected to the cassette grounding. The grounding of the HV module was also connected to that of the LV grounding at the power supplies level. Applying this scheme improved the grounding significantly but did not eliminate all the related problems and events related to coherent noise, although very rare, are still present. It is clear that a global grounding of the detector is to be necessarily taken into account at early

stages of the design to minimize loops and ensure an effective and reliable grounding of all the system. The coherent noise leads to high local multiplicity and could be responsible for fake estimation of the shower start. However, the coherent noise events are easily identified and removed since the hits of such events are concentrated in one or two layers, as shown in the Fig. 5.7.

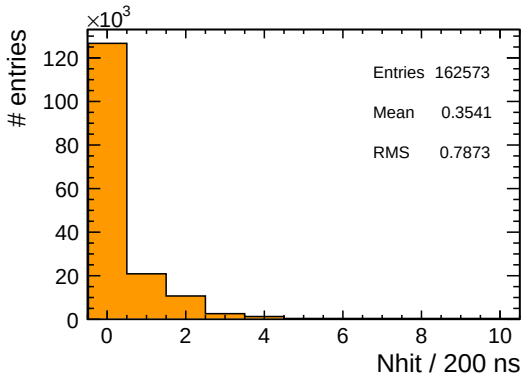

FIGURE 5.6: Distribution of the number of noise hits in a time slot of 200 ns (one clock tick) for the whole detector. An average of 0.35 hits / 200ns is found for the complete detector.

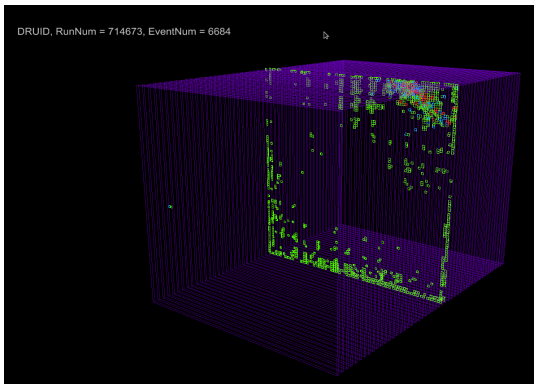

FIGURE 5.7: Example of a coherent noise events display.

5.3 SDHCAL Performance

The SDHCAL performance can be characterized by the measurement of the detection efficiency and pad multiplicity. These parameters are of high interest for the detector calibration and for the online monitoring [1] of the data taken during beam test. They are also needed for simulation purposes of the detector.
The reconstruction of the tracks left by the muons in the calorimeters, thanks to its tracking capability, allows the measurement of these quantities.

5.3.1 Muon track reconstruction

In order to measure the detector efficiency and multiplicity muon-like events were used. These events are defined as such events with $N_{hit} < 200$. To build a muon track, hits of a given layer are grouped in clusters using a nearest-neighbour clustering algorithm. It consists in merging in each plane the hits sharing a common edge. Fig. 5.8 shows two examples of possible clusters configurations. For a clusters cleaning purpose, some selections were applied. A minimum number of clusters is required and corresponds to 5 clusters per muon event. Events with a cluster having more than 5 hits ($N_{hit} < 5$) are removed in order to exclude any possible hard muon interaction in the calorimeter. Besides that, isolated clusters, which are at least 12 cm away from other clusters of the same layer but also of those of other layers, are dropped. Tracks are then built from all the selected clusters excluding the layer under study. The position of the cluster(x_c,y_c) is determined as the unweighted average of the position of the centers of the fired pads. A linear regression, with four parameters ($p_i \in$0,1,2,3), is applied to fit the muon track with a straight line in the space. Its parametric equation is defined by :

$$\begin{cases} x(z_i) = p_0 + p_1 z_i \\ y(z_i) = p_2 + p_3 z_i \end{cases}$$

Finally, tracks are reconstructed by performing χ^2 minimisation using MINUIT [89] and only tracks with a $\chi^2 < 20$ are selected. The reduced χ^2/ndf of the fit is calculated as :

$$\chi^2/ndf = \sum_{i}^{N_{plate}} \left(\frac{x(z_i) - x_{c,i}}{\sigma_{x_{c,i}}} \right)^2 + \left(\frac{y(z_i) - y_{c,i}}{\sigma_{y_{c,i}}} \right)^2 \tag{5.1}$$

1. The online monitoring makes a fast preliminary analysis by computing the efficiency and multiplicity as well as a noise estimation which allows the control of the data quality

The errors on the position σ_{x_c} and σ_{y_c} are calculated as X and Y spread of each cluster divided by $\sqrt{12}$ (estimated assuming a uniform distribution ranges in x $\in$ [0,1]).

One example of muon event is displayed in Fig. 5.9 representing a cosmic ray muon.

FIGURE 5.8: Example of two cluster configurations. In the left the hits are joined by a vertex, yielding two clusters. In the right configuration, all hits are sharing edges two by two, forming an unique cluster.

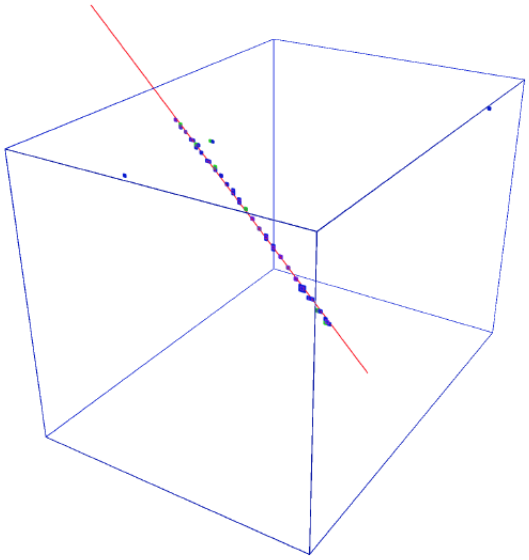

FIGURE 5.9: Example of Mip event in the SDHCAL prototype (cosmic muon).

5.3.2 Efficiency and multiplicity

The efficiency ϵ_i of a given layer 'i' of the detector is defined as the probability to find at least 1 hit within 3 cm of the muon reconstructed track. The studied layer 'i' is removed from the track reconstruction to avoid any bias on the efficiency calculation. The multiplicity μ_i is defined as the mean number of hits matched on the studied layer within 3 cm of the track impact. Fig. 5.10 shows the efficiency and multiplicity of each of the 48 layers measured for the beam test period in H2 line. An efficiency of $\approx$96% observed over most of the chambers with an average multiplicity for the lowest threshold of about 1.7. The low efficiency noticed for layer 42 ($\epsilon_{42} \approx 77\%$) is due to a problem on one of its three electronic slabs. The dispersion around the average multiplicity is mainly due to an irregular resistivity of the painting coating of the layers.

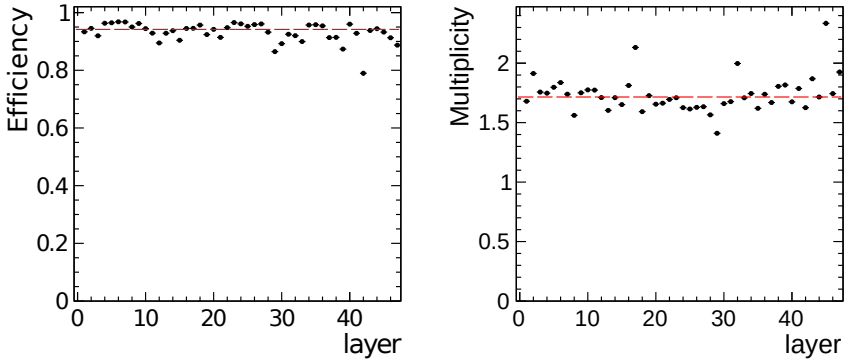

FIGURE 5.10: Efficiency (left) and particle multiplicity (right) of the 48 layers used in 2012 H2 run. The dashed black line is the average efficiency (left) and average multiplicity (right).

5.4 Event selection

As stated previously, the pion beam is contaminated by different kind of particles (electrons, muons, comics, neutrals...). To study the hadronic energy resolution, it is necessary to develop analysis strategies in order to optimize the identification and the selection of the pion events.

For this purpose, a traditional event selection method has been performed and will be discussed in the following.
In both cases, it is required to determine variables capable of discriminating between signal and background processes. Indeed, the segmentation of SDHCAL offers good particle identification capabilities. Thanks to its high longitudinal and transverse granularity, it is possible to measure precisely many physical variables of interest related to the shower topology, usefull to distinguish between incoming pions from elctrons, cosmics and muons as shown in Fig. 5.11.
Such variables, exploited in particle identification in SDHCAL, are the following :

– **First interaction plate :** P_{start}(also called shower start) : corresponds to the plane where the first inelastic interaction of an incoming hadron occurs. To recognize the start of the shower, we look for the first layer containing more than 4 fired pads followed by three consecutive layers with more than 4 fired pads each [2].
– **The number of fired layers** N_{layers} : A fired layer contains at least one fired pad. This variable is of interest in cosmic muons and electrons rejection and can be made more powerful when combined with other variables. It is also useful for shower properties measurments with respect to the starting point.
– **The longitudinal center of gravity of the shower** : It can also be used as an electron/pion separation criterion. The longitudinal centre of gravity, Z_{cog}, is a spatial observable computed for each event with an unweighted sum of the longitudinal hit coordinate with respect to the calorimeter first plane and is defined as :

$$Z_{cog} = \frac{1}{N}\sum_{i=1}^{N} Z_i \tag{5.2}$$

where N is the total number of hits in the calorimeter for event i. Z_i is the distance from the hit layer to the calorimeter front.
In order to avoid fluctuations in the starting point of the longitudinal shower, we used the longitudinal centre of gravity, Z_{cog}^{start}, calculated with respect to the shower start.
Unlike Z_{cog}, the Z_{cog}^{start} is independent of the shower start position and hence considered as an intrinsic variable to describe the longitudinal shower development which is useful for the comparison of different kinds of particles of different nuclear interaction lengths. Z_{cog}^{start} is calculated for each event as :

$$Z_{cog}^{start} = \frac{1}{N_{shower}} \sum_{i=1}^{N_{shower}} (Z_i - Z_{start}) \tag{5.3}$$

where N_{shower} is the number of hits in the calorimeter counted from the shower start layer, and Z_{start} is the distance from the shower start layer to the calorimeter

2. Three consecutive layers are required to eliminate fake shower starts due to accidental noise or local high multiplicity effect.

front. Besides Z_{cog}^{start}, the standard deviation $\sigma_{Z_{cog}^{start}}$ of the longitudinal hit position arround the longitudinal centre of gravity is calculated for each event using the following formula :

$$\sigma_{Z_{cog}^{start}} = \sqrt{\frac{1}{N_{shower}} \sum_{i=1}^{N_{shower}} (Z_i - Z_{start})^2 - (Z_{cog}^{start})^2} \tag{5.4}$$

– **The shower axis** : The shower axis is defined using the coordinates of the transversal centre of gravity of the shower defined by :

$$X_{cog} = \frac{1}{N} \sum_i^N X_i \text{ and } Y_{cog} = \frac{1}{N} \sum_i^N Y_i$$

$$\sigma_{X_{cog}} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} X_i^2 - (X_{cog})^2} \tag{5.5}$$

$$\sigma_{Y_{cog}} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} Y_i^2 - (Y_{cog})^2} \tag{5.6}$$

– **The mean shower radius** : it describes the transversal shower development and is defined as the mean radial distance of the hits to the shower axis.

$$\langle R \rangle = \frac{1}{N} \sum_{i=1}^{N} R_i \tag{5.7}$$

$R_i = \sqrt{(X_i - X_{cog})^2 + (Y_i - Y_{cog})^2}$ is the distance from the hit position with coordinates (X_i, Y_i) to the shower axis with coordinates (X_{cog}, Y_{cog}).
In addition to $\overline{R}$, the standard deviation, $\sigma_{\overline{R}}$, of the hits around the shower radius is defined for each event as :

$$\sigma_R = \sqrt{\langle R^2 \rangle - \langle R \rangle^2}$$

$$= \sqrt{\frac{1}{N_{shower}} \sum_{i=1}^{N_{shower}} (R_i - \langle R \rangle)^2} = \sqrt{\frac{1}{N_{shower}} \sum_{i=1}^{N_{shower}} R_i^2 - \langle R \rangle^2} \tag{5.8}$$

With $\langle R^2 \rangle = \frac{1}{N} \sum_{i=1}^{N} R_i^2$

– $\boldsymbol{N_{neut}}$: The number of hits in the first fifth layers. This variable helps to eliminate neutral contamination.

– N_{IP} : corresponds to the number of planes for which the standard deviation of the hit position in X and Y directions (σ_X, σ_Y) exceed 5 cm. Combined to other variables, it is used for radiative muon rejection.

$$\sigma_X = \sqrt{\frac{1}{N_{plane}} \sum_{i=1}^{N_{plane}} X_i^2 - (\overline{X_{plane}})^2} \tag{5.9}$$

$$\sigma_Y = \sqrt{\frac{1}{N_{plane}} \sum_{i=1}^{N_{plane}} Y_i^2 - (\overline{Y_{plane}})^2} \tag{5.10}$$

N_{plane} is the number of hits in the corresponding plane.

– N_{edge} : the number of hits situated in the edge of the calorimeter layers, of 10 cm thickness. This variable can be used to reduce the lateral leakage.

The distribution of some of this topological variables are shown in Appendix G.1.

5.4.1 Proton contamination

According to the ATLAS Collaboration, the H6 pion beam is contaminated by protons for momenta above 20 GeV/c and which becomes larger in the momentum range from 50 to 100 GeV/c (45-60%) [105].
Unfortunately, it would not be possible to distinguish the two species in the abscence of a Cherenkov detector in front of SDHCAL protoype. Consequently, the results on the hadronic energy resolution using H6 data samples will concern both species.
However, the H2 beam is free of protons and hence the results of the hadronic energy resolution obtained in our prototype using H2 collected data will be assigned only to pions and will be then considered as a reference.

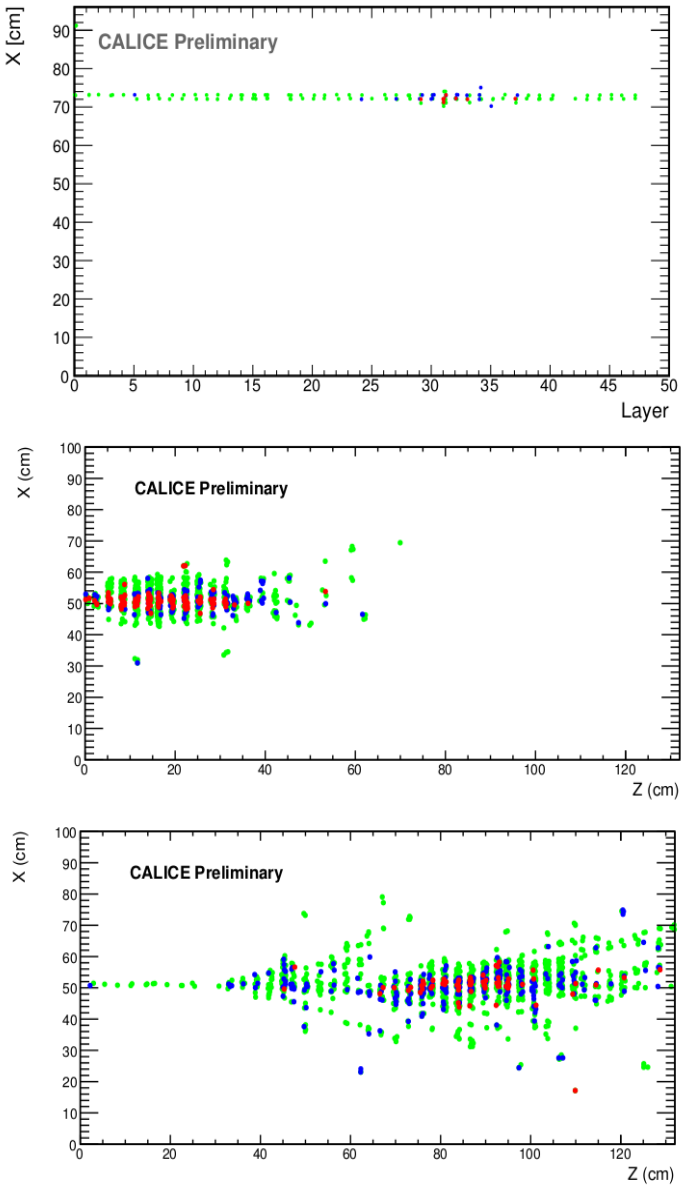

FIGURE 5.11: Hit patterns of 50 Gev muons, 70 GeV electrons and 80 GeV pions recorded in SDHCAL prototype. Red color indicates the highest threshold fired pads, blue color indicates the middle threshold, and green color is for the lowest one.

5.4.2 Electron contamination

The pion runs suffers from electron contamination. Although a 4 mm lead filter was used to reduce this contamination, a non negligible amount of electrons is still present among the pion samples.

Since the nuclear interaction length (λ_I = 20.4 cm for pions in steel) is much larger than the radiation length (X_0 = 1.76 cm for electrons in steel), electrons interact earlier in the calorimeter compared to pions, particularly in the first plates. Therefore, electrons deposit the largest fraction of their energy in the first half of the calorimeter and hence, pion cascades should be much longer compared to the relatively small penetrating electromagnetic showers in the calorimeter. According to this, electromagnetic showers in the energy range between 5 and 80 GeV are longitudinally contained in less than 30 layers of SDHCAL prototype. Besides that, requiring that the shower starts in the fifth layer or after should in principle remove almost all of the electrons. This was validated by simulation studies as shown in Fig. 5.12 and Appendix G.1. In addition, Fig. 5.13 shows the simulated distributions of the selection variables for 50 GeV electrons and provides the proportion of events not rejected by the cuts defined by $\frac{N_{events}|Var>cut}{N_{events}}$. As we can notice the proportion of electromagnetic showers identifed as hadronic showers is low. However, one must take into account, that these electron rejection criteria should be correlated to minimize the loss of true pion hadronic shower, so that high energy pions starting their shower before the first fifth layers and where the number of fired layers exceeds 30 are not rejected. Except for the showers generated by hadrons of 5 GeV, low energy hadrons starting showering before the first fifth plates with less than 30 fired layers, are rejected by the selection criteria and thus lost but no bias is introduced on the energy reconstruction. Fig. 5.14 demonstrates the high identification capability reached with this selection applied to 10, 40 and 70 GeV electron runs.

The selection cuts used to reject electrons are recapitulated in the following :
– The shower start ≥ 5 planes.
– $N_{layers} > 30$, The number of fired planes should be larger than 30.

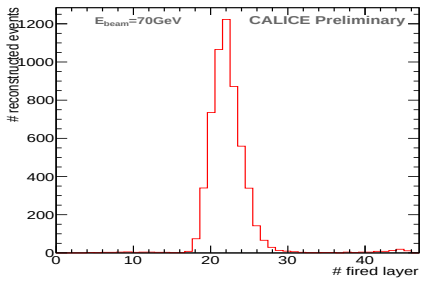

FIGURE 5.12: Distribution of number of fired layers by 70 GeV electrons as given by the simulation.

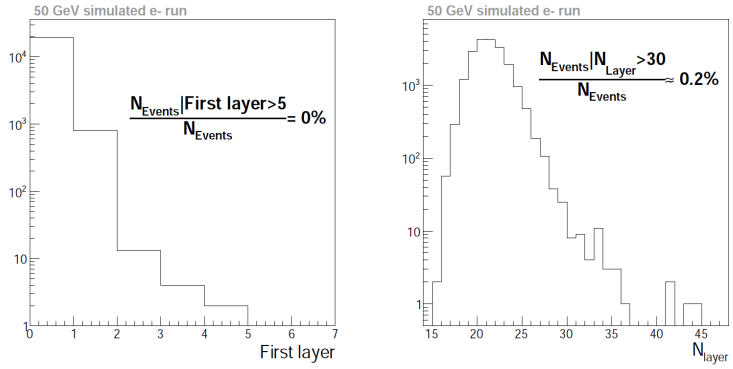

FIGURE 5.13: Distributions of the first plane of interaction ($showerstart$) (left) and the number of fired planes (N_{layers}) (right) for simulated 50 GeV electrons [80].

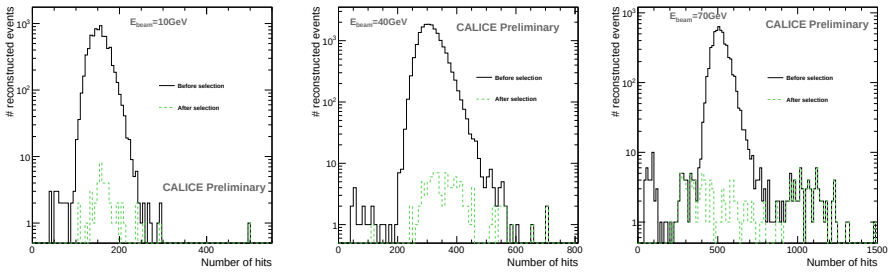

FIGURE 5.14: Distribution of number of hits for 10, 40 and 70 GeV electron runs before (solid black line) and after (dashed green line) electron rejection.

5.4.3 Muon contamination

The muon component is an admixture of the cosmic-ray muons and the beam ones produced by pions stopped in the collimator or those decaying before reaching the prototype.
Indeed, muons can be distinguished from pions and from electrons by their small rate energy loss in the calorimeter which allow them to travel much longer without showering. Muons generate straight tracks almost without interacting in the detector, while electrons and pions produce electromagnetic and hadronic showers with many particle production as shown in Fig. 5.11, which exhibit a distinctly different pattern in the calorimeter considered as an efficient mean for muon identification. Thus, the quantity of muons is greatly reduced by requiring an average number of hits per fired layer greater than 2.2, which is a value higher than the muon multiplicity estimated in SDHCAL measurements. For the case of radiative muons(Fig. 5.15), we require that the ratio between the number of layers for which the standard deviation of the hit position in both X and Y directions (σ_X, σ_Y) exceed 5 cm and the total number of layers with at least one fired hit is more than 20%. This selection could impact particularly the pions which start showering in the last layers of the calorimeter. Although this process is quite rare, it can nevertheless lead to a small loss of this pion category.
An excellent muon rejection is ensured by applying the following selections :

– $\frac{N_{hits}}{N_{planes}} > 2.2$, N_{hits} is the number of hits produced in the calorimeters per event and N_{planes} is the number of fired planes.
– $\frac{N_{LP}}{N_{planes}} > 0.2$ for radiative muon rejection.

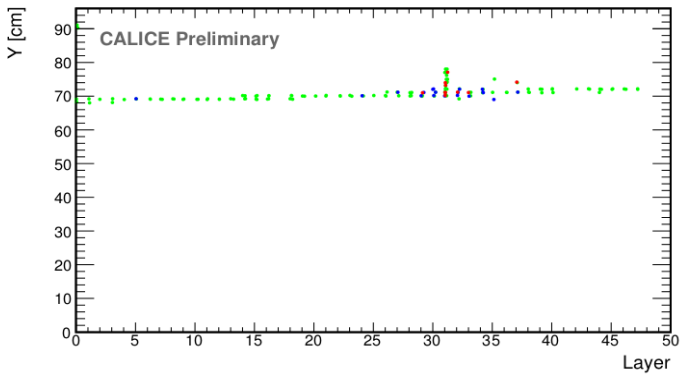

FIGURE 5.15: An example of radiative muon of 50 GeV.

5.4.4 Additional selections

The energy resolution can be deteriorated by the leakage, caused mostly by hadrons interacting later and thus escaping the calorimeter at the back (longitudinal leakage) or at the sides (lateral leakage) where an event contains many particles.
In order to make the longitudinal energy leakage negligible, the number of fired planes with at least one hit each should be less than 42 plates.
For the transversal leakage case, a selection using the transversal components of the shower center of gravity is applied to ensure a transversal containment of the shower : $-20cm < X_{cog} < 20cm$ and $-20cm < Y_{cog} < 20cm$. It is also asked that the ratio of N_{edge} to the total number of hits in the calorimeter should be less than 20%.
In addition, to these selections, to eliminate neutral contamination, events with less than 4 hits in the first 5 layers are not considered ($N_{neut} \geq 4$).
To enhance the purity of our data samples, a special care was given to avoid the events with more than one incoming particle. Such events for which the first five layers have hits separated by more than 5 cm are eliminated.

The selection criteria are summarized in Table. 5.2. Table. 5.1 shows the efficiency of the selection of the hadronic showers provided by simulation and corresponding to the ratio of events identified as hadronic showers to the total number of events. The result of the application of such selection in test beam data is shown in Appendix I.1 and in Fig. 5.16 for three energies : 10, 40 and 80 GeV where the total number of hits of the collected events is drawn.

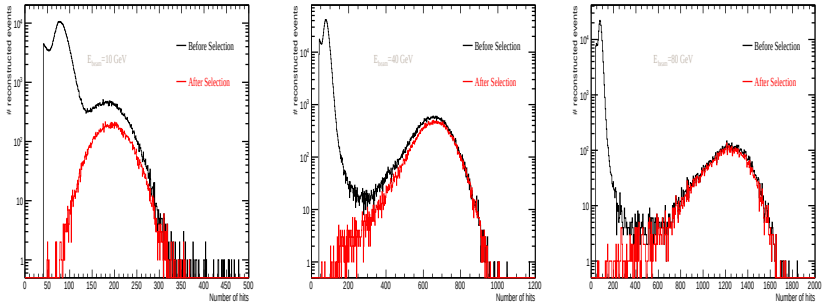

FIGURE 5.16: Number of hits for 10, 40 and 80 GeV pion runs before (black line) and after (red line) full selection.

$E_{beam}(GeV)$	Efficiency
5	$57.9 \pm 0.3\%$
10	$85.8 \pm 0.2\%$
15	$90.5 \pm 0.2\%$
20	$92.3 \pm 0.2\%$
25	$93.9 \pm 0.2\%$
30	$94.3 \pm 0.2\%$
40	$95.1 \pm 0.2\%$
50	$94.9 \pm 0.2\%$
60	$95.1 \pm 0.2\%$
70	$94.5 \pm 0.2\%$
80	$94.3 \pm 0.2\%$

TABLE 5.1: Efficiency of the selection of the hadronic showers with respect to the incident energy[80].

Electron rejection	Shower start ≥ 5 or $N_{layer} > 30$
Muon rejection	$\frac{N_{hit}}{N_layer} > 2.2$
Radiative muon rejection	$\frac{N_{layers}\|RMS>5cm}{N_{layer}} > 20\%$
Neutral rejection	$N_{hit \in First\ 5\ layers} \geq 4$

TABLE 5.2: Summary of the different cuts applied to select the pions.

5.4.5 Time Correction

Because the resistivity of the electrodes is quite high ($\rho = 10^{12}\Omega cm$), the detection efficiency can dramatically decrease when the particle rate increases during the spill. Although the beam was configured to ensure a rate of less than 1000 particles/spill, this effect was observed for some runs with a decrease of the number of hits associated to the hadronic showers during the spill. The decrease is accentuated for high energies and especillay noticed for the number of hits related to the second and third thresholds as demonstrated in Fig. 5.17 and Appendix H.1. This behaviour should be corrected otherwise it yields a deterioration of the energy resolution.
For this purpose, a linear fit calibration was applied. In this method, for each hadronic shower, the average number of hits related to each of the three thresholds is plotted as a function of their time occurence with respect to the spill starting time. Then a linear fit to the hits distribution is performed and the slope of the fit is used to correct the number of hits N_j for each threshold j and for each run .

The corrected number of hits N_{corrj} is defined as :

$$N_{corr,j} = N_j - S_{lope_j} * T_{imeInSpill} \tag{5.11}$$

where S_{lope_j} is the correction slope for threshold j and $T_{imeInSpill}$ is the occurrence time.
Fig. 5.18 shows the effect of the application of this correction. As can be seen the conditions of low particle rate at the beginning of the spill is restored. The results after the time correction for 60 and 70 GeV runs for the total number of hits can be seen in Fig. 5.19. Fig. 5.20 shows the distribution of the number of hits associated to the three thresholds for 60 GeV pion run before and after the time calibration.

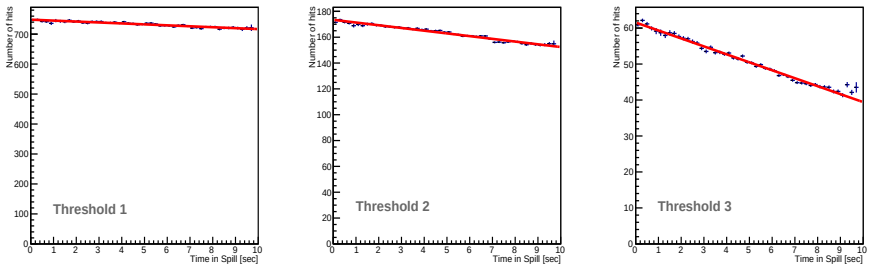

FIGURE 5.17: From left to right : mean number of hits as a function of spill time for first, second and third threshold at 60 GeV pion run from H6. Linear fit is also shown.

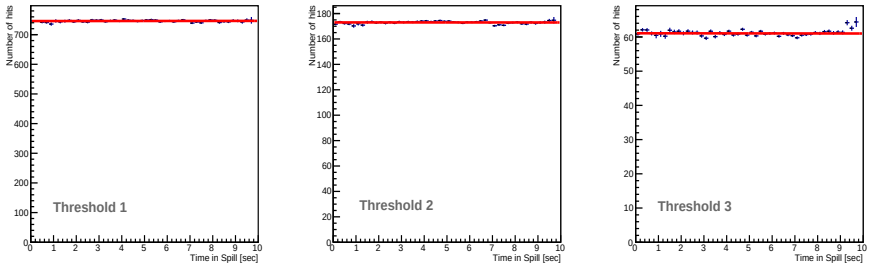

FIGURE 5.18: From left to right : corrected mean number of hits as a function of spill time for first, second and third threshold at 60 GeV pion run from H6. Linear fit is also shown.

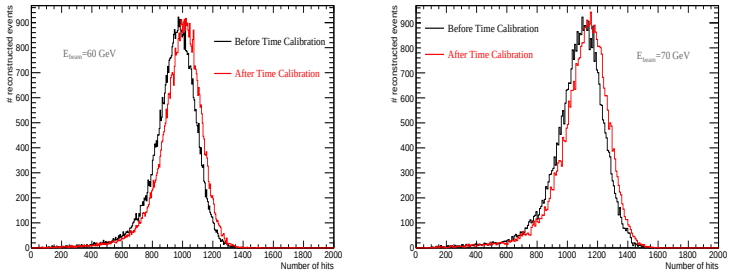

FIGURE 5.19: Distribution of number of hits for pion runs 60 and 70 GeV before (black line) and after (red line) time correction.

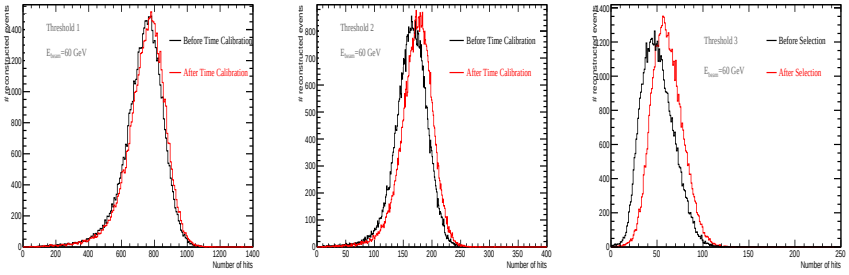

FIGURE 5.20: Distribution of number of hits of the three thrsholds for pion run 60 GeV before (black line) and after (red line) time correction.

5.5 Energy reconstruction and intrinsic energy resolution

After the identification of the hadronic showers using the approches presented in section 5.4, relevant quantities may be determined. These informations are particularly well suited for a study of the energy resolution and linearity of SDHCAL.

The knowledge of the total number of hits N_{hit} is of particular importance in such studies. The distribution of the total number of hits N_{hit} is plotted for each energy for H2 and for H6 runs. Fig. 5.21 shows the distributions of N_{hit} for two energies of the H6 runs. The distributions indicate an asymetric shape and a tail at low number of N_{hit}, more apparent for high energies and which is due to the shower leakage.
To reduce the effect of the tail, there are two possibilities of fit performed to estimate the average number of hits for a given energy, used later to study the linearity of the

SDHCAL.
The first is a Gaussian fit applied in a range limited to 1.5 standard deviations around the mean value as shown in Fig. 5.21.
The second fit uses the Crystal Ball (CB) function [106](see Appendix J.1) to take into consideration the presence of this tail. This latter kind of fit has been used for the estimation of systematic uncertainties.
Fig. 5.22 shows the mean number of hits as a function of the beam energy. The plots are fitted with a linear function between [5,20] GeV in case of the 2012 H6 runs and those between [10,30] GeV for the 2012 H2 runs. One can notice for energies above 40 GeV deviations from expected values of the number of hits obtained by fit, which may indicate a saturation effect at high energies.
The knowledge of further informations, such as the number of hits associated to the three thresholds, is necessary, if one is interested in the correction of the saturation effect and thus improving the energy resolution.

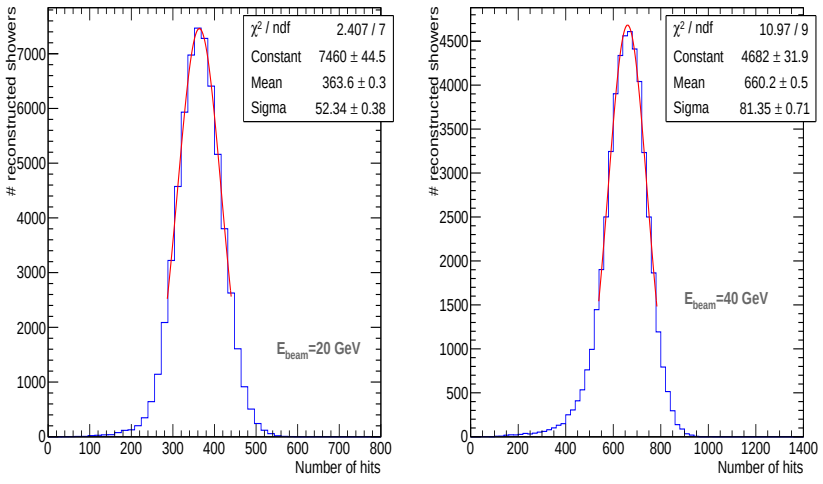

FIGURE 5.21: Total number of hits for pion showers of 20 GeV (left) and 40 GeV (right) of the 2012 H6 runs. The distributions are fitted with a Gaussian function in a $\pm 1.5\sigma$ range around the mean.

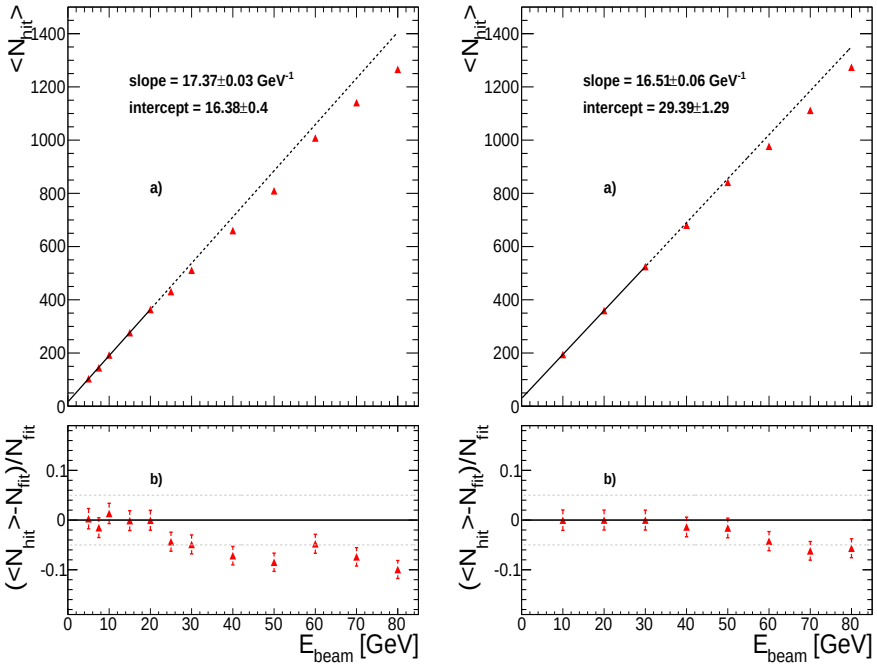

FIGURE 5.22: (a) Mean number of hits as a function of the beam energy for reconstructed hadron showers in the 2012 H6 runs (left) and for pion showers in the 2012 H2 runs (right). The line indicates the result of a linear fit of the form $< N_{hit} >$=slope $\cdot E_{\mathrm{beam}}+$ C, for energies up to 20 GeV for H6 line and up to 30 GeV for H2 line(solid section of the line). (b) Relative deviation of the observed mean number of hits to the fitted line shown in (a) as a function of the beam energy for reconstructed hadronic showers.

5.5.1 Binary mode

In order to correct the saturation effect observed in the behaviour of the mean number of hits as a function of the energy, a quadratic parametrization of N_{hit} has been proposed for the energy estimation :

$$E_{rec} = A_1 N_{hit} + A_2 N_{hit}^2 + A_3 N_{hit}^3 \tag{5.12}$$

The parameters A_1, A_2 and A_3 are determined through a minimization based on the χ^2 function :

$$\chi^2 = \sum_{i=1}^{N} \frac{(E^i_{\text{beam}} - E^i_{\text{rec}})^2}{E^i_{\text{beam}}} \tag{5.13}$$

Where N is the total number of events used in minimization.
The reconstructed energy distributions are then plotted and fitted with two step Gaussian fit. A first Gaussian fit was performed over the full range of the distribution, followed by a second Gaussian fit used only in the range of $\pm 1.5\sigma$ of the mean value of the first fit as shown in Fig. 5.23.
The results using this approach are summarized in Fig. 5.24 and Fig. 5.25 as well as in Table 5.3.
Fig. 5.24 shows the mean reconstructed energy as a function of the beam energy. One can see from these plots that, as expected, this method restores the linearity.
The sigma of the second fit is used to estimate the energy resolution shown in Fig. 5.25.

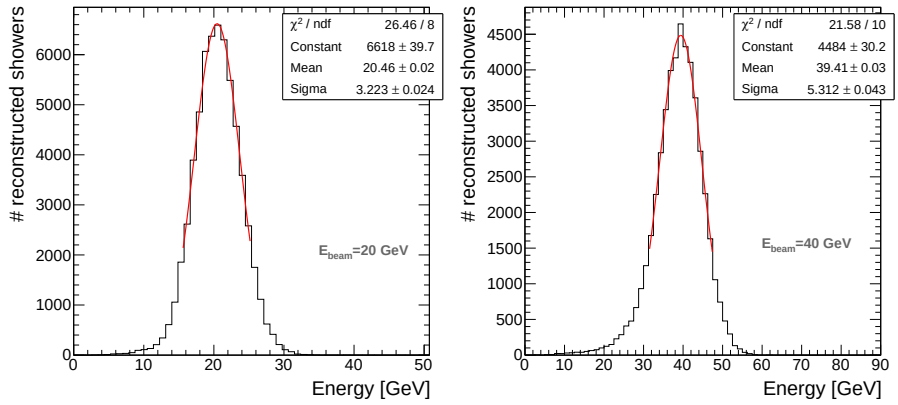

FIGURE 5.23: Reconstructed energy for pion showers of 20 GeV (left) and 40 GeV (right) for 2012 H6 runs using only the total number of hits (binary mode using quadratic Nhit function coefficients). The distributions are fitted with a Gaussian function in a $\pm 1.5\sigma$ range around the mean.

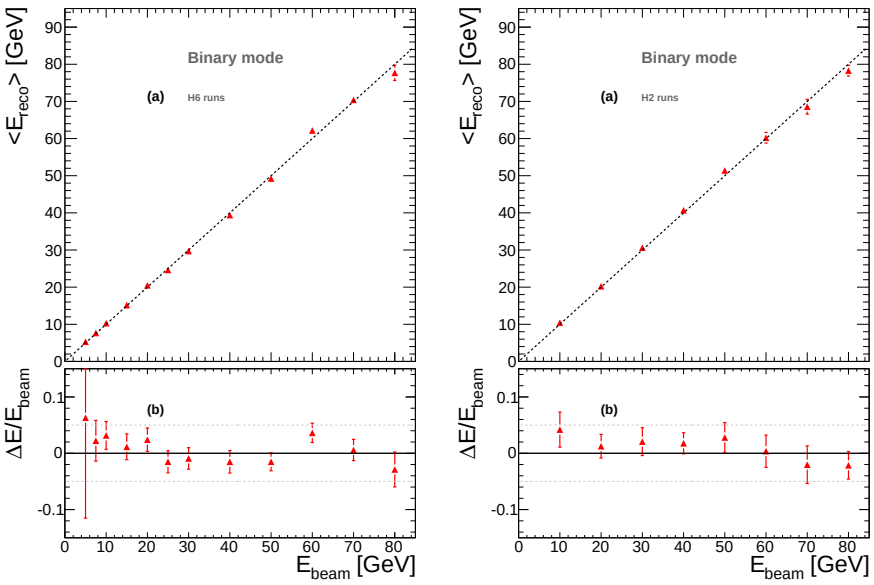

FIGURE 5.24: (a) Mean reconstructed energy for pion showers as a function of the beam energy in the 2012 H6 runs (left) and for pion showers in the 2012 H2 runs (right). Dashed line is the y=x one. (b) Relative deviation of the pion mean reconstructed energy relative to the beam energy as a function of the beam energy. The reconstructed energy is computed using only the total number of hits (N_{hit})

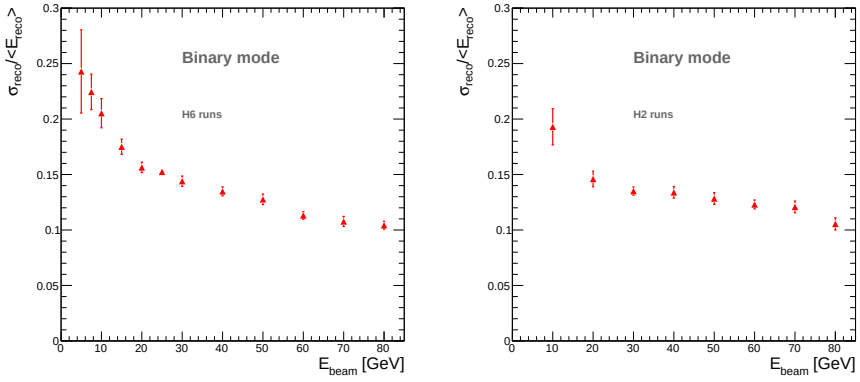

FIGURE 5.25: $\frac{\sigma_{reco}}{E_{reco}}$ of the reconstructed pion energy as a function of the beam energy in the 2012 H6 runs (left) and the 2012 H2 runs (right). The reconstructed energy is computed using only the total number of hits (N_{hit})

H_6			H_2	
$E_{beam}(GeV)$	$E_{reco}(GeV)$	$\frac{\sigma_{reco}}{E_{reco}}$	$E_{reco}(GeV)$	$\frac{\sigma_{reco}}{E_{reco}}$
5	5.32 ± 0.91	0.243 ± 0.037		
7.5	7.67 ± 0.25	0.224 ± 0.016		
10	10.32 ± 0.18	0.205 ± 0.013	10.42 ± 0.27	0.193 ± 0.016
15	15.17 ± 0.24	0.175 ± 0.006		
20	20.48 ± 0.26	0.156 ± 0.005	20.25 ± 0.26	0.146 ± 0.007
25	24.62 ± 0.29	0.153 ± 0.002		
30	29.73 ± 0.32	0.144 ± 0.005	30.62 ± 0.54	0.134 ± 0.004
40	39.41 ± 0.48	0.135 ± 0.004	40.71 ± 0.38	0.134 ± 0.005
50	49.25 ± 0.16	0.128 ± 0.005	51.40 ± 1.05	0.128 ± 0.005
60	62.17 ± 0.24	0.113 ± 0.003	60.22 ± 1.31	0.123 ± 0.004
70	70.41 ± 0.61	0.107 ± 0.005	68.58 ± 1.80	0.120 ± 0.005
80	77.69 ± 2.07	0.104 ± 0.004	78.29 ± 1.47	0.105 ± 0.006

TABLE 5.3: Mean reconstructed energy E_{reco} and energy resolution $\frac{\sigma_{reco}}{E_{reco}}$ observed and associated uncertainties for the Binary mode for H6 and H2 beam lines.

5.5.2 Multi-threshold mode

The semi-digital mode of SDHCAL provides informations of the number of hits associated to the three thresholds which may be used to improve the energy reconstruction as predicted by the simulation studies presented in chapter 4. In this case, the energy of the hadronic shower is expressed as a weighted sum of the number of hits related to the three thresholds N_1, N_2 and N_3 according to the following equation :

$$E_{\rm rec} = \alpha(N_{\rm hit}) \cdot N_1 + \beta(N_{\rm hit}) \cdot N_2 + \gamma(N_{\rm hit}) \cdot N_3 \tag{5.14}$$

α, β and γ are parametrized with a quadratic function of the total number of hits, such as :

$$\alpha(N_{\rm hit}) = A_1 + A_2 \cdot N_{\rm hit} + A_3 \cdot N_{\rm hit}^2 \tag{5.15}$$

The optimum values of these parameters is determined through a minimization based on equation 5.13. The procedure was applied to only few energy points using only about a third of the collected data in H2 runs where there is no proton contamination. The parametrization of α, β and γ as a function of Nhit is presented in Fig. 5.26.

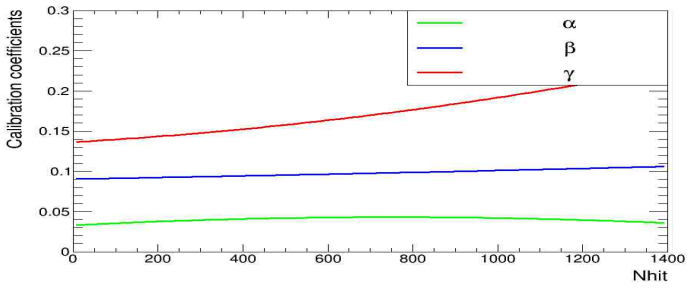

FIGURE 5.26: Evolution of the coefficient $\alpha(green)$, $\beta(blue)$ and γ in terms of the total number of hits.

The parameters are then used to estimate the energy of all collected data (H2 and H6 runs). The energy distributions obtained in this way are fitted with a two step Gaussian fit as described before and the standard deviation σ and the mean reconstructed energy are then extracted to calculate the energy resolution and the linearity.
Fig. 5.27 shows the energy distributions reconstructed using the multi-threshold mode and fitted with the two step Gaussian fit.
The results obtained with the multi-threshold mode are summarized in Table 5.4.

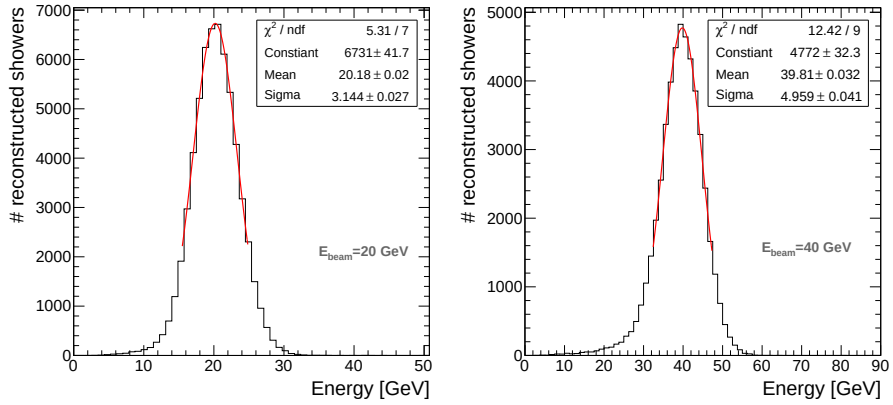

FIGURE 5.27: Reconstructed energy for pion showers of 20 GeV (left) and 40 GeV (right) for 2012 H6 runs using information from three thresholds (multi-threshold mode). The distributions are fitted with a Gaussian function in a $\pm 1.5\sigma$ range around the mean.

Fig. 5.28(a) shows the mean reconstructed energy of pions as a function of the beam energy obtained with the multi-threshold mode. As expected the linearity has been restored on large energy range going from 5 GeV up to 80 GeV. The relative deviation of the reconstructed energy with respect to the beam energy is below 5% for all the enrgy points as shown in Fig. 5.28(b). As predicted from the simulation studies, the use of the three thresholds information provides a very good improvement of the energy resolution at energies higher than 30 GeV and reaching 7.8% at 80 GeV as shown in Fig. 5.29.

These results are promising since the data were collected without using any electronics gain correction to improve the homogeneity of the detector's response. Also, the results obtained for the two periods of test beam with the same energy points are in a good agreement especially at low energy where the proton contamination of the H6 pion beam is low. For energies higher than 60 GeV, the presence of protons in the H6 data can explain that the reconstructed energies are higher than those of the H2 as explained before. It is important to mention here that having the same response to hadrons in the energy range for which the proton contamination is low shows clearly that the behavior of SDHCAL prototype is stable between the two periods.

	H_6		H_2	
$E_{beam}(GeV)$	$E_{reco}(GeV)$	$\frac{\sigma_{reco}}{E_{reco}}$	$E_{reco}(GeV)$	$\frac{\sigma_{reco}}{E_{reco}}$
5	5.19 ± 0.67	0.240 ± 0.044		
7.5	7.33 ± 0.19	0.235 ± 0.028		
10	9.96 ± 0.23	0.214 ± 0.013	10.06 ± 0.26	0.196 ± 0.017
15	14.73 ± 0.38	0.178 ± 0.006		
20	20.2 ± 0.22	0.155 ± 0.005	19.80 ± 0.29	0.144 ± 0.004
25	24.56 ± 0.37	0.150 ± 0.005		
30	29.75 ± 0.69	0.144 ± 0.005	30.39 ± 0.48	0.132 ± 0.003
40	39.8 ± 0.55	0.124 ± 0.005	40.51 ± 0.46	0.121 ± 0.005
50	50.00 ± 0.83	0.110 ± 0.004	51.34 ± 1.08	0.111 ± 0.005
60	61.84 ± 0.37	0.106 ± 0.005	60.15 ± 1.29	0.102 ± 0.004
70	71.7 ± 0.91	0.088 ± 0.003	67.20 ± 1.15	0.101 ± 0.005
80	80.06 ± 1.08	0.078 ± 0.002	78.26 ± 1.33	0.079 ± 0.004

TABLE 5.4: Mean reconstructed energy E_{reco} and energy resolution $\frac{\sigma_{reco}}{E_{reco}}$ observed and associated uncertainties for multi-threshold mode for H6 and H2 beam lines.

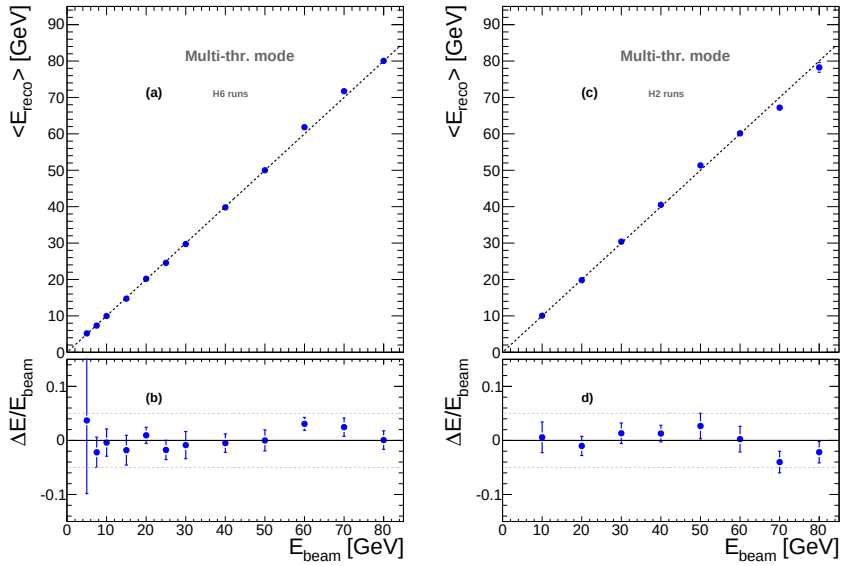

FIGURE 5.28: (a) Mean reconstructed energy for pion showers as a function of the beam energy in the 2012 H6 runs (left) and for pion showers in the 2012 H2 runs (right). Dashed line is the y=x one. (b) Relative deviation of the pion mean reconstructed energy relative to the beam energy as a function of the beam energy. The reconstructed energy is computed using the three thresholds information (multi-threshold mode).

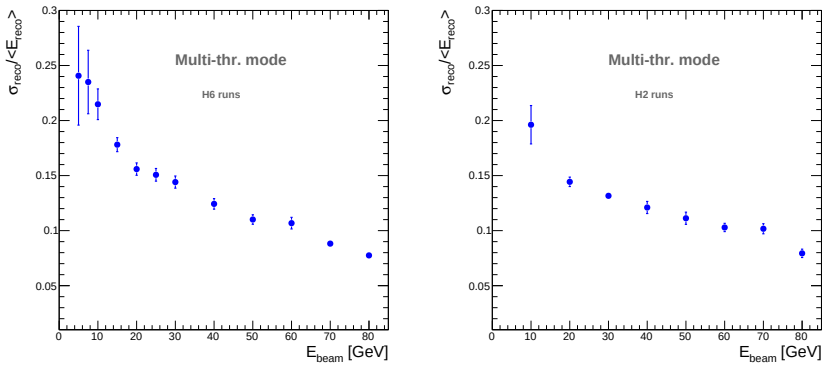

FIGURE 5.29: $\frac{\sigma_{reco}}{E_{reco}}$ of the reconstructed pion energy as a function of the beam energy in the 2012 H6 runs (left) and the 2012 H2 runs (right).The reconstructed energy is computed using the three thresholds information (multi-threshold mode).

5.5.3 Binary vs. Multi-threshold modes

The energy reconstruction using a quadratic parametrization of the total number of hits has been realized for both binary and multi-thresholds modes of SDHCAL.
This method has demonstrated its efficiency to restore linearity over a large energy scale for both modes. However, the energy resolution achieved with the multi-threshold mode is the most improved. Although both of them achieve comparable energy resolution at low energies, the energy resolution obtained with the multi threshold mode is much better for energies higher than 30 GeV. This may indicate that the thresholds information seems to provide additional informations useful to correct for the saturation which starts to show up at energies higher than 30 GeV. A direct comparison of the two results is presented in Fig. 5.30.

Fig. 5.31 shows the energy distributions for both binary and multi-thresholds modes for 80 GeV, 70 GeV and 20 GeV pions of 2012 H6 runs. At 20 GeV, the two modes have similar energy distributions and consequently a comparable energy resolution. However, at 70 GeV and 80 GeV we observe a difference in the energy distributions between the two modes whith a wider distribution for the binary mode and thus a better energy resolution for the multi-threshold mode.

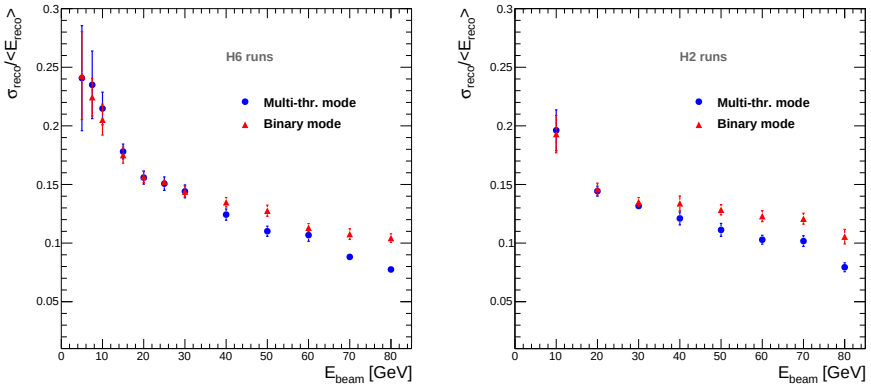

FIGURE 5.30: $\frac{\sigma_{reco}}{E_{reco}}$ of the reconstructed pion energy as a function of the beam energy in the 2012 H6 runs (left) and the 2012 H2 runs (right). For the red triangles graph, the reconstructed energy is computed using only the total number of hits (binary mode). For the blue circles graph, the reconstructed energy is computed using the three thresholds information (multi-threshold mode). For both modes, the energy is reconstructed using quadratic functions of the total number of hits.

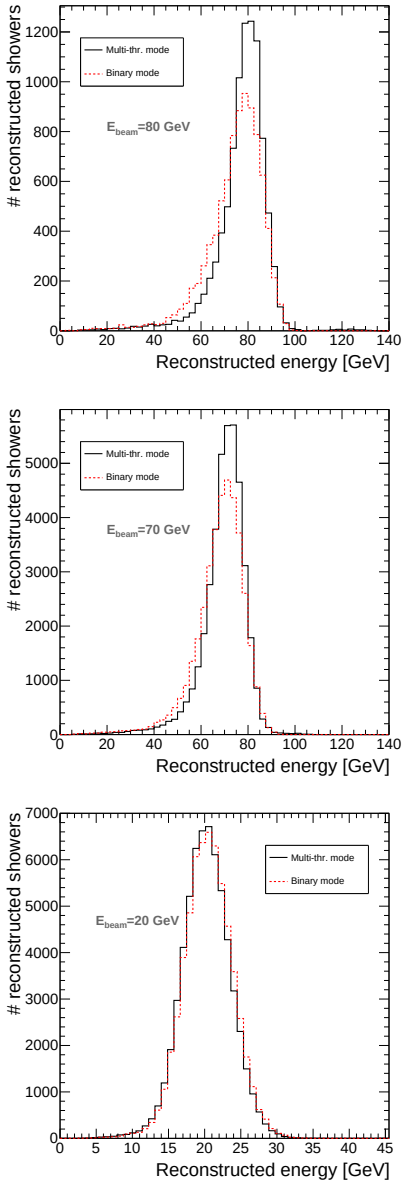

FIGURE 5.31: Distribution of the reconstructed energy with the SDHCAL binary mode (red dashed line), and with the SDHCAL multi-threshold mode (solid black line) for pions of 80 GeV (top), 70 GeV (middle) and 20 GeV (bottom) of 2012 H6 runs.

5.5.4 Systematic uncertainties

To compute the uncertainty errors represented in the linearity and energy resolution plots presented previously, contributions to both statistical and systematical uncertainties were added quadratically.
First, the statistical uncertainties arise from the Gaussian fit of the energy distributions. In addition, different sources of systematics included in our study are taken into consideration and they are the following :

– A Crystal Ball fit performed for all energy points to estimate $\overline{E}_{CB}$and σ_{CB}, was used to take into account the difference in the estimation of the mean and standard deviation of the reconstructed energy for Crystal Ball and Gaussian fits.
– Each of the selection criteria was varied by $\pm 5\%$. For each energy point the fitting procedure was applied to estimate the energy. Then the contribution of all the cuts were added quadratically to estimate ($\overline{E}_{-5\%}$, $\sigma_{-5\%}$) and ($\overline{E}_{+5\%}$,$\sigma_{+5\%}$).
– The effect of the selection criteria was also estimated using the simulation samples. The difference between the estimated values of ($\overline{E}_{sim,NoCut}$, $\sigma_{sim,NoCut}$) and ($\overline{E}_{sim,Cut}$, $\sigma_{sim,Cut}$) obtained respectively before and after the selection, applied to the simulation samples, are included.
– Finally 1.6% uncertainty on the beam energy was added [102].

The results are summarized in Tables 5.3 and 5.4. The contributions of these systematics for different energy points are given in Appendix K.1.

The mean reconstructed error is defined as :

$$\Delta^2\overline{E}_{reco} = \Delta^2\overline{E}_{stat} + (\overline{E}_{reco} - \overline{E}_{CB})^2 + (\overline{E}_{reco} - \overline{E}_{+5\%})^2 + (\overline{E}_{reco} - \overline{E}_{-5\%})^2 + (\overline{E}_{sim,Cut} - \overline{E}_{sim,NoCut})^2 \quad (5.16)$$

The error on the standard deviation is defined as :

$$\Delta^2\sigma_{reco} = \Delta^2\sigma_{stat} + \Big(\frac{\sigma_{reco}}{\overline{E}_{reco}}\Big)^2.\Delta^2\overline{E}_{stat} + (\sigma_{reco} - \sigma_{+5\%})^2 + (\sigma_{reco} - \sigma_{-5\%})^2 + (\sigma_{reco} - \sigma_{CB})^2 + (\sigma_{sim,Cut} - \sigma_{sim,NoCut})^2 \quad (5.17)$$

The error on the relative deviation is defined as :

$$\Delta^2 Error = \Big(\frac{\Delta\overline{E}_{stat}}{E_{beam}}\Big)^2 + 0.016 * 0.016\Big(\frac{\overline{E}_{reco}}{E_{beam}}\Big)^2 + \Big(\frac{\overline{E}_{reco} - \overline{E}_{CB}}{\overline{E}_{reco}}\Big)^2 + \Big(\frac{\overline{E}_{reco} - \overline{E}_{+5\%}}{\overline{E}_{reco}}\Big)^2 + \Big(\frac{\overline{E}_{reco} - \overline{E}_{-5\%}}{\overline{E}_{reco}}\Big)^2 + \Big(\frac{\overline{E}_{sim,NoCut} - \overline{E}_{sim,Cut}}{\overline{E}_{sim,Cut}}\Big)^2 \quad (5.18)$$

5.6 Energy reconstruction using Neural Network Technique

In order to further improve the energy resolution obtained within SDHCAL, the use of a neural network in pion energy reconstruction has been studied and applied for test beam data. Indeed, a first application of the neural network in energy reconstruction to simulation presented in section 4.2.4 of chapter 4, indicates this algorithm as a promising technique combining the advantages of achieving both good linearity and energy resolution in comparison to the standard weighting method. The present analysis will include a description of the neural network approach, the results of its application to the experimental data as well as the algorithm performance and its comparison with the traditional weighting method of energy reconstruction.

5.6.1 Training and Testing of the neural network

The pion data obtained with the SDHCAL technological prototype during the 2012 test beam runs has been used to test the neural network technique. Since the test beam data was available for only some specific energy points, the training of the neural network has been realised using simulation data sets where only odd energies, from 1 to 99 GeV, were used for this task.
However, in the test phase, the hadronic showers belonging to runs of the H2 and H6 data are exploited separetly to study the energy resolution and linearity that can be achieved with the neural network.
The topology of the neural network used here consists of one hidden layer with 8 neurons and one output neuron corresponding to the reconstructed energy estimated using the numbers of hits N_1, N_2 and N_3 as input informations directly fed into the network input nodes.

5.6.2 Results

Fig. 5.32 shows some energy distributions, corresponding to 2012 H6 runs, reconstructed with the neural network. The distributions of the reconstructed energies were fitted with a Gaussian fit over the full range except for 70 and 80 GeV energies where a two step Gaussian fit has been applied to account for the tail observed in the left side of their distributions as described previously.
Figures 5.34 and 5.33 show the results of the energy reconstruction of H6 and H2 test beam data with the neural network trained with the simulation pion data.
As expected from our preliminary simulation studies presented previously in chapter

3, this method of energy reconstruction using the neural network provides a significant improvement in energy resolution especially at high energies in comparison with the traditional weighting method. The energy resolution reaches the value of 6.3% at 80 GeV which is a promising result.
In addition, one can observe a good linearity on a large energy scale at the level of $\pm 5\%$.
The results using this approach are summarized in Table 5.5.

	H_6		H_2	
$E_{beam}(GeV)$	$E_{reco}(GeV)$	$\frac{\sigma_{reco}}{E_{reco}}$	$E_{reco}(GeV)$	$\frac{\sigma_{reco}}{E_{reco}}$
5	5.52 ± 0.46	0.238 ± 0.029		
7.5	7.74 ± 0.26	0.179 ± 0.006		
10	10.35 ± 0.27	0.169 ± 0.009	10.07 ± 0.32	0.195 ± 0.006
15	15.86 ± 0.28	0.155 ± 0.005		
20	20.98 ± 0.57	0.144 ± 0.008	19.25 ± 0.37	0.174 ± 0.009
25	23.92 ± 0.60	0.142 ± 0.004		
30	30.28 ± 0.87	0.136 ± 0.006	29.64 ± 0.53	0.146 ± 0.005
40	40.42 ± 1.12	0.120 ± 0.005	40.24 ± 0.28	0.131 ± 0.003
50	50.9 ± 0.48	0.114 ± 0.006	51.09 ± 0.45	0.110 ± 0.005
60	61.59 ± 1.58	0.093 ± 0.003	59.45 ± 1.04	0.088 ± 0.004
70	68.38 ± 1.10	0.092 ± 0.005	69.31 ± 0.44	0.078 ± 0.005
80	78.79 ± 1.9	0.068 ± 0.008	78.19 ± 1.15	0.063 ± 0.007

TABLE 5.5: Mean reconstructed energy E_{reco} and energy resolution $\frac{\sigma E_{reco}}{E_{reco}}$ observed and associated uncertainties obtained with neural network method for H6 and H2 beam lines.

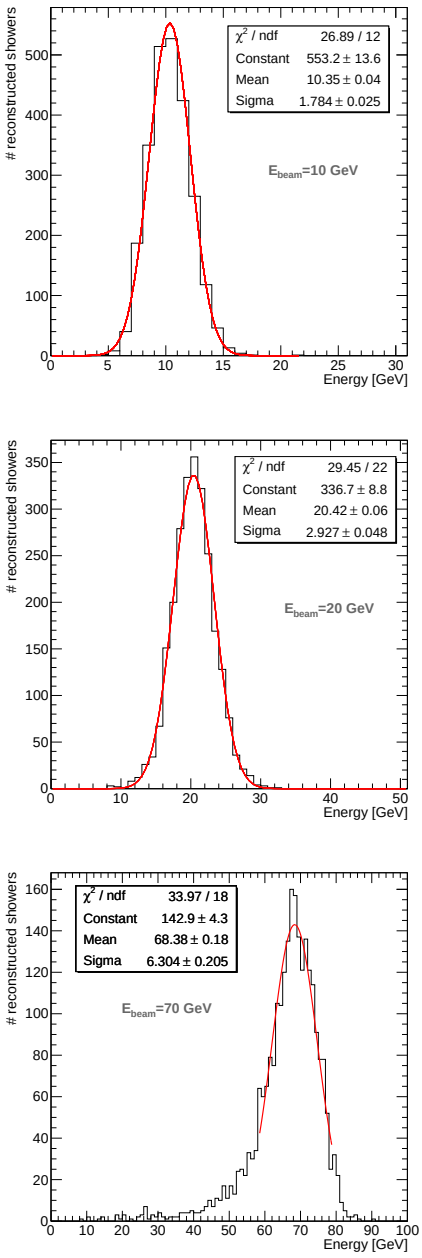

FIGURE 5.32: Distribution of the reconstructed energy performed with the neural network for pions of 10 GeV (top), 20 GeV (middle) and 70 GeV (bottom) of 2012 H6 runs.

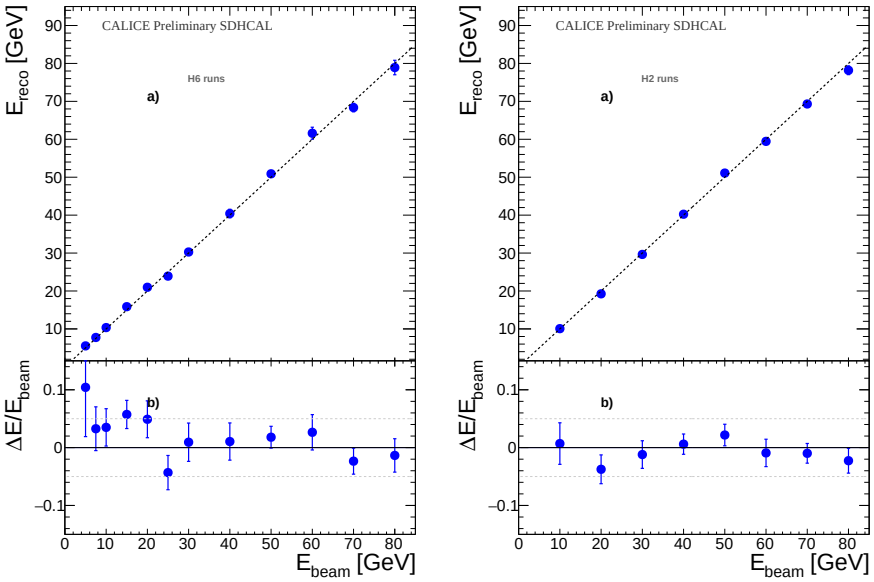

FIGURE 5.33: (a) Mean reconstructed energy for pion showers as a function of the beam energy in the 2012 H6 runs (left) and for pion showers in the 2012 H2 runs (right). Dashed line is the y=x one. (b) Relative deviation of the pion mean reconstructed energy relative to the beam energy as a function of the beam energy. The reconstructed energy is computed using ANN

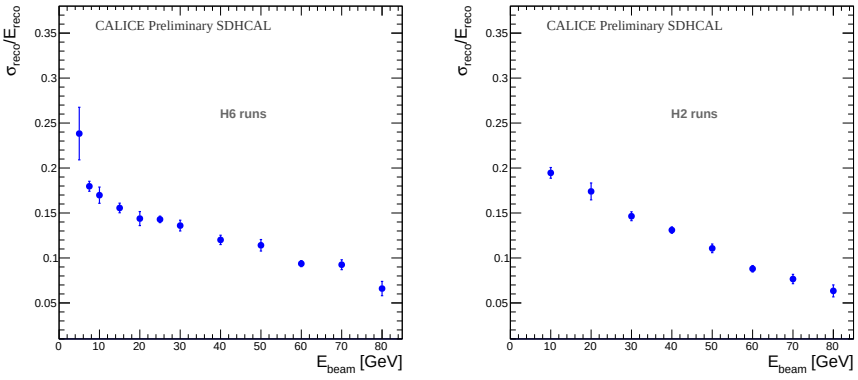

FIGURE 5.34: $\frac{\sigma_{reco}}{E_{reco}}$ of the reconstructed pion energy as a function of the beam energy in the 2012 H6 runs (left) and the 2012 H2 runs (right). The reconstructed energy is computed using ANN

5.6.3 Systematics uncertainties

The computation of the uncertainty errors is done in a similar way presented in 5.5.4, where both statistical and systematical uncertainties were added quadratically. The results are summarized in Table 5.5.

The mean reconstructed error is defined by :

$$\Delta^2\overline{E}_{reco} = \Delta^2\overline{E}_{stat} + (\overline{E}_{reco} - \overline{E}_{CB})^2 + (\overline{E}_{reco} - \overline{E}_{sim})^2 \tag{5.19}$$

The error on the standard deviation is defined by :

$$\Delta^2\sigma_{reco} = \Delta^2\sigma_{stat} + \Big(\frac{\sigma_{reco}}{\overline{E}_{reco}}\Big)^2.\Delta^2\overline{E}_{stat} + (\sigma_{reco} - \sigma_{CB})^2 + (\sigma_{reco} - \sigma_{sim})^2 \tag{5.20}$$

The error on the relative deviation is defined by :

$$\Delta^2 Error = \Big(\frac{\Delta\overline{E}_{stat}}{E_{beam}}\Big)^2 + 0.016 * 0.016\Big(\frac{\overline{E}_{reco}}{E_{beam}}\Big)^2 + \Big(\frac{\overline{E}_{reco} - \overline{E}_{CB}}{\overline{E}_{reco}}\Big)^2 + \Big(\frac{\overline{E}_{reco} - \overline{E}_{sim}}{\overline{E}_{reco}}\Big)^2 \tag{5.21}$$

5.7 Conclusion

The SDHCAL has been exposed to particle beams during different beam periods in 2012. The results obtained using the trigerless and power-pulsing mode show a good data quality. The energy reconstruction, energy resolution and linearity has been studied for both digital and semi-digital modes.

Different parametrization used to linearize the calorimeter response and convert it to energy were presented. The energy resolution is clearly improved with the multi-threshold mode for energies higher than 30 GeV. The improvement in energy resolution reaches 30% at 80 GeV and is achieved thanks to the information related to the three thresholds.

Furthermore, a neural network has been applied for reconstruction of the pion energy in SDHCAL and has shown a good impact on the energy resolution as was predicted from the simulation studies. Neural network technique provided a significant improvement of the energy resolution and linearity in comparison with the ones acquired with the weighting method. The improvement is more obvious for the high energies reaching 20% at 80 GeV.

Chapitre 6

Pion Identification in SDHCAL by using a Neural Network

Particle identification is an important stage in data analysis in High energy physics. What makes it challenging is the fact that experimental data are most often a mixture of several types of particles due to the particle beam generation process. Hence, the choice of a selection criteria is mandatory. However, while choosing a selection procedure, there are some relevant requirements to fullfil, such achieving a high efficiency for signal as well as a high background rejection.

Different methods can be used for event selection. The simplest method is to apply cuts on various physical parameters available for analysis as described and used for instance in section 5.4 of Chapter 5. However, there exist sophisticated methods, based most of the time on multivariate techniques (MVA), capable to discriminate particles with high sensitivity.

Indeed, the present analysis describes a particle classifier based on Multivariate methods using informations provided by our high granularity calorimeter thanks to its imaging capabilities. The final goal is to determine if by using MVA instead of traditional selection, the particles can be better identified and the response of SDHCAL to pions showers will be improved.

This chapter is divided in four sections. The purpose of the first part is to study different MVA methods, using only one simulated energy point (80 GeV), and find the best among them able to separate pions from background. A description of the procedure to separate electrons and muons from pions, as well as a comparison of the performance of the used MVA methods will be presented. In the second part of this chapter, further

simulated energy points have been used to investigate the best MVA method chosen in the first section. In addition, since the goal is to apply the result of this analysis to experimental data, the MVA method has been trained with a mixed background file of different energies, in order to be in a similar condition as in test beam. The third section consists in the application of the MVA method, studied previousely, to 2012 test beam data taken at CERN in order to separate electrons and muons from pions. Finally, the selected pions performed with MVA, are used to study the energy resolution and linearity in the fourth section.

6.1 The MVA methods

The aim of this study is to evaluate different MultiVariate methods in order to determine which one offers better results in Particle Identification than the traditional cut on the response of SDHCAL to particle showers as presented in section 5.4. The particles considered in the analysis were pions, electrons and muons. The MVA methods used were the Boosted Decision Tree (BDT) and Multilayer perceptron(MLP) with two learning methods : the Broyden, Fletcher, Goldfarb, Shanno (which I called BFGS) and the MLP using the Stochastic minimization (which I called here MLP), implemented in the ROOTs MVA toolkit called TMVA. More details about the MVA methods and the TMVA tools can be found in [107]. Their application allows to distinguish pion events considered as signal from muons and electrons regarded as background in the SDHCAL case. These methods will be compared among them using efficiency and contamination.

A basic concept of statistical test, formulated by a decision to accept or reject a given hypothesis, is adopted to calculate these two quantities. This can be achieved based on the distributions of the response of the neural network associated to the signal and background processes. In general, this response can be illustrated as in Fig. 6.1.
Let's assume that g($t \mid \pi$), is the p.d.f. that corresponds to the hypothesis that the particle is a pion and g($t \mid e$) corresponds to the hypothesis that the particle is an electron.
In order to select the signal process, we define a critical region for t on which depends the acceptance or the rejection of a hypothesis. By requiring $t \geq t_{cut}$, signal events are accepted and thus considered as pions, otherwise they are rejected and thus considered as electrons associated to background.
In this context, the pion efficiency $\epsilon_{\pi\to\pi}$ and the electron contamination [1] $\eta_{e\to\pi}$ are

1. The muon contamination is similarly defined.

defined, respectively, as :

$$\epsilon_{\pi\to\pi} = \frac{n_{\pi\to\pi}}{N_\pi} \qquad\qquad \eta_{e\to\pi} = \frac{n_{e\to\pi}}{N_e}$$

where $n_{\pi\to\pi}$ ($n_{e\to\pi}$) is the number of pions (electrons) classified as pions by the NN and N_π (N_e) is the total number of true pions (electrons) sent to the NN.

$n_{\pi\to\pi}$ and $n_{e\to\pi}$ are determined according to :

$$n_{\pi\to\pi} = \int_{t_{cut}}^{+\infty} g(t \mid \pi)dt \qquad n_{e\to\pi} = \int_{t_{cut}}^{+\infty} g(t \mid e)dt \tag{6.1}$$

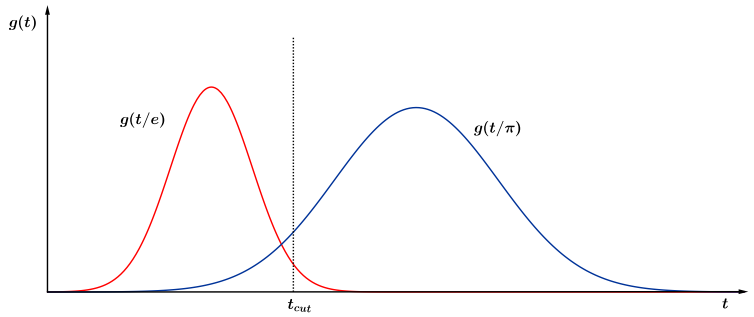

FIGURE 6.1: Probability density for the test statistic t under assumption of the hypotheses e and π. The hypothesis e is rejected if $t > t_{cut}$

The usage of TMVA consists in 2 steps : Classification and Application.
In this section we will focus on the Classification part where the MVA methods mentioned previously are trained, tested and evaluated using Monte Carlo simulation. Its Application to the experimental data will be discussed later in the next sections. The data used in this part consist in three samples generated with Monte Carlo simulation for three types of particles : pions, electrons and muons all for an energy of 80 GeV. The data sets are divided into three sets used independently for training, testing and finally application of the NN. The discriminating variables used as inputs to TMVA are shown in Fig. 6.2 where red represents background (electrons and muons) and blue represents signal(pions) and most of them were defined previously in section 5.4 :
– $\boldsymbol{N_{tot}}$: the total number of hits.

– **Start** : is the first interaction plate.
– **Last** : is the last plate with at least one fired pad.
– **Width** : is the width of the shower.
– **Zcount** : is the number of fired layers.
– **MzStart** : is the longitudinal centre of gravity calculated with respect to the shower start.
– **R** : The mean shower radius

At the end of the Classification phase, several plots are produced for each of the three methods as shown in Figs. 6.3, 6.4, 6.5. These plots show the architecture of the NN, its convergence test and the efficiencies plots as a function of the classifier response. In addition, TMVA provides a plot illustrating the separation between background and signal as a function of the classifier response. Several architectures of the neural networks have been tested and compared according to their capability to separate the background and signal distributions of the classifier response. These graphs are used to chose the appropriate cut on the classifier response for an optimal separation between background and signal. Indeed, the cut depends on the analysis needs and in our case a high pion identification efficiency($\epsilon_{\pi\to\pi} > 90\%$) combined with a low electron and muon contamination ($\eta_{e\to\pi} < 1\%$) is required.

The basic features of the three methods have been investigated in this step, and the results are shown in Fig. 6.6 and Table 6.1. Fig. 6.6 represents the ROC(Receiver Operating Characteristic) curves constructed with the testing set and are used to compare the performance of the MVA methods used in this analysis. As it can be seen all the methods show similar results where a pion identification efficiency larger than 97% is obtained with a contamination from electron and muon misidentification lower than 1%. However the MLP method have been chosen for additional studies. Table 6.1 summarizes the results, for each of the three methods, of efficiency, contamination and significance with respect to the chosen value of the cut on the classifier reponse.

Indeed, in order to apply the multivariate techniques to test beam data, the analysis was extended to include further energy points.

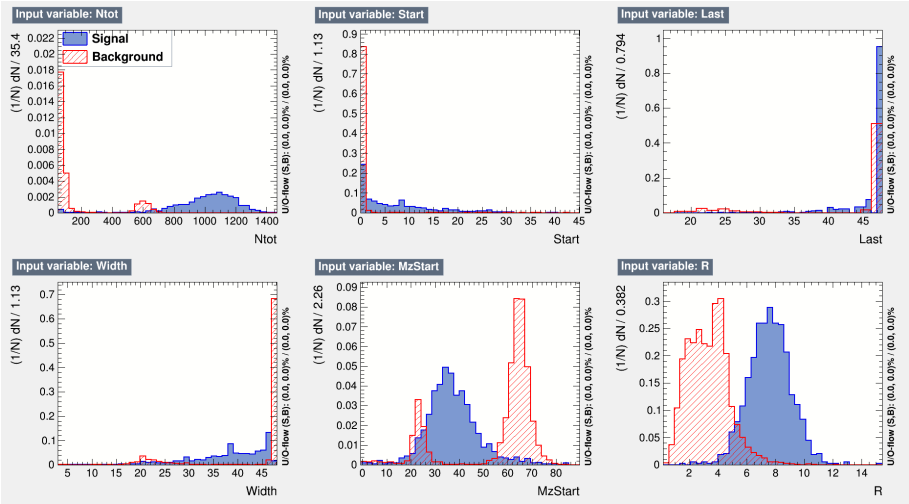

FIGURE 6.2: Discriminating variables.

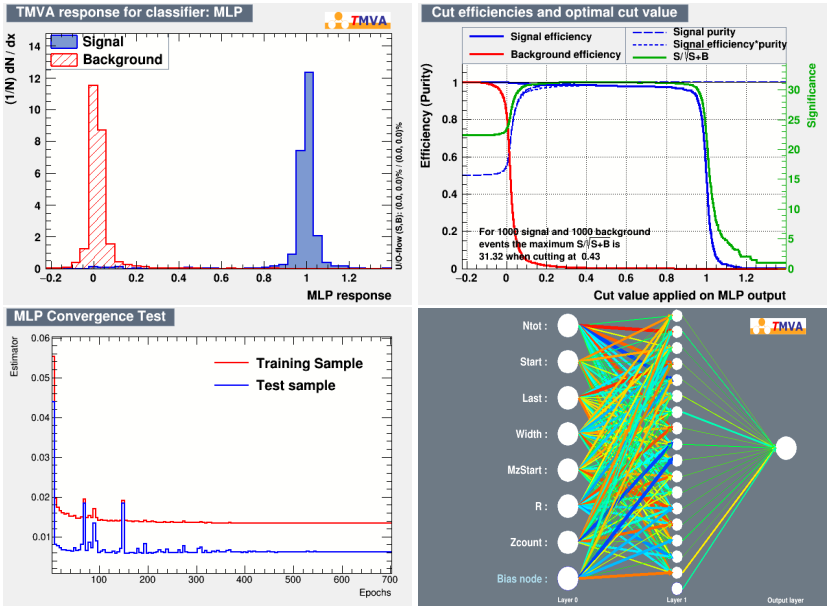

FIGURE 6.3: TMVAGui for MLP.

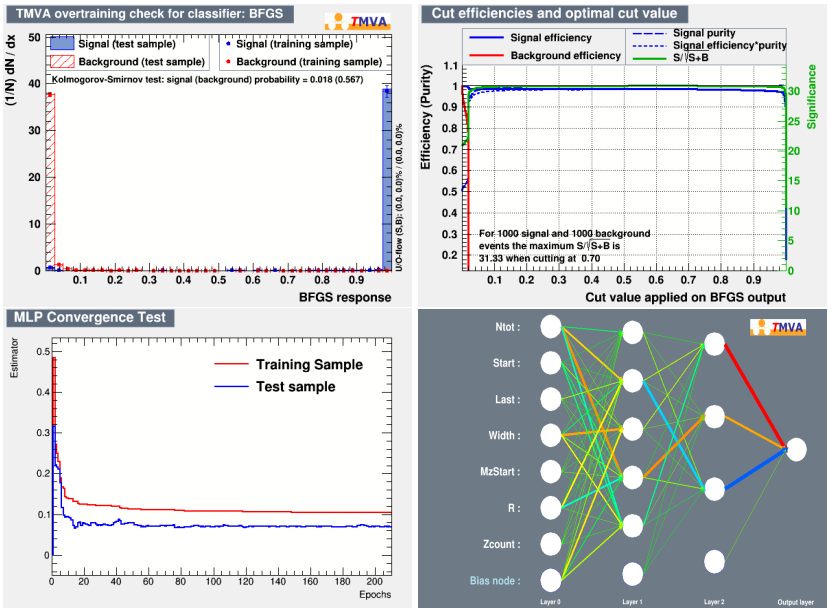

FIGURE 6.4: TMVAGui for BFGS.

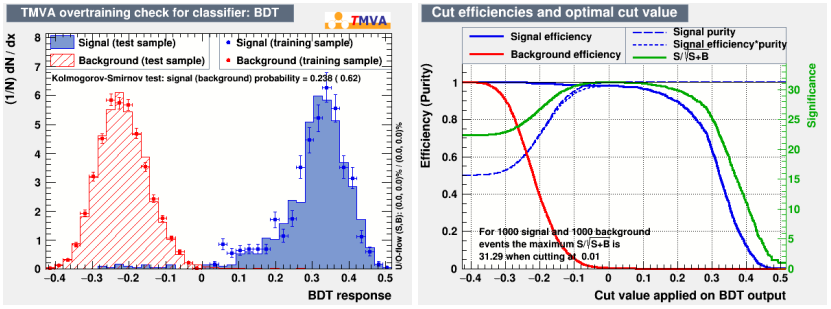

FIGURE 6.5: TMVAGui for BDT.

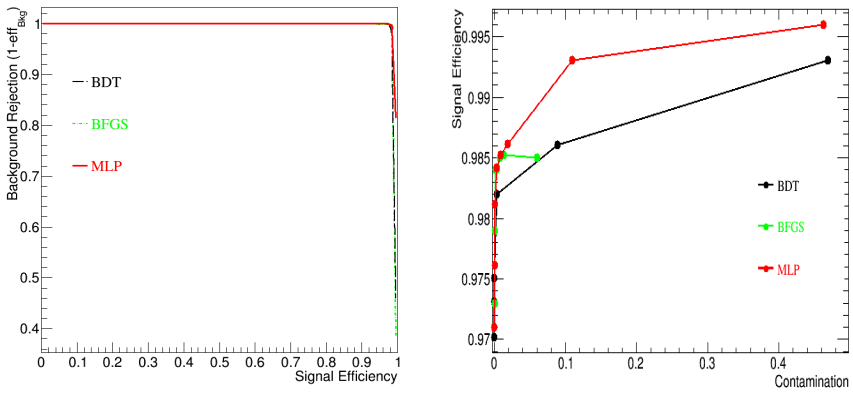

FIGURE 6.6: Left : 80 GeV Monte Carlo simulation of the background rejection versus pion efficiency for the different MVA methods. Right : 80 GeV Monte Carlo simulation of the pion efficiency as a function of the electron and muon contamination for the different MVA methods.

Method	Cut	Efficiency(%)	Contamination(%)	Significance : $S/\sqrt{S+B}$
MLP	0.87	97.1	0.008	31.03
BFGS	0.98	97.3	0.01	31.14
BDT	0.1	97	0.007	31.03

TABLE 6.1: TMVA methods comparison : The cut on the NN response, pion efficiency, electron and muon contamination and significance for simulated 80 GeV pions, electrons and muons, obtained with the different MVA methods.

6.2 Results of the Particle identification in Monte Carlo Simulation

In this section we used additional energies, obtained with simulation, for both background and signal to study the performance of the MLP method.

The goal of this analysis is to study the response of the NN when trained with a mixed background file with different energies. The idea behind this study, comes from the test beam data conditions where the data files are an admixture of differents types of particles with unknown energies.
For this task we used, successevely, two neural networks. The first one has been trained

and tested with only mixed muons as backgound and pions as signal. The second NN has been trained and tested with only mixed electrons as backgound and pions as signal. After the training and testing, the first NN will be applied to remove muons and in a second step the second NN will be used to remove the electrons.

In practice, there are two backgound files depending on the particle type : muon or electron. The first one contains 9000 muons events with energies 0.5, 2 and 10 GeV, all mixed together. The second one contains 28000 electrons events with energies of 2, 3, 5, 8, 10, 15, 20, 25, 30, 40, 50, 60, 70 and 80 GeV which are also mixed together. Each background file is then splited into two background data sets used for training and testing separately its associated NN.
The signal events correspond to pions with energies of 5, 7.5, 10, 15, 20, 25, 30, 40, 50, 60, 70 and 80 GeV.

The NNs have been trained separately for each pion energy with the mixed background file.

During the first step, since N_{tot} and $Start$ are expected to be powerfull variables for muons discrimination, they have been used as input variables fed to the first NN. The second NN, dedicated to electrons discrimination, used all the input variables defined in the previous section except N_{tot} since this last one yields a confusion between high energy electrons and low energy pions, having similar N_{tot}.

Once the NNs have been trained and tested, pion efficiencies and electron and muons contaminations have been computed for the simulated monochromatic pion beams. The results on the behaviour of the NNs responses as a function of the energy as well as the pions identification efficiency and on the muons and electrons contaminations obtained are presented in Figs. 6.7 and 6.8.
In Fig. 6.7 we show the profile of the response of the NN as a function of the energy. The NN has a good (electron/muon)-pion separation capability for the whole energy range. At 5 GeV energy, the performance is slightly worse where pions are misidentified with muons. This limitation is due to the multiple scattering undergone in this energy range by pions when interacting in the absorber.
Fig. 6.8(left) indicates that a pion identification efficiency larger than 90% can be achieved over the whole energy range with an electron contamination of about 1%. Similar results are indicated in Fig. 6.8(right) except for 5 GeV energy where the observed $\epsilon_{\pi\to\pi}$ is much lower and $\eta_{\mu\to\pi}$ is much higher. The results are sumarized in Table 6.2.

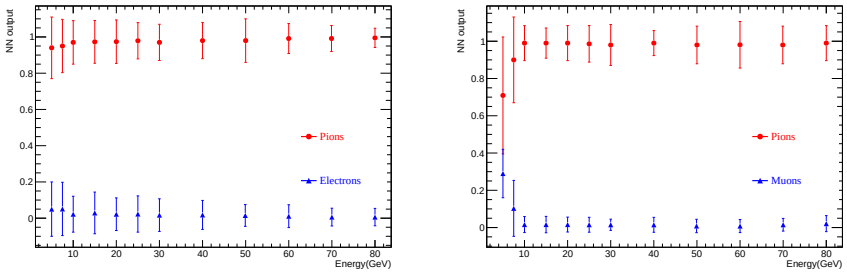

FIGURE 6.7: Monte Carlo data. NNs outputs as a function of the energy.

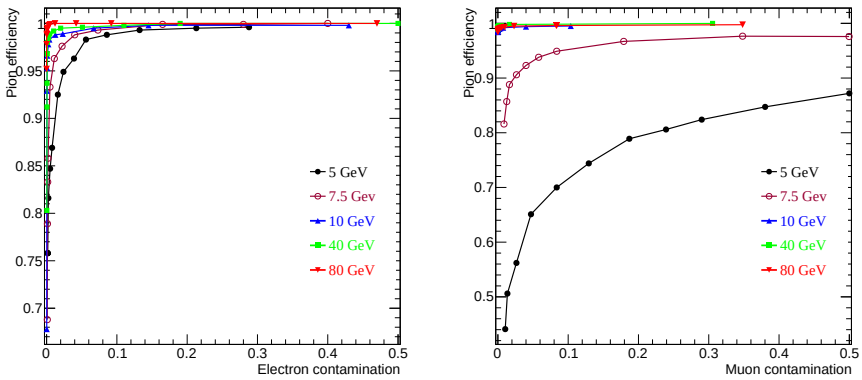

FIGURE 6.8: Monte Carlo simulation of the pion efficiency as a function of the electron contamination (left) and as a function of the muon contamination (right) for different energies.

E(GeV)	$\epsilon_{\pi\to\pi}(\%)$	$\eta_{e\to\pi}(\%)$	$\epsilon_{\pi\to\pi}(\%)$	$\eta_{\mu\to\pi}(\%)$
5	92.5 ± 1.6	1.6 ± 0.8	56.2 ± 3.1	2.6 ± 1.0
7.5	96.3 ± 1.2	1.1 ± 0.7	88.8 ± 1.9	1.6 ± 0.8
10	96.6 ± 1.1	0.1 ± 0.2	98.9 ± 0.6	0.02 ± 0.09
40	98.0 ± 0.9	0.1 ± 0.2	99.5 ± 0.5	0.04 ± 0.13
80	99.3 ± 0.5	0.1 ± 0.1	98.8 ± 0.7	0.02 ± 0.09

TABLE 6.2: Pion efficiency, electron contamination and muon contamination for simulated data using the two NNs as described in the text. The cuts on the outputs NNs are fixed to 0.7.

6.3 Application of the Particle identification in Data

As stated before, in order to train the NN under conditions similar to the test-beam, we added all the pions files of the different energies in only one mixed signal file.
This is the only difference with the training and testing conditions of the NNs used in the previous section.

In this way, we used two different NNs as previously. The first NN is trained and tested with 9000 mixed muons events of different energies as background and 24000 mixed pions events of different energies as signal.
The second NN is trained and tested with 28000 mixed electrons events of different energies as background and 24000 mixed pions events of different energies as signal.
The results of these two steps are then used in the application phase for testbeam data.

The input variables are used in the same way as described in the previous section. Fig. 6.9 shows the comparison of the discriminating variables for both electrons (blue) and pions (red). The distributions of the variables are ploted for mixed electrons of different energies in the same time (2, 3, 5, 8, 10, 15, 20, 25, 30, 40, 50, 60, 70 and 80) and for mixed pions of different energies (5, 7.5, 10, 15, 20, 25, 30, 40, 50, 60, 70 and 80 GeV).

The results on the pions identification efficiency and on the muons and electrons contaminations obtained with the simulated events are presented in Fig. 6.10. We can see that when the NNs are trained with mixed background and signal files, a pion identification efficiency of 98% can be achieved over the whole energy range with an electron and muon contaminations of about 1%.
The statistical errors shown in Fig. 6.10 has been computed at 95% C.L, assuming a binomial distribution [108].

Indeed, the purpose from this analysis is to exploit the MVA technique to better identify the pion events in experimental data which is contaminated with muons and electrons. In this context, the trained and tested NNs are applied to 2012 test beam data taken at H2 and H6 lines at CERN. The obtained results are shown in Fig. 6.11 for some energies of H6 runs where the total number of hits of the collected events is drawn before the selection (black) and after the cut performed with the neural network. The selected events with the NN technique are exploited for a study of the energy resolution and linearity of SDHCAL. This study will be presented in the next section.

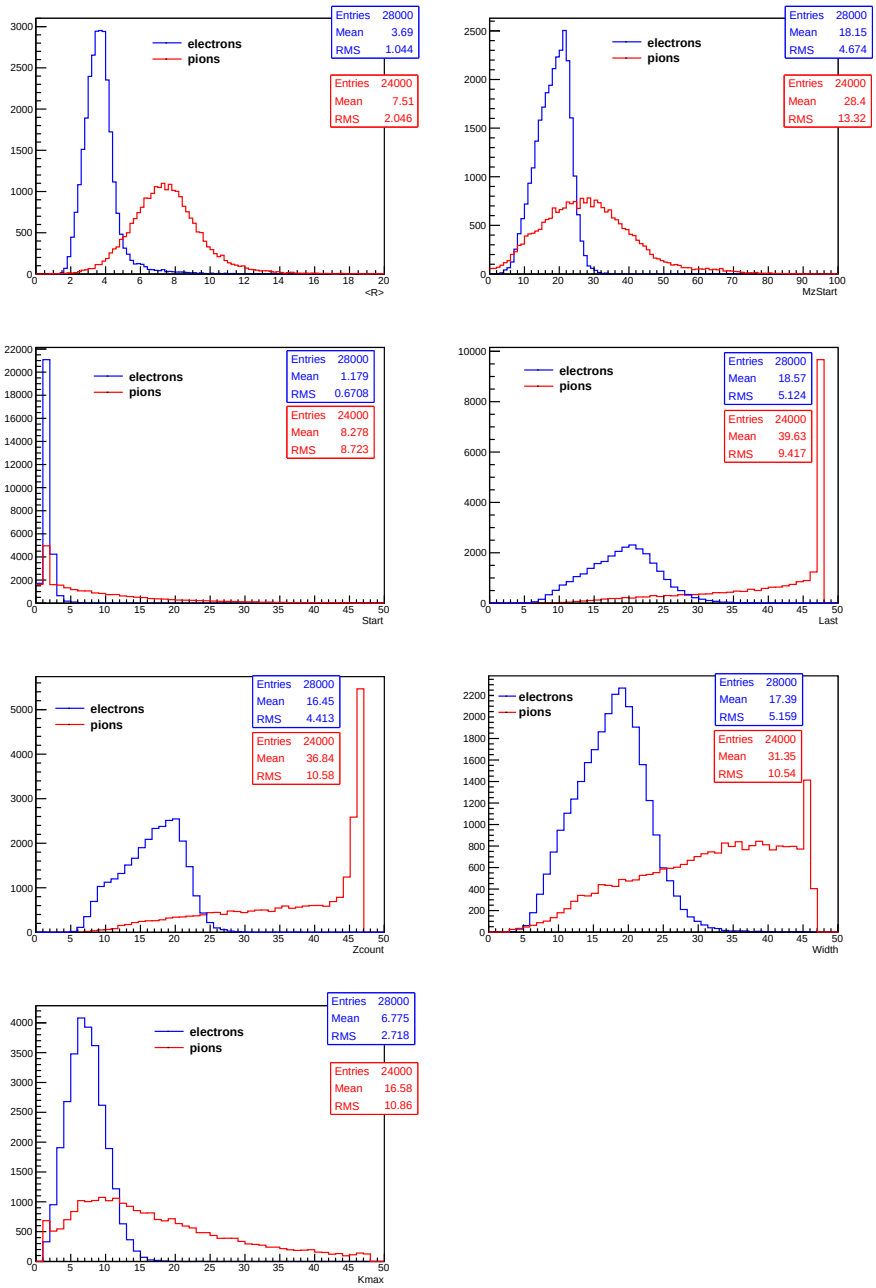

FIGURE 6.9: Monte Carlo simulation : Comparison of the discriminating variables used as inputs to the NN for both electrons (blue) and pions (red) of different energies mixed together as described in paragraph 6.3 .

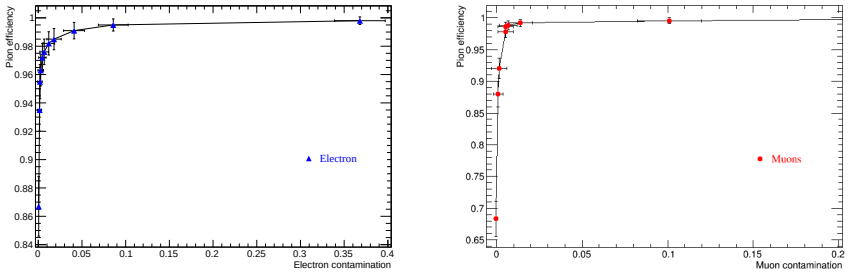

FIGURE 6.10: Monte Carlo simulation of the pion efficiency as a function of the electron contamination (left) and as a function of the muon contamination (right).

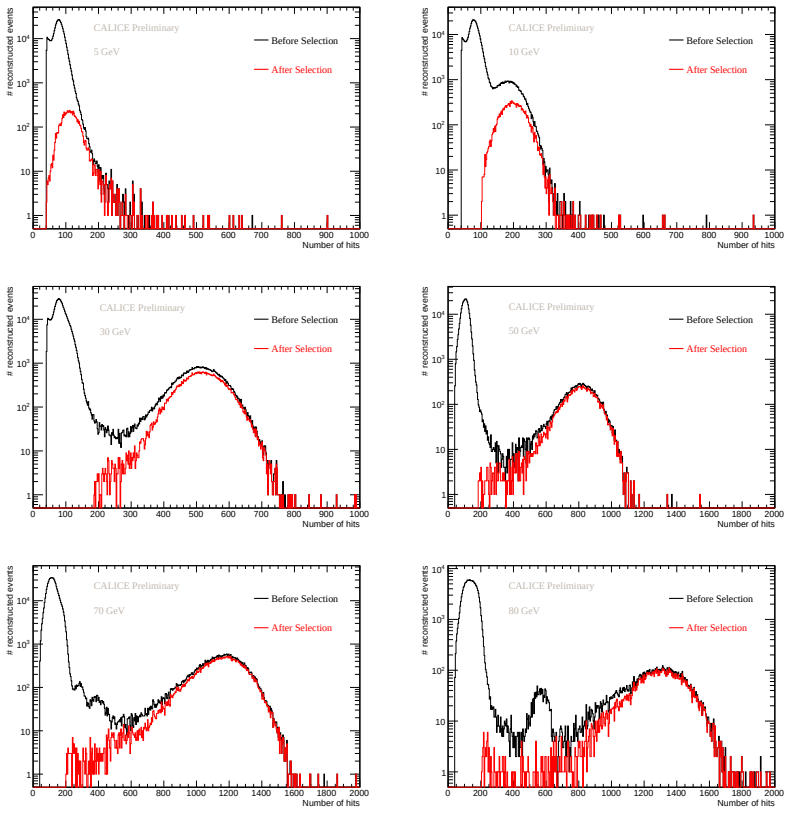

FIGURE 6.11: Number of hits for 5, 10, 30, 50, 70 and 80 GeV pion runs before (black line) and after (red line) NN selection.

6.4 Results of the energy reconstruction using the pion identification with TMVA

The energy reconstruction using a quadratic parametrization of the total number of hits has been realized for the multi-thresholds mode of SDHCAL using the event selection performed with MVA.
The procedure for the energy reconstruction is the same as the one described in 5.5.3 of Chapter 5.

Fig. 6.13(a) shows the mean reconstructed energy of pions as a function of the beam energy obtained with the multi-threshold mode using the event selection performed with TMVA. One can observe a good linearity on a large energy scale with a relative deviation of the reconstructed energy with respect to the beam energy below 5% for all the energy points, except 5 GeV, as shown in Fig. 6.13(b). The energy resolution is shown in Fig. 6.12 and presents a similar behaviour to the results obtained with the classic event selection.
The obtained results are summarized in Table 6.3.

The computation of the uncertainty errors is done in a similar way presented in 5.5.4, where both statistical and systematical uncertainties were added quadratically.
The statistical uncertainties generated by the Gaussian fit of the energy distributions. In addition, different sources of systematics included in our study are taken into consideration and they are the following :

– A Crystal Ball fit performed for all energy points, was used to take into account the difference in the estimation of the mean and standard deviation of the reconstructed energy for Crystal Ball and Gaussian fits.
– The neural network cut was varried by $\pm 5\%$ to take into account its effect. For each energy point the fitting procedure was applied to estimate the energy and then ($\overline{E}_{-5\%}$, $\sigma_{-5\%}$) and ($\overline{E}_{+5\%}$,$\sigma_{+5\%}$).
– The effect of the neural network cut was also estimated using the simulation samples. The difference between the estimated values of ($\overline{E}_{sim,NoCut}$, $\sigma_{sim,NoCut}$) and ($\overline{E}_{sim,Cut}$, $\sigma_{sim,Cut}$) obtained respectively before and after the selection performed with the neural network, applied to the simulation samples, are included.
– Finally 1.6% uncertainty on the beam energy was added.

H_6			H_2	
$E_{beam}(GeV)$	$E_{reco}(GeV)$	$\frac{\sigma_{reco}}{E_{reco}}$	$E_{reco}(GeV)$	$\frac{\sigma_{reco}}{E_{reco}}$
5	5.49 ± 0.49	0.254 ± 0.017		
7.5	7.57 ± 0.36	0.215 ± 0.009		
10	10.09 ± 0.22	0.205 ± 0.014	9.71 ± 0.21	0.239 ± 0.013
15	14.67 ± 0.14	0.167 ± 0.010		
20	19.79 ± 0.70	0.152 ± 0.007	19.50 ± 0.29	0.144 ± 0.004
25	24.26 ± 0.52	0.139 ± 0.007		
30	30.61 ± 0.30	0.128 ± 0.007	30.26 ± 0.10	0.128 ± 0.005
40	40.96 ± 0.47	0.117 ± 0.004	40.67 ± 0.12	0.113 ± 0.002
50	50.98 ± 0.96	0.106 ± 0.005	51.33 ± 0.17	0.102 ± 0.004
60	60.33 ± 0.17	0.106 ± 0.004	61.26 ± 0.21	0.091 ± 0.003
70	71.99 ± 0.10	0.088 ± 0.004	68.66 ± 0.28	0.087 ± 0.004
80	80.64 ± 0.16	0.073 ± 0.003	78.71 ± 0.35	0.074 ± 0.004

TABLE 6.3: Mean reconstructed energy E_{reco} and energy resolution $\frac{\sigma E_{reco}}{E_{reco}}$ observed and associated uncertainties for multi-threshold mode for H6 and H2 beam lines. The reconstructed energy is computed using the event selection performed with TMVA.

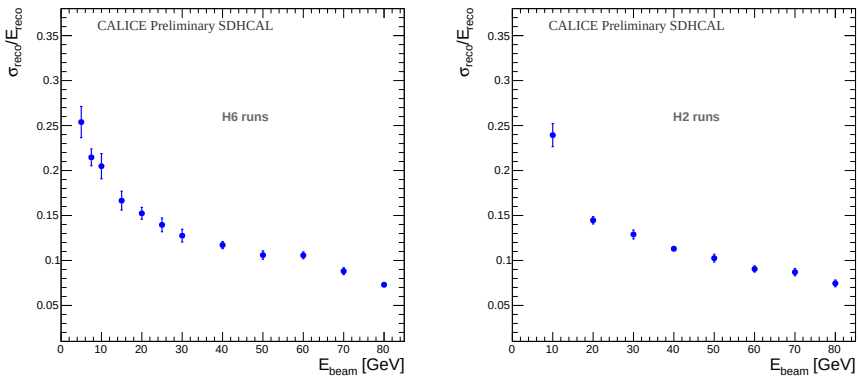

FIGURE 6.12: $\frac{\sigma_{reco}}{E_{reco}}$ of the reconstructed pion energy as a function of the beam energy in the 2012 H6 runs (left) and the 2012 H6 runs (right) and for pion showers in the 2012 H2 runs (right). The reconstructed energy is computed using the event selection performed with TMVA.

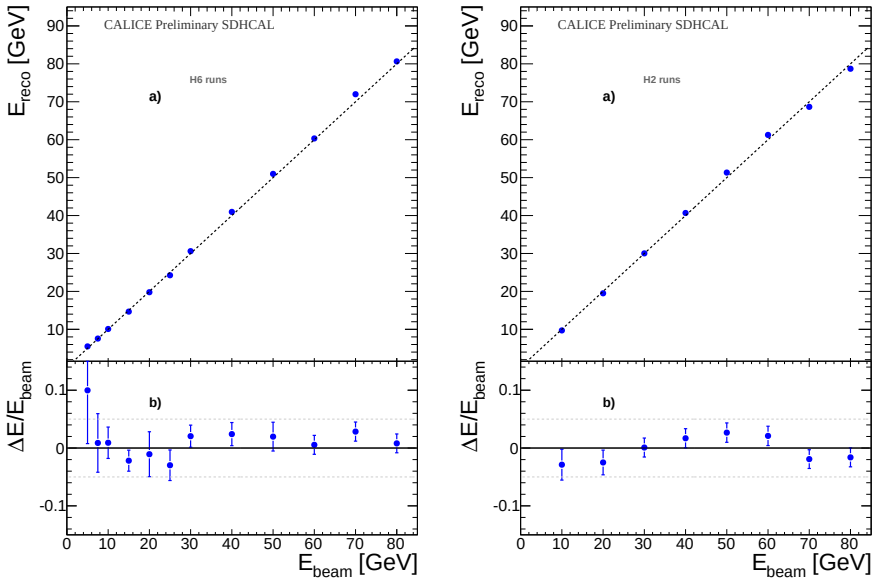

FIGURE 6.13: (a) Mean reconstructed energy for pion showers as a function of the beam energy in the 2012 H6 runs (left) and for pion showers in the 2012 H2 runs (right). Dashed line is the y=x one. (b) Relative deviation of the pion mean reconstructed energy relative to the beam energy as a function of the beam energy.The reconstructed energy is computed using the event selection performed with TMVA.

6.5 Conclusion

A particle classifier based on Multivariate technique and using informations provided by our high granularity calorimeter is developed and presented in this work.
A comparison of different MVA methods has been presented and shows similar results with a slight improvement in the performance presented by the NN method using MLP.
The performance of MLP method is further invesitigated using Monte Carlo simulated data and show promising results with a high pion efficiency ($\epsilon_{\pi\to\pi} > 90\%$) while keeping the muon and electron contamination very low ($\eta_{e\to\pi} < 1\%$ and $\eta_{\mu\to\pi} < 1\%$) for most of the energy points, although some deterioration has been observed at 5 GeV.
The MVA has been applied to 2012 CERN test beam data and the resulted selection of pion events is exploited to study the hadronic energy resolution and linearity of the response of SDHCAL. A good linearity has been observed in the results. However, no improvement has been observed in energy resolution compared to the

results achieved in chapter 5, but the results obtained using TMVA are in agreement with the ones obtained with the multi threshold mode where the event selection was performed with the classic cuts.

Conclusion

This thesis presents studies to improve the energy reconstruction and resolution of hadronic showers within a Semi-Digital Hadronic CALorimeter (SDHCAL) proposed for the future electron-positron collider ILC.
Today ILC is the most advanced project to succeed LHC in the future program of elementary particle physics. ILC is expected to extend our understanding of particle physics beyond the information delivered by the LHC, while providing high precision measurements. To reach such unprecedented precision, ILC places a premium on the innovative technologies of its detectors and novel reconstruction techniques like the Particle Flow Algorithm (PFA), which has been a guide for the calorimeter R&D.
The CALICE collaboration (CAlorimetry for LInear Collider Experiments), has developped several prototypes of electromagnetic and hadronic calorimeters of different technologies to be used in ILC. SDHCAL as one of those prototypes for a hadronic calorimeter option is outlined in this thesis. It is a highly granular and sampling calorimeter with the challenge to cope with millions of electronic channels needed to read out such a high-granularity. The prototype is the first of a series of new-generation detectors equipped with a power-pulsing mode intended to reduce the power consumption. The SDHCAL has been exposed to particle beams since 2012 at CERN (European Council for Nuclear Research).

The results obtained using the trigerless and power-pulsing mode have shown stability and a good data quality. SDHCAL has proved that a high granularity, compact and homogenous calorimeter with PFA capability can be acheived using cheap and simple but robust and efficient Glass Resistive Plate Chamber (GRPC) detectors.

One of the main analysis of this thesis is the development of energy reconstruction techniques used to improve the energy resolution of hadronic showers. An exhaustive study on the optimization of the energy reconstruction using Monte Carlo

simulation was presented and detailed in chapter 4. A comparison between the binary and multi-threshold modes within SDHCAL have been discussed. The multi-threshold capabilities of SDHCAL at high energies have been demonstrated and confirm the important impact on energy resolution of a semi-digital readout. The improvement in energy resolution at high energies is related to a better treatment of the saturation effect thanks to the information provided by the second and third thresholds. Different analytic methods were developed for the optimization of the energy reconstruction within the multi-threshold mode of SDHCAL. A further approach of energy reconstruction based on Neural Network was presented. The Neural Network technique led to the largest improvement in energy resolution at high energies.
The study of energy resolution presented in this work corresponds to the single hadron energy resolution which is different from the jet energy resolution. The jet energy resolution depends on the single hadron energy resolution and $\sigma_{Jet} > \sigma_{HCAL}$.

A similar analysis investigating the energy resolution of pion showers recorded in SDHCAL during 2012 beam tests at CERN, has been presented. The energy reconstruction, energy resolution and linearity has been studied for both digital and semi-digital modes. Different parametrization used to linearize the calorimeter response and convert it to energy were presented. The energy resolution is clearly improved with the multi-threshold mode for energies higher than 30 GeV. The improvement in energy resolution reaches 30% at 80 GeV thanks to the information related to the three thresholds. Furthermore, a neural network has been applied for reconstruction of the pion energy in SDHCAL and has shown a good impact on the energy resolution as was predicted from the simulation studies. Neural network technique provided a significant improvement of the energy resolution and linearity in comparison with the ones acquired with the weighting method. The improvement is more obvious for the high energies.

Finally, a particle classifier based on Multivariate technique and using informations provided by our high granularity calorimeter is developped and presented in this work. The performance of MLP method is further invesitigated using Monte Carlo simulated data and shows promising results with a high pion efficiency ($\epsilon_{\pi\to\pi} > 90\%$) while keeping the muon and electron contamination very low ($\eta_{e\to\pi} < 1\%$ and $\eta_{\mu\to\pi} < 1\%$) for most of the energy points, although some deterioration has been observed at 5 GeV. The MVA has been applied to 2012 CERN beam test data and the resulted selection of pion events is exploited to study the hadronic energy resolution and linearity of the response of SDHCAL. A good linearity has been observed in the results. However, no improvement has been observed in energy resolution compared to the results achieved in chapter 5, but the results obtained

using TMVA are in agreement with the ones obtained with the multi threshold mode where the event selection was performed with the classic cuts.

Outlook The SDHCAL studies are continuing. The results presented in this thesis with beam tests data are obtained without corrections. The application of different types of correction, such as gain correction or correction based on efficiency and multiplicity information, will lead to an improvement in energy resolution.

Currently, SDHCAL data taken at different beam tests campaigns at CERN are being analyzed. This will allow several interesting studies, such as the study of the response of the prototype to low energy and to hadrons with different inclination angles. A special interest is given to the study of the optimization of the threshold values but also for the study of the topology of the hadronic showers in order to achieve a realistic simulation of the SDHCAL.

Additionally, the SDHCAL studies will be complemented by the studies with the other CALICE calorimeters. Indeed a combined beam test with SiWECAL has been realized in October 2016 and will allow to test the performance in the presence of both electromagnetic and hadronic calorimeters.

In addition, Efforts are deployed to improve the SDHCAL concept by the development of new and more robust mechanical structure and the construction of larger GRPCs equiped with sophisticated electronics readout, in order to validate their ability for the final detector.

Currently, new R&D are ongoing to achieve high rate capability and timing precision measurement that could be of big interest for other colliders(CLIC, LHC, CEPC,FCC...).

Finally, the SDHCAL technology can be adopted in different fields of application such as in the medical imaging technique of positron emission tomography (PET) and single-photon emission computed tomography (SPECT) and is used in muography to perform density imaging using atmospheric muons of the Puy de Dôme, a volcano 2 km wide close to Clermont-Ferrand in France (TOMUVOL project).

Annexe A

More simulation studies

A.1 Effect of the absorber material

The response of the SDHCAL to pions is studied for two different absorber materials, steel and tungsten, having both a thikness of 2 cm. The results obtained are shown in Figs A.1 and A.2. The steel absorber gives a better linearity and energy resolution.

A.2 Effect of the pad size

The effect of the granularity on the energy resolution and the linearity of SDHCAL has been studied. We compared between different pad sizes having respectively $[1 \times 1]cm$, $[3 \times 3]cm$ and $[4 \times 4]cm$. Figs A.3 and A.4 show the effect of the pad size on the linearity and the energy resolution. We deduce that the fine granularity has a crucial effect on the linearity and the energy resolution.

A.3 Validation of Geant4 Models

These simulations were realized using version 4.9.3 of GEANT4. We aimed to test different GEANT4 physics models[81]. The results are shown in Figs A.5 and A.6.

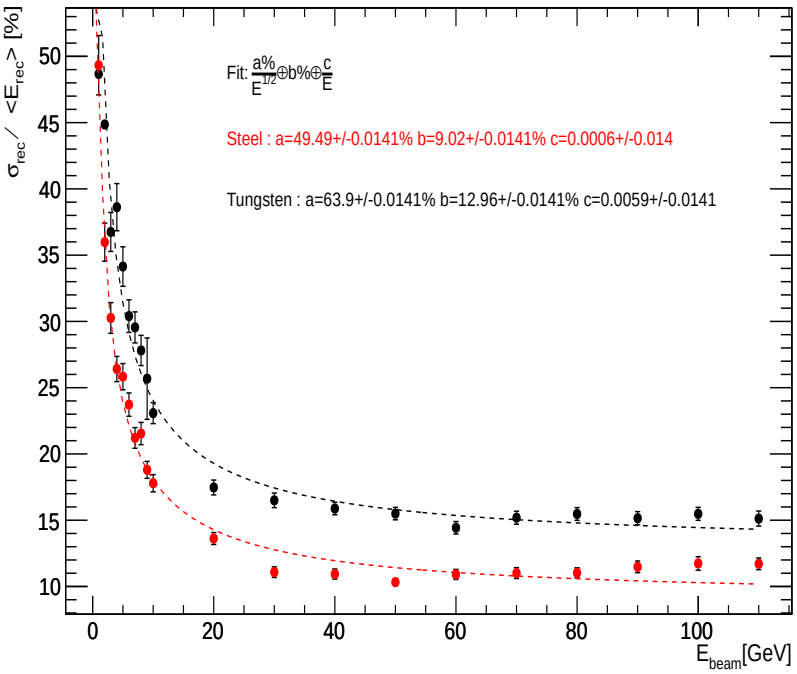

FIGURE A.1: Pion energy resolution as a function of the beam energy obtained for two absorber materials : Steel(red circles graph) and Tungsten(black circles graph).

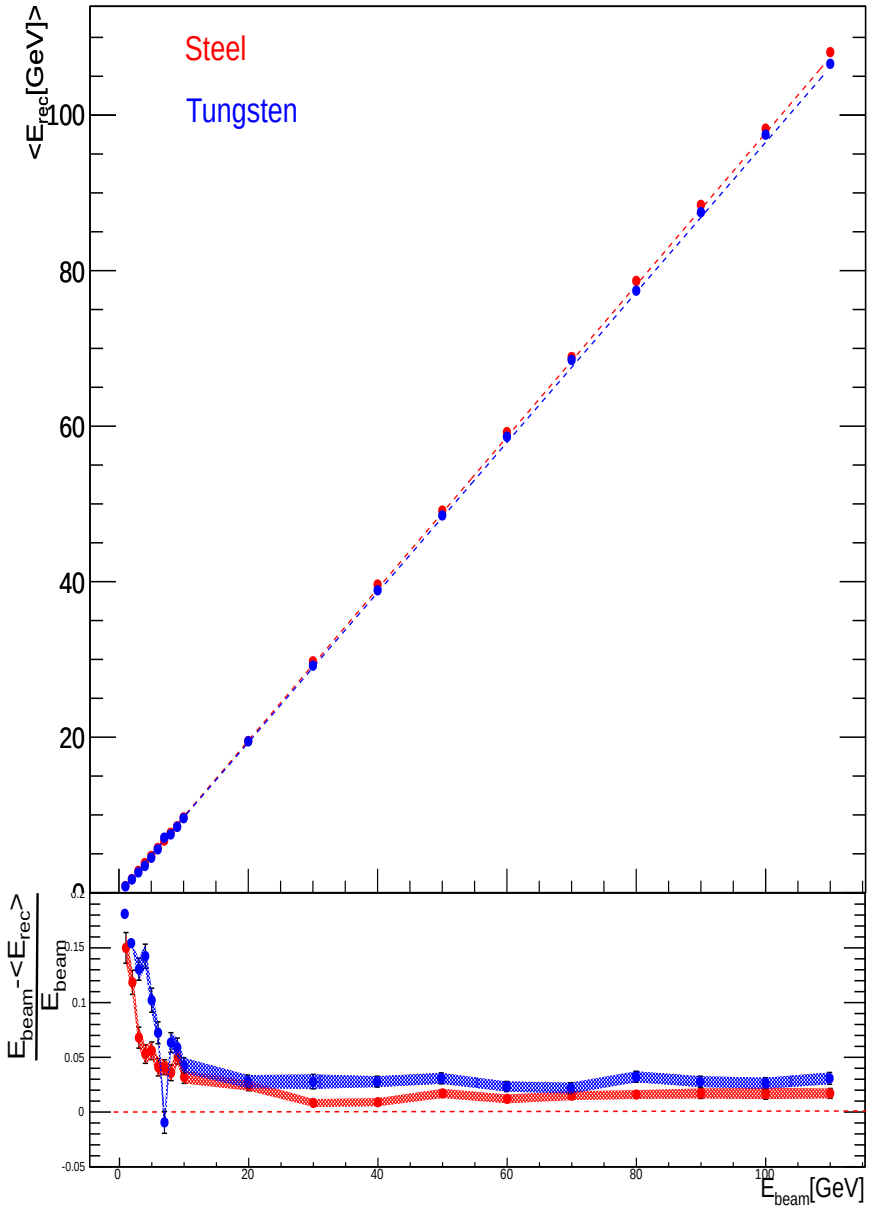

FIGURE A.2: Mean reconstructed energy as a function of the beam energy for two Absorber materials :Steel(red circles graph) and Tungsten(blue circles graph).

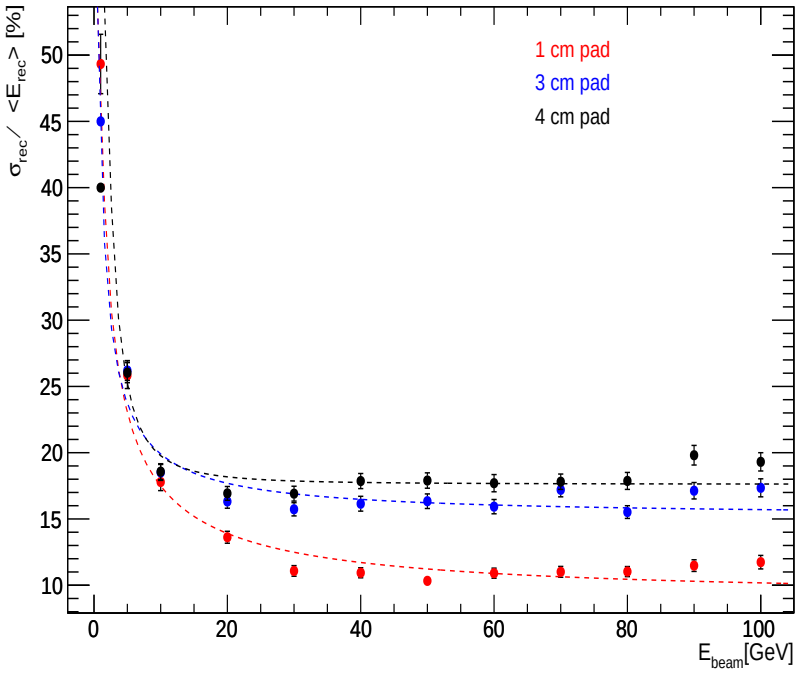

FIGURE A.3: Pion energy resolution as a function of the beam energy obtained for different pad sizes : $1cm$(red circles graph), $3cm$(blue circles graph) and $4cm$(black circles).

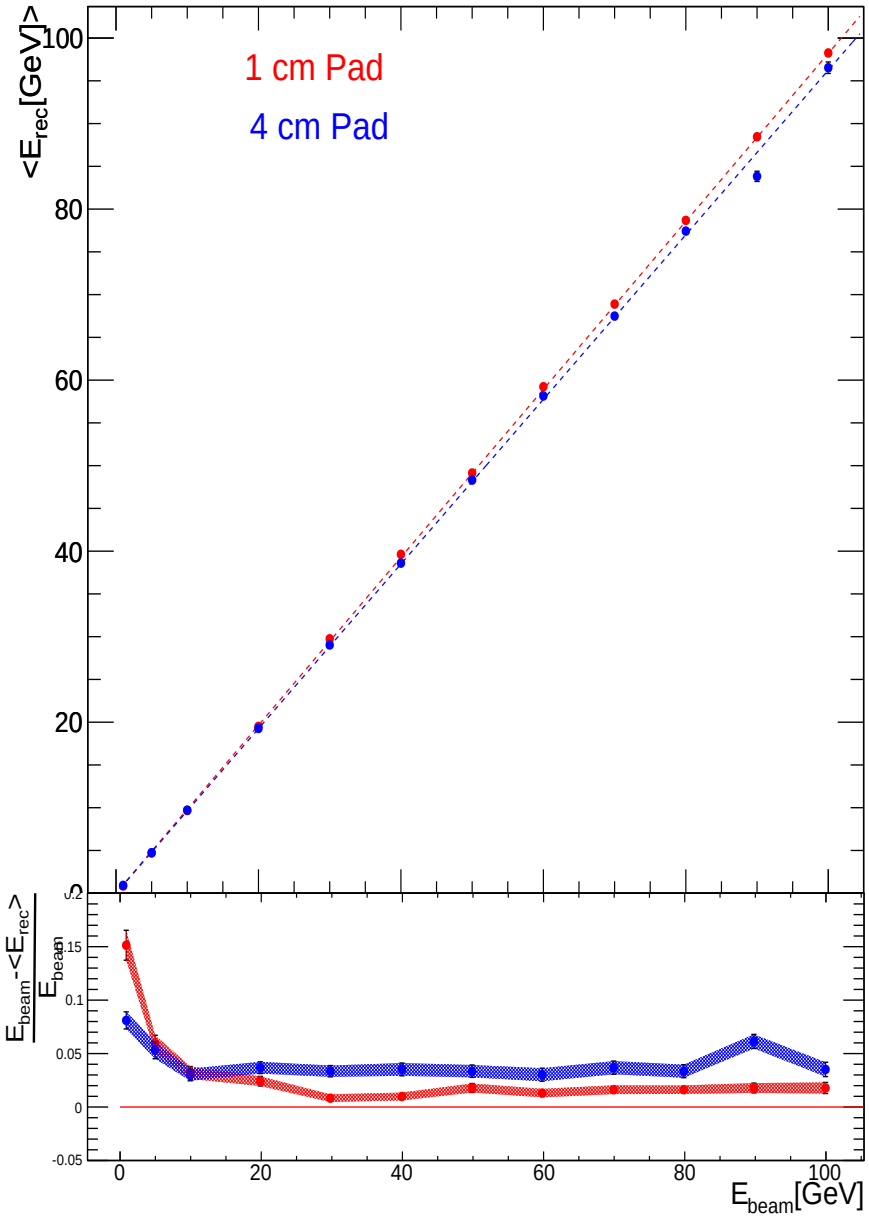

FIGURE A.4: Mean reconstructed energy as a function of the beam energy for different pad sizes : $1cm$(red circles graph) and $4cm$(black circles).

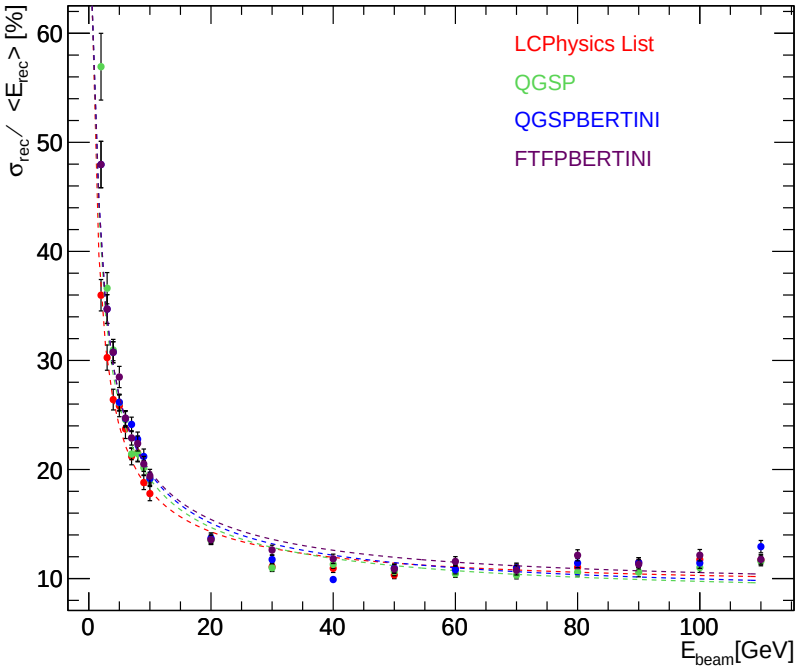

FIGURE A.5: Pion energy resolution as a function of the beam energy obtained for different GEANT4 physics lists.

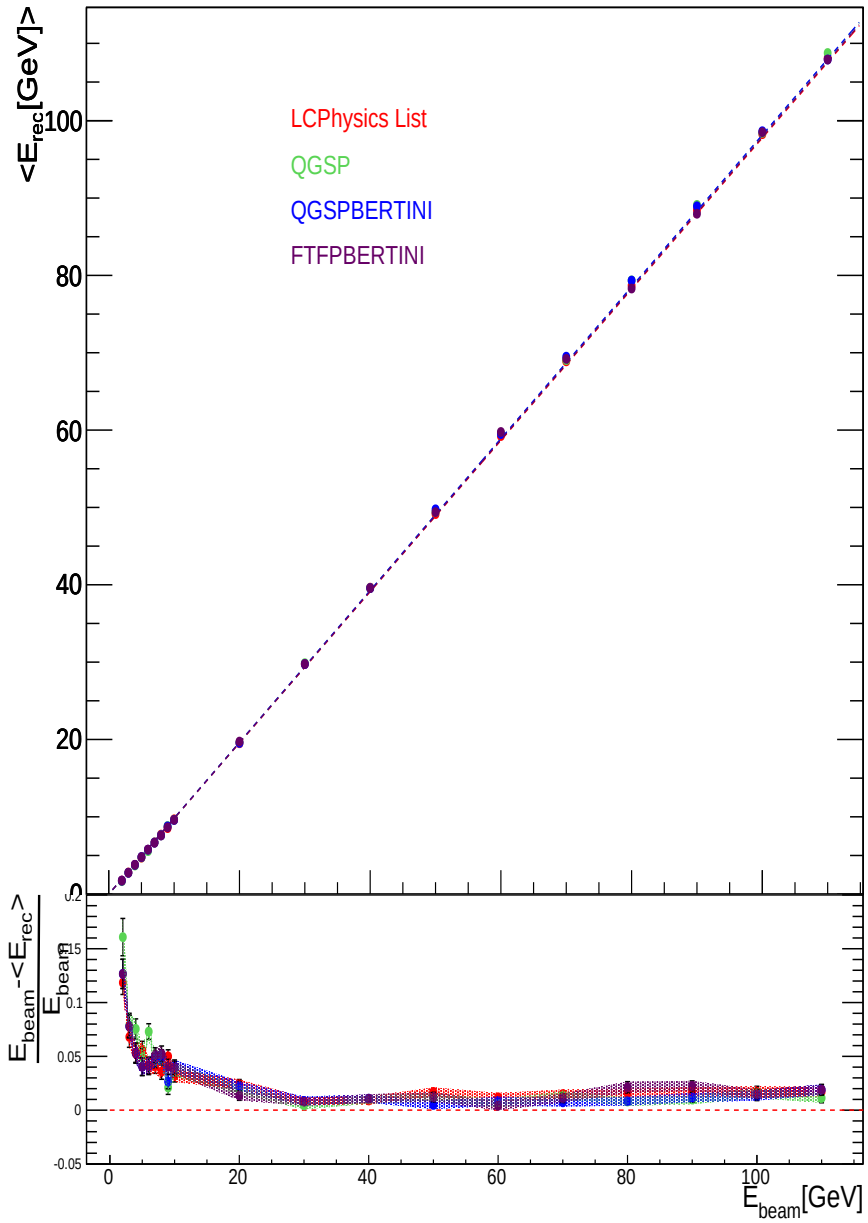

FIGURE A.6: Mean reconstructed energy as a function of the beam energy for different GEANT4 physics lists.

Annexe B

Energy dependent weights

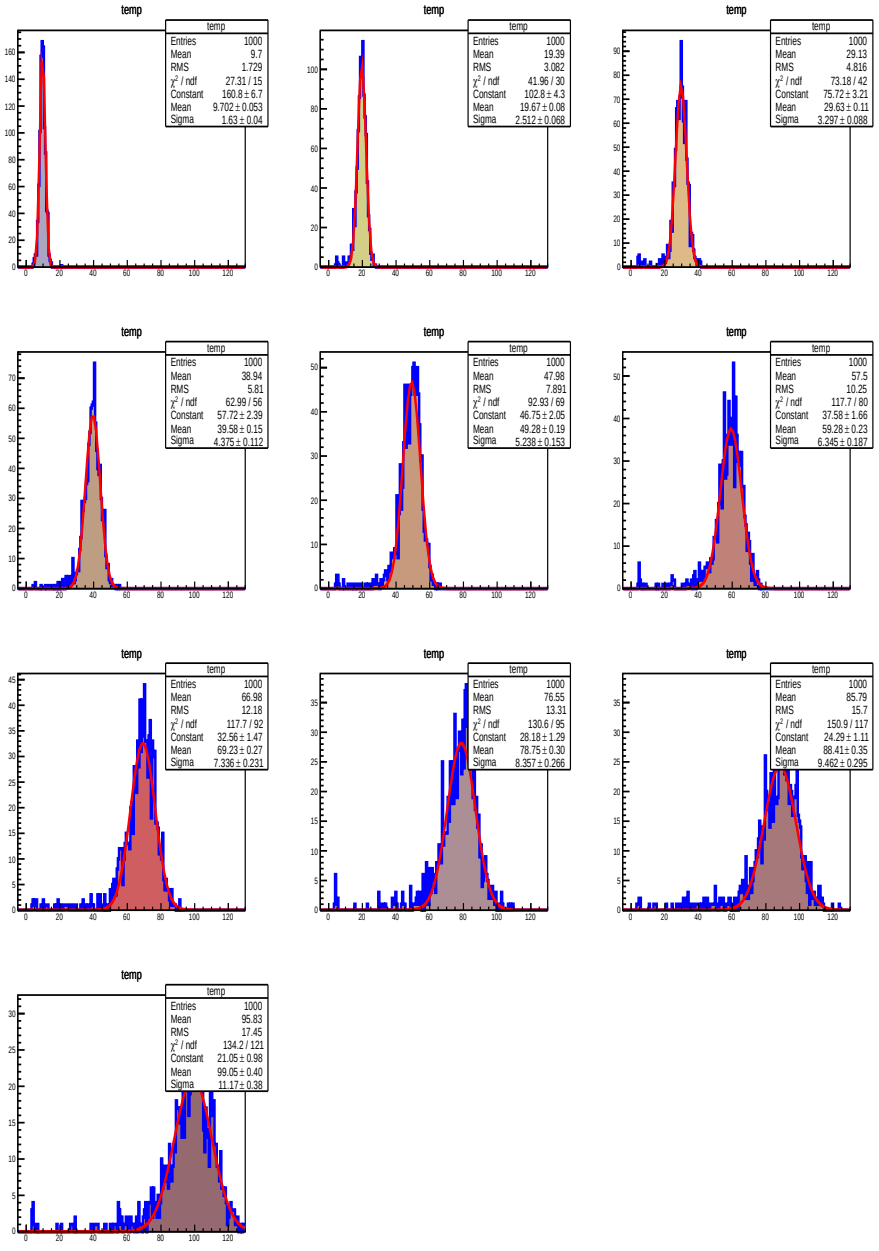

FIGURE B.1: Reconstructed energy distributions obtained with energy dependent weights method, using TMinuit minimization, discussed in section 4.2.3.1(10,20,30,40,50,60,70,80,90,100GeV).

Annexe C

Free Energy dependent constant Calibration by TMinuit

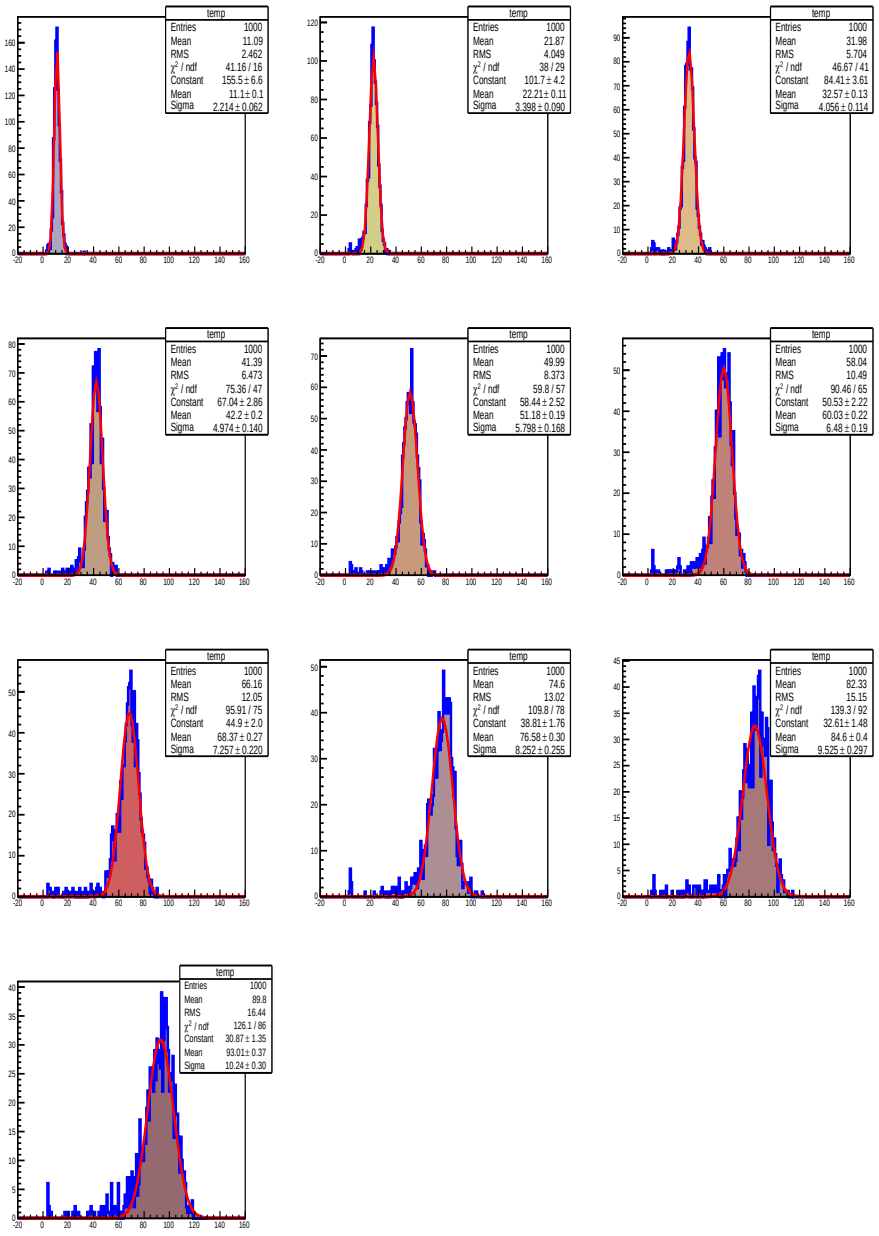

FIGURE C.1: Reconstructed energy distributions using TMinuit with the constant weights method discussed in section4.2.3.2(10,20,30,40,50,60,70,80,90,100GeV).

Annexe D

Calibration weights with exponential parametrization

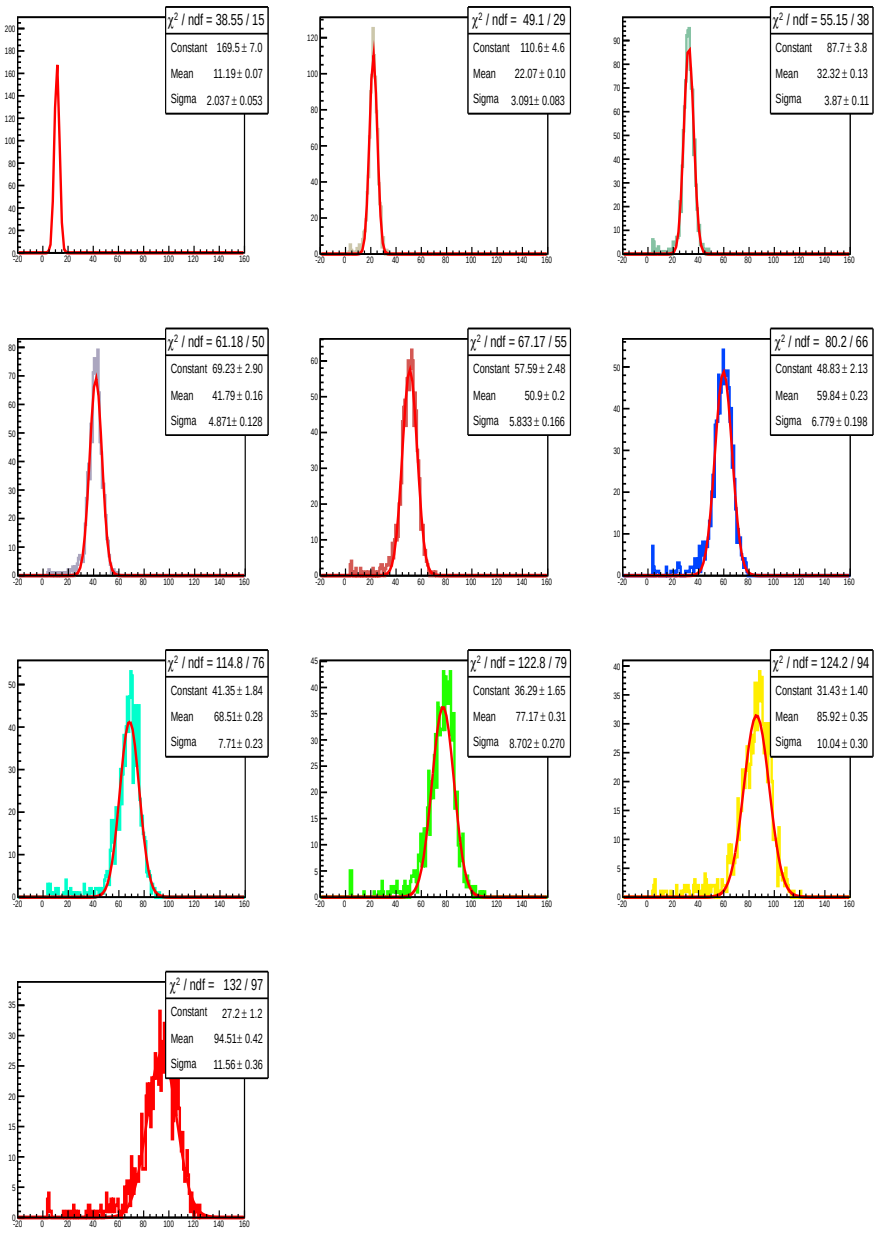

FIGURE D.1: Reconstructed energy distributions obtained with the exponential and polynomial parametrisations method discussed in section4.2.3.3(10,20,30,40,50,60,70,80,90,100GeV).

Annexe E

Quadratic calibration weights

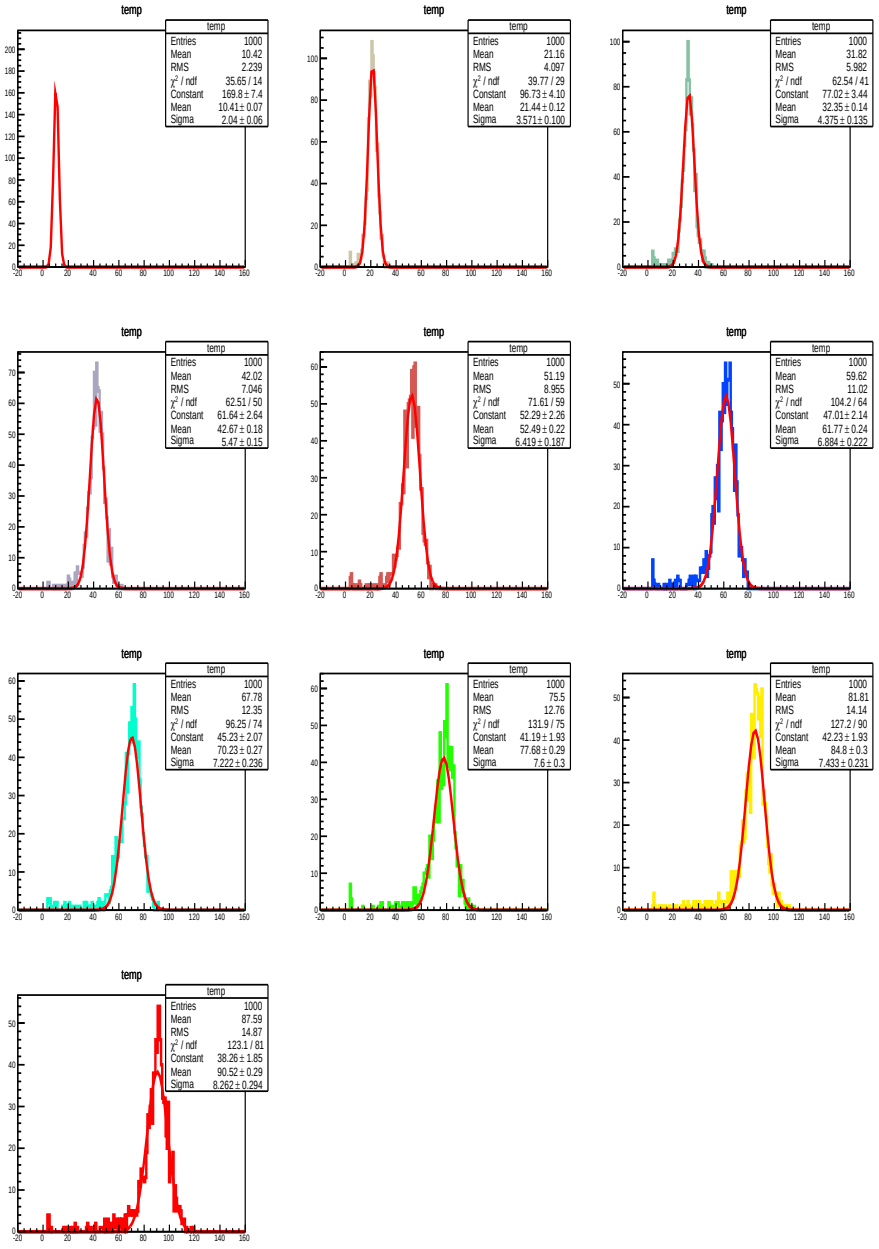

FIGURE E.1: Reconstructed energy distributions obtained with the quadratic parametrisation method discussed in section4.2.3.3(10,20,30,40,50,60,70,80,90,100GeV).

Annexe F

Energy reconstruction using Neural Network

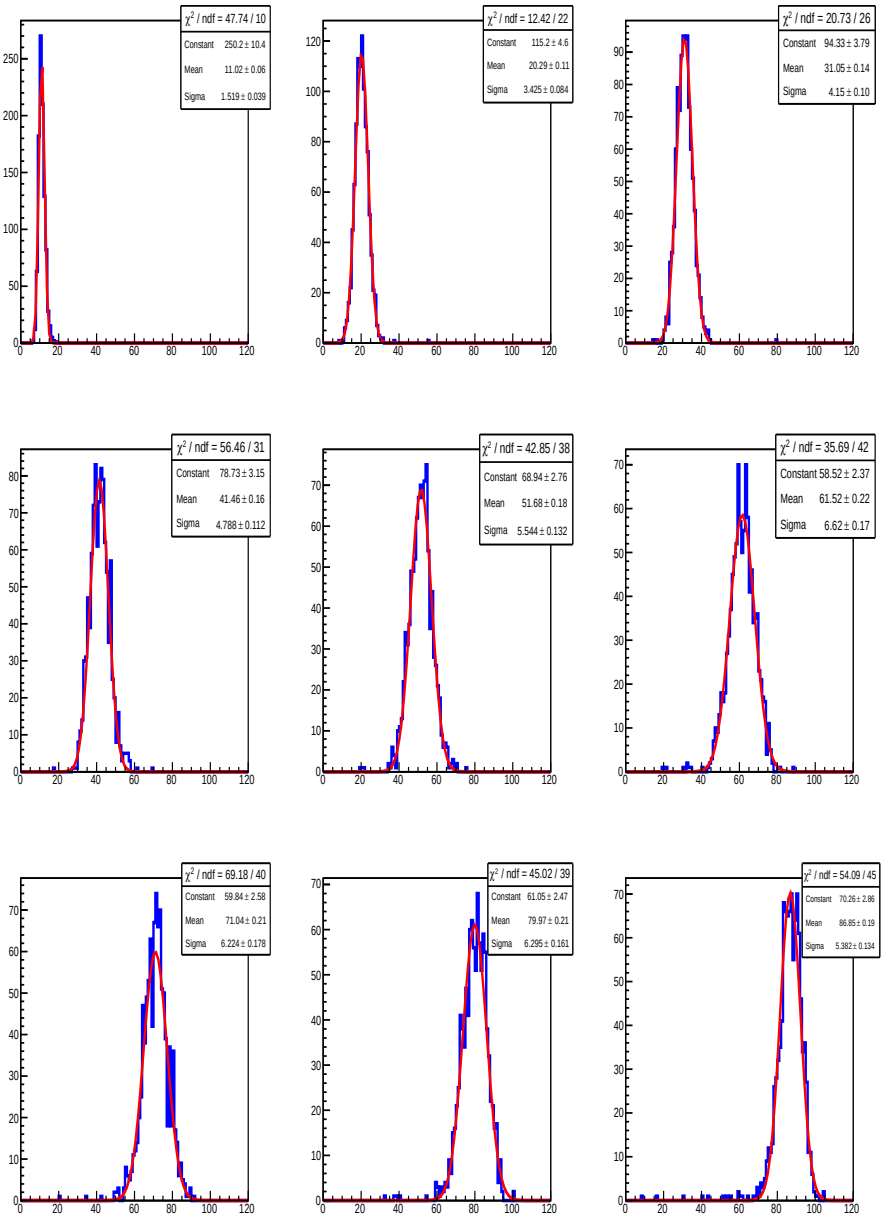

FIGURE F.1: Reconstructed Energy distributions obtained with ANN method discussed in section 4.2.4(10,20,30,40,50,60,70,80,90GeV).

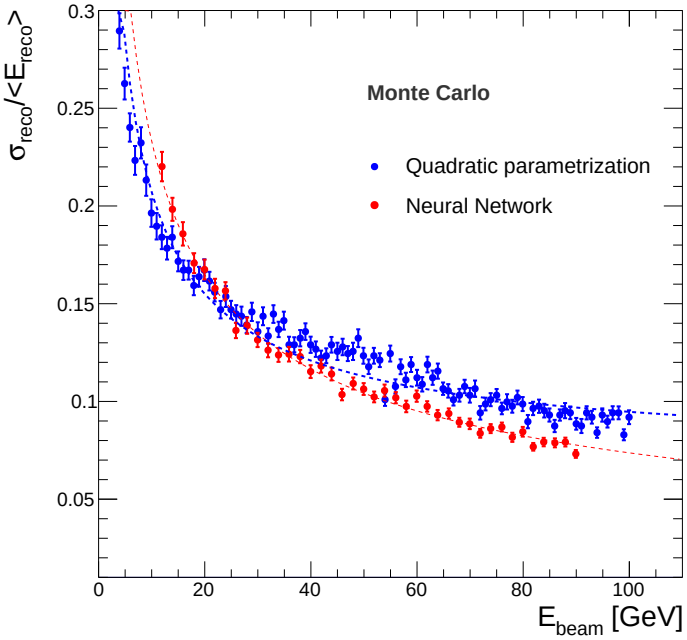

FIGURE F.2: $\frac{\sigma_{reco}}{E_{reco}}$ of the reconstructed pion energy as a function of the beam energy obtained in simulation. For the red circles graph, the reconstructed energy is estimated with neural network method. For the blue circles graph, the reconstructed energy is estimated using the standard method.

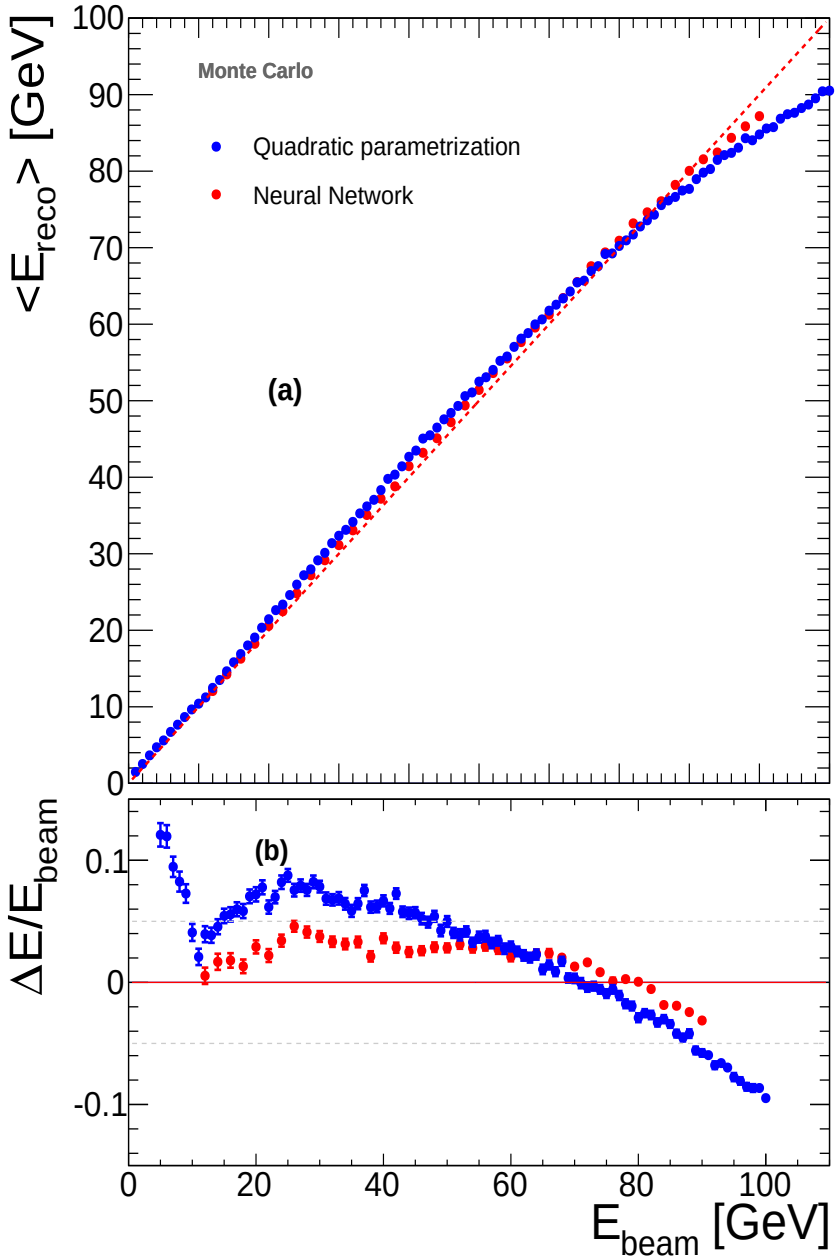

FIGURE F.3: (a) Mean reconstructed energy for pions as a function of the beam energy obtained in simulation. For the red circles graph, the reconstructed energy is estimated with neural network method. For the blue circles graph, the reconstructed energy is estimated using the standard method. The line corresponds to x=y. (b) Relative deviation of the mean reconstructed energy with respect to the beam energy as a function of the beam energy for pions obtained in simulation.

Annexe G

Energy reconstruction using Data

G.1 Selection variables

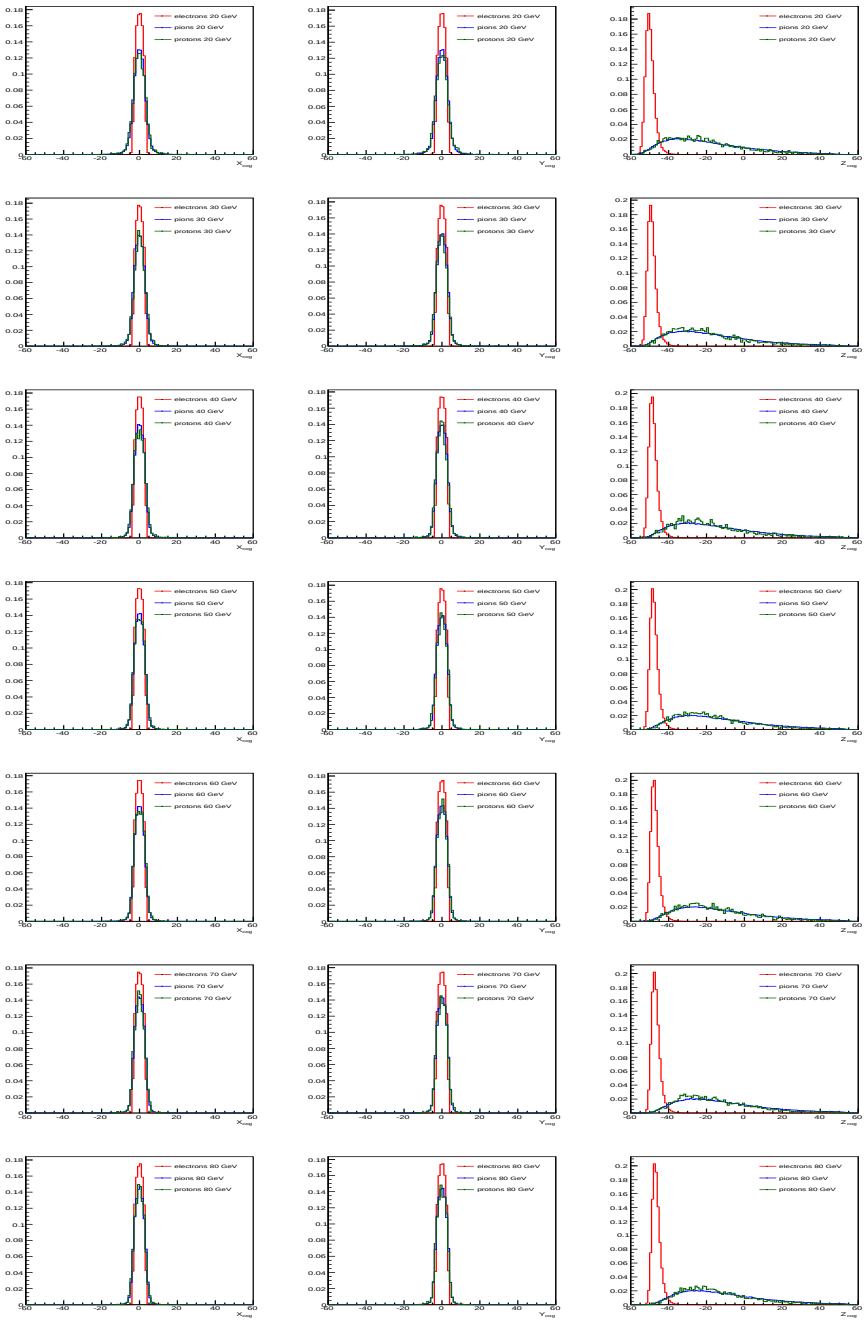
electrons 20 GeV
pions 20 GeV
protons 20 GeV
electrons 30 GeV
pions 30 GeV
protons 30 GeV
electrons 40 GeV
pions 40 GeV
protons 40 GeV
electrons 50 GeV
pions 50 GeV
protons 50 GeV
electrons 60 GeV
pions 60 GeV
protons 60 GeV
electrons 70 GeV
pions 70 GeV
protons 70 GeV
electrons 80 GeV
pions 80 GeV
protons 80 GeV

FIGURE G.1: .

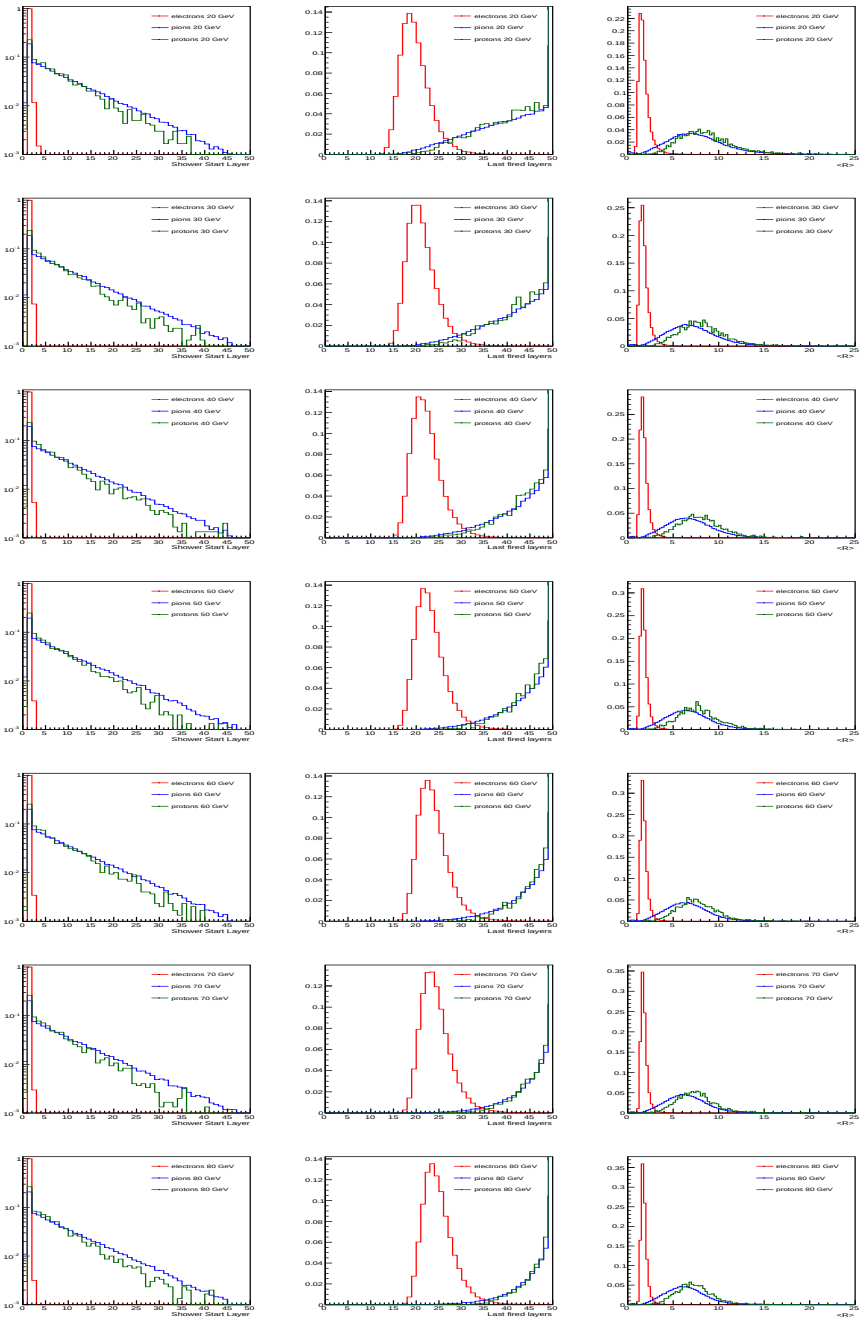
electrons 20 GeV
pions 20 GeV
protons 20 GeV
Shower Start Layer
Last fired layers
<R>
electrons 30 GeV
pions 30 GeV
protons 30 GeV
Shower Start Layer
Last fired layers
<R>
electrons 40 GeV
pions 40 GeV
protons 40 GeV
Shower Start Layer
Last fired layers
<R>
electrons 50 GeV
pions 50 GeV
protons 50 GeV
Shower Start Layer
Last fired layers
<R>
electrons 60 GeV
pions 60 GeV
protons 60 GeV
Shower Start Layer
Last fired layers
<R>
electrons 70 GeV
pions 70 GeV
protons 70 GeV
Shower Start Layer
Last fired layers
<R>
electrons 80 GeV
pions 80 GeV
protons 80 GeV
Shower Start Layer
Last fired layers
<R>

FIGURE G.2: .

Annexe H

Energy Reconstruction in SDHCAL using Data

H.1 Time calibration

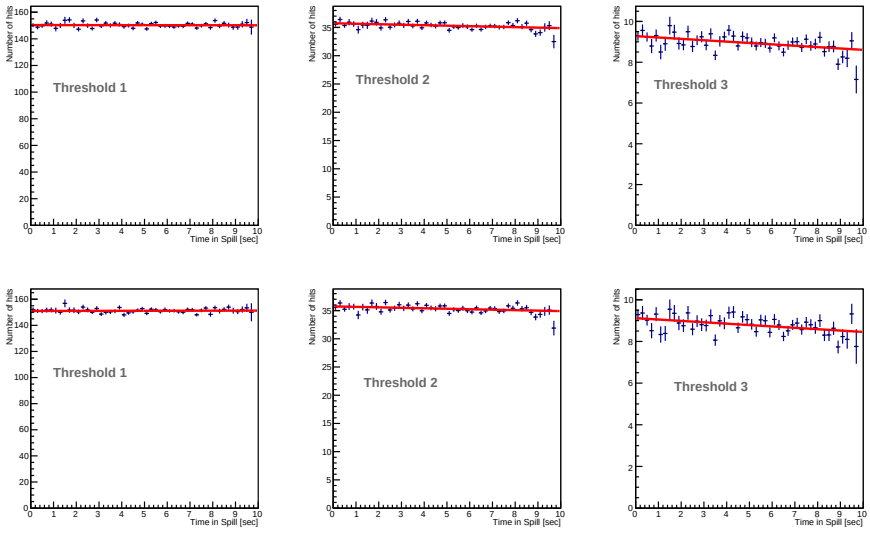

FIGURE H.1: Time calibration for pion run 10 GeV. Before time calibration(top) and after time calibration(bottom).

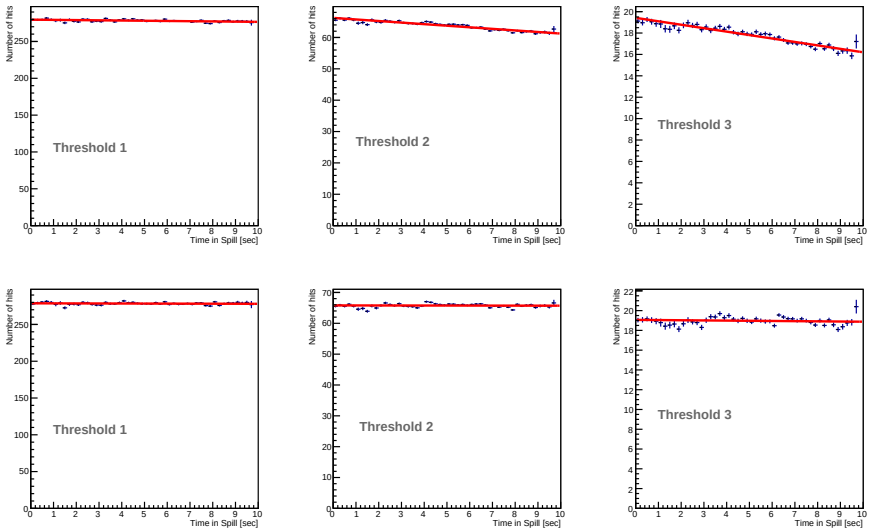

FIGURE H.2: Time calibration for pion run 20 GeV. Before time calibration(top) and after time calibration(bottom).

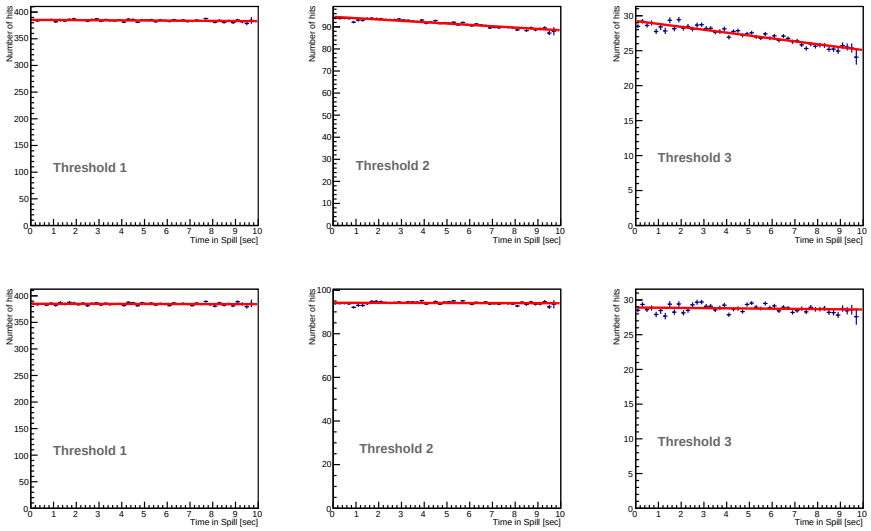

FIGURE H.3: Time calibration for pion run 30 GeV. Before time calibration(top) and after time calibration(bottom).

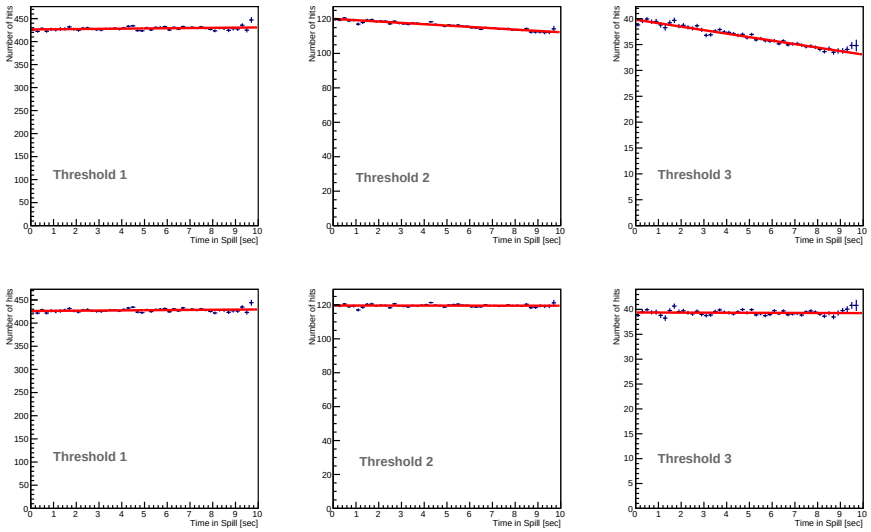

FIGURE H.4: Time calibration for pion run 40 GeV. Before time calibration(top) and after time calibration(bottom).

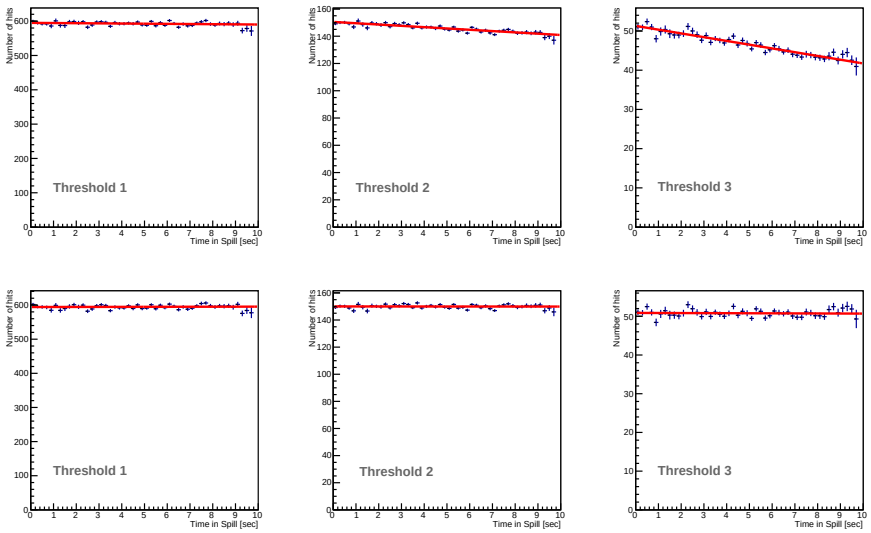

FIGURE H.5: Time calibration for pion run 50 GeV. Before time calibration(top) and after time calibration(bottom).

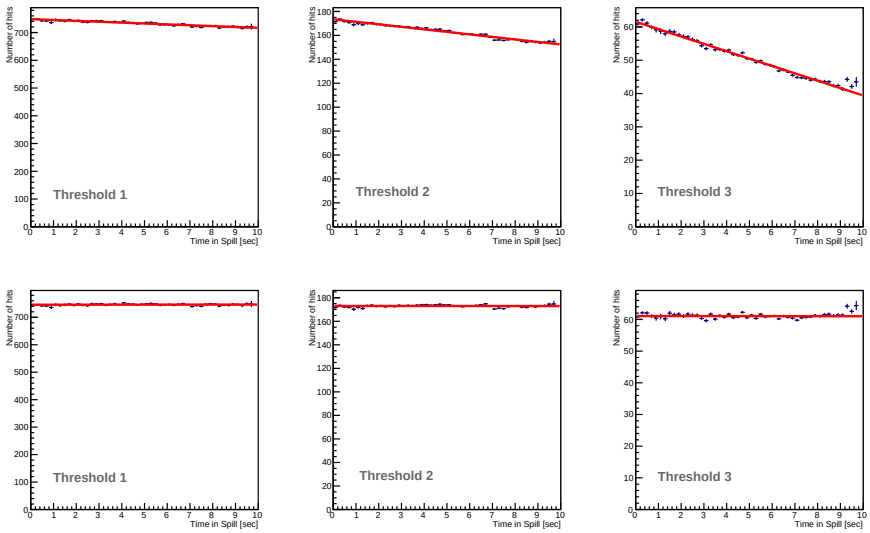

FIGURE H.6: Time calibration for pion run 60 GeV. Before time calibration(top) and after time calibration(bottom).

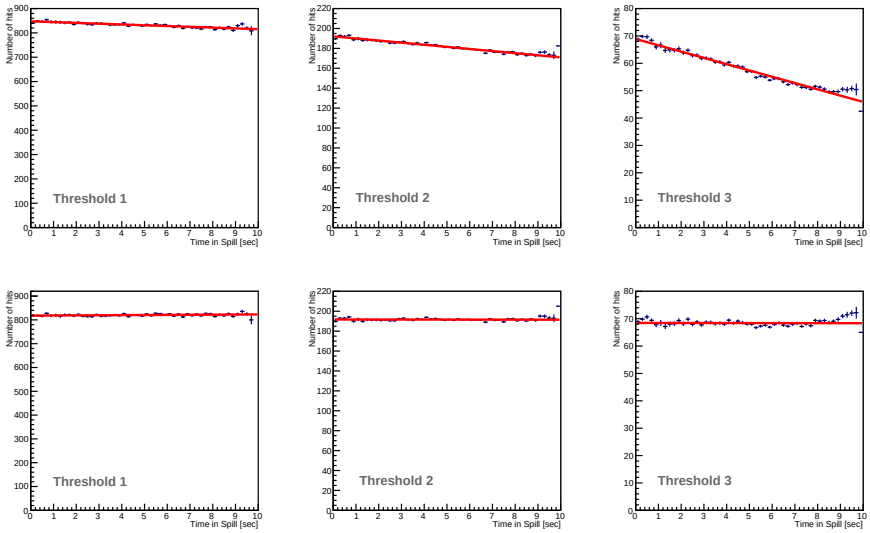

FIGURE H.7: Time calibration for pion run 70 GeV. Before time calibration(top) and after time calibration(bottom).

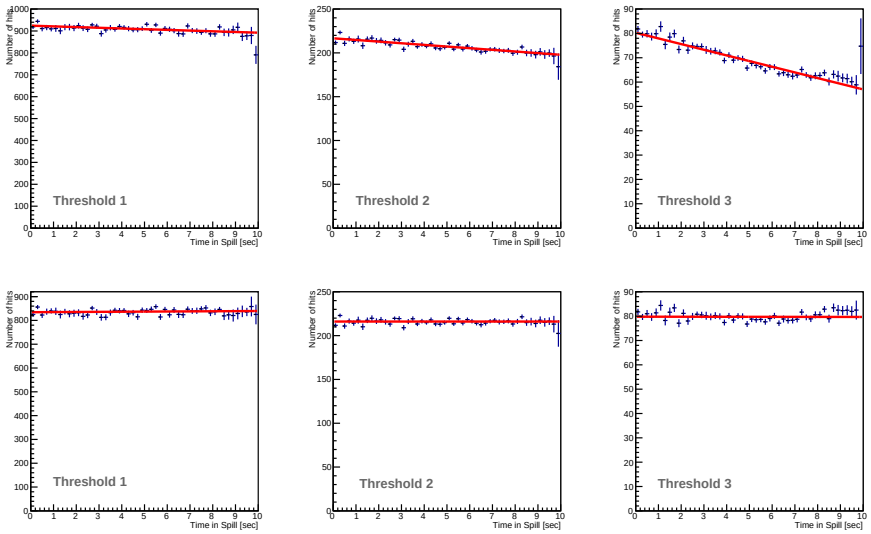

FIGURE H.8: Time calibration for pion run 80 GeV. Before time calibration(top) and after time calibration(bottom).

Annexe

Energy Reconstruction in SDHCAL using Data

I.1 Shower selection

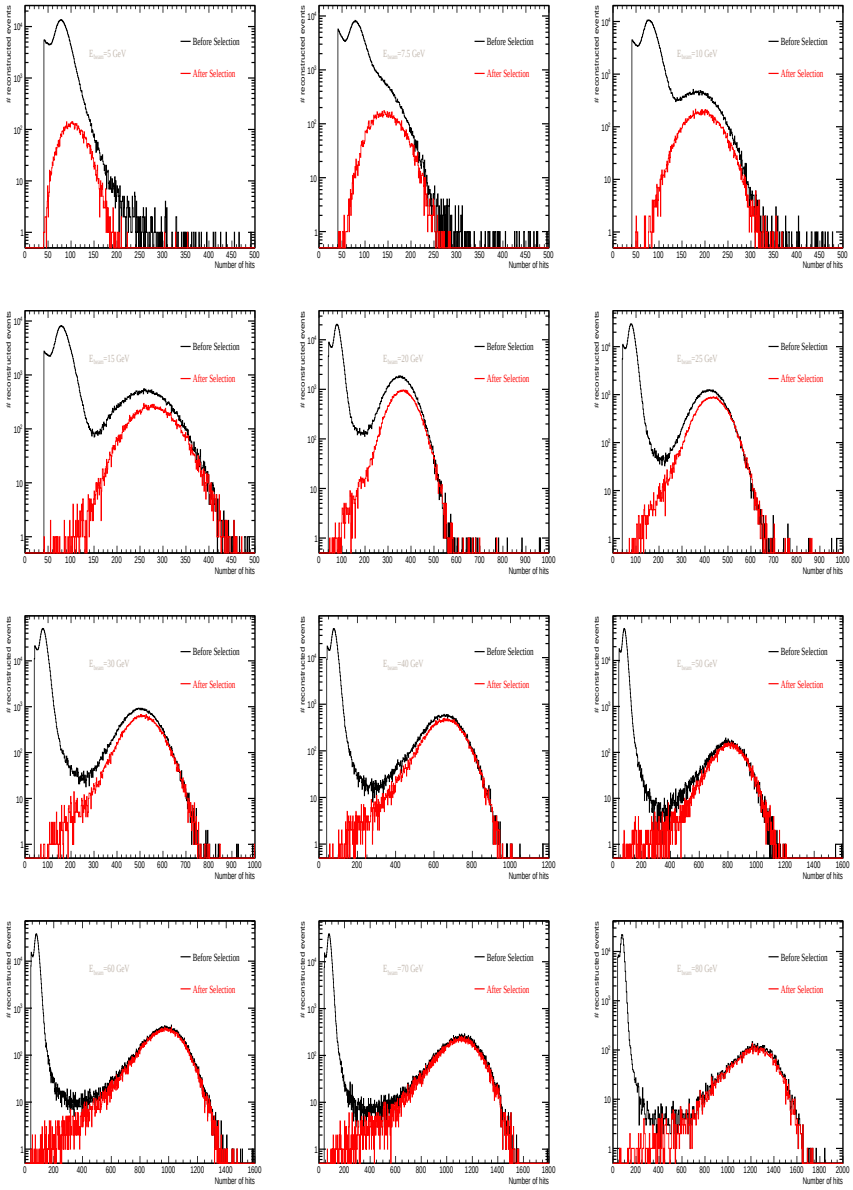

FIGURE I.1: Distribution of number of hits for pion runs from 5 (top left) to 80 GeV (bottom right) before (black line) and after (red line) selections.

Annexe J

Energy Reconstruction in SDHCAL using Data

J.1 Crystal Ball function

The Crystal Ball function [106], used first by the Crystal Ball Collaboration, is a probability density function that can be used to model a detector Gaussian response with low-end tails. In SDHCAL case, it allowed to take into consideration the saturation effect as well as the hadronic shower leakage at high energy. It consists of a Gaussian core portion and a power-law low-end tail, below a certain threshold. The Crystal Ball function is defined by :

$$f(x;\alpha,nth,\bar{x},\sigma) = N \cdot \begin{cases} \exp(-\frac{(x-\bar{x})^2}{2\sigma^2}) & \text{for } \frac{x-\bar{x}}{\sigma} > -\alpha \\ A \cdot (B - \frac{x-\bar{x}}{\sigma})^{-nth} & \text{for } \frac{x-\bar{x}}{\sigma} \leq -\alpha \end{cases} \tag{J.1}$$

where :

$$A = \left(\frac{nth}{|\alpha|}\right)^{nth} \cdot \exp\left(-\frac{|\alpha|^2}{2}\right) \tag{J.2}$$

$$B = \frac{nth}{|\alpha|} - |\alpha| \tag{J.3}$$

N is a normalization factor.

Annexe K

Systematic uncertainties

K.1 Multi-threshold mode : Quadratic parametrization for 2012 beam test data

TABLE K.1: Systematic uncertainties contributions used in paragraph 5.5.4.

E(GeV)	$\frac{\Delta^2\sigma_{stat}}{E^2_{reco}}$	$(\frac{\sigma_{reco}}{\overline{E}_{reco}})^2 . \frac{\Delta^2\overline{E}_{stat}}{E^2_{reco}}$	$\frac{(\sigma_{reco}-\sigma_{CB})^2}{E^2_{reco}}$	$\frac{(\sigma_{reco}-\sigma_{cut})^2}{E^2_{reco}}$	$\frac{(\sigma_{sim,cut}-\sigma_{sim,NoCut})^2}{E^2_{reco}}$
5	3.7055×10^{-5}	1.41095×10^{-6}	1.59161×10^{-5}	0.00022796	0.00121343
15	4.82267×10^{-6}	8.41966×10^{-8}	4.7646×10^{-8}	5.04036×10^{-6}	8.04015×10^{-6}
30	1.29225×10^{-6}	1.40869×10^{-8}	2.96896×10^{-6}	8.48758×10^{-7}	7.19897×10^{-6}
40	1.06281×10^{-6}	9.3339×10^{-9}	1.45294×10^{-5}	6.8428×10^{-7}	5.38408×10^{-7}
60	4.92284×10^{-7}	2.93931×10^{-9}	1.87891×10^{-6}	1.297348×10^{-6}	5.38408×10^{-7}

E(GeV)	$\Delta^2\overline{E}_{stat}$	$(\overline{E}_{reco} - \overline{E}_{CB})^2$	$(\overline{E}_{reco} - \overline{E}_{cut})^2$	$(\overline{E}_{sim,Cut} - \overline{E}_{Sim,NoCut})^2$
5	0.000690638	0.000072243	0.00387194	0.45526
15	0.000644652	0.00000841072	0.0084499	0.0467424
30	0.000802022	0.000193178	0.01656186	0.00465113
40	0.00113232	0.0104447	0.0486732	0.00396897
60	0.00141903	0.000161171	0.1934382	0.0640082

Bibliographie

[1] A. Salam, "Weak and Electromagnetic Interactions", *Originally printed in Svartholm : Elementary Particle Theory, Proceedings Of The Nobel Symposium Held 1968 At Lerum, Sweden, Stockholm 1968, 367-377.*

[2] S. L. Glashow, "Partial Symmetries of Weak Interactions", *Nucl. Phys.* ***22*** *(1961) 579-588.*

[3] S. Weinberg, "A Model of Leptons", *Phys. Rev. Lett.* ***19*** *(1967) 1264-1266.*

[4] UA1 Collaboration, "Experimental observation of isolated large transverse energy electrons with associated missing energy at $\sqrt{s}$=540 GeV", *Phys. Lett. B* ***122*** *(1983) 103-116.*

[5] UA2 Collaboration, "Observation of single isolated electrons of high transverse momentum in events with missing transverse energy at the CERN $p\bar{p}$ collider", *Phys. Lett. B* ***122*** *(1983) 476-485.*

[6] F. Abe et al. (CDF Collaboration), "Observation of Top Quark Production in $p\bar{p}$ Collisions with the Collider Detector at Fermilab", *Phys. Rev. Lett.* ***74*** *(1995) 2626-2631.*

[7] The CMS Collaboration, "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC", *Phys. Lett. B* ***716*** *(2012) 30.*

[8] The CMS Collaboration, "Observation of a new boson with mass near 125 GeV in pp collisions at $\sqrt{s}$=7 and 8 TeV", *JHEP* ***06*** *(2013) 081.*

[9] The ATLAS Collaboration, "Observation of a new particle in the search for the standard model Higgs boson with the ATLAS detector at the LHC", *Phys. Lett. B* ***716*** *(2012) 1.*

[10] V. Rubakov, "Large and infinite extra dimensions : An Introduction", *Phys. Usp.* ***44*** *(2001) 871-893, arXiv :hep-ph/0104152 [hep-ph].*

[11] M. B. Green, J. H. Schwarz, and E. Witten, *Superstring Theory. Vol. 1 : Introduction. Vol. 2 : Loop Amplitudes, Anomalies & Phenomenology*. WILEY-VCH Verlag, 1988.

[12] S. P. Martin, "A Supersymmetry Primer", *arXiv :hep-ph/9709356*.

[13] CALICE *collaboration web site*, https ://twiki.cern.ch/twiki/bin/view/CALICE/.

[14] The ATLAS, CDF, CMS and D0 Collaborations, "First combination of Tevatron and LHC measurements of the top-quark mass", *arXiv :1403.4427*.

[15] K. Seidel, F. Simon, and M. Tesar, " Prospects for the Measurement of the Top Mass in a Threshold Scan at CLIC and ILC".

[16] "CLIC Conceptual Design Report 2011", *http ://lcd.web.cern.ch/LCD/CDR/CDR.html*.

[17] Howard Baer, Tim Barklow, Keisuke Fujii, Yuanning Gao, Andre Hoang, Shinya Kanemura, Jenny List, Heather E Logan, Andrei Nomerotski, Maxim Perelstein, Michael E Peskin, Roman Pöschl, Jürgen Reuter, et al., " The International Linear Collider Technical Design Report - Volume 2 : Physics. Technical Report", *http ://www.linearcollider.org/ILC/TDR*, *arXiv :1306.6352*.

[18] H. Li, F. Richard, R. Poeschl, and Z. Zhang, "Precision Measurements of SM Higgs Recoil Mass and Cross Section for $\sqrt{s}$ of 230 GeV and 250 GeV at ILC", *arXiv :0901.4893*.

[19] C. Grefe, T. Lastovicka, J. Strube, F. Teubert, and B. Pie Valls, "Measurement of the Cross Section Times Branching Ratio of Light Higgs Decays at CLIC", *arXiv :1202.5647*.

[20] C. Grefe, T. Lastovicka, and J. Strube, "Prospects for the Measurement of the Higgs Yukawa Couplings to b and c quarks, and muons at CLIC", *arXiv :1208.2890*.

[21] ILD Design Study Group Collaboration, H. Li et al., " HZ Recoil Mass and Cross Section Analysis in ILD", *arXiv :1202.1439*.

[22] Ties Behnke, James E. Brau, Philip Burrows, Juan Fuster, Michael Peskin, Marcel Stanitzki, Yasuhiro Sugimoto, Sakue Yamada and Hitoshi Yamamoto " The International Linear Collider Technical Design Report - Volume 4 : Detectors. Technical Report", *http ://www.linearcollider.org/ILC/TDR*, *arXiv :1306.6329*.

[23] Chris Adolphsen, Maura Barone, Barry Barish, Karsten Buesser, Phil Burrows, John Carwardine, Jeffrey Clark, Hélène Mainaud Durand, Gerry Dugan, Eckhard Elsen, Atsuto Enomoto, Brian Foster, Shigeki Fukuda, Wei Gai, et al., " The International Linear Collider Technical Design Report - Volume 3.I : Accelerator : Part I, RD in the Technical Design Phase. Technical Report", *http ://www.linearcollider.org/ILC/TDR*, *arXiv :1306.6353*.

[24] B. Aune et al., *Phys. Rev. ST Accel. Beams **3** (2000) 092001. arXiv :physics/0003011, DOI :10.1103/PhysRevSTAB.3.092001.*

[25] Ties Behnke, James E. Brau, Brian Foster, Juan Fuster, Mike Harrison, James McEwan Paterson, Michael Peskin, Marcel Stanitzki, Nicholas Walker, Hitoshi Yamamoto. " The International Linear Collider Technical Design Report - Volume 1 : Executive Summary. Technical Report", *http ://www.linearcollider.org/ILC/TDR*, *arXiv :1306.6327*.

[26] ILD Concept Group, T. Abe et al., "The International Large Detector : Letter of Intent", DESY-2009-87, FERMILAB-LOI-2010-03, FERMILAB-PUB-09-682-E, KEK-Report-2009-6, 2009. *arXiv :1006.3396 [hep-ex]*.

[27] GLD Concept Study Group, K. Abe et al., "GLD Detector Outline Document : Version 1.2", *arXiv :physics/0607154 [physics]*.

[28] LDC Working Group, D. Kisielewska et al., " Detector Outline Document for the Large Detector Concept", *http ://ilcild.org/documents/miscdoc/ldc/outlinedoc/view*.

[29] ILC Global Design Effort and World Wide Study, T. Behnke, C. Damerell, J. Jaros, and A. Miyamoto, eds., " ILC Reference Design Report Volume 4 - Detectors. 2007", *arXiv :0712.2356 [physics.ins-det]. ILC-REPORT-2007-001*.

[30] F. Sauli, "GEM : A new concept for electron amplification in gas detectors", *Nucl. Instrum. Meth. A **386** (1997) 531-534*.

[31] Y. Giomataris, P. Rebourgeard, J. P. Robert, and G. Charpak, "MICROMEGAS : A high-granularity position-sensitive gaseous detector for high particle-flux environments", *Nucl. Instrum. Meth. A **376** (1996) 29-35*.

[32] J. Kaminski, " Status of European Pixel Modules (LCTPC collaboration meeting presentation)", *http ://ilcagenda.linearcollider.org/getFile.py/access ? contribId=18sessionId=9resId=0materialId=slidesconfId=5504*.

[33] C. Adloff, Y. Karyotakis, J. Repond, J. Yu, G. Eigen, C.M. Hawkes, Y. Mikami, O. Miller, N.K. Watson, J.A. Wilson, T. Goto, G. Mavromanolakis, et al. " Response of the CALICE Si-W electromagnetic calorimeter physics prototype to electrons", *Nucl. Instrum. Meth. A **608(3)** (2009) 372-383*.

[34] K. Francis, J. Repond, J. Schlereth, J. Smith, L. Xia, E. Baldolemar, J. Li, S.T. Park, M. Sosebee, A.P. White, J. Yu, G. Eigen, et al. "Performance of the first prototype of the CALICE scintillator strip electromagnetic calorimeter", *Nucl. Instrum. Meth. A **763** (2014) 278-289*.

[35] The CALICE collaboration, C Adloff, Y Karyotakis, J Repond, A Brandt, H Brown, K De, C Medina, J Smith, J Li, M Sosebee, A White, et al. " Construction and commissioning of the calice analog hadron calorimeter prototype", *Journal of Instrumentation, **5** (2010) P05004*

[36] C Adloff, J Blaha, J J Blaising, C Drancourt, A Espargilière, R Gaglione, N Geffroy, Y Karyotakis, J Prast, G Vouters, K Francis, J Repond, et al. " Hadronic energy resolution of a highly granular scintillator-steel hadron calorimeter using software compensation techniques", *Journal of Instrumentation, 7(09) (2012) P09017*

[37] V.L.Morgunov, " Calorimetry design with energy-flow concept (imaging detector for high energy physics)", *in Proceedings of 10th International Conference on Calorimetry in High Energy Physics(CALOR2002), Pasadena, CA, 2002*.

[38] J.-C.Brient, " Improving the jet reconstruction with the particle flow method : an introduction", *in Proceedings of 11th International Conference on Calorimetry in High-Energy Physics(Calor2004), Perugia, Italy, 2004.*

[39] M. A. Thomson. "Particle flow calorimetry and the pandorapfa algorithm", *Nuclear Instruments and Methods in Physics Research Section A,* ***611*** *(2009) 25-40, arXiv :0907.3577*.

[40] M. Aicheler, P. Burrows, M. Draper, T. Garvey, P. Lebrun, K. Peach, N. Phinney, H. Schmickler, D. Schulte, et al. "A Multi-TeV Linear Collider Based on CLIC Technology", doi :10.5170/CERN-2012-007.

[41] M. Bicer, H. Duran Yildiz, I. Yildiz, G. Coignet, M. Delmastro, T. Alexopoulos, C Grojean, S. Antusch, T. Sen, H.-J. He, K. Potamianos, S. Haug, et al. "First look at the physics case of tlep", *Journal of High Energy Physics, 2014(1) :164,2014* doi :10.1007/JHEP01(2014)164. URL http ://dx.doi.org/10.1007/JHEP01%282014%29164

[42] Wang Dou, Gao Jie, Xiao Ming, Geng Hui-Ping, Guo Yuan-Yuan, Xu Shou-Yan, Wang Na, An Yu-Wen, Qin Qing, Xu Gang, and Wang Sheng. "Optimization parameter design of a circular e^+e^- higgs factory", *Chinese Physics C, 37(9) :097003, 2013*, URL http ://stacks.iop.org/1674-1137/37/i=9/a=097003.

[43] LCTPC Collaboration, "Report to the DESY PRC 2010", *http ://www-flc.desy.de/lcnotes/notes/LC-DET-2012-067.pdf. ILC NOTE 2012-063, LC-DET-2012-067.*

[44] W. R. Leo, "Techniques for nuclear and particle physics experiments : A how to approach", newblock *second edition, 1994.*

[45] R. Wigmans, "Calorimetry Energy Measuremenr in Particle Physics", *Oxford University Press, 2000.*

[46] Particle Data Group Collaboration, K. Nakamura et al., "Review of particle physics", *J.Phys.G* ***37*** *(2010) 075021*.

[47] C. Grupen and B. Shwartz, "Particle Detectors", Particle Detectors, Cambridge University Press, 1996.

[48] J. E. Turner "Atoms, Radiation, and Radiation ProtectionParticle Detectors", J. Wiley sons, third edition 2007.

[49] U. Amaldi, "Fluctuations in calorimetry measurements", *Physica Scripta **23** (1981), no. 4A 409*.

[50] F. Sefkow, A. White, K. Kawagoe, R. Pöschl and J. Repond, "Experimental Tests of Particle Flow Calorimetry", *Rev. Mod. Phys. **88** (2016) 015003*, *arXiv :1507.05893*.

[51] CALICE Collaboration, J. Repond et al., "Design and Electronics Commissioning of the Physics Prototype of a Si-W Electromagnetic Calorimeter for the International Linear Collider", *Journal of Instrumentation **3** (2008) P08001, arXiv :0805.4833*.

[52] R.M. Brown and D.J.A. Cockerill, "Electromagnetic calorimetry", *Journal of Instrumentation **666** (2012) 47-79*.

[53] The CALICE Collaboration, K. Francis, et al., "Performance of the first prototype of the CALICE scintillator strip electromagnetic calorimeter", *Nuclear Instruments and Methods in Physics Research Section A, **06** (2014) 39, arXiv :1311.3761*.

[54] The CALICE Collaboration, C. Adloff, et al., "Construction and Commissioning of the CALICE Analog Hadron Calorimeter Prototype", *Journal of Instrumentation **5** (2010) P05004, arXiv :1003.2662*.

[55] Bilki, Burak, John Butler, Tim Cundiff, Gary Drake, William Haberichter, et al., "Calibration of a digital hadron calorimeter with muons", *Journal of Instrumentation **3** (2008) P05001, arXiv :0802.3398*.

[56] The CALICE collaboration, "Construction and commissioning of the CALICE analog hadron calorimeter prototype", *Journal of Instrumentation **5** (2010) P05004*.

[57] The CALICE collaboration, "Design and electronics commissioning of the physics prototype of a Si-W electromagnetic calorimeter for the International Linear Collider", *Journal of Instrumentation **3** (2008) P08001*.

[58] J. A. Ballin et al., "Design and performance of a CMOS study sensor for a binary readout electromagnetic calorimeter", *Journal of Instrumentation **6** (2011) P05009*.

[59] K. Francis et al., "Performance of the first prototype of the CALICE scintillator strip electromagnetic calorimeter", *Nuclear Instruments and Methods in Physics Research Section A **763** (2014) pp.278-289*.

[60] C. Adloff et al., "Construction and test of a $1 \times 1m^2$ Micromegas chamber for sampling hadron calorimetry at future lepton colliders", *Nuclear Instruments and Methods in Physics Research Section A **729** (2013) pp.90-101*.

[61] B. Bilki et al., "Hadron showers in a digital hadron calorimeter", *Journal of Instrumentation **4** (2009) P10008.*

[62] G. Baulieu et al., "Construction and commissioning of a technological prototype of a high-granularity semi-digital hadronic calorimeter", *Journal of Instrumentation **10** (2015) P10039.*

[63] M. Bedjidian et al., "Performance of Glass Resistive Plate Chambers for a high-granularity semi-digital calorimeter", *Journal of Instrumentation **6** (2011) P02001.*

[64] R. Santonico and R.Cardarelli, "Development of Resistive Plate Counters", *Nuclear Instruments and Methods in Physics Research Section A, **187** (1981) pp. 377-380.*

[65] R. Cardarelli and R. Santonico, A. di Biagio and A. Lucci, "Progress in resistive plate counters", *Nuclear Instruments and Methods in Physics Research Section A, **263** (1988) pp. 20-25.*

[66] R. Cardarelli, A. Di Ciaccio and R. Santonico, "Performance of a resistive plate chamber operating with pure CF3Br", *Nuclear Instruments and Methods in Physics Research Section A, **333** (1993) pp. 399-403.*

[67] E. Cerron Zeballos, I. Crotty, D. Hatzifotiadou, J. Lamas-Valverde, S. Neupane, M. C. S. Williams, A. Zichichi, "A new type of resistive plate chamber : the multigap RPC", *Nuclear Instruments and Methods in Physics Research Section A, **374** (1996) pp. 132-135.*

[68] P. Fonte, R. Ferreira-Marques, J. Pinhão, N. Carolino, A. Policarpo, "High-resolution RPCs for large TOF systems", *Nuclear Instruments and Methods in Physics Research Section A, **449** (2000) pp. 295.*

[69] P. Fonte, N. Carolino, L. Costa, Rui Ferreira-Marques, S. Mendiratta, V. Peskov, A. Policarpo, "A spark-protected high-rate detector", *Nuclear Instruments and Methods in Physics Research Section A, **431** (1999) pp. 154-159.*

[70] V. Peskov and P. Fonte, "Gain, Rate and Position Resolution Limits of Micropattern Gaseous Detectors", *presented at the PSD99-5th International Conference on Position-Sensitive Detectors, London, England, 1999. Available : http ://arxiv.org/pdf/physics/0209072.pdf.*

[71] S. Callier et al., "HARDROC1, readout chip of the Digital HAdronic CALorimeter of ILC", *Nuclear Science Symposium Conference Record, NSS´ 07. IEEE **3** (2007) 1851-1856.*

[72] S. Callier et al., "ROC chips for imaging calorimetry at the International Linear Collider", *Journal of Instrumentation **9** (2014) C02022.*

[73] R. Kieffer, *Caractérisation d'un calorimètre hadronique semi-digital pour le futur collisionneur. ILC. PhD thesis, Université Claude Bernard-Lyon I, 2011.*

[74] C. Adolphsen et al., "The International Linear Collider Technical Design Report - Volume 3.II :Accelerator Baseline Design., 2013".

[75] S. Aplin, J. Engels, F. Gaede, N. A. Graf, T. Johnson, et al., "LCIO : A Persistency Framework and Event Data Model for HEP".

[76] H.R. Band, "Experience with the BaBar resistive plate chambers", *In Nuclear Science Symposium Conference Record, 2003 IEEE,* ***5*** *(2003) 3735-3739*, ISSN 1082-3654. doi :10.1109/NSSMIC.2003.1352718.

[77] J.Va'vra, "Some comments about the BaBar RPC experience. What to do next ?", *http ://www-off-axis.fnal.gov/workshops/stanford03/transparencies/vavra_RPC_summary_2003_talk.pdf*.

[78] S. Agostinelli et al, "Geant4-a simulation toolkit", *Nuclear Instruments and Methods in Physics Research Section A* ***506*** *(2003) 250-303*.

[79] J. Allison et al, "Geant4 developments and applications", *Nuclear Science, IEEE Transactions* ***53*** *(2006) 270-278*.

[80] A. Steen, *Étude des gerbes hadroniques à l'aide du prototype du calorimètre hadronique semi-digital et comparaison avec les modèles théoriques utilisés dans le logiciel GEANT4. Ph.D. thesis, Université Claude Bernard Lyon-1, 2015*.

[81] Physics Lists in Geant4, `http://geant4.cern.ch/support/proc_mod_catalog/physics_lists/physicsLists.shtml`

[82] The CALICE collaboration, "Resistive Plate Chamber Digitization in a Hadronic Shower Environment", *CALICE Analysis note CAN-053. 2016*.

[83] M. Abbrescia et al, "The simulation of resistive plate chambers in avalanche mode : charge spectra and efficiency", *Nuclear Instruments and Methods in Physics Research Section A* ***431*** *(1999) 413-427*.

[84] C. Lippmann and W. Riegler, "Space charge effects in resistive plate chambers", *Nuclear Instruments and Methods in Physics Research Section A* ***517*** *(2004) 54-76*.

[85] W. Riegler, "Induced signals in resistive plate chambers", *Nuclear Instruments and Methods in Physics Research Section A* ***491*** *(2002) 258-271*.

[86] W. Riegler, C. Lippmann, and R. Veenhof, "Detector physics and simulation of resistive plate chambers", *Nuclear Instruments and Methods in Physics Research Section A* ***500*** *(2003) 144-162*.

[87] R. Han, "Induced charge signal of a glass RPC detector", *Chinese Physics C* ***4*** *(2014) 38*.

[88] S. Mannai et al, "High granularity Semi-Digital Hadronic Calorimeter using GRPCs", *Nuclear Instruments and Methods in Physics Research Section A* ***718*** *(2013) 91-94*.

[89] F. James, "MINUIT â A system for function minimization and analysis of the parameter errorsand correlations", *Computer Physics Communications **10** (1975) 343-367*.

[90] The CALICE Collaboration, "First results of the CALICE SDHCAL technological prototype", *CALICE Analysis Note CAN-037. 2012*.

[91] The CALICE Collaboration, "First results of the CALICE SDHCAL technological prototype", *Journal of Instrumentation **11** (2016) P04001*.

[92] S. Haykin, "Neural NetworksâA Comprehensive Foundation", *Prentice-Hall (1999)*.

[93] B. Denby et al, "Neural Networks and Cellular Automata in Experimental High Energy Physics", *Computer Physics Communications **49** (1988) 429*.

[94] C. Peterson, "Track Finding with Neural Networks", *Nuclear Instruments and Methods in Physics Research Section A **279** (1989) 537*.

[95] V.M. Abazovw et al (D$\emptyset$ Collaboration), "Search for single top quark production at $D\emptyset$ using neural networks", *Physics Letters B **517** (2001) 282-294*.

[96] J. Freeman et al, "An Artificial Neural Network Based b-Jet Identification Algorithm at the CDF Experiment", *Nuclear Instruments and Methods in Physics Research Section A **663** (2012) 27-37*.

[97] P. Abreu et al (DELPHI Collaboration), "Classification of the hadronic decays of the Z0 into b and c quark pairs using a neural network", *Phys. Lett. B. **295** (1992) 383-395*.

[98] S. Abachi et al, "Direct measurement of the top quark mass", Phys. Rev. Lett. **79** (1997) 1197.

[99] TMultiLayerPerceptron : Designing and using Multi-Layer Perceptrons with ROOT, `http://cp3.irmp.ucl.ac.be/~delaere/MLP/`

[100] S. Mannai "Energy Reconstruction in a High Granularity Semi-Digital Hadronic Calorimeter for ILC Experiments", *IEEE Transactions on Nuclear Science **63** (2016) 2880-2886*.

[101] Y. Haddad, *A highly granular semi-digital hadron calorimeter for a future linear e^+e^- collider and a model independent Higgs boson measurement in the $ZH \rightarrow qq + X$ channel. Ph.D. thesis, L'école polytechnique, LLR - Laboratoire Leprince-Ringuet, 2014*.

[102] `https://sba.web.cern.ch/sba/`

[103] Instruction for using CESAR the new control system of the SPS secondary beam lines `http://sba.web.cern.ch/sba/cesar/Cesar_manual.htm`

[104] B. Bilki et al., "Measurement of the rate capability of Resistive Plate Chambers", *Journal of Instrumentation **4** (2009) P06003*.

[105] E. Abat et al., "Study of energy response and resolution of the ATLAS barrel calorimeter to hadrons of energies from 20 to 350 GeV", *Nuclear Instruments and Methods in Physics Research Section A* ***621*** *(2010) 134-150*.

[106] J. E. Gaiser, *Charmonium Spectroscopy from Radiative Decays of the J/ψ and ψ', SLAC-R-255, PhD thesis, Stanford University, Stanford, California 94305, 1982*.

[107] A. Hoecker P. Speckmayer J. Stelzer J. Therhaag E. von Toerne and H. Voss, "TMVA - Toolkit for Multivariate Data Analysis with ROOT - Users Guide", *CERN-OPEN-2007-007*

[108] G. D'Agostini, "Probability and Measurement Uncertainty in Physics - a Bayesian Primer", arXiv :hep-ph/9512295.