OPTIMIZATION OF ENERGY RESOLUTION IN THE DIGITAL HADRON CALORIMETER USING LONGITUDINAL WEIGHTS

by

JACOB RUSSELL SMITH

Presented to the Faculty of the Graduate School of
The University of Texas at Arlington in Partial Fulfillment
of the Requirements
for the Degree of

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF TEXAS AT ARLINGTON
May 2013

To my family.

ACKNOWLEDGEMENTS

I owe a debt of gratitude to my supervising professor Dr. Jaehoon Yu. His continued encouragement, mentorship and commitment throughout the years has motivated me to pursue excellence in all my work. His excitement for Physics has developed a deeper joy for research in me as well. I also wish to thank Dr. José Repond for his role and support in developing me as a scientist and providing an opportunity to pursue research with the DHCAL group at Argonne National Laboratory. I'm thankful for the work and interest that the members of my committee Dr. Andrew Brandt, Dr. Ali Koymen and Dr. Andrew White have put in to my work and times as a student at UTA.

I want to thank all the scientists, students, engineers, technician and administrative staff at Argonne National Laboratory for also helping me make such great progress. Thanks to the HEP Division Director Harry Weerts for his leadership and support during all stages of the DHCAL project. I would like to thank the scientists in the DHCAL Detector R&D group that provided insightful assisted and guidance in my data analysis; Lei Xia, Kurt Francis, and Burak Bilki. In addition I would like to highlight key contributions from these scientists. Lei Xia has made tremendous contributions at all stages including RPC design, construction and testing; glass painting; test beam data taking; and all aspects of data analysis. Kurt Francis' work on RPC construction and cassette assembly provided quality RPCs for the DHCAL. His contributions to GEANT4 simulation imparted great insight into the data anslysis. Burak Bilki's implementation of the ROOT data structure afforded simplicity to a complex analysis. I am grateful to Dave Underwood and Qingmin Zhang's initial

work on R&D on RPC glass surface resistive layers. Thanks to Jim Schlereth for providing the excellent event displays and direction in my writing of the event builders for a data analysis format.

A great appreciation is due to the mechanical and electronics groups at Argonne that contributed to this project. The excellent work and support by the electronics group, led by Gary Drake, was invaluable to this project. Many people contributed to various aspects of the DHCAL project in ways I am unaware but I would like to highlight a few things. Thanks to Pat De Lurgio for his work on the low voltage distribution system for the front-end electronics. Thanks to Andrew Kreps for his contributions to designing and testing the FPGA's on the FEBs. Thanks to Tim Cundiff for his design and development of the layout of the FEBs. Additional electronics support and maintenance was provided by Carol Adams, Mike Anthony, and Bill Habericter.

I would also like to highlight the work led by Ken Wood and Vic Guarino in the mechanical and engineering groups in the ANL HEP division. Thanks to Frank Skrzecz for building fixtures for DHCAL construction and providing a multitude of additional structural components. Thanks to Robert Furst and Zeljko Matijas for their work in their detailed excellent production of RPC chambers. Thanks to Vic Guarino for designing the paint booth and motor control for the glass and spray gun. Thanks also to Allen Zhao for designing and implementing the glueing robot for FEB assembly. Dave Northacker has furnished the DHCAL with a robust and flexible gas supply system for RPC testing and test beam use.

Additional institutions have provided exceptional work to the DHCAL project. The Fermilab ASIC group designed and tested ASICs used in the FEBs and also provided Timing and Trigger Modules for the DAQ system. Furthermore, Boston University designed and tested Data Collectors. The University of Iowa has provided a flexible and robust HV power supply system. Their contributions enriched the quality of the data taken during beam tests.

I would additionally like to stress the contributions of the Fermilab Test Beam Facility staff. Thanks to Aria Soha for her invaluable coordination of test beam activities at FTBF. Thanks to Todd Nebel for his mechanical and electronics support to install the DHCAL. Thanks also to Erik Ramberg for his oversight and advisement.

Additionally I would like to address my gratitude to the CALICE collaboration. They have furnished a first-class steel calorimeter structure and provided support for data taking at the FNAL testbeam during all data taking campaigns. They provided excellent support during

Finally, I would like to thank my family; my mother Janet Smith, father Russell Smith, brother Jeremy and sister Taryn. They have individually and collectively supported, encouraged, motivated, inspired, and cheered me on through every stage of graduate studies.

March 28, 2013

ABSTRACT

OPTIMIZATION OF ENERGY RESOLUTION IN THE DIGITAL HADRON CALORIMETER USING LONGITUDINAL WEIGHTS

JACOB RUSSELL SMITH, Ph.D.

The University of Texas at Arlington, 2013

Supervising Professor: Jaehoon Yu and José Repond

Particle physics seeks to describe matter in its most elementary structure. With lepton colliders; couplings of gauge bosons and heavy quarks, physics beyond the Standard Model, and properties of a recently discovered Higgs boson can be studied with very high precision. Particle Flow Algorithms (PFA), able to achieve necessary jet-energy and di-jet mass resolutions, require fine transverse and longitudinal segmentation from calorimeters. To validate digital imaging calorimetry in this context, a Digital Hadron Calorimeter (DHCAL) with single-bit (digital) readout on imbedded electronics from $1 \times 1\ cm^2$ pads throughout its volume has been constructed and exposed to particle beams as a large prototype with 54 layers of Resistive Plate Chambers and nearly 500k readout channels. With this data, I report on a sophisticated statistical technique to improve the single particle energy resolution using weights that take advantage of correlations of the energy deposit between layers. Limitations to resolution improvement are identified and possible modifications are discussed. Simulation is used to verify particle identification techniques applied to the data.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

LIST OF TABLES

CHAPTER 1

INTRODUCTION

Humanity aims to describe what exists, where it's from and how it works. In general, science is a systematic endeavor to experimentally measure and mathematical describe nature. Particle physics seeks to predict and measure the existence and interactions of the fundamental constituents of matter. The goal of this thesis is to report on recent progress in the development of novel detectors for future high energy particle physics experiments.

In order to measure sub-structure, experiments need an observational mechanism, or probe, with similar spatial resolution [1]. If a beam of point-like particles is used as a probe, the spatial resolution Δr, defined as the measurable distance between two points in an object, is limited by the de Broglie wavelength of the beam particle,

$$\lambda = h/p \tag{1.1}$$

where p is the beam momentum, h is Plank's constant and λ is the wavelength. Therefore, high momentum beams result a in shorter wavelength and thus higher spatial resolutions. This means that increasingly high energy probes could allow discover of new physics at increasingly smaller scales. This motivates experiments that collide high energy beams of various types of particles.

Some particles require high energy collisions to generate them in experiment. The energies of colliding beams depends on the contents of the beam and the desired reach of the experiment. For instance, hadron (particles with quark constituents such as pions and neutrons) colliders have energies in the multi-TeV range to probe fundamental particles while lepton colliders can access similar scales with less than

Table 1.1. SM elementary particles

Particle	Flavor			$Q/\|e\|$
leptons	e	μ	τ	-1
	ν_e	ν_μ	ν_τ	0
quarks	u	c	t	+2/3
	d	s	b	$-1/3$

1 TeV collisions [2]. Particle physics experiments that collide beams of particles at high energies are classified as High Energy Physics (HEP).

1.1 The Standard Model

The Standard Model (SM), formed in the 1960s [3, 4, 5] presents quarks and leptons as fundamental constituents and provides a theoretical description of the three forces; strong, weak and electromagnetic (EM). There are six quarks and six leptons that are fermions, having spin-1/2.

Leptons have integral electric charge, relative to the charge, e, of an electron. There are three charged leptons; electron e, muon μ and tauon τ. There are also three neutral leptons called neutrinos ν associated with the charged leptons in a classification called flavor as in Table 1.1. An example of this association is in nuclear β-decay, $^6He \rightarrow {}^6Li + e + \bar{\nu_e} + 3.5MeV$ [cite Csikay and Szalay 1957], where an electron is emitted with electron-type anti-neutrino $\bar{\nu_e}$.

Quarks carry fractional charge relative to e, +2/3 or -1/3. Quarks are also divided by charge where the three $+2/3e$ are called up, charm and top while the three -1/3e are called down, strange and bottom. Furthermore, leptons can exist as free particles but quarks are only found in combinations of two or three. For example, protons consist of two ups and a down, uud, and neutrons consist of two downs and one up, ddu. Another distinction of quarks from leptons is that each quark has three

Table 1.2. SM boson mediators

Interaction	Mediator	Spin/parity	Strength
Strong	gluon, g	1^-	1
EM	photon, γ	1^-	10^{-2}
Weak	W^+, W^-, Z^0	$1^-, 1^+$	10^{-7}
Gravitation	-	-	10^{-39}

types of charge called color for the strong interaction. The total charge from all fermions is zero.

Fermions interact with each other via the exchange of force mediators that are bosons with integral spin. There are four types of interactions; strong, EM, weak and gravitational. The strong force binds the quarks inside hadrons and is mediated by the massless gluon. EM interactions describe a wide range of physics from bound states of electrons with atoms to inter molecular forces in matter. EM force is mediated by exchanges of photons. Weak interactions are involved in processes such as nuclear β-decay, and its mediators are the W^+, W^- and Z^0 bosons. Gravitational interactions are not yet described by the SM. The interactions and their associated mediators, spins and relative strengths are listed in Table 1.2

The SM has been experimentally verified to very high precision. A group of physicists called GFitter [6] perform indirect measurement of SM observable with a complex set of global fits [7] using input data from direct measurements. Figure 1.1 [8] shows deviations of indirect measurements from experimental values. Deviations are given by the difference between the result of the fit, O_{fit}, and the input data, O_{meas}, normalized by the data uncertainty, σ_{meas}. The deviations here are called pull values because a particular observable's prediction omits input data from direct measurements for only that observable. Pull values are all within 3σ showing that the SM is a self-consistent theory.

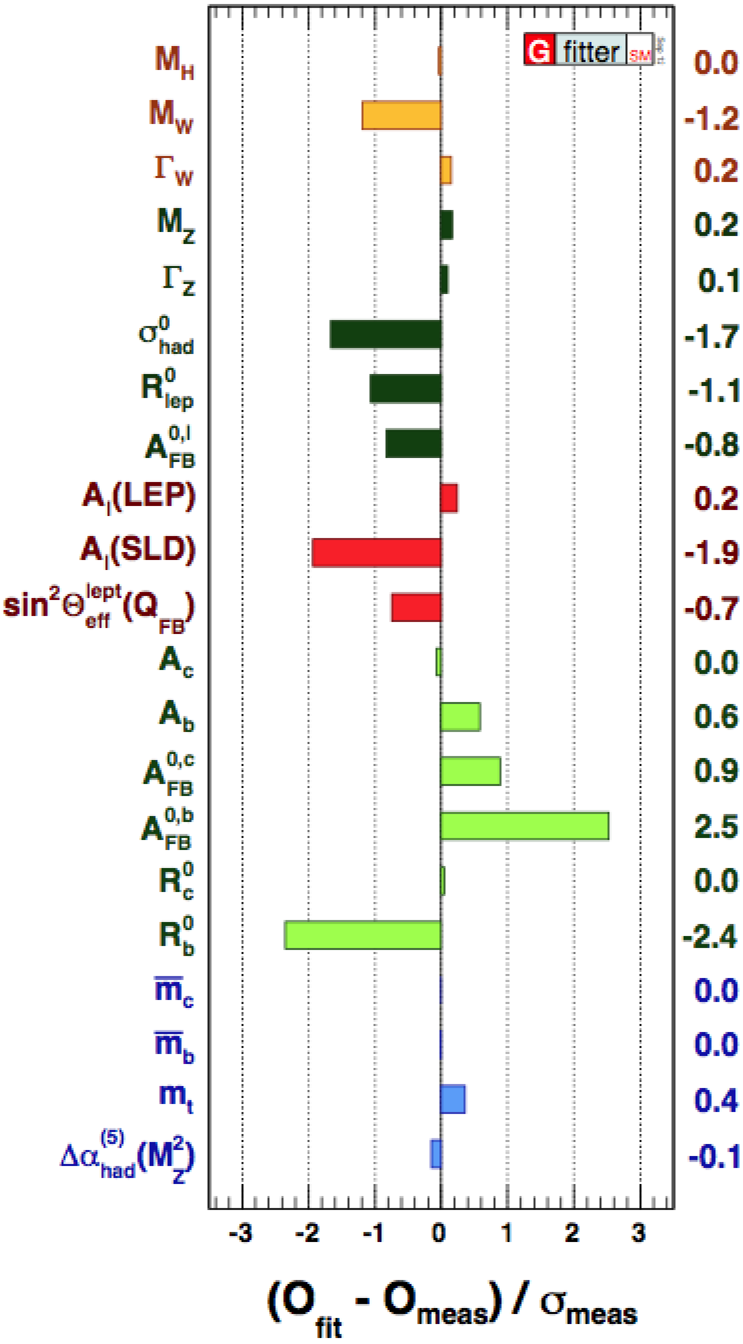

Figure 1.1. Pull values, $O_{fit} - O_{meas}/\sigma_{meas}$, from global electroweak fits of SM observables.

Even though the SM is shown to be a consistent theory, it does not completely describe all of nature. A simple list of some limitations of the Standard Model follows as

- there is no candidate field for cold dark matter
- doesn't account for dark energy or dark matter
- lacks any description of gravity

New experiments and accelerators are needed to precisely study properties of a newly discovered Higgs boson, other SM particles, and explore physics beyond the SM (BSM). One such accelerator currently proposed is the International Linear Collider (ILC) [2]. The ILC is designed to collide electrons and positrons at the energies between 0.5 and 1.0 TeV. Detectors under development for the ILC are the Silicon Detector (SiD) [9] and the International Large Detector (ILD) [10]. Each of these detectors are designed to improve the jet position and energy resolutions by optimizing them with particle flow algorithms [11] using different detector technologies.

1.2 Particle Flow Algorithms

Future experiments in HEP will explore physics at higher energy and with greater precision. The goal of ILC experiments is to make unprecedented precision measurement of the fundamental particle properties using the collisions between electrons and positrons.

The collision of leptons is different from the collision of hadrons. One difference is that leptons are fundamental particles whereas hadrons are composite. In the collision of hadrons at high energies, each quark shares some fraction of the hadron energy. Additionally, the high momentum of the beam serves as a probe of the hadron's constituent quarks. Therefore, the energies available for interactions are reduced and fluctuate between collisions. This is a disadvantage for precision measurements of properties of new particles. Conversely, a low momentum transfer between colliding hadrons is a useful tool to study properties of new particles and explore BSM physics [12, 13]. The advantage of using leptons is that, as a fundamental particle, the energy is not shared among constituents, and thus is available for entire energy of collision of e^+ and e^- collision, and in general, have a cleaner environment. Additionally the energy spread between collisions is less for lepton colliders.

Precision required in the lepton-lepton collider physics goals further require unprecedented spatial and energy resolution in the detector. Some of the motivations are to be able to separate the jet final states of W and Z bosons resulting from the collision. Jet-energy resolution describes the detectors ability to measure the energy of jets, closely bound collections of particles resulting from collision. The particle flow strategy helps accomplish the required precision in jet energy measurement. PFAs improve jet energy resolution by replacing charged particle energy measured in the calorimeter with that of the momentum measured in the tracker, and the calorimeter measures the neutral particle energies.

In a PFA optimized detector, the jet-energy resolution is optimized by measuring the energy of particles in a hadronic jet individually. The detector system with the best energy resolution for a particles is used to measure the energy. The tracking system measures the momentum of charged particles. The EM calorimeter (ECAL) measures the energy of photons. The hadronic calorimeters (HCAL) and ECAL measure the energy of neutral, charged, and EM particles and the muon system measures the muon momentum [14]. The jet-energy resolutions without PFAs for these groups of particles are shown in Table 1.3 [15]. About 60% of the jet-energy is carried by charged particles, 30% by photons, and 10% by neutral hadrons. PFAs under development for the ILC seek to attain a jet-energy resolution of $30\%/\sqrt{E_j}$ [2]. In order to optimize PFA performance, the detectors themselves must be be optimized to the

Table 1.3. Particle, detector, and jet-energy resolutions within a PFA

Component	Detector	Energy frac	Energy res	Jet energy res
Charged particles ($X^{\pm}$)	Tracker	$\sim 0.6E_j$	$10^{-4}E^2_{X^{\pm}}$	$< 3.6 \times 10^{-5}E^2_j$
Photons (γ)	ECAL	$\sim 0.3E_j$	$0.15\sqrt{E_\gamma}$	$0.08\sqrt{E_j}$
Neutral hadrons (h^0)	HCAL	$\sim 0.1E_j$	$0.55\sqrt{E_{h^0}}$	$0.17\sqrt{E_j}$

PFA strategy. Each sub-detector in the system must be designed to maximally utilize the advantages of a PFA.

The hadronic calorimeter design is a good illustration of the relationship between detector design and particle flow algorithm performance. Hadronic jets resulting from a beam collision will leave tracks in the trackers and initiate showers in the EM and hadronic calorimeters. The PFA matches tracks in the tracker to energy deposits in the calorimeters. If the track is associated to an energy cluster in the calorimeter then the tracker momentum is used to replace it. On the other hand, if energy deposits of the jet are not associated to any tracks then it is counted as particle of neutral shower components. In general, the hadronic calorimeter has a worse resolution that the tracker so that errors in matching energy deposits as charged or neutral adversely effects the jet-energy resolution in the PFA. Therefore, a finely segmented calorimeter is needed to have a better match between the charged particle shower and the corresponding tracks.

A PFA optimized calorimeter must have a high level of segmentation. This high level of segmentation provides the opportunity to identify and separate particle showers within a jet. It also provides the opportunity to distinguish jets from one another. However, fine segmentation requires many readout channels possibly increasing the cost of electronic readout.

1.3 Calorimetry

Calorimetry is the process of measuring the energy in the absorption of particles as they interact with matter. A calorimeter is a detector that measures the energy as a particle interacts within it. This can be achieve, either by requiring an incoming particle to be completely absorbed within a homogeneous material (homogeneous total absorption calorimeter) or by inserting minimally interacting detectors in alternating layers sandwiched between absorbing materials. Absorbing materials are often referred to as the passive media and are inactive. The calorimeter with alternating layers of active and passive materials is known as a sampling calorimeter. For example, the ZEUS calorimeter was a sampling calorimeter using uranium as absorber and scintillator as active media [16]. The D∅ calorimeter [17] employs a sampling hadron calorimeter with liquid argon as the active medium where the passive medium consists of two sections; a fine-hadronic section with uranium plates and a course-hardronic section with copper or stainless steel plates.

The readout of calorimeters used in HEP experiments mentioned so far record signals that are proportional to the ionization energy loss in active media using analog (multi-bit) electronics. With increasing granularity of calorimeter readout it becomes possible to minimize electronics cost and complexity recording signals using digital (single-bit) electronics. A fundamental assumption for digital calorimetry is that the number of hits in a finely segmented calorimeter is proportional to the total energy deposited by incident particles. Digital refers to the bit resolution of recorded signals. Readout electronics digitize charge induced in the active medium and records signals as a single bit determined by a threshold value. Analysis performed in this thesis will assist to further validate digital calorimetry.

In a sampling calorimeter, an incident particle will primarily lose energy as it interacts with the absorbing material. Figure 1.2 is an example of a hadronic shower

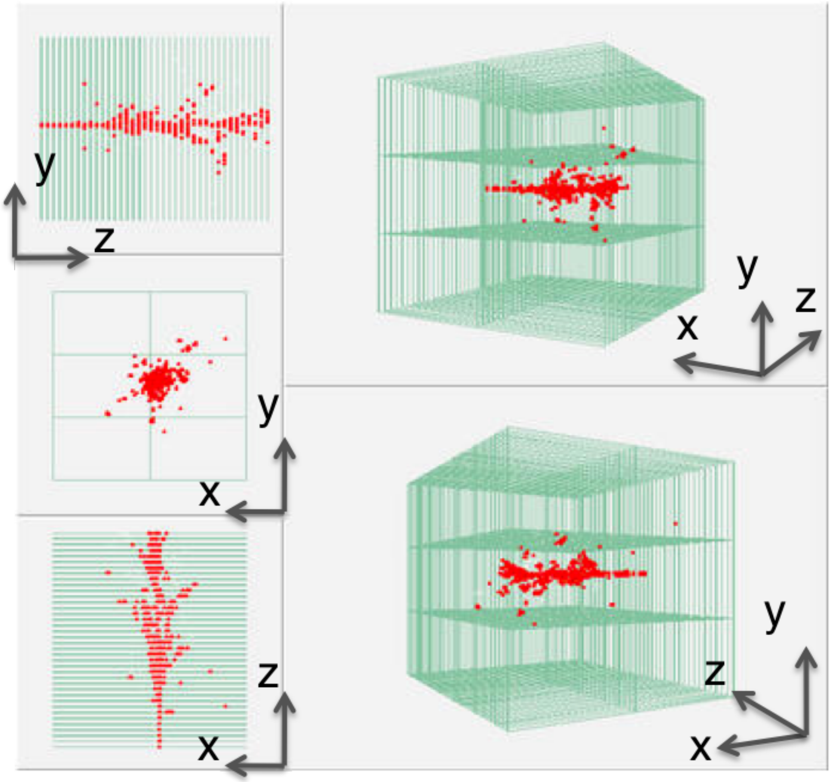

Figure 1.2. Event display of a 60 GeV/c π^+ showering in a digital hadron calorimeter instrumented with resistive plate chambers as active medium. Each red point represents a fractional loss of energy by the π^+ and its shower components..

cascade within a sampling calorimeter. The active detector then, in terms of analogy, takes pictures of the interactions in repeating intervals. This means that after a layer of absorber, the secondary particles in the nuclear interactions traverse through the active detector and loses energy via ionization. The sample of energy is proportional to the incident particle energy. This process repeats until the shower particle energy is smaller than the critical energy in them to continue through the calorimeter. At this point the incident particle has been completely absorbed in the calorimeter, and its energy can be measured with the sampled energy in the active medium. In this context, longitudinal refers to the direction perpendicular to the sampling layers and transverse refers to the direction parallel to sampling layers.

The reason why a particle's energy measurement has limited resolution is because of the fact that in the interactions with the absorber there are many low energy interactions that do not get detected in the active medium, in addition to the statistical fluctuations of shower particles. There are also types of collisions that are invisible to the active detectors [18], such as neutrons, neutrinos and photons. The components of a shower induced by hadron collisions include hadronic and EM components. The hadronic components consist of other charged hadrons, nuclear fragments, neutrons and nuclear breakup. The EM components consist of photons, positrons, electrons, and neutral pions (that further decay to di-photon pairs). These effects contribute to the calorimeter energy resolution. The energy resolution of a calorimeter depends on the type of absorbing material, the active detectors involved, and the sampling fractions determined by the energy loss ratio in passive and active media. For example, particles incident to the DØ central hadron calorimeter generate secondary charged particles in of 16 layers uranium and 16 layers of stainless steel and copper to further generate signals in the liquid argonne active medium. The liquid argonne samples ionizations produced in showers [17].

The energy resolution also depends on the type of incident particle. Some particles, such as electrons, positrons and photons primarily interact via EM interactions while hadrons primarily interact with the nuclei of the absorber material via nuclear collisions. In general, calorimeters do not respond equally to EM and hadronic interactions. That is, for the same incident particle energy, the energy measured from positrons (EM) is different from the energies measured from pions (hadronic). In turn, this degrades the resolution of a hadronic shower energy measurement. Additionally, hadronic showers have EM and hadronic components that fluctuate between showers. Energy resolution is typically written as

$$\frac{\sigma(E)}{E} = \frac{a}{\sqrt{E}} \oplus b \oplus \frac{c}{E} \tag{1.2}$$

where the first term describes statistical fluctuations in shower propagation, the second term is a constant describing calibration, non-linearity, etc. and the third term accounts for noise. Statistical fluctuations are the main source of resolution degradation in hadron showers. The $\oplus$ symbol signifies that these sources of resolution are added in quadrature. In order to equalize the response, techniques are employed in the physical construction and in offline software to optimize the energy resolution.

Another aspect of calorimeters is the response of their signal as a function of energy of incident particles, known as the response linearity. The most desirable is for the detector to respond without deviation from a straight line over an extended energy range of the incident particle. In other words, if the calorimeter is designed to measure pions between the energies of 10 - 300 GeV. Then a linear response will result in the measured energy for 200 GeV particles to be twice that for incoming particles of 100 GeV. This is desired because in an experiment the energies of particles created in collisions vary in a wide range, and the energy's of particles incident in the calorimeter sections are not known *a priori*.

1.4 Outline of Dissertation

In chapter 2, I will discuss the general aspects of the Digital Hadron Calorimeter (DHCAL) protoype. This will include a description of the Resistive Plate Chamber (RPC) technology and the electronics used to read out the signal from the detector. Following this in chapter 3, I will elaborate on the various components and steps involved in constructing the DHCAL and testing during the production stage.

Construction of the DHCAL consists of painting glass, assembling RPC chambers, integrating the readout components and assembling cassettes. Then the details of the test beam facilities used to collect data will be discussed in chapter 4.

The remainder of this thesis focuses on data analysis. There is an extensive description of particle identification algorithms used to select events for analysis in chapter 5. Simulation of the detector are described in chapter 6 in order validate particle identification from test beam data. Finally the technique to improve single particle energy resolution using a longitudinal calibration is developed and applied to data and simulation in chapter 7, along with the performance of the method. The work is summarized in the concluding chapter as well as a few comments on future work.

CHAPTER 2

DIGITAL HADRON CALORIMETER

The DHCAL prototype is the first large-scale hadron calorimeter with embedded front-end electronics with digital readout. Additionally, it is the first use of pad-readout with RPCs. This chapter further develops the description on the design of digital hadron calorimetry with RPCs, the technological details of these RPC chambers, the electronics readout systems and relevant software. The design, construction and commissioning of the DHCAL are documented in [19]. Construction and quality control are deferred for further discussion in chapter 3.

2.1 Resistive Plate Chambers

The active detector technology employed in the DHCAL is a two-glass design Resistive Plate Chamber (RPC). Its basic design was originally introduced as a“Resistive Plate Counter” by R. Santonico and R. Cardarelli [20] and used plates of Bakelite instead of glass.

A typical modern RPC chamber, as shown in figure 2.1, consists of two sheets of resistive electrodes separated by a volume of inert gas with a potential difference applied to maintain a uniform electric field between the plates. The gas mixture is the source of ionization caused by through going charged particles. Surfaces on the exterior of the electrodes are coated with a resistive paint that the HV is applied to. Spacers on the interior maintain uniform gap size and gas flow (see section 2.1.1 for further details). Signals generated in the the RPC induce charge on conductive electrodes in contact with the resistive paint. These conductive electrodes are a part

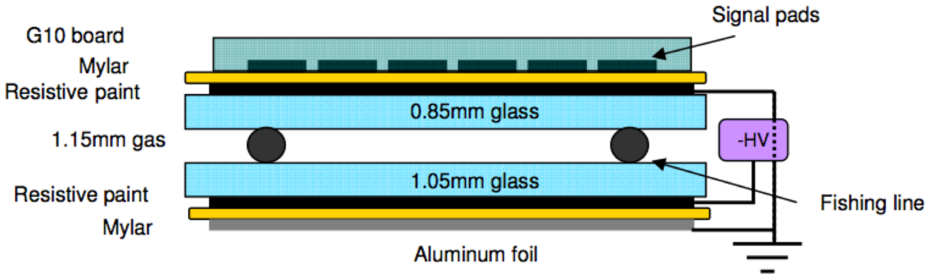

Figure 2.1. Profile of generic 2-glass design RPC. Dimensions do not reflect those used in DHCAL nor is the diagram to scale..

of the readout electronics placed on the same side of the RPC where the HV is connected to ground. Some options for conductive electrodes are pads or strips of a conductive metal such as copper or gold. Front-end electronics (signal shaping, amplification and digitization) are typically not embedded with the chambers. The DHCAL uses pads and integrated electronics for readout.

Generally, chambers can operate in either of two modes, avalanche or streamer. Streamer mode refers to a larger electric field in the gas gap that produces discharges localized near the track of an ionizing particle [21]. Streamers produce charges in the 10-100 pC (or greater) range. Some disadvantages of streamer mode for digital imaging calorimetry are as follows,

- rate problems
- large charge distributions
- after pulses
- possible aging

Avalanche mode uses a smaller potential difference and the signals recorded are due solely to avalanche multiplication from the original ionizations within the gas. This is the mode used in the DHCAL. The signals area in the few pC range, are more suitable for a PFA based calorimeter readout system and have reduced effects of aging

[22]. Rate dependence in a double glass RPC in avalanche mode is fully described in [23]. A charged particle incident on a RPC ionizes gas on the interior where it is amplified and an avalanche of charge is developed. With a uniform charge particle exposure the surface charge densities are modified generating potential differences in the gas gap and through each glass. The voltage across the glass induces a current that dissipates the charge accumulated on the inner surfaces of the glass electrodes, very close to the avalanches. Simultaneously, the potential in the gas gap drops exponentially with some time constant to find a stable value. This process affects the efficiency of the RPC as a function of integrated time exposed to a charge particle beam, shown in figure 2.2.

This shows the efficiency of RPCs as a function of time exposed to beams of muons for various intensities, 91, 346, 588 and 1795 Hz/cm^2. At the beginning of exposure RPC efficiency is relatively unaffected (except at very high rates). The efficiency drops exponentially with time at a rate that depends on the intensity of the beam The DHCAL RPCs are measured to be efficient up to $\sim 100Hz/cm^2$ particle flux. The beam intensity is kept below this value in beam tests of the DHCAL described in this thesis .

2.1.1 RPC Design for the DHCAL

In the DHCAL, a high voltage of approximately 6.3 kV is applied across the chamber via a thin layer of paint applied to the outer layers of the glass sheets. The paint is a mixture of green and black artists paint along with water. The exact mixture ratio is determined at the time of application. Paint is further described in Chapter 3.

A mixture of 94.5% Freon ($C_2H_2F_4$), 5% isobutane and 0.5% sulfer hexaflouride (SF_6) [24] is used as the inert gas. As previously stated, DHCAL RPCs operate in

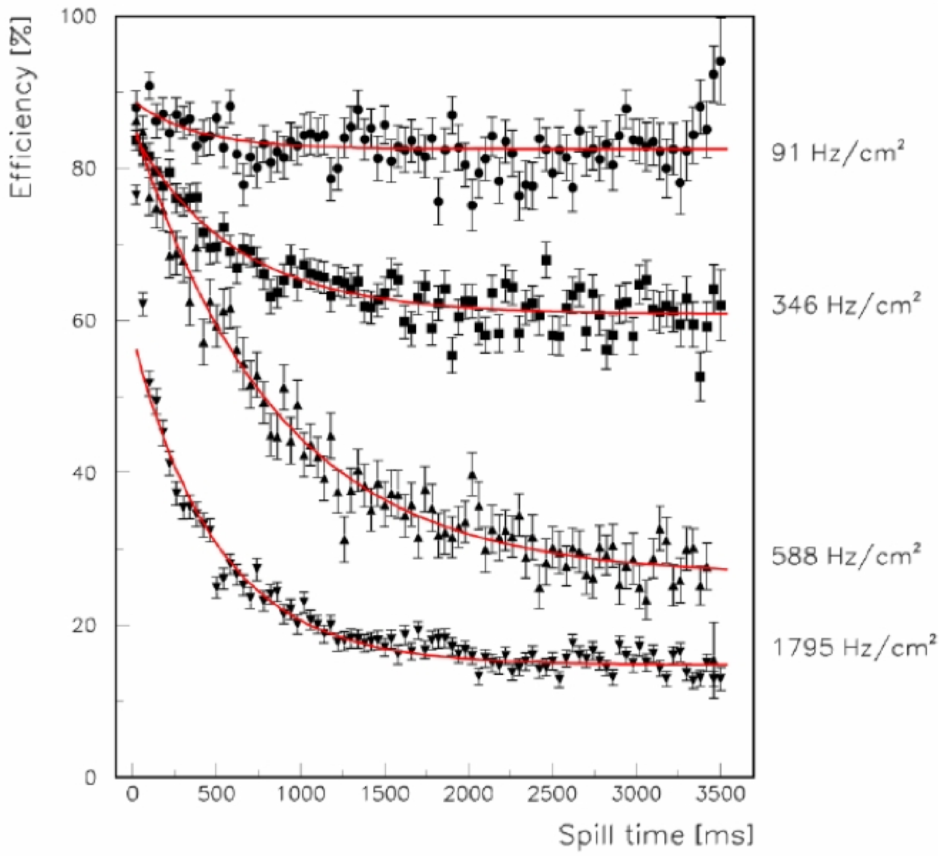

Figure 2.2. Efficiency as a function of exposure time to a beam of muons. Red lines are fits from rate calculations previously described. Each set of points is associated with a different beam intensity.

avalanche mode. This reduces the high voltage required, minimizes the multiplicity and provides opportunity to take data in a higher flux of particles.

Each RPC chamber is 32 cm x 96 cm. This size is determined by the size of the readout board of the DHCAL, and feasibility of handling large area thin glass sheets. Glass is fragile so that handling square meters requires additional complicated machinery. Paint is more easily applied to smaller areas of glass and results in more uniform surface resistivity given the method of painting for production.

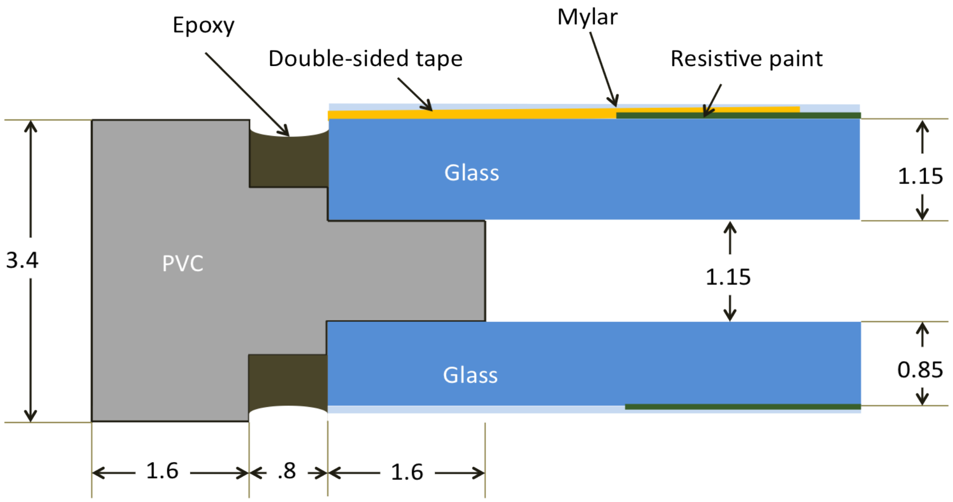

Figure 2.3. Profile of generic 2-glass RPC with details of electronics readout.

The structure of an RPC chamber consists of a frame along the border and interior fishing lines. The frame, illustrated in figure 2.3, of each chamber sets the size of the gas volume, maintains a seal for the gas and provides structural support for the chamber. The frame is a custom made with PVC material extruded to a precision 25 50 μm. There is a section of frame for each of the four side of the RPC.

The interior fishing line serves two purposes. One purpose is to maintain a uniform gap size throughout the chamber where possible bowing of the glass could occur due to the external pressure applied by the cassettes, described later in this section). This is accomplished by a thin PVC sleeve with an external diameter to match that set by the frame. The advantage in using a sleeve is that it provides a natural boundary to produce a chicane for the gas flow.

The other purpose of using fishing line is to create a path for the gas to flow uniformly through the interior volume of the chamber. Each fishing line is partially sleeved so that the gas follows a path through the chamber as illustrated in Figure 2.4. Starting from the inlet tube, gas enters the chamber and flows lengthwise toward

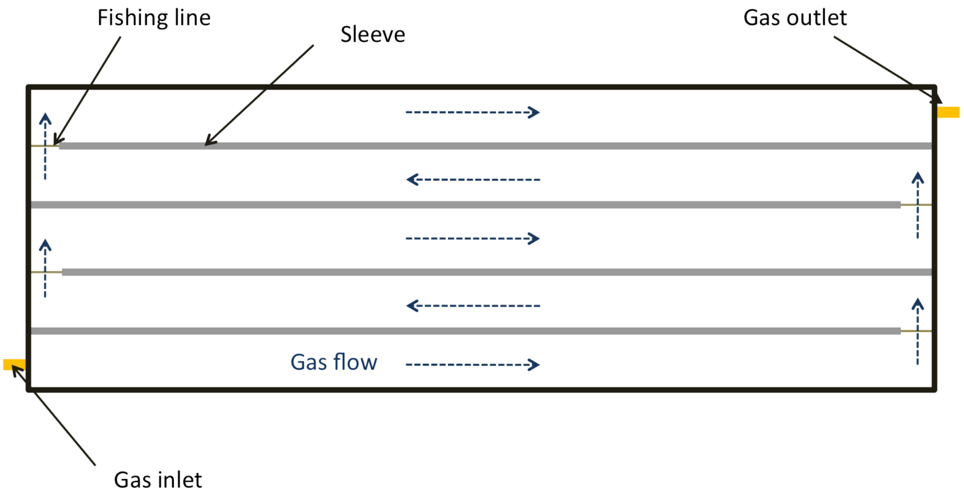

Figure 2.4. Diagram of the interior of the RPC showing the flow of gas around partially sleeve covered fishing line..

the end of the fishing line where there is no sleeve and continues in this manner until reaching the outlet tube. The pressure from the gas flow is smaller than the external pressure the the E-field when the chambers are under high voltage and from pressure applied by cassettes.

2.2 Electronics and Readout

The data acquisition system for the DHCAL consists of a front-end component and a back-end component. The front-end electronics process signals directly from the RPC chambers, collects data for ouput transmission and interfaces with the slow controls. Back-end electronics collect data from the front-end interfaces with the trigger and clock signals to the front-end and passes data to the Data Acquisition (DAQ) system. Figure 2.5 is a block diagram of the readout system. The front-end electronics is divided into two boards to avoid blind and buried vias. One component is the electronics board that contains the 24 ASIC chips, each with 64 channels. The other component is called the pad board. It contains $1cm x 1cm$ pads that are in

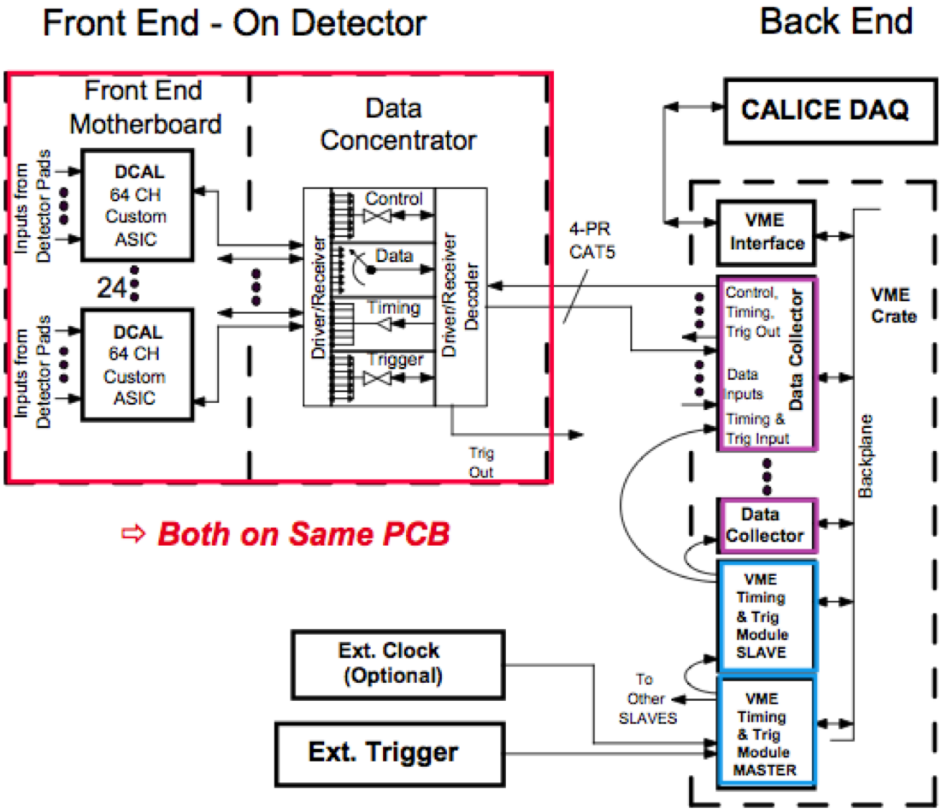

Figure 2.5. Block diagram of the electronic readout system.

contact with the thin-glass side of the RPC chamber. These two boards have smaller matching contacts that are glued together with conductive epoxy to make a single readout board referred to as a front-end-board, or FEB. Figure 2.6 shows the profile of an RPC chamber with details about the front-end electronics readout. Each ASIC chip has 64 channels for an 8 x 8 area of pads. The chips were designed by Fermi National Laboratory ASIC design group. Tests of these chips were performed at Fermilab and at Argonne National Laboratory (ANL). Each channel in the ASIC consists of input signal processing and threshold electronics. Input signal processing is performed for each of the 64 channels in either a high gain mode or low gain mode. High gain mode can be used to read out detectors with low charge signals (10-200 fC)

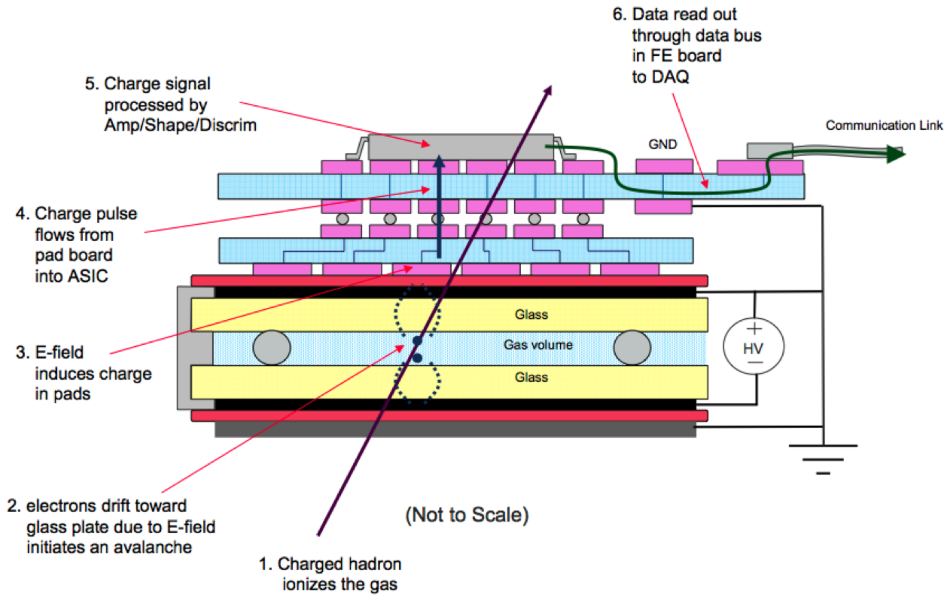

Figure 2.6. Profile of generic 2-glass RPC.

such as a Gas Electron Multiplier (GEM) or Micromegas detector. Low gain mode is used for the DHCAL since RPC produce large charge signals (100 fC - 10 pC). Threshold electronics consists of an 8-bit digital-analog converter up to 600 pC. A single threshold value is applied to each chip to determine which pads have a large enough signal to be recorded as a hit.

The FEB is currently the most complex electronics board designed by the ANL HEP electronics support group. Each front-end-board also contains a Data Concentrator (DCON) that is used to process ASIC signals. There are a total of 6 FEBs used in a single DHCAL layer and an average of 264 boards for the detector. The number of layers used depends on the test beam configuration. One data cable is needed to transfer signals between each FEB and the back-end electronics.

The back-end systems consist of Data Dollectors (DCOL) sitting in two VME crates and a Trigger and Timing Module. The DAQ computer contains all the software used to configure all detector components, collect data and store data to disk.

The DCOLs are custom VME modules designed and built by Boston University. These DCOL modules serve as a hub for front end boards. A total of 12 front end boards connects to a single DCOL module. Each VME crate holds a total of 10-13 DCOLs and so as many as two VME controllers were needed. Each data collector also forward trigger and clock signals to the front end electronics from the trigger system.

The trigger system consists of another custom VME module called the Timing and Trigger Module (TTM). The TTM receives timing and trigger systems from peripheral subsystems and interfaces with DCOLs to send data to the FEBs. Each timing and trigger module can accommodate up to 16 DCOLs. In the test beam configuration there is one TTM in each of the the VME crates. One TTM is the master that accepts external triggers and passes data to the second, slave TTM. The slave TTM passes timing and control signals to the DCOLs. The TTM can accept two inputs for a trigger signal and two inputs for additional tagging, such as that from a Cherenkov detector.

2.3 Cassette Design

Each active layer in the DHCAL consists of three chambers and their electronics readout, see section 2.2, enclosed in a cassette. A cassette consists of one large, $1m x 1m$, sheet of copper and one large sheet of steel and a rectangular rod at the top and bottom to set the interior distance. Chambers are lengthwise placed within a cassette for each layer resulting in a top, middle and bottom chamber. Two readout board are required for each chamber. The cassettes serve as a means of structural support and heat dissipation.

Since the readout boards are simply placed against the thin-glass side of each chamber, the cassettes provide a means to maintain a constant and uniform pressure

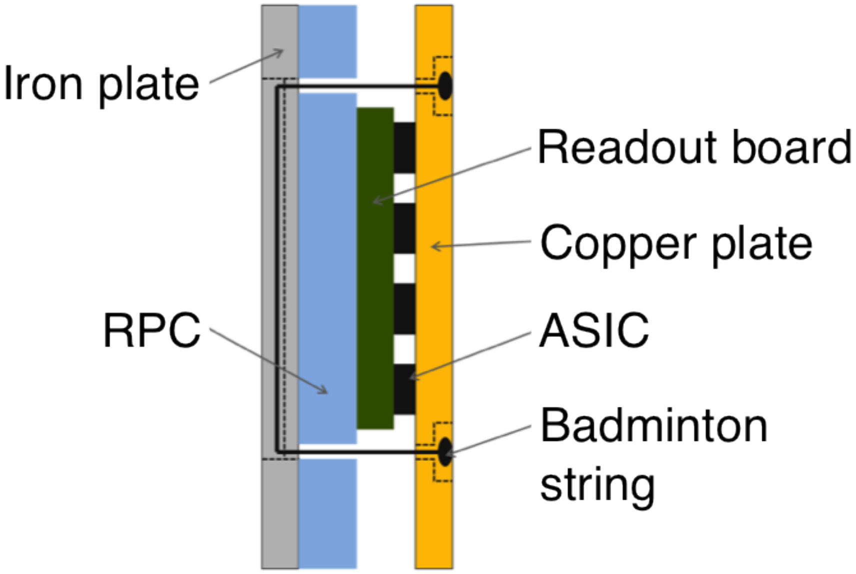

Figure 2.7. Side view of the cassette design.

to physically secure the readout electronics to the chambers. In general the cassette is also structural support for a group of three chambers and serve as an easy way to transport layers. In addition to structural support the cassettes assist in cooling the front-end readout electronics. The chips housing the amplifying electronics are in contact with the copper side of the cassette to maximize heat dissipation. Badminton strings maintain constant pressure against the RPCs and FEB. A badminton stringing machine is used to set the tension to 10 lbs. A schematic of the above description is illustrated in figure 2.7.

2.4 Software and the Event Builder

The data acquisition system (DAQ) software is based on CALICE software [25] and modified by James Schlereth of Argonne National Laboratory. Hit patterns and electronics addressing are stored in a dense binary format designed for quick data recording. In addition to hit patterns these binary files store additional information about the trigger and other detector systems.

The event builder processes the complicated binary files to write a simple ASCII [26] file containing the trigger timestamp and source as well as the hit timestamps and positions. This ASCII file contains four entries per line. Events are written serially in the order stored. The first line of each event contains the timestamp of the trigger, the interaction layer (IL and only written for simulated data), a trigger value that identifies the source of the trigger, and a place holder to identify this line in the ASCII file as the first line of an event. For data the interaction layer value is written as -1 and a value of -1 is used in the fourth position to indicate this line contains event trigger information. The lines following this event trigger information contain the timestamp and coordinates for every pad that generated a hit in the event.

The following is a simple representation of the ASCII format.

0^{th} event timestamp	IL(sim) or -1(data)	trigger value	-1
hit timestamp	x-coordinate	y-coordinate	layer number
hit timestamp	x-coordinate	y-coordinate	layer number
⋮			
⋮			
k^{th} event timestamp	IL(sim) or -1(data)	trigger value	-1
⋮			

CALICE DAQ software stores data in the following hierarchy; binary file, Record headers, Locations, detector dependent binary content. Record headers are simply a data segment to describe a following set of binary data. These include data from the triggers and DHCAL. Locations are binary data objects containing data from slots in the VME crate, which are DCOLs for the DHCAL. Finally DHCAL specific data is stored in objects called FeHitData. This is the data associated with a FEB. These FeHitDatas contain the various 16-byte data packages. Each 16-byte package contains the timestamp in 3 bytes, and the hit patterns in 8 bytes. The other bytes contain address data, error flags and check sums. There are data packages for each DCON and for each of the 24 chips on the FEB. The DCON packages contain trigger information.

The raw binary data is complicated and machine dependent requiring a complex set of software to read and process it so that an event builder is necessary. The event builder is written based on the software used to generate and process raw binary data. To convert raw binary data into ASCII files, the event builder processes data only from the DHCAL, filters out corrupt data, decodes binary hit patterns, maps addresses into position coordinates and determines and processes trigger data.

Since the binary data contains more information than needed for an analysis of DHCAL events, the event builder selects and processes the parts of the data containing DHCAL hits and trigger information pertaining to hits. Filtering corrupt data involves identifying errors and data corruption. Data corruption is identified by values in the data that refer to unphysical addresses, e.g. a DCOL or chip address that doesn't exist.

The data for a single DCAL chip includes a timestamp, electronics addresses and hit patterns. The timestamp in the DCAL data is the time that data from the chip was recorded and has a resolution of 100ns. The address information is a set of bits that uniquely identifies each DCAL chip in the DHCAL. The address classifies the location of the chip within its associated FEB, the FEB's location in the DCOL and the DCOL's location in the VME. The address of the VME is stored in the header information in the binary data. The hit pattern stores a binary code describing which pads have hits above threshold.

To decode binary hit information the event builder uses DAQ software to read the bits that hits are stored in and maps individual hits to a particular coordinate in the DHCAL. In this way a spatial coordinate is generated for each hit. Mapping hits to coordinates requires decoding the address bits. Coordinates are related to a VME crate, a DCOL, a FEB, a chip and the address for the channel in the chip.

Trigger information is stored with the data from a VME and includes signals from scintillator counters and Cherenkov detectors. The event builder generates a simple code to identify all the sources of trigger signals. The following is the definition of the trigger value store in ASCII data.

1. no Cherenkov
2. inner Cherenkov
3. outer Cherenkov
4. both Cherenkov
5. last muon scintillator, no Cherenkov
6. last muon scintillator, inner Cherenkov
7. last muon scintillator, outer Cherenkov
8. last muon scintillator, both Cherenkov

2.5 Calibration Strategies

As mentioned in 1.3 the total number of hits, referred to as nhits, in the DHCAL is assumed to be proportional to the energy of incident particles. Calibration performs two tasks. One is to convert nhits to the energy of the incident particle. A second task is to identify and correct inhomogeneity in response within and between RPCs due to various reasons such as construction variations, differences between RPC gas gap sizes, temperature, surface resistivity of the paint, gas purity and gas flow.

Particles behaving as minimum ionizing particles (MIP) traversing through the DHCAL lose a very small amount of energy in the absorber layers and produce nearly energy independent signals in the RPCs forming a track of signals. The response of a RPC to a minimizing particle is quantified by efficiency and multiplicity. Efficiency is the measure of the fraction of signals generated from particles that should produce a signal. Ideally values should be 1 but sometimes the charge generated by a traversing charged particle does not result in a hit. Multiplicity is the average number of pads that fire when only a single MIP passes.

Efficiency and multiplicity are measured in the DHCAL with test beam muons, see chapter 4. Efficiency is a ratio of the number of events with hits to the number of tracks that should produce hits. The number of tracks is determined by fitting the hits in a clean sample of muons to a straight line while excluding the layer to be measured. The efficiency is then calculated for an RPC by dividing the number of tracks with hits matching the track in that RPC to the total number of tracks going through that RPC. Multiplicity considers efficient RPC events (at least one hit is present at the track location) and calculates the average number of hits present. Perfect multiplicity has a value of 1 but realistic RPCs have multiplicities typically between 1-3.

With efficiency and multiplicity calculated in this way it is possible to correct the number of hits in each RPC to match a uniform response for all RPCs. The task of converting nhits to the energy of the incident particle is left to the analysis of pions and positrons. Equation 2.1 describes one possible strategy for calibrating nhits to an incident hadron, for instance. This particular calibration strategy, and that currently taken for data analyzed in this thesis, is previously documented in [27].

$$E_{hadron} = \alpha_s^p \left(\sum_i H_i \right) \cdot \left\{ \sum_i \frac{\epsilon_0 \mu_i}{\epsilon_i \mu_i} H_i \right\} \tag{2.1}$$

where

- i is an index running over all RPCs in the DHCAL
- H_i is the sum of hits in each RPC with values of 1(0) for pads with hits (no hits)
- ϵ_i is the efficiency of the RPC
- μ_i is the multiplicity of the RPC
- $\alpha_s^p \left(\sum_i H_i \right)$ is the sampling term associated with particle type p.

To find the sampling term, pions and positrons with interactions mostly contained in the DHCAL are needed. The sampling term for muons is essentially the multiplicity since it is gives the average number of hits per layer for a MIP and a muon is always a MIP. Since pions and positrons result in a different number of hits as a function of beam energy, they have unique sampling terms.

CHAPTER 3

DHCAL CONSTRUCTION AND TESTING

Construction, assembly, testing and analysis of a large-scale DHCAL physics prototype is a multi-stage process spanning several years and institutions. Construction began in 2008 and data analysis is expected to continue into 2014. Stages of construction of the DHCAL consists of

- glass painting
- RPC chamber and frame assembly
- front-end electronics
- cassette assembly
- back-end electronics
- low and high voltage
- gas system
- DAQ software
- event builder and display
- data analysis
- simulation

Glass painting is the process used to apply a resistive layer of paint to the surface of the outward face of each glass sheet. Frame assembly and RPC chamber assembly is the procedure to construct a single RPC chamber. In this process two sheets of glass are glued in a frame and the gas routing is completed. Assembly of the front-end-boards consists of gluing a FEB to a corresponding pad board. It is during cassette assembly that individual RPC chambers and readout boards are secured between two sheets of metal.

Testing of the DHCAL is a mixture of quality control and performance measurements. Quality control describes the standard of construction and provides a means to qualitatively define acceptance. The goal of quality control is to test various aspects of a particular construction stage. Quality control was performed after each of the aforementioned construction stages and is described in further detail in the upcoming sections.

3.1 Painting Glass

RPC chambers require a large electric field to be evenly applied throughout the gas volume. In the DHCAL this is done by applying a thin layer of paint on the outermost faces of a single chamber. This means that every sheet of glass in the DHCAL must have one side painted. To complete the painting of all glass a dedicated booth as shown in Figure 3.1 was constructed. Each glass requires about 2-3 minutes of painting.

Due to a high sensitivity to environmental conditions such as temperature and humidity painting became an empirical process. The mixture of black and green paints to water changed each day depending on the results of the previous day. A large batch of paint mixture was made to last for 3 days of painting so to account for the environmental sensitivities the first day of painting was often not successful. To minimize inefficiencies of painting only 4 sheets of glass were painted on these days. We were unable to identify environmental or weather conditions that affected each day's painting.

Another source of uncertainty was due to the sensitivities involved in using a large commercial paint sprayer. The paint sprayer controls the amount of paint

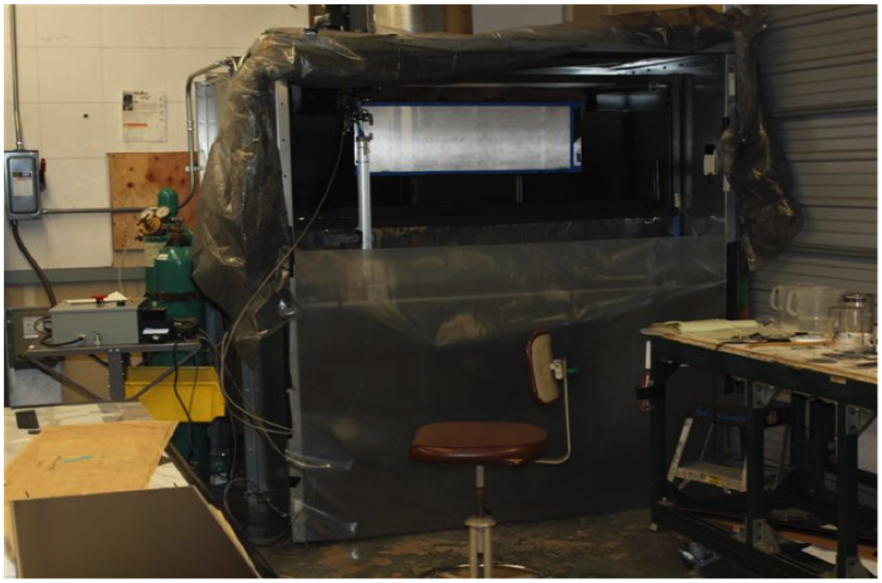

Figure 3.1. Photo of the paint booth used to apply paint to RPC chambers.

applied to a glass through control of the fluid and nitrogen flow. Nitrogen was used in place of compressed air since compressed air has an unknown amount of water vapor and alter the water-paint ratio in the paint mixture.

Paint applied to RPC glass was a mixture of black paint, green paint and tap water. Both paint components are Badger Air-Opaque paints typically used for artistic in airbrush applications [28]. Black 10-01 and chrome oxide green 10-40 are the colors from the Air-Opaque style of paint. Water is used to dilute the paint and contributes to consistent spraying with the gun.

The spray gun used to apply the paint mixture is an industrial conventional spray gun from Binks Model 2100 [insert citation] . This spray gun is selected due to its flexibility in air and fluid control, uniform results and rugged construction necessary for continuous daily spraying during the production stage of the DHCAL. The spray gun can control the mixture of paint and air via the air nozzle assembly, air adjusting valve, fluid control knob and slide port control. These controls are

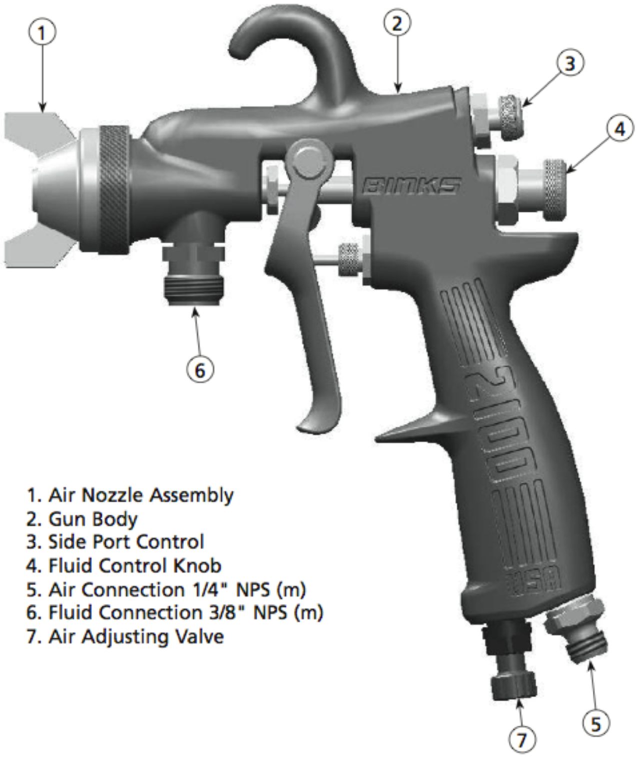

Figure 3.2. Pictture of the Binks Model 2100 spray gun with controls for painting.

illustrated in figure 3.2. To adjust the amount of paint sprayed only the fluid control knob is used.

3.1.1 Painting Procedure

To complete the painting of RPC glass the procedure consists of cleaning, painting, drying and quality control of the glass. Each glass must be thoroughly cleaned on both sides. Although only one side of a glass will be painted the other side must remain clean as it will be part of the interior gas volume of the chamber. First each side of a glass was cleaned with a soft damp lint-free cloth. Then a small amount of alcohol with another soft cloth was used for fine cleaning. After the glass is cleaned it is secured to an aluminum fixture with painters masking tape applied along the edges and placed into the paint booth and then painted by the automated system. The masking tape also defines the area to be painted which leaves about 3 mm of blank space along the edges. After the glass is painted it is taken out of the paint booth but left on the aluminum fixture. The tape is immediately removed and the edges cleaned with a razor blade to remove excessive paint. It is then allowed to sit overnight to dry. Quality control can then be complete on the following day.

Spraying glass was performed in a dedicated spray booth. The booth is an indoor ventilated structure with user controlled robots that move glass as a paint spray gun applies paint onto the glass. Glass is supported on a thick aluminum plate and fastened into the booth at the start of spraying and removed after. A paint mixing area is used to mix and weigh enough paint to spray 10 glass sheets to account for extra losses of paint during painting. The following procedure to paint glass is followed to paint 8 sheets of glass a day.

1. place a mixing magnet into the beaker containing pre-mixed paint and place on a mixer
2. check pressure levels of nitrogen tanks in the paint booth.
3. clean 4 thin and 4 thick sheets of glass. First use a clean white lint-free damp cloth on both sides to remove dirt. Then squirt some isopropyl alcohol onto one surface and wipe with a new white lint-free dry cloth to remove any oils. Place each clean glass into the storage shelf.
4. clean an aluminum backplate with alcohol and a white lint-free dry cloth
5. tape a glass sheet to a backplate with blue painters tape along each edge of the glass so that about 1/8 " of glass is covered. fasten this plate to the painting booth in the starting position. repeat taping glass to the remaining 3 backplates and place nearby with glass side facing up

6. remove pre-mixed paint from mixer and pour into air gun paint reservoir and assemble gun for painting in the paint booth. connect air supply line from nitrogen to air gun
7. turn on paint booth's ventilation fan, open nitrogen tank valve and start air gun. Power on the paint booth robot electronics and press start button.
8. When glass has been completely painted stop the air gun and turn off nitrogen tank. allow the robot to complete motions bringing glass up to starting position.
9. remove painted glass backplate from paint booth and replace with one of the pre-clean and taped backplates containing a glass sheet.
10. repeat steps 7 - 9 with new glass sheet.
11. meanwhile remove blue painters tape from previously painted glass and place glass in storage shelf
12. remove another clean glass and tape it to the same backplate as in step 5
13. repeat steps 5 - 12 until all 8 glasses are painted. note that some steps are not necessary once the last clean glass is taped to a backplate.
14. thoroughly clean all parts of the air gun with water and bristles. pay special attention to the paint passages to ensure to obstructions remain.

The following procedure is used for mixing paint during the production phase of DHCAL construction.

1. clean beakers and mixing magnets with soap and water
2. vigorously shake each gallon bottle of black and green paints for several minutes to disrupt any settling that may have occurred from storage
3. fill two 350 mL beakers with green and black paint separately.
4. fill one 350 mL beaker with clean water
5. place a clean 1000 mL beaker with mixing magnet on weight scale to measure empty mass

6. pour black paint into clean 1000 mL beaker until a mass of 288 g is achieved.
7. pour green paint into partially filled beaker until a mass of 212 g is achieved.
8. Pour water into partially filled beaker until the mass necessary to reach the desired water mixture ratio is reached.
9. place a piece of double-overed plastic wrap over the top of the filled beaker and secure with a rubber band.
10. place filled beaker on mixing machine and allow paint to mix until a homogenous mixture is achieved. stop mixing machine in no less than 5 minutes.
11. clean used beakers and equipment thoroughly

The following procedure is used top operate the painting robot in manual mode.

Horizontal Slide (Velmex)

- Press the red ”stop” button on the front of the Velmex controller (this takes the controller out of program/auto mode
- Use the ”jog” buttons on the front of the Velmex controller to move the slide as needed.

Vertical Slide

- Place the toggle switch to ”manual”
- Use the panel to control the motion

The following procedure is used to operate the painting robot in automatic mode

- Simultaneously twist and lift red ”stop” button
- Place the toggle switch to ”program”
- Press the green ”run” button on the front of the Velmex controller
- Press and hold for 2 seconds the green button on the PLC box to start the motion

- Press the red button on the PLC box to stop the motion

  Additional info:

- When the red stop button is hit the vertical motion stops immediately but the horizontal slide must complete its cycle
- The red stop button must be twisted to release. Pressing the green start button again will begin automatic motion from the previous point.
- When the vertical slide makes contact with the bottom limit switch it will automatically rise back to the home position and stop after making contact with the top limit switch.

3.1.2 Quality Control

The resulting surface resistivity mostly depends on the ratios of black and green paint with water in the final mixture but other factors contribute to fluctuations of these values. Given two mixtures of paint with the same ratios of black paint, green paint and water, the measured surface resistivity could vary at most by an order of magnitude. Additional variables monitored as possible causes of resistivity fluctuations are temperature, relative humidity, paint gun settings. The amount of visibly wet paint on the glass immediately after painting is a consequence these causes.

It was found that the amount, in units of inches, of visibly wet on the glass was the most accurate indicator of the final surface resistivity. The combination of paint gun settings and environmental conditions seemed to be coupled in such a way to affect results. The degree of evaporation is a combination of the amount of paint sprayed onto the glass and the environmental conditions. Also, given the same mixtures in paint, extremely large amounts of paint applied to glass result in less surface resistivity. This is due to density of paint on the surface. Less paint approaches the surface resistivity of clean glass which is relatively infinite. About 1 - 4" of wet paint at the top of a glass indicates a potentially successful painting job.

After 24 hours of drying a painted glass can be tested for quality control. In order for a glass to pass as acceptable it must meet requirements for surface resistivity measured in units of MΩ per square. A probe was constructed to measure the surface resistivity in a 3 in x 3 ind area of the glass. Multiple measurements were made across the surface at equidistant point. The average value of all measurements made on a single glass is taken as the surface resistivity of the glass. The root-mean-square was also recorded to determine the uniformity. Although uniformity is important it was not used as a quality control parameter because the uniformity is usually good enough.

Requirements for surface resistivity were classified according to the thickness of the glass. The resistivity limits for thin glass were 1-5 MΩ while the limits for thick glass were 1-10 MΩ. As noted earlier, towards the end of production the upper limit was relaxed to 10 and 20 MΩ respectively.

Charge spread and hit multiplicity is a function of surface resistivity. In order to get small enough hit multiplicity we need R $\geq 1\ M\Omega$, where the dependence is pretty weak.The upper limit is due to rate considerations. At high rate, current increases and we don't want a voltage drop along the paint. The surface resistivity maintains the potential difference across the chamber and defines the potential at every point on the surface. The HV was applied to the thick glass with the ground side on the thin glass.

3.1.3 Results

The average surface resistivity and RMS surface resistivity were recorded for each glass that passed acceptance. Over 700 sheets of glass were painted but due to

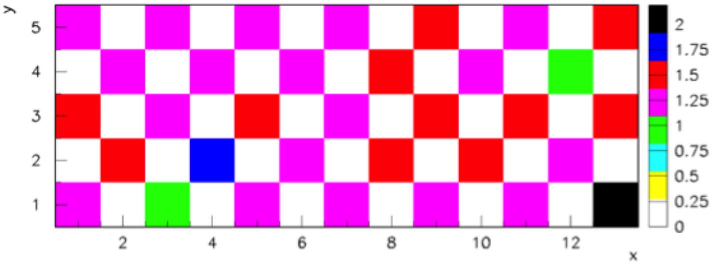

Figure 3.3. Typical surface resistivity measurements on one side of a RPC glass.

quality control standards only 415 passed resulting in 60% efficiency. The mean of average surface resistivity for accepted glass is 2.5 MΩ with an RMS of 0.2 MΩ.

Surface resistivity of the painted glass was measured in a checkered pattern over the entire surface of each glass. A total of 33 values were recorded for each glass. The tool for measuring the surface resistivity was a plastic square with two pieces of adhesive copper tape place on rubber strips located on opposite sides of the same side of the plastic square. Each copper strip has a thin wire soldered directly onto it that is then connected to a standard multimeter capable of reading resistance up to 2 GΩ. As indicated in Figure 3.3 the values on individual glasses became stable.

During the beginning phases of construction the measured values of every painted glass were recorded regardless of whether a particular glass was accepted. This was done to observe the results of various painting procedures and techniques. Toward the middle of painting only glass that was accepted was recorded. Glass that was not accepted was cleaned and repainted until it passed. This procedure is reflected in figure 3.4, showing surface resistivity as a function of glass index number. The large fluctuations for smaller index values are a result of glass that was not accepted. Only glass in or near the yellow band was accepted.

As shown in figure 3.5, most accepted glass has a surface resistivity between 1 and 10 MΩ with an average value of 5 MΩ. There appears to be no sensitivity

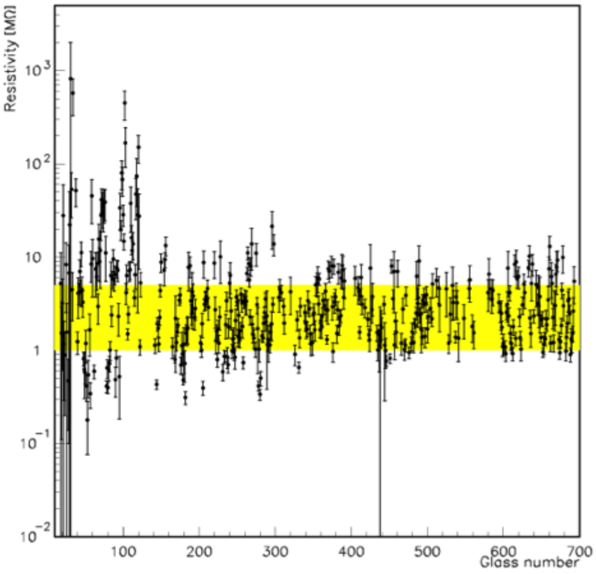

Figure 3.4. A graph of the average surface resistivity for each RPC glass with the yellow band showing the accepted region..

of the paint surface resistivity to the thickness of the glass. A correlation between the uniformity and central value is easily seen in figure, 3.6 as the larger values have a greater fluctuation. Uniformity refers to the standard deviation of all the measurements on a glass. A histogram of each glass' standard deviation divided by the mean represents the average uniformity of the surface resistivity for all the glass painted during construction. Painting was consistent and the accepted plates have less than 25% fluctuation, as shown in figure 3.7.

3.2 RPC Chamber Assembly

The assembly of RPC chambers for the DHCAL is a staged process that require coordination between technicians and implementation of quality control at each stage. Three stations were assembled for a single technician to simultaneously work on different stages of various RPC chambers allowing up to 3 RPCs to be completely assembled each day. Quality control was performed at each stage of RPC assem-

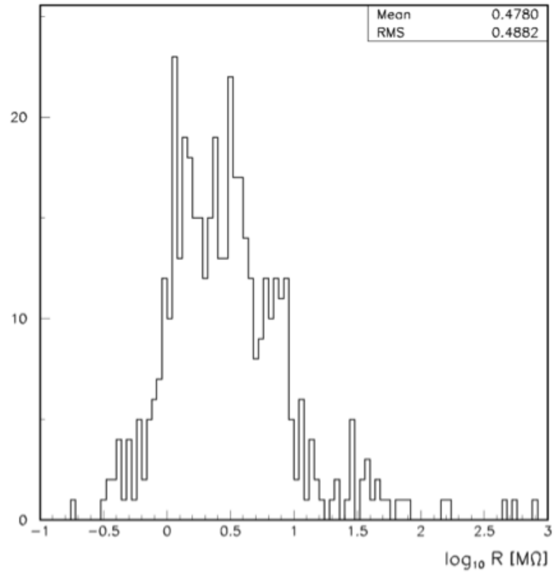

Figure 3.5. The distribution of the average surface resistivity of each painted glass sheet.

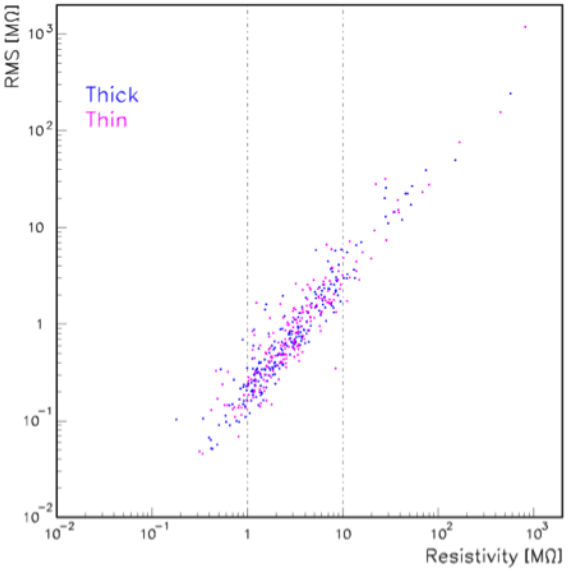

Figure 3.6. RMS vs surface resistivity for thick(thin) glass in blue(magenta). Recall thin glass is on the readout side of the chamber..

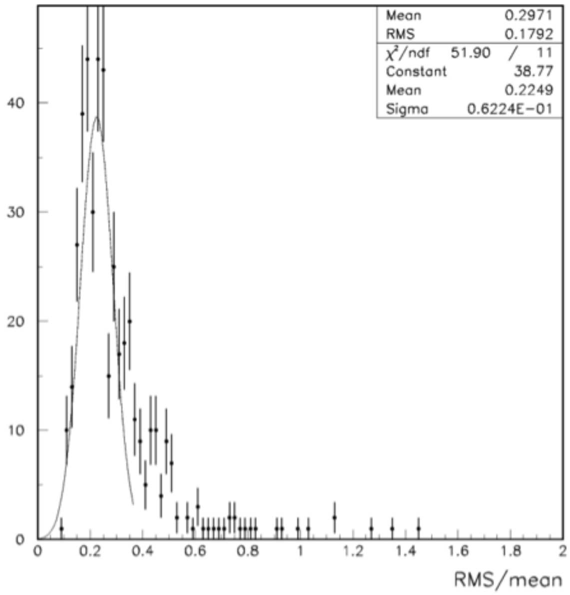

Figure 3.7. Distribution of RMS/mean values..

bly. Figure 3.8 is a photo of the assembly station used to construct individual RPC chambers. A total of 205 RPCs were produced.

3.2.1 Procedure

Assembling RPCs consist of the following steps and a photo of a single completed RPC is shown in figure 3.9

- PVC frames were cut to size with a dedicated cutting fixture. frames are glued together in a precisely machined gluing fixture
- with the glued frame in the same gluing fixture, a thick glass is place paint side up and glued to the frame with epoxy and allowed to set overnight
- Glass and frame are taken out of the gluing fixture,flipped to add four nylon fishing lines with plastic sleeves covering about 95% of the interior length
- A thin glass is placed on top paint side up and glued to frame with epoxy and set overnight

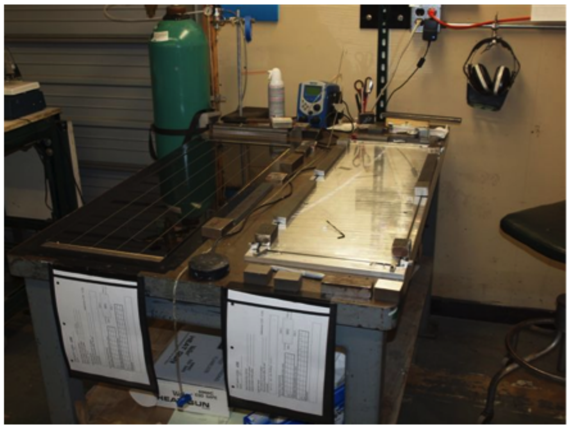

Figure 3.8. Distribution of RMS/mean values..

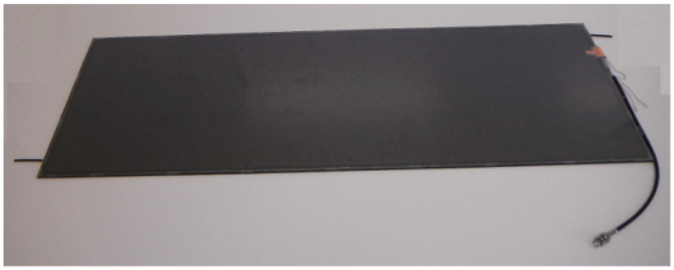

Figure 3.9. Photo of a completed RPC.

- Gas inlet tubes, high voltage mounting fixtures and fishing lines are glued in place
- HV connection and insulation are installed after leakage tests

3.2.2 RPC Chamber Testing and Quality Control

Gas tightness was tested for each constructed RPC to ensure proper gluing of the glass, gas tubes and fishing lines to the frames. A dedicated pressure test was used to precisely measure the leak rate. The RPCs were pressed with weights (to imitate the effects of being housed in cassettes and to avoid glass bowing) and slightly pressurized to approximately 0.3 to 0.4 inches of water and a decrease in pressure over 30 seconds was measured. RPCs losing less than 15% pressure were accepted.

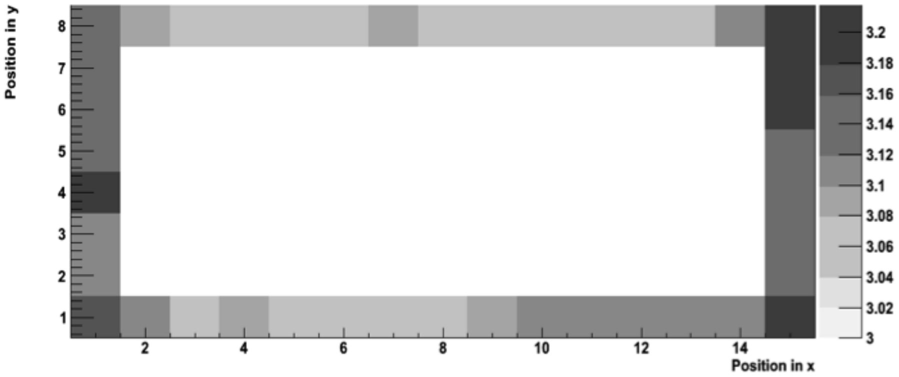

Figure 3.10. Example of thickness measurement around the frame of an RPC after construction.

In order to maintain functionality each chamber must have the following requirements

- Maintain gas flow. We have a bubbler in each gas path, which introduces 1 in water pressure. Each bubbler channel serves 2-3 RPCs so gas tightness is needed to allow output gas to go through bubbler.
- Roughly uniform gas flow to assume minimum rate of gas exchange
- Reduce gas usage

The performance of individual chambers is sensitive to the thickness of the gap between the glasses and can change within an RPC. The frame and fishing lines are responsible for defining and maintaining the size of the gas gap. One unforeseen problem is that the corner of the frames sometimes do not fit precisely together so that a slightly larger separation is created in the corners. RPC thickness at many points along the outside frame around an entire chamber are measured with a micrometer to record the fluctuations of the total chamber thickness. Chambers resulted in a fluctuation of about 0.1 mm along the sides and 0.3 - 0.4 mm in the corders. This only affected the performance of the chambers in a very small area near the corners.

Figure 3.11. Photo of the CRTS used to validate chamber assembly and measure RPC performance.

To test the HV installation each chamber was tested at 7.0 kV in a cassette-like environment, well above the operational voltage of 6.3 kV. Chambers that maintained a current less than 0.5 μA over approximately 24 hours were accepted.

3.2.3 Cosmic Ray Test Stand

The Cosmic Ray Test Stand, CRTS, is a dedicated sytsem to test the operational performance of RPC. It consists of 9 layers of RPCs each laying flat on plywood boards, one over the other approximately 10 cm away from each other. Figure 3.11 shows a picture of the CRTS.

Efficiency and multiplicity for each chamber are measured by using tracks of cosmic muons. These performance tests served as validation for the chamber construction procedures and techniques. As a result of measuring efficiency and multiplicity, a calibration factor was also calculated using the proposed calibration strategy for the DHCAL to normalize RPC response. Each of these three variable are plotted as a function of the incident angle of cosmic rays, shown in figure 3.12

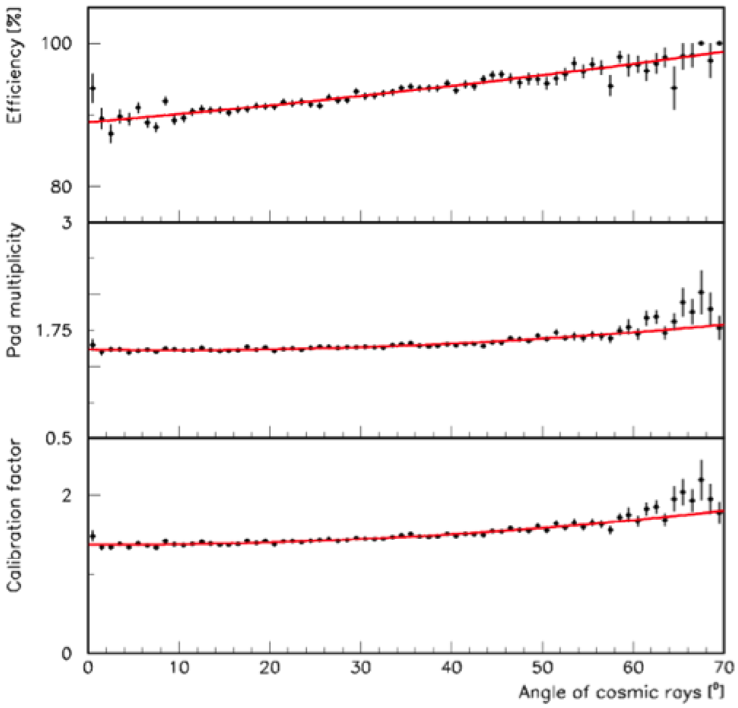

Figure 3.12. Performance results of RPCs. Starting from the top, efficiency, multiplicity, and calibration constant as a function of the incident cosmic ray muons..

Only the first few batches of RPC are tested in order to measure their performance, test electronics and validate the production procedure. The remaining RPC are tested after cassette assembly.

3.3 Front-end Board Gluing

In order to glue the pad boards to the electronics front-end board an automated glue machine was developed by the mechanical support group at ANL HEP division. The machine consists of a robot that holds a syringe of mixed glue and depresses a small dot onto each of the connection pads where the pad board and the front-end board will connect. The glue used to provide electrical connection between gluing pads on the two boards is a two part conductive epoxy, consisting of silver paste and

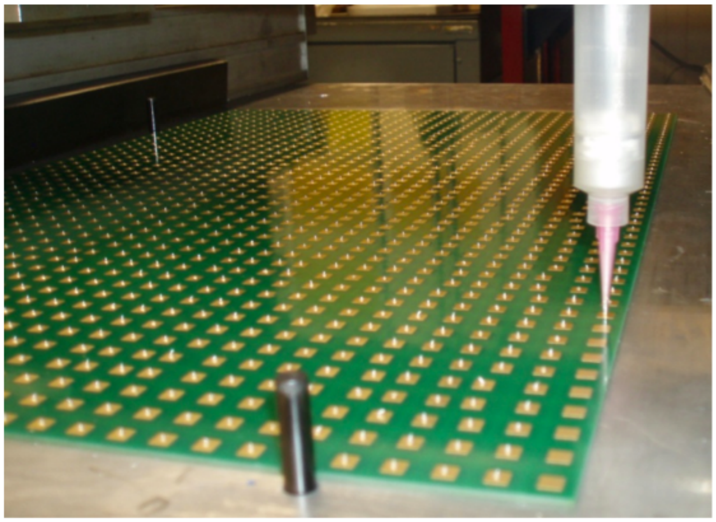

Figure 3.13. Photo of the gluing machine as it is placing placing glue dots onto the pad board..

hardener. The conductive epoxy doesn't provide good enough bonding strength, so additional structural epoxy was used to assure mechanical strength of the final glued board.

After the glue robot finished small pieces of 0.001" thick plastic film were placed uniformly between the boards as spacers to keep a small gap so that the glue didnt squeeze over to the other pads. The strong structural epoxy was then applied between pads before placing, by hand, the boards together. Immediately after placing the boards together a weight was uniformly applied and each board was placed in an oven at 40 degrees C for 24 hours to cure.

CHAPTER 4

BEAM TESTS of the DHCAL

The DHCAL collected data at Fermi National Accelerator Laboratory (FNAL) and European Organization for Nuclear Research (CERN) over multiple time periods and different detector configurations. Only data collected at FNAL is used for the analysis discussed in this thesis.

This chapter describes the beam test facility at FNAL and the available beams. A description of the setup of the DHCAL is also given. This chapter concludes with a summary of the various data taking campaigns taken with the DHCAL at FNAL and some displays of events from different types of particles.

4.1 Fermilab Test Beam Facility

Fermilab Test Beam Facility (FTBF) is located on the FNAL site in Illinois, USA on the west side of the Meson Detector Building. FTBF provides test beams for multiple types of particles in a range of energies from two main beam lines, MTest and MCenter. MTest is used for DHCAL data taking and enters from the south coming to completion within the MT6 areas of the building. Figure 4.1 shows an overhead diagram of the beam lines going into the Meson Detector Building.

Creation of the beams starts with the extraction of 750 keV H-minus ions into the Linac where they are accelerated to 400 MeV and extracted to the Booster Accelerator. Electrons are stripped off during injection into the Booster leaving 400 MeV protons. The Booster accelerates captured protons to 8 GeV in 1 batch of 84 bunches each 19 ns long. A full batch contains approximately 2×10^{11} protons and is 0.2 -

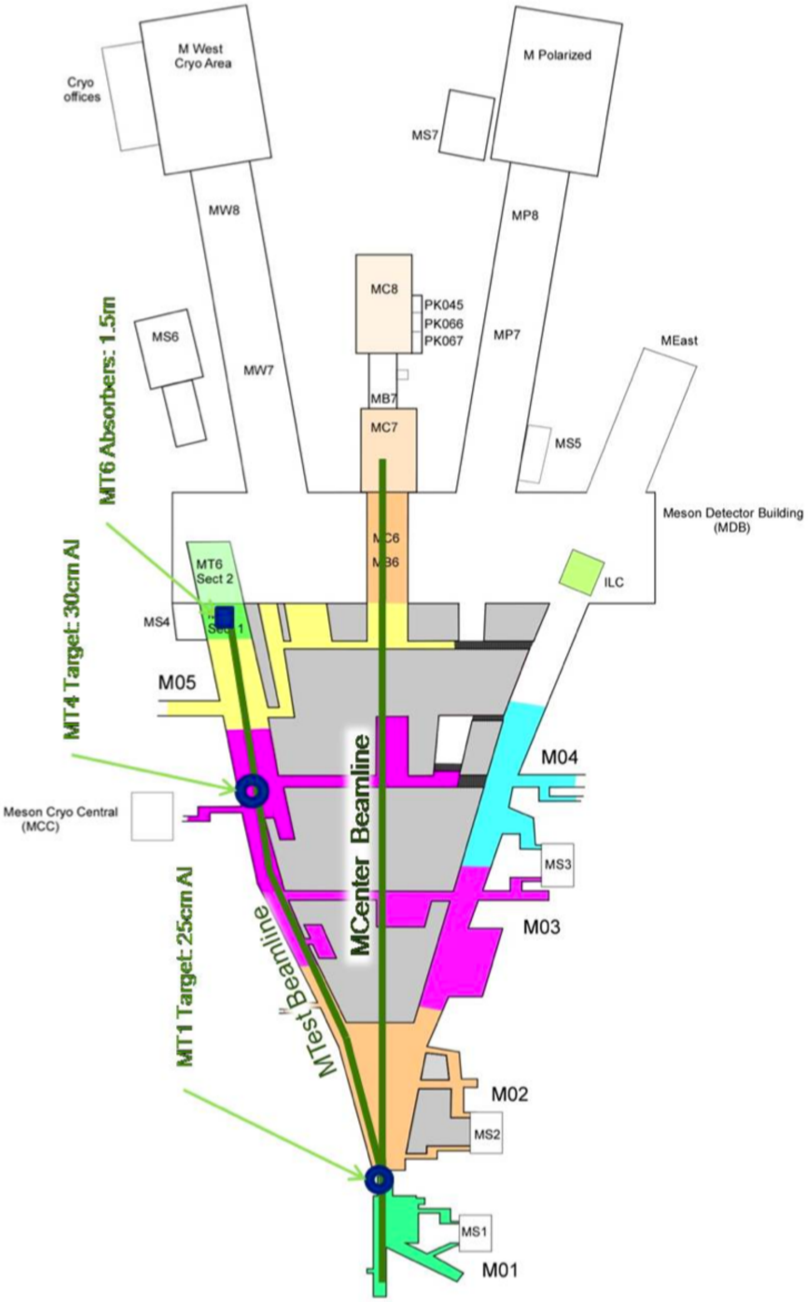

Figure 4.1. A diagram of the available beam lines showing MCenter and MTest.

1.6 μs in length. In a process called partial-batching, 8 - 30 bunches are extracted to the Main Injector (MI) for operation in the test beam. The MI accelerates this beam to 120 GeV at 53 MHz. A fraction of the beam is resonantly extracted each rotation over 4.2 s to the Switchyard (SY120) defining the duty cycle for test beams. At SY120 a Septa Magnet splits the beam to the meson lines, MCenter and MTest. MTest concludes in the MT6 areas where the DHCAL is located.

Some consequences of extracting the beam from the MI affect the beam composition. A smooth extraction would provide a single particle per MI RF (radio-frequency) bucket per rotation. But for intensities up to 100 kHz, double occupancy occurs 35% of the time resulting in two particles per rotation. This means that two particles, within 19 ns from each other, can generate triggers in the DHCAL. Double occupancy is accounted for in the trigger electronics and offline event selection so that this is a negligible effect for the analysis.

The MTest beam line, in the MT6 areas, consists of a 120 GeV/c proton primary beam. This beam can be targeted to create secondary particle beams consisting of pions, muons, and electrons between the energies of 1-66 GeV. The secondary beam can be tuned and configured using several targets, absorbers and collimators to achieve pions ($\pi^{\pm}$), electrons ($e^{\pm}$), or broadband muons at single energies in the range 1 - 66 GeV. The beam can be selected for positively or negatively charged particles. A momentum spread of 2 - 2.5% is reported for a wide variety of tunes [29]. Various triplet, doublet and dipole magnets downstream of the targets are used to focus the beam, select beam momentum and fine tune the position.

Downstream and upstream refer to the direction of the beam. Downstream is in view of the direction of the beam, away from the Main Injector. Upstream is in view opposite the direction of the beam.

In the MTest beam line, there are two targets, a collimator and two absorbers as follows:

- MT1 Target: 435 m upstream (before) of MT6; 25 cm Aluminum (Al)
- MT4 Target: 145 m upstream of MT6; 30 cm Al
- pinhole ($1m \times 1m$) collimator (MT3CH) in MT3
- two absorbers, 1.5 m long, at the end of MT6.2 enclosure (F:MT6AB1,2)

The following beam modes are generated implementing various configurations of these instruments and momentum tuning with beam line magnets;

- Proton mode (Primary): 120 GeV protons - MT1 is IN, pinhole collimator and MT4 is OUT
- Pion mode: 1 - 66 GeV MT1 is OUT, MT4 is IN, momentum tuned
- Muon mode: absorbers are IN; typically using 32 GeV Pion mode

Pion mode is tuned for single particle energies in the range described above. Table 4.1 lists some of the tunings and associated purities, spot size and momentum spread where applicable [30]. Spot size is the cross-sectional spread of the beam. Purity is the percentage of pions in the beam. The remaining particles in the beam are electrons (or positrons) and muons. In muon mode there is an undefined spot size due to scattering in the 3 m of absorber. Its broadband momentum is due to energy loss of the pion beam through the absorber and the mixture of the pion beam.

4.2 DHCAL Setup at FTBF

The DHCAL prototype at FTBF consists of the RPC cassettes installed in the Main Stack (MS) and the Tail Catcher and Muon Tracker (TCMT). The MS has 38 layers of 17.4 mm thick $1m \times 1m$ wide S235 class steel separated 3.17 cm apart. Two sets of steel plates comprise the 14 layer absorber structure of the TCMT; eight 2.54 cm thick plates followed by six 10.00 cm thick plates. The TCMT is 40.1

Table 4.1. Data table describing a few available tunings at FTBF. Data provided by FTBF.

Energy	Mode	Purity	Spot Size	Momentum Spread
120	Proton	100%	6 mm	-
16	Pion	87%	10 mm	-
8	Pion	55%	12 mm	2.3%
4	Pion	31%	13 mm	2.7%
2	Pion	<30%	13 mm	2.7%
1	Pion	<30%	-	2.7%

Table 4.2. Summary of absorber structure including interaction properties.

Structure	Number of Layers	Radiation Length in X_0	Pion Interaction Length in λ_I
Main	38	48.0	4.50
TCMT	14	43.8	3.90
Total	52	91.8	8.40

cm downstream from the MS to allow for rotation of the CALICE Movable Stage. Thickness refers to the direction parallel to the MTEST beam line.

The CALICE Stage is a moveable structure offering horizontal, vertical and rotational movements of the MS. These movements allow for alignment with the beam line and test of the DHCAL for non-perpendicular incident particles and shower development. Figures 4.2 and 4.3 show the DHCAL in the MS and TCMT, respectively. Table 4.2 lists properties of these absorber structures including thickness in terms of radiation and pion interaction lengths with RPC cassettes installed.

4.3 Test Beam Campaigns at FTBF

The DHCAL collected data at FTBF in five separate test beam campaigns using the primary proton beam, secondary pion beams and broadband muon beam.Table 4.3 briefly accounts the various run periods, configurations and statistics of events

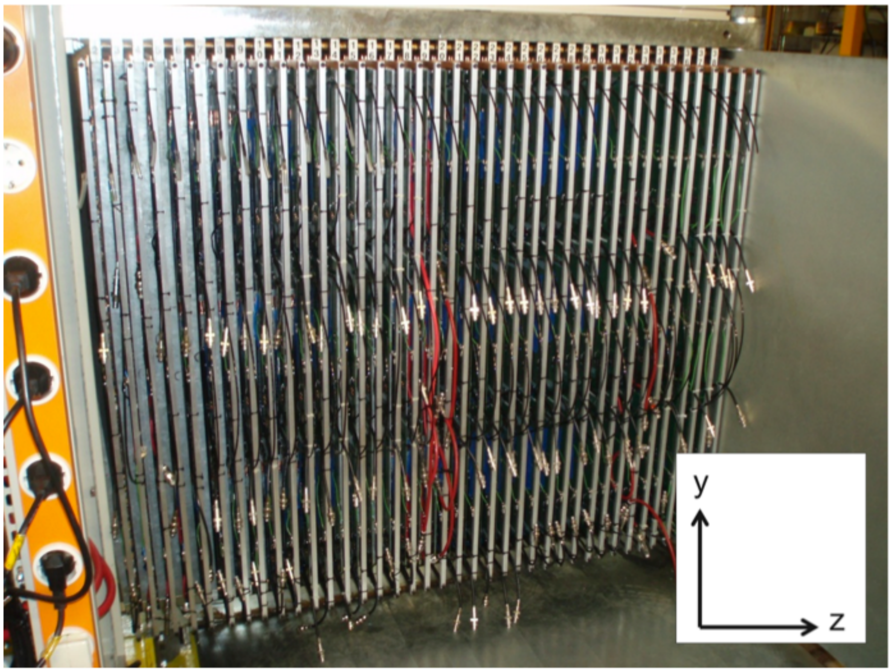

Figure 4.2. The main structure provided by CALICE at FNAL with cassettes installed before cabling.

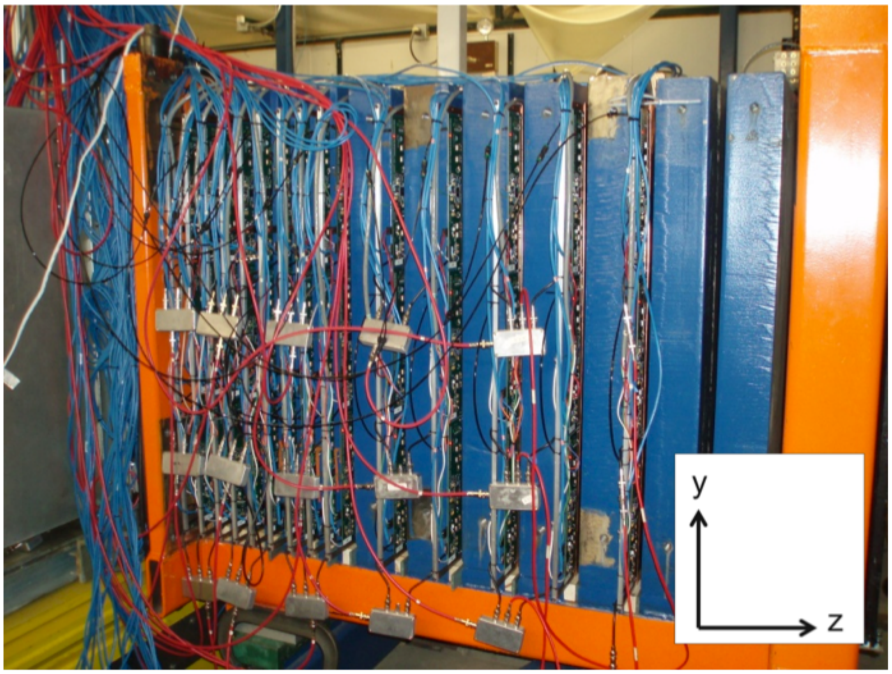

Figure 4.3. The TCMT with RPCs installed and instrumented to 14 layers.

Table 4.3. Summary of beam test campaigns at FNAL. Analysis described in this thesis did not use data from April or November.

Run period	Configuration	Detector layers	Collected μ events	Collected secondary events
October 2010	DHCAL	38	1.4 M	1.7 M
January 2011	DHCAL+ TCMT	38 + 13 = 51	1.6 M	3.6 M
April 2011	Si-Tungsten ECAL+ DHCAL+ TCMT	30 + 38 + 14 = 82	2.5 M	5.1 M
June 2011	DHCAL+ TCMT	38 + 14 = 52	3.3 M	2.7 M
November 2011	Minimal absorber	50	0.6 M	1.3 M
Total			9.4 M	14.4 M

collected during DHCAL testing at FNAL. Analysis described in this thesis uses data from October 2010, January 2011 and June 2011 campaigns to study a longitudinal calibration technique with pions and positron from the secondary pion mode beams.

Čerenkov detectors are also available at FTBF for use in the trigger system. These reside upstream from the DHCAL in the MT6.1 enclosure as shown Figure 4.4. Two Čerenkov signals provide a signal to the TTM that is recorded in the trigger information in the raw binary data. The offline event builder reads this data to tag events belonging to electrons or positrons.

The broadband muon beam is used for calibration, tuning RPC simulations and alignment. Two large $1m \times 1m$ scintillators, provided by the CALICE group at DESY, one upstream and one downstream of the DHCAL trigger muons in the broadband muon beam. Figure 4.4 also shows the location of the two 1.5 m (3 m total) absorbers placed in the beam line for a broadband muon beam. The muon beam was created from colliding the 32 GeV/c pion mode with the 3 m absorber

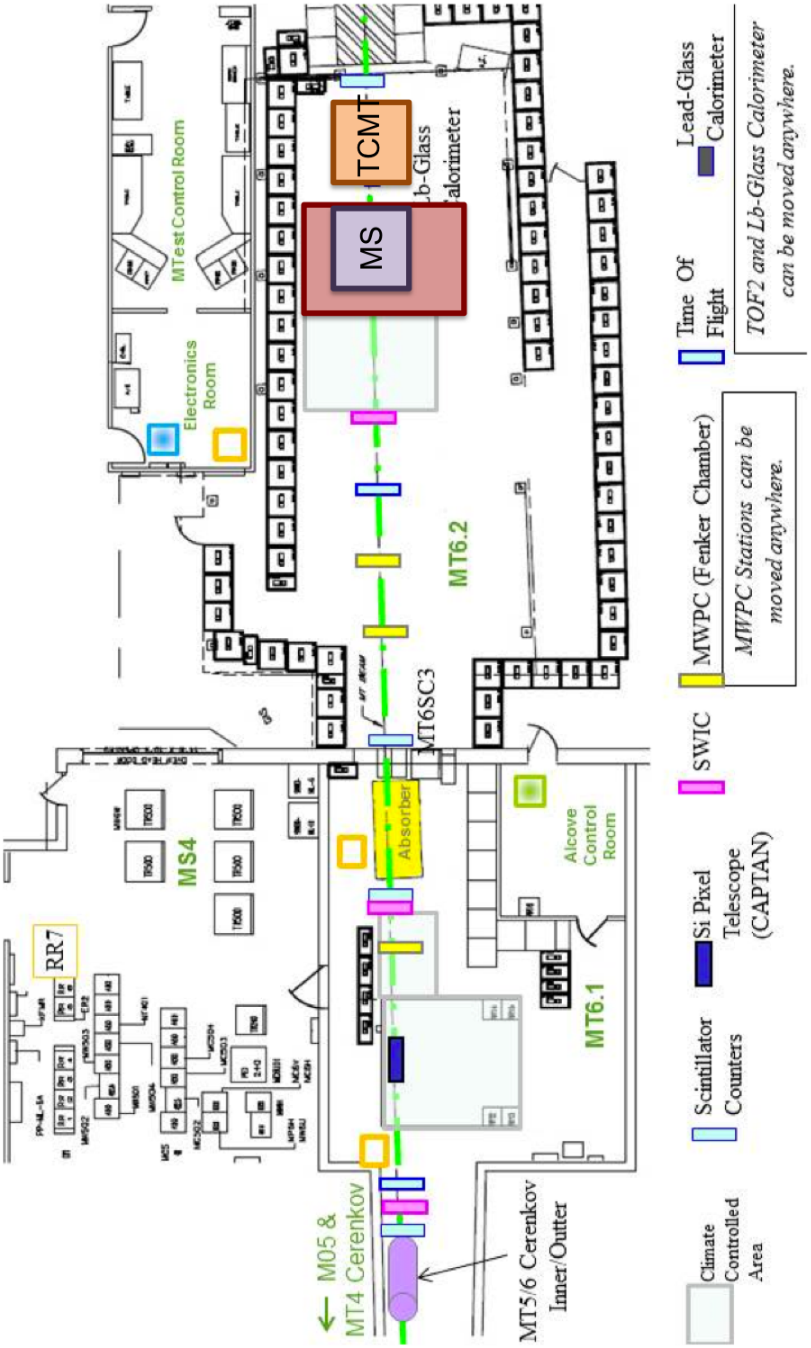

Figure 4.4. A schematic map of the MTest beam line. The MS and TCMT are in the downstream part of MT6.2.

along the beam line. The DAQ rate recorded approximately 500 - 1000 muons over a 4.5 s spill.

The coordinate system used is a slight modification of CALICE standard. Positions of hits are (x,y,layer number). Layer number is synonymous with the z-axis. Facing downstream (in view with the direction of the beam), the origin is located in the first RPC layer, at the bottom right pad, as 0,0,0. The z-axis (i.e. layer number) is parallel to the nominal beam direction. The y-axis points upward and the x-axis completes a right-handed coordinate system, i.e. it points left when view downstream.

Figure 4.5 is an event displays illustrating the imaging capabilities of the DHCAL. Note absence of isolated hits separate from the muon track. The implication of this absence is highlighted in event displays of pions and positrons.

Alignment of the DHCAL consists of correcting the (x,y) coordinates of the hits throughout the MS due to misalignment of the cassettes and FEBs. The resolution in x is corrected to 52 μm and to 14 μm in y. Since the FEB and cassettes slide along the x direction individually but the entire MS is supported in y by the CALICE movable stage, the resolution in y is much better. Tuning of the RPC simulation with muon data is discussed in Chapter 6. A general calibration scheme employed to smooth the response of RPCs in the DHCAL can be found in Section 2.5

The secondary pion mode beam provides positively charge pions and positrons. Since the beam spot size for this mode of operation is typically less than 15 mm the large $1m \times 1m$ scintillators are not used to trigger. To trigger the DHCAL in the pion mode, two $19cm \times 19cm$ scintillator paddles are placed nearly 30 cm apart along the beam line upstream from the DHCAL. A coincidence of these signals is the input to the TTM. The TTM also collects signals from the Čerenkov and the rear 1 m x 1 m scintillator paddle to tag positrons and muons, respectively.

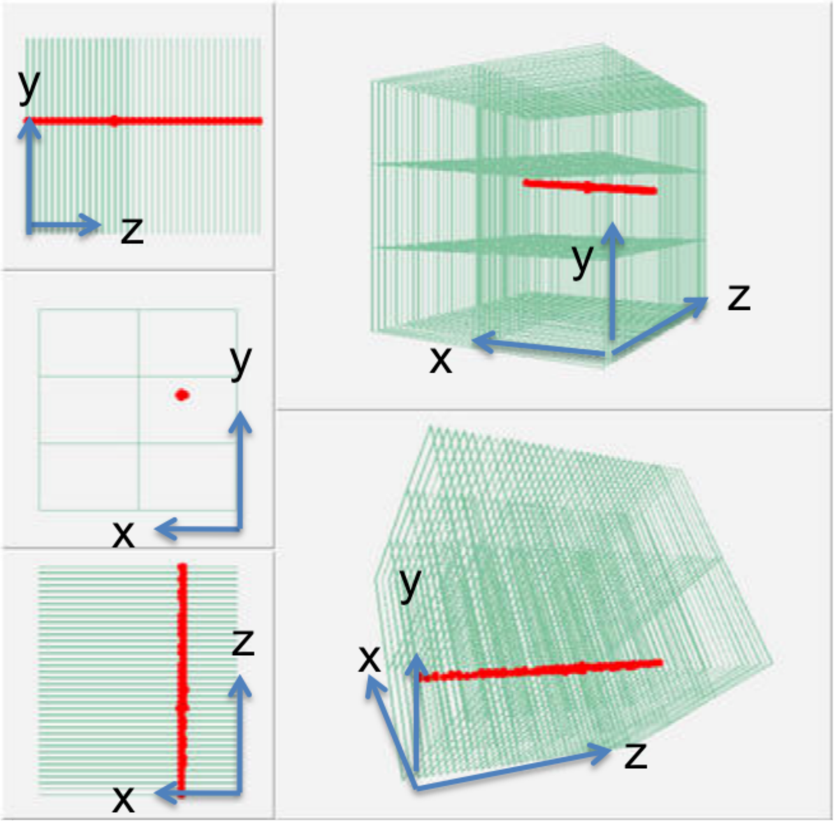

Figure 4.5. Event display picture for a typical muon in the DHCAL.

An event display picture is shown for pions and positrons in Figures 4.6 and 4.7, respectively. An absence of isolated hits in muon events implies that isolated hits in pion mode events are part of the event, possibly low-energy neutrals (photons or neutrons). These are the first large scale digital pictures of hadron showers in data.

The long track in the first few layers of the pion event is the pion behaving as a minimum interacting particle before interacting. The first layer in the shower is called the interaction layer. Off-line algorithms identify this layer as discussed in Chapter 5. Positrons interact in the first layer and only extend a few layers deep. Notice how the shape of the positron shower is defined better than the pion shower. A review of many pion showers reveals that the shape of pion showers have a high degree of fluctuation. Conversely, positron showers do not differ as much in shape from event to event. Also notice that density of hits between the two types of events is very different. Topological particle identification, also in Chapter 5 takes advantage of the various differences in shower shapes.

Data from the pion mode is the primary mode to validate digital hadron calorimetry for future high energy physics experiments. Hadronic shower shapes are studied for protons, pions and positrons. The energy linearity and resolution of the DHCAL are measured using the pion mode beams. This data provides the first opportunity to develop analysis tools for large scale digital calorimeters. Pion and positron data are also useful for tuning RPC simulations. The ultimate goal is to validate the DHCAL concept and measure hadronic showers in great detail.

Particle identification and event selection using the properties of hadronic and electromagnetic showers are described in Chapter 5. A first attempt to improve the energy resolution of the DHCAL, a step towards validating the DCHAL concept is described in great detail in Chapter 7.

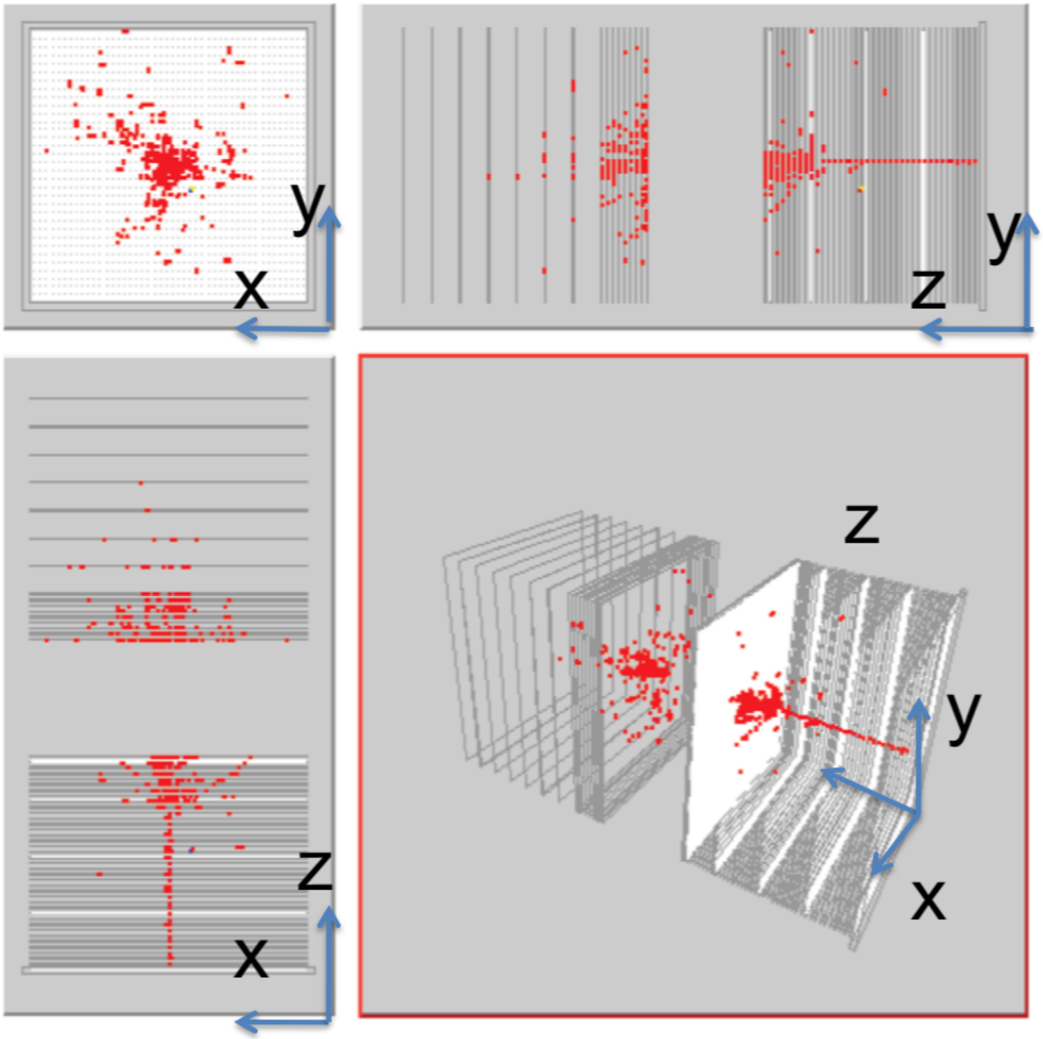

Figure 4.6. Event display of a 60 GeV pion in the DHCAL and TCMT. Red points are hits, green lines are the RPCs.

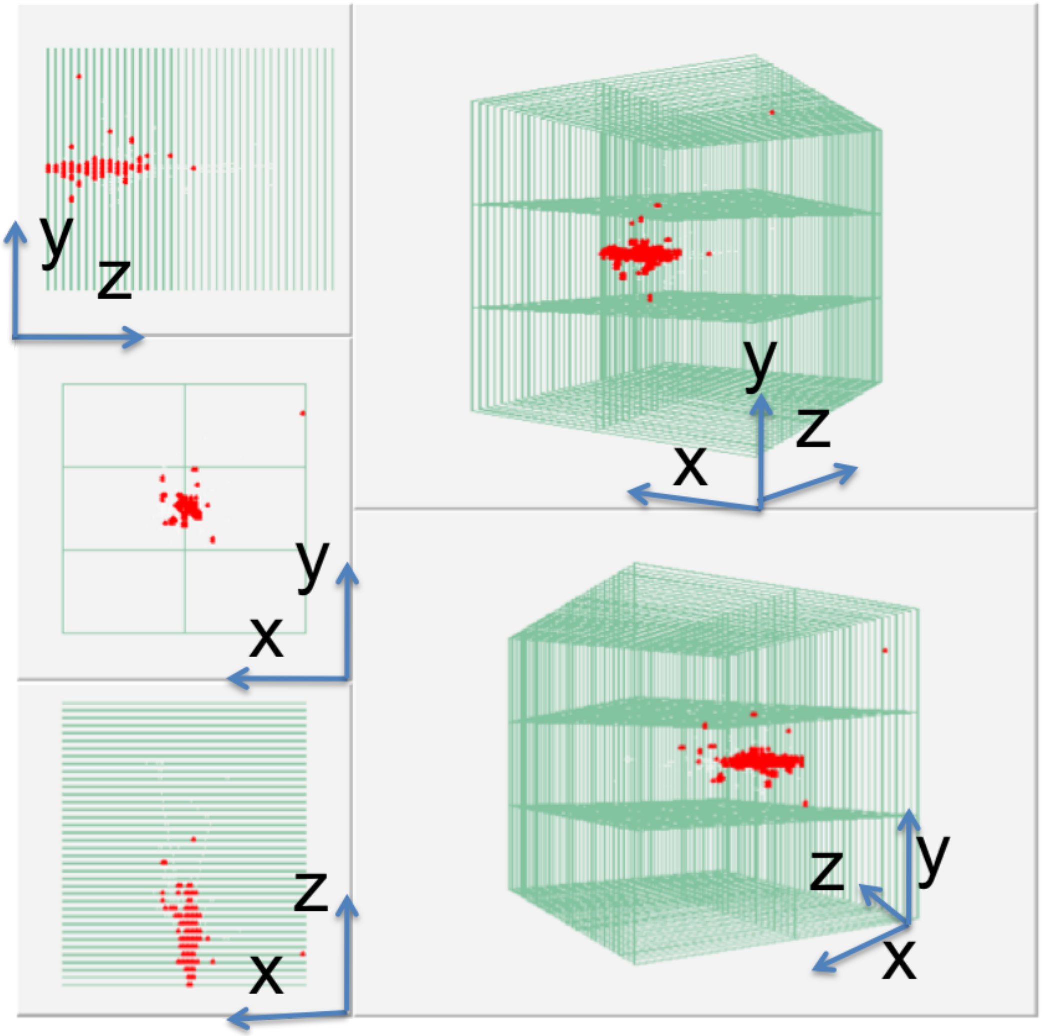

Figure 4.7. Event display of a 20 GeV positron in the DHCAL. Red points are hits, green lines are the RPCs.

CHAPTER 5

PARTICLE IDENTIFICATION AND EVENT SELECTION

Particle identification (PID) is the software classification of events as pion, positron or muon. Event selection identifies events for analysis. PID is most important for secondary beam runs since there are various mixtures of pions, positrons and muons in the beam. Although a Čerenkov detector is used in the trigger system to tag a particle as a positron, it was very inefficient for most runs or not working at all. Therefore, the PID is essential as the primary means to identify the positrons. Event selection pertains to removing events that are neither pions, positrons or muons. The events not passing the selection criteria are un-useable for analysis due to noise, too few total hits or early interactions.

Another component of PID is leakage. Leakage refers to the loss of energy when the shower extends beyond the last physical layer of the DHCAL and is discussed in further detail in section 5.3. So the role of leakage PID is to identify and reject those events with an unacceptable amount of leakage.

5.1 The Local and Global Response Variables

Response variables are measurable quantities in the detector resulting from test beam particles' interactions in the DHCAL. In general, local variables are those associated with individual events, and global variables for a set of events. Practically most variables are calculated for each event, and are local. Input to both local and global response variables can come from a single test beam run at a single energy, a set of runs at the same energy or a set of runs in a set of energy ranges.

5.1.1 The Interaction Layer

The total number of hits, nhits, produced by a showering particle that deposits all or most of its energy in the DHCAL is proportional to its initial energy. The first active layer that a particle interacts hadronically (with a nucleus) and generates secondary particles is the interaction layer (IL) or shower start. Before initiating a shower, charged particles lose energy through ionization in the gas gap of the RPCs. A particle that primarily loses energy through ionization is called a minimum ionizing particle (MIP). The layers before the IL are referred to as a MIP stub.

Muons do not have an interaction layer because they behave like a MIP and don't initiate showers as they pass through the DHCAL. As a result, the total number of hits produced by a MIP does not represent the initial energy of the MIP. On the other hand, pions and positrons generate showers and have an interaction layer. Positrons interact electromagnetically in the DHCAL and generate showers very early so that the interaction layer is usually the first layer. The interaction layer for pions, and other hadrons, fluctuates between events. Before the start of the shower pions behave like MIPs. The result is that muons, positrons and pions each have a different sampling fraction, the ratio between nhits and initial energy. For the DHCAL, sampling fractions are given in units of nhits/GeV.

For this analysis, interaction layer is reconstructed as the first of two consecutive layers containing no less than 4 hits each. To verify this definition, the difference between the known interaction layer of simulated pions and the reconstructed interaction layer is calculated and stored into a histogram. Different definitions are tested by minimizing the RMS of the resulting distributions, where the RMS represents the uncertainty to correctly identify the IL. The definitions differ by varying the number of hits required in two consecutive layers. This work was originally done in [31] and

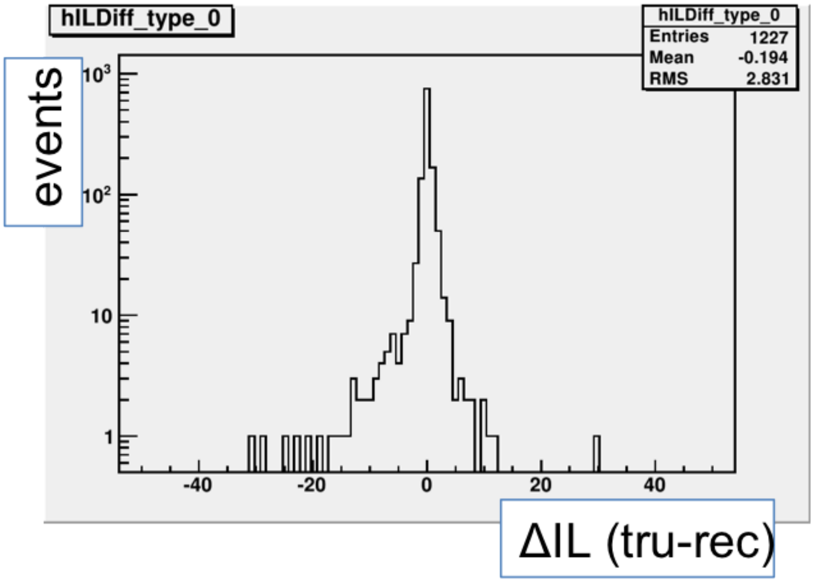

Figure 5.1. Histogram of the difference between true IL from simulation (tru) and reconstructed IL from data(rec) using 20 GeV pions. RMS of reconstructed IL in data is 2.8 layers.

again repeated by the author as shown in Figure 5.1. Misidentification of the IL is a source of systematic uncertainty due to particle identification.

5.1.2 The Hit Density

Hit density is defined for each hit in an event as the sum of hits surrounding it. The volume that hits are counted in has a volume of 5 cm x 5 cm x 5 layers, centered on each hit. The range of values for hit-density is 0-124 hits since the central hit does not contribute to the density. The unit used for density in this definition is number of hits. For use as a local variable, an average density, defined as a statistical average of hit densities, is calculated for each event. Positron events typically have a higher average density than pions so that this variable is used as input for more complicated particle identification algorithms. Figure 5.2 shows an example of average density for pions and positrons at 8 GeV. The sample of events used in Figure 5.2 is a mixture

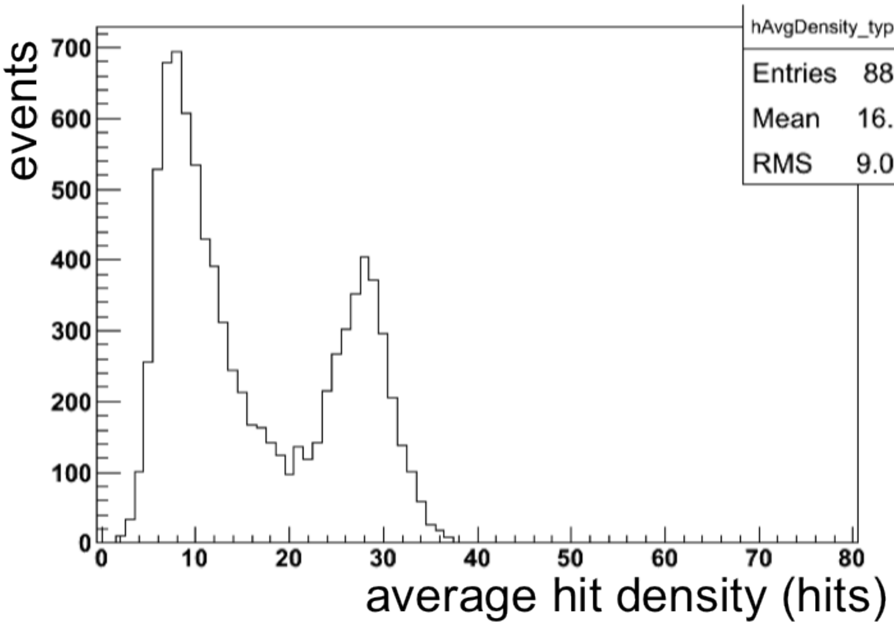

Figure 5.2. Histogram of average density for an 8 GeV set of pions and positrons. Positrons occupy the right peak while pions occupy the left peak.

of pions and positrons. Pions populate the first peak on the left, near 10 hits, while positrons populate the second peak near 28 hits.

This definition of density takes advantage of the DHCAL's response to positrons. Positrons have a higher average density because the secondary particles produced in a shower can be close enough together so that multiple ionizations are produced within 1 cm from each other. Readout pads with 1 cm x 1 cm size cannot precisely resolve the numbers of ionizing tracks in electromagnetic interactions resulting in fewer total hits. This means that there are more hits for positrons than pions, within a given unit volume.

5.1.3 The Number of Active Layers

Another useful local variable is the number of layers containing hits, or active layers, AL. It has been observed that positrons deposit most of their energy early in

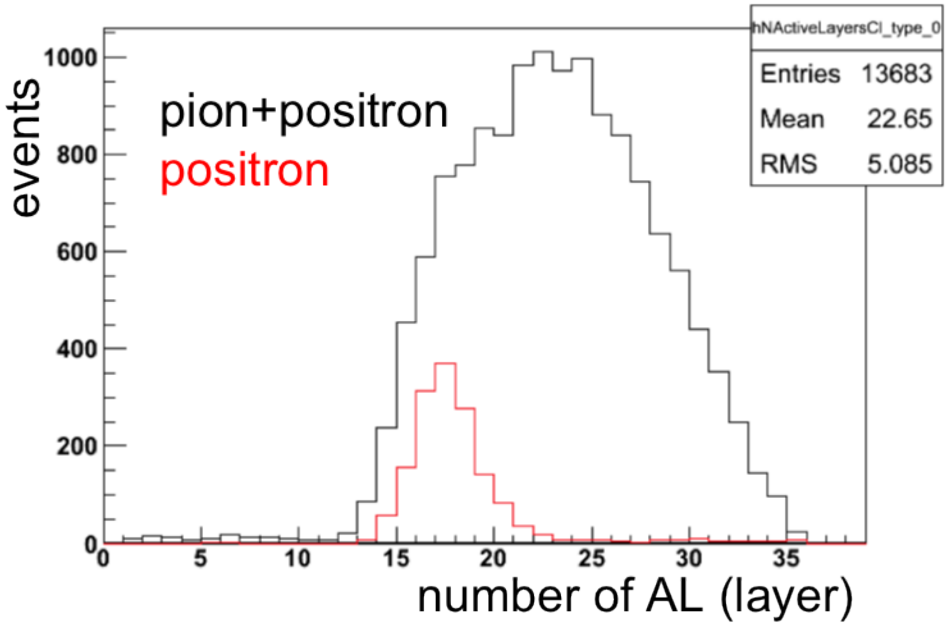

Figure 5.3. Distribution of number of active layers, nactivelayers, for a sample of pions and positrons (black) and positrons (red, unnormalized). The bump on the left of the black histogram are the positrons in the pion sample.

the calorimeter in denser showers. Conversely, pions are less dense longitudinally and therefore occupy more layers. This variable is a simple counting of the layers that have hits. The number of active layers is an input to the particle identification algorithms. An example of active layers is shown in Figure 5.3. The large, wide histogram in black is a mixture of pions and positrons at 25 GeV. The smaller red histogram is just positrons. The advantage of using this variable in the PID to separate positrons and pions is that more than half of the events, above 22 layers in this example, in the mixed sample are only pions.

5.2 Particle Identification

Particle identification is described in further detail in this section to distinguish positrons from pions. The strategy of the PID used in analyses discussed in this thesis is to first identify events as either muons or positrons. The remaining events are taken as those of pions. Additional event selections to remove noisy events are applied after PID.

PID Efficiency is defined as

$$\epsilon = \frac{N_{gen} \cap N_{id}}{N_{gen}} \tag{5.1}$$

and purity is defined as

$$P = \frac{N_{gen} \cap N_{id}}{N_{id}} \tag{5.2}$$

where N_{gen} is the number of events generated and N_{id} is the number of events identified by the PID. Efficiency is a measure of the number of events identified from the provided sample. On the other hand, purity is a measure of how many events in the sample are the true type of particle desired.

5.2.1 Muons

Muons have a consistent response in nhits throughout all beam energies. The difficulty identifying muons occurs in low energy tirtiary beam runs when it is necessary to remove muons from the pion and positron samples. For beam energies under 6 GeV pions and positrons have a response very near to that of muons. Some events have a MIP stub Any reference to the muon or MIP like properties of pions and positrons is left for later in this chapter.

Two local variables are used to distinguish muons, interaction layer and the number of hits in each layer. Muons do not initiate particle showers in the DHCAL. As a result, an IL is not found using the algorithms described in Section 5.1.1. If an IL is not identified, then the event is label as a muon. This step alone distinguishes muons that may be in the pion and positron samples. In order to clean the muon sample further, the number of hits in each layer must be less than 4. This limit on the number of hits in each layer is justified because muons are MIPs in the DHCAL and should on average generate a number of hits equal to the average pad multiplicity of approximately 2.

5.2.2 Positrons

Identification of positrons is necessary in offline analysis of FNAL beam data because there was no reliable external tagging of positrons in the trigger system for all runs. Therefore the strategy taken for analyses described in this thesis is to fully rely on offline PID to identify positron events. The PID for positrons is also used to purify the pion sample. Positrons exist in beam data for energies up to 32 GeV. There is a negligible amount of positrons at 40 GeV and higher so there is no need to identify them. Three local variables used as 3-dimensional cut to identify positrons are average hit density, number of active layers and the interaction layer.

A Čerenkov detector located upstream of the DHCAL was included in the trigger system to tag events as possible positrons. The test beam runs that include the Čerenkov are used to characterize positrons. The purity of positron samples tagged by the Čerenkov is very good as evident from a distribution of each event's total hits, shown in Figure 5.4. This histogram fits very well to a gaussian function. Positrons tagged with the Čerenkov are used to characterize the behavior of positrons' average density, number of AL and IL.

At each energy, positrons are characterized with average density, number of active layers and IL. First the average hit density of positrons is fit to a gaussian, as in Figure 5.4, in the range of mean ±3sigma. The boundary between positrons and pions is taken as mean - 3sigma so that positrons are greater than this value. For a pure sample of positrons, events with an average hit density within 3sigma of the mean contains 99.6% of the sample. This fitting procedure is applied to positron

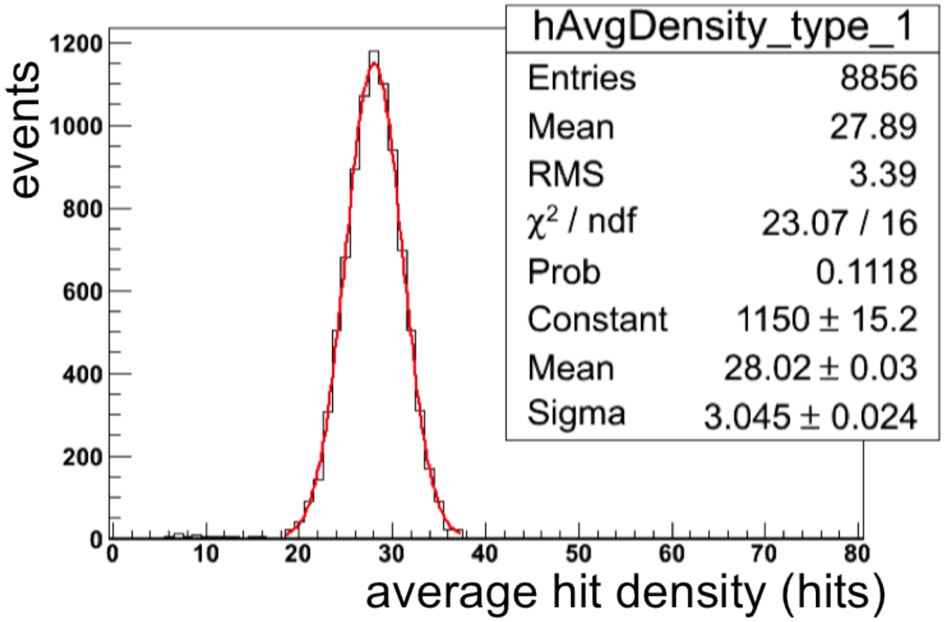

Figure 5.4. Histogram of average density for an 8 GeV set of positrons. A comparison of the peak of this distribution to the rightmost peak in the sample containing both pions and positrons identifies the right peak as positrons.

samples for 8, 10, 12, 20, and 25 GeV; and the corresponding average densities are fit to a log function as shown in Figure 5.5. Each point is the lower limit, mean-3sigma, for positrons. This limit is used in the pion sample to identify possible positrons. The red point is 6 GeV but is not included in the fit. The logarithmic fit parameters are also shown in the graph. Since the Čerenkov detectors were not working for 32 GeV test beam runs, a limit is extrapolated using the fit parameters.

The next component of positron ID comes from the number of active layers. Using the same sample of positrons from data as for average hit density, positrons are fit with a gaussian but in the range of mean-2sigma. This range is chosen because the positrons have a narrow spread, the red histogram in Figure 5.3, whereas the pions have a very wide spread and its more important to identify the peak of the positrons. An upper limit for positrons is mean+3sigma. Even though this limit extends well

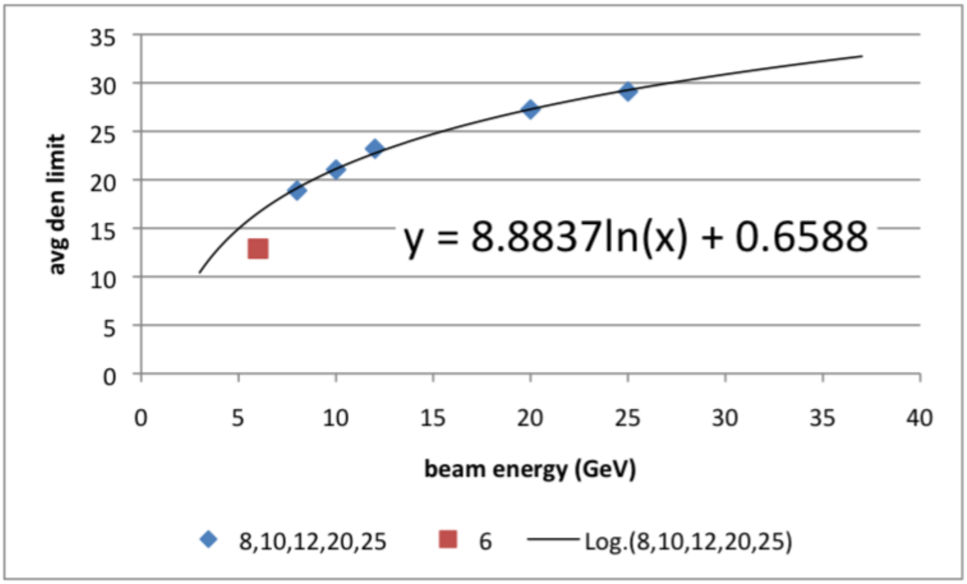

Figure 5.5. Graph of lower limit of average hit density for positrons.

into the range where there are more pions, this limit will maintain a higher purity of pions when removing positrons. The limits on number of AL for positrons is plotted in Figure 5.6. Points are fit to a log function and a limit at 32 GeV is extrapolated.

The final positron identification variable is IL. Positions consistently interact in the first layer and very rarely in the second layer. So if an event is classified as a positron by both average density and number of active layers, and the interaction layer is less than 4, it is finally analyzed as a positron.

To estimate the purity of a positron sample identified with the positron ID, simulated pions are selected with the positron ID. The fraction of pions remaining represents an estimate of the purity of the positron ID. Simulated pions at 6 GeV were selected with the positron PID and nearly 4% pions remained. This only estimates the purity of the positron PID since it will remove 96% of the pions that are mixed in

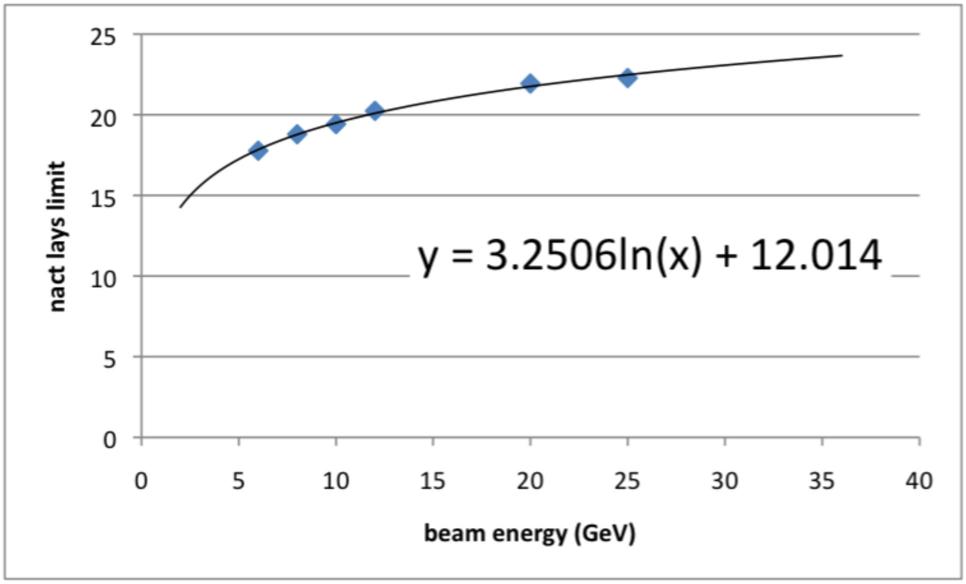

Figure 5.6. Graph of upper limit of number of active layers for positrons.

a sample of positrons, but it the number of positrons not identified is a combination of the positron PID efficiency and the pion PID efficiency and purity.

5.2.3 Pions

The jobs of identifying pions is actually that of removing positrons, muons, junk events and noise. In order for an event to be analyzed as a pion it must simultaneously fail to be identified as a muon, positron or junk event. The focus of data analysis in this dissertation is to improve the energy resolution of pions so the purity of muons and positrons is a lower priority. In simulation, removing positrons and muons resulted in an average of 95% for pion PID efficiency. This translates into a high purity in the pion sample in data. The purity of the pion sample is a source of systematic error and described in Section 7.4.2.

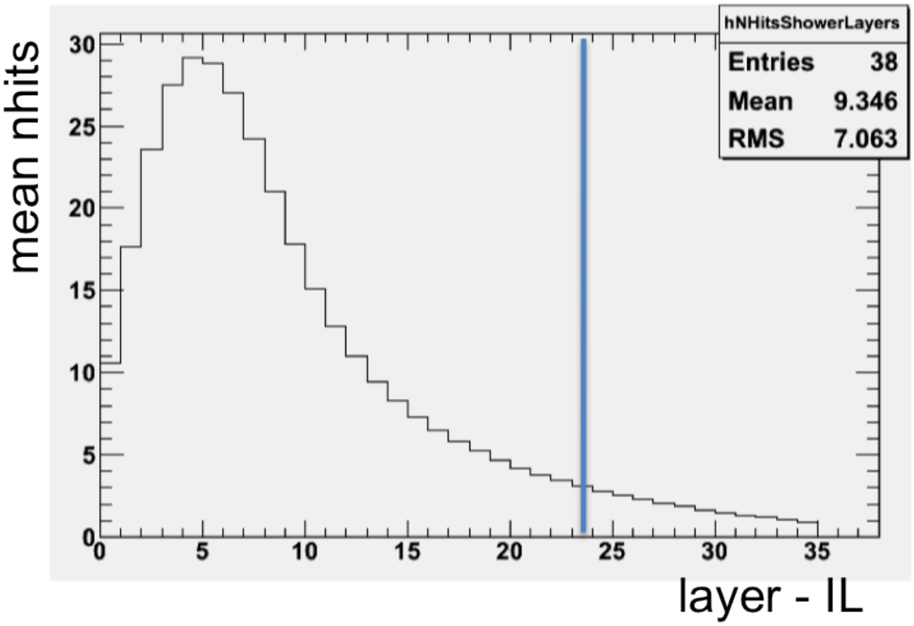

Figure 5.7. Longitudinal profile of 25 GeV pions from data.

5.3 Leakage

Leakage refers to the loss of measured energy when a particle's shower extends beyond the last layer of the calorimeter. Since the DHCAL is finite some amount of the energy is not measured so that the measured energy is less than that of the incident particle. This loss of energy out of the back of the calorimeter, or leakage, can possibly distort measurements of various properties of shower development.

In order to maintain a consistent requirement on leakage only the interaction layer (IL) is used as a cut. Interaction layer has the advantage that it is a local event variable. The IL cut value on leakage also depends on energy and is determined by the number of layers required to contain most of the event. The number of layers needed to contain most of each events energy is determined by the IL-subtracted longitudinal profile shown in 5.7 for 25 GeV pions.

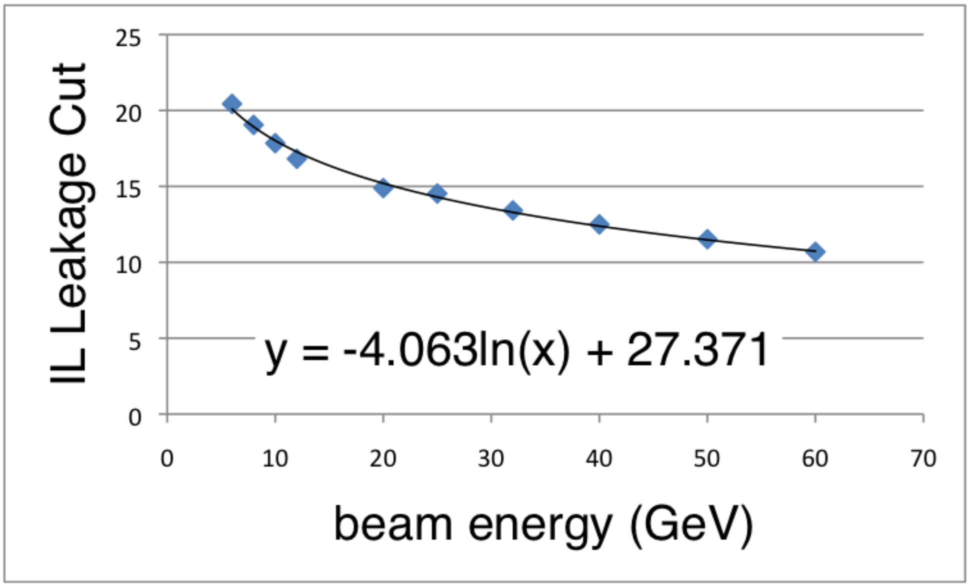

Figure 5.8. IL Leakage Cuts as a function of beam energy.

The statistical mean plus twice the standard deviation is used as the measure of the number of layers most 25 GeV pions require to be fully contained in the DHCAL. Since the DHCAL has 38 layers the IL cutoff value follows as.

$$ILLeakageCut = 38 - (Mean + 2 * RMS) \tag{5.3}$$

where RMS is actually Root's notation for the statistical standard deviation and Mean is the statistical mean. Both Mean and RMS are shown in the key of figure 5.7. Therefore, if the IL of a particular event is greater than the IL Leakage Cut for that energy, then it is classified to have leakage and is not analyzed. A summary of the IL Leakage Cut values as a function of energy is shown in Figure 5.8. This means that the increase in depth of pions showers slows with increasing energy. The fit is used as a sanity check to make sure there is no bias.

Leakage cuts are only applied to pions since positrons show no evidence of leakage. Parameterizing the leakage with energy defines a consistent set of cuts to eliminate bias. Events at each energy point contain a similar fraction of leakage.

The disadvantage is that many events at high energy are rejected, thus reducing the statistical sample size.

5.4 Additional Event Selection

Event selection algorithms separate events to use for analysis and are typically the same for each kind of beam. In general event selection specifies the type of analysis. For instance event selection for a noise analysis in the DHCAL will be different from selections that analyze properties of pions, positrons and muons. Events useful for particle analysis are those that interact within the DHCAL and whose physical or shower properties are not effected by sources other than the physics of the shower. These additional sources of bias are beam and trigger effects and noise in the calorimeter.

Beam and trigger effects can cause false triggers in the trigger system. False triggers with the beam at FTBF come from particles that interact upstream of the calorimeter. If a particle interacts in the beam line upstream of the trigger system and produces a signal in the trigger then data in the DHCAL is recorded. Event selection also removes events that contain multiple particles due to multiple occupancy in the beam line.

Early interactions are identified by analyzing properties of the first layer. For an interaction to occur in the DHCAL, hits in the first layer must be characteristic of a MIP, with no more than 4 hits. The DHCAL cannot distinguish pions and positrons from muons before an interaction occurs because each behave as a MIP. Also, if there is more than one cluster in the first layer then it is probable that the beam interacted with something prior to the DHCAL. It may also be that other particles are carried along the beam or that a spill has multiple occupancy.

False triggers are also identified by inactivity in the DHCAL. Muons generate hits in most layers, and if pions or positrons interact in the DHCAL, as is required by event selection, then these generate hits in more than 4 layer.

5.4.1 Noise

Noise identification is a subset of event selection that removes events containing specific kinds of noise. Hits that can be identified by sources other than the physics of a particular event are identified as possible noise. The sources of noise identified for removal are the HV ground connectors on each RPC , individual DCAL chips or other electronics noise associated with a FEB. The codes used to identify noise hits, clusters, and events used as part of noise identification [32] of event selection in the analyses described in this thesis.

The HV ground connectors on each RPC chamber can generate hits in the nearby pads. Instead of omitting an entire event only hits identified as ground hits are ignored. These pads are on the boarder, far from most showers. The x and y coordinates for ground pads are described in table 5.1. Any hit with a pair of x, y components from this list is labeled as a ground hit and is ignored. Clusters containing ground hits are also ignored. So each chamber has a 2 cm x 8 cm area near the edge where hits are ignored, essentially masking these areas of each chamber.

Table 5.1. y-coordinates of ground pads if x = 0 or 1.

chamber position	y-coordinate (inclusive)
top	82 - 90
middle	50 - 57
bottom	19 - 26

The FEBs and ASIC chips can produce noise hits in complicated geometric patterns that require various subsets of algorithms [32] to identify. Events identified by these algorithms are called box events and are ignored by the analyses discussed in this thesis. Algorithms to identify these electronics noise compare the number of hits and their geometries. For FEBs these algorithms look at total hits on a FEB, hits on a FEB compared to hits in the event, hits on the boarder pads of a FEB compared to total FEB hits and hits on boarder pads of FEB compared to neighbor FEB boarders. In all cases, the patterns are not associated with any physical process of the incident particle. For the ASIC chips, the algorithms look at total hits on a ASIC, hits on an ASIC boarder compared to total ASIC hits, and hits on an ASIC compared to neighbor ASIC hits.

As an example of FEB noise, simulated data is compared to test beam data. In Figures 5.9 and 5.10 highest board occupancy (the number of hits in the FEB with the most hits) is plotted versus the total number of hits in the event, nhits, respectively. If an event has a board occupancy greater than 120 it is ignored.

Another characteristic of the data that can be used to remove noise is the time structure. Since raw data is stored in 7 time bins of 100 ns by the DAQ, it is not possible for an incident particle's interaction in the DHCAL to span this whole time. It was found that only one of these time bins contains most hits relevant to physical processes. But due to effects of the trigger and beam, the specific time bin occupied varies from event to event so that, given a particular configuration, hits in two consecutive time bins are finally reconstructed into a single event. Therefore, the noise identification algorithms find the time bins containing the physics event and ignore the hits from the other time bins. It was found that noise identification removes approximately 3% events [32].

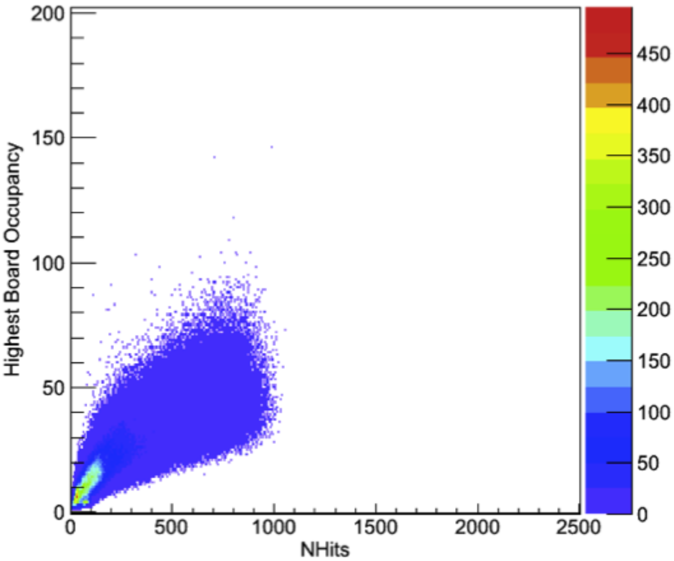

Figure 5.9. Maximum board occupancy vs. total nhits for simulated data.

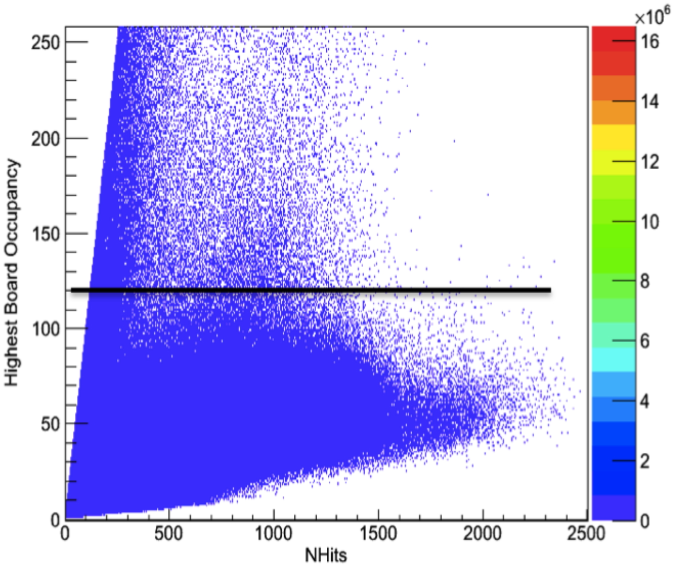

Figure 5.10. Maximum board occupancy vs. total nhits for test beam data before noise ID.

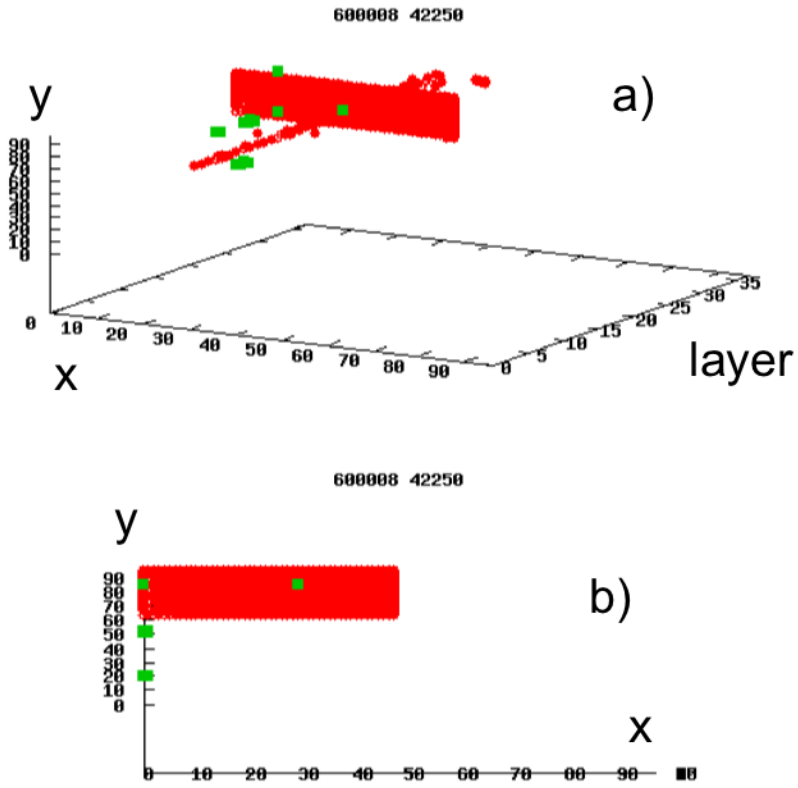

Figure 5.11. Event display illustrating hit patterns with FEB noise.

An example of how hits are arranged due to noise from a FEB and ASIC is shown in Figures 5.11 - 5.12 , respectively. Each point is a hit from a pad where the red points are hits in the two final time bins and the green points are hits in other time bins (out of time hits are omitted from event reconstruction). Here it seems that almost an entire FEB generated hits. The straight line is a muon.

The ASIC event display 5.12 shows a pion interacting near the middle of the DHCAL. The green points are generated by an ASIC and are also out of time. In this particular case, only the algorithms to identify out of time hits could be used. It

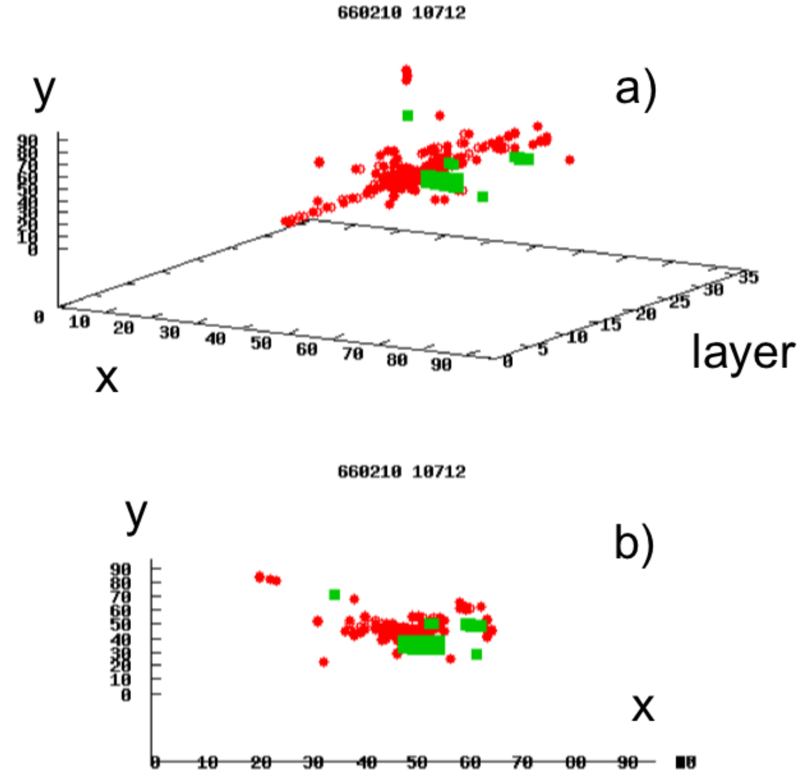

Figure 5.12. Event display illustrating hit patterns with ASIC noise.

was found that most ASIC and FEB noise with these patterns is in the least occupied time bin of the two reconstructed events.

CHAPTER 6

SIMULATION

Simulation of DHCAL data in test beam experiments is processed in two stages, the passage of particles through matter and the production of signals generated by RPCs. Mokka [33] and GEANT4 [34] [35] programs are used to simulate the passage of particles through various material in DHCAL test beam experiments. RPCSIM [36], is a stand-alone program to simulate the response of the RPCs and front-end readout. More specifically, GEANT4 produces locations within a RPC's gas volume where ionization occurs. RPCSIM is then required to simulate how RPCs treat ionizations in the gas volume while HV is applied and a front-end electronics readout system is connected.

Figure 6.1 shows a flow chart of the simulation strategy employed for DHCAL modeling. A geometric description of the DHCAL in the test beam consists of the dimensions and orientations of materials, and components in the beam line. Together with various Physics Lists, that contain phenomenological models of hadron interactions, GEANT4 produces points. Points are coordinates of energy deposition in the RPC gas gap. Functional charge distributions from previous experiments [37] [38] along with tunable parameters are used by RPCSIM to generate hits from the points provided by GEANT4. Conclusions about GEANT4 event generation, and RPC digitization and tuning can be made by comparing hits from data and simulation.

There are many goals for comparing test beam data to Physics Lists in GEANT4. Comparisons reveal inaccurate geometric (materials and dimensions) descriptions of the detector and test beam, and are a useful tool for validating the digitization process

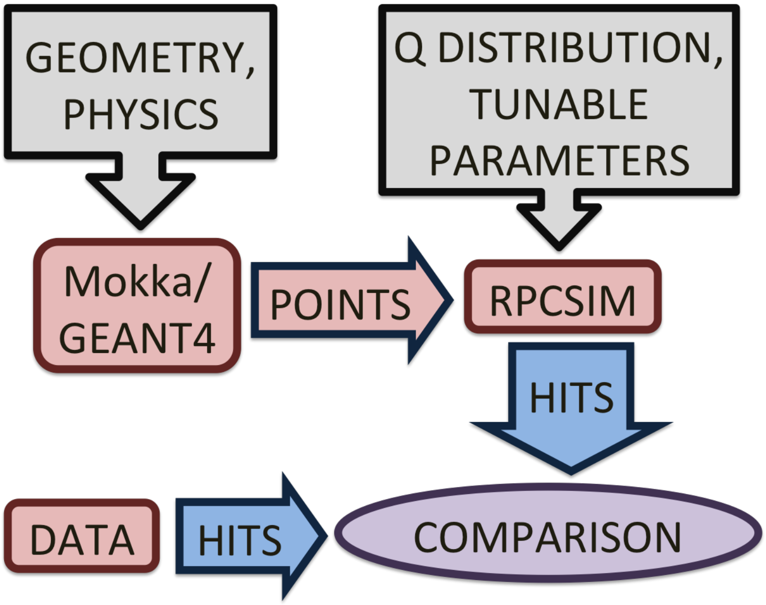

Figure 6.1. Flow chart of the DHCAL simulation strategy. Grey boxes are inputs to the simulation software in red boxes. Arrows are representations of events; points are from Mokka/GEANT4 and hits are from data and RPCSIM. Comparison generally describes any combined use of data and simulation.

(RPCSIM). Then with validation of geometry and digitization, along with detailed analysis, comparisons show possible deficiencies in the various Physics Lists and possibly systematic uncertainties in data. Since the DHCAL with glass based RPCs has the first fine-grain digital readout of a large scale hadron calorimeter, data from test beam can be a powerful tool for testing the accuracy of particular Physics List and provide possible corrections to the modeling of the propagation of high energy hadrons through matter.

This chapter separately describes these, Mokka/GEANT4 and RPCSIM, stages of simulation. Attempts to validate RPCSIM are discussed in detail. It is not in the scope of this thesis to validate physics models used in GEANT4. For this thesis, simulated data validates estimations of systematic uncertainties and serves as a sanity check for the average response of the DHCAL to longitudinal weighting.

6.1 Event Generation in GEANT4

Pion, positron and muon events are generated using GEANT4 with a DHCAL geometry matching the test beam configurations. Internally, GEANT4 uses Physics Lists to generate the development of hadron showers in the detector. The default Physics List for DHCAL simulations is called QGSP_BERT [39]. Previous experiments [40] have found this model to agree with data at the 2-4% level. QGSP_BERT is a compound model that applies the quark gluon string (QGS) model [41], GEANT4 precompound model [42] and Bertini cascade model [43] [44].

Quark-Gluon String Precompound (QGSP) combines QGS and precompound models of high energy collisions with matter between ∼10 - 60 GeV. QGS models the formation of strings in the initial hadron collision with a nucleon. De-excitation of the remnant nucleus is described by the precompound component. The Bertini cascade component is used for primary hadron interactions below 10 GeV.

A few different GEANT4 Physics Lists, including QGSP_BERT have been compared to similar test beam data by another calorimeter prototype, the Analog Hadron Calorimeter (AHCAL)[45]. The AHCAL uses scintillator, rather than RPC, as the active element in the same 38-layer steel structure for their test beam experiments [40]. Figure 6.2, shows comparisons between data and simulation (labeled as MC for Monte-Carlo) measurements of pion energy response and pion interaction length over a wide range of energies. Energy response of QGSP_BERT is within 4% of data below 10 GeV and between 7-10% at 100 GeV. QGSP_BERT agrees with data for pion interaction length near 9% for at low energy but within 4% at all other energies. These data show that QGSP_BERT is a fair model to use for the description of pion showers in the DHCAL.

A beam source (particle gun) simulates the primary and secondary (muon, positron and pion) beams at FTBF. The same energy range of test beam particles

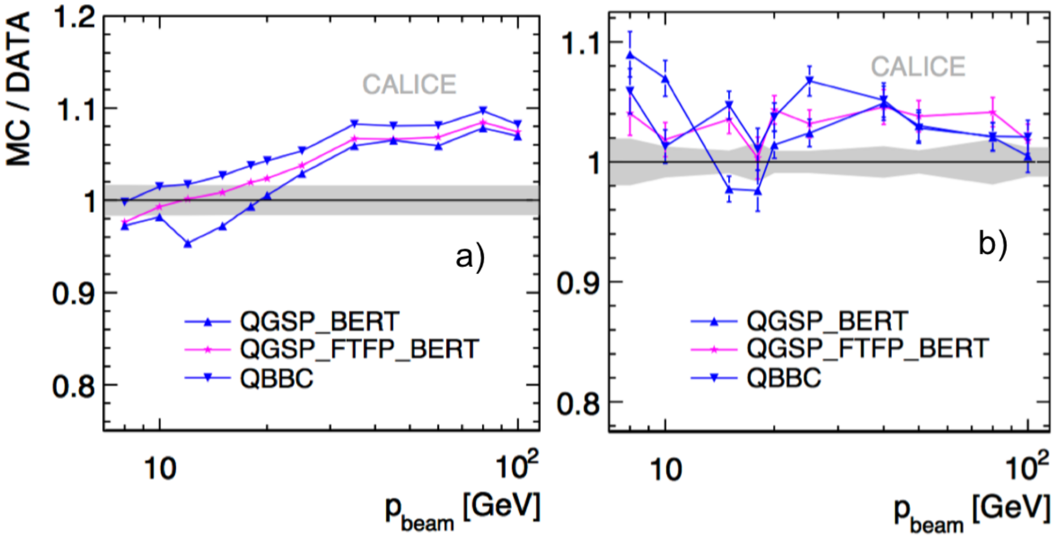

Figure 6.2. Comparison of GEANT4 Physics Lists to pion data as a ratio of MC/DATA for a) energy response and b) interaction length. Energy response is the mean value of gaussian distribution of each event's reconstructed energy for each beam energy of pions. Interaction length comes from a fit to a distribution of each events interaction layer for each beam energy of pions with the usual probability function $P_I = 1 - e^{-x/\lambda_\pi}$, where x is the longitudinal distance and λ_π is the interaction length. Data is from beam tests with pions in the energy range shown.

are generated in simulation. The spot size of simulated beams is designed to match that of the test beam.

Mokka provides the framework to implement the geometry of the DHCAL configuration at the FNAL test beam. The geometry includes scintillators (as a plastic mixture) and air as beam line components. The alternating structure of steel and RPC cassettes in the MS and TCMT is also simulated. The order of material in cassettes (in the downstream direction) is as follows,

- copper
- mylar
- thick glass
- RPC gas volume (freon, isobutane, SF_6)

- thin glass
- 2 mylar layers
- PCB layer (electronics boards)
- steel

As stated in the introduction, GEANT4 generates points of energy deposition in the gas gap of the RPC as a function of position coordinates in the layers (x, y, layer number). Energy deposition is in the form of primary, secondary and tertiary ionization of the gas, without HV applied. RPCSIM then simulates the response of RPCs to these ionizations.

6.2 RPC Digitization and Tuning

In general, digitization is a process in simulation that produces detector signals from active calorimeter elements out of the events generated by GEANT4. This process is not named after the digital nature of the RPC's single-bit resolution readout electronics. Specifically in RPCSIM, digitization generates DHCAL hits out of GEANT4 points, named points, by simulating the propagation of ionization induced avalanches within the RPCs and the front-end electronics readout.

Digitization is a multi-step process that involves sampling glass-RPC charge spectrum, modeling spatial charge distributions, converting charge to hits and tuning model parameters in order to most accurately describe RPC response in data. RPCSIM generates a charge, Q, and spatially distributes it over the readout pads. Then a threshold, in units of pC, is applied to the sum of charges on each pad, from nearby points, identifying pads as hits.

RPCs do not generate individual avalanches from distinct ionizations that are too close to one another. This translates to a possible inefficiency to distinguish ionizations, i.e. points. To accommodate this effect, RPCSIM applies a tunable distance threshold, d_{cut}, to remove multiple points within a distance d_{cut} cm from another given point. This parameter is referred to as "d cut".

Charge spectrums have been measured with analog readout in RPCs using muons from cosmic rays [37]. The charge spectrum is fit with the following functional form:

$$N(Q) = \alpha Q^{\beta} exp^{-\gamma Q} \tag{6.1}$$

where the parameters α, β and γ depend on HV and are fixed in simulation. Equation 6.1 is sampled to generate a charge, Q, for each point. To account for possible differences in charge distributions, such as noise, between the laboratory and test beam an additional tunable parameter, Q_0 is introduced such that

$$Q' = Q + Q_0 \tag{6.2}$$

The induced charge on readout pads is assumed to decrease exponentially with slope a according to:

$$\frac{dQ}{dR} = \left(\frac{Q'}{a}\right) \exp^{-Ra} \tag{6.3}$$

where R is the lateral radial distance extending up to 4 cm from a given point and a is the tunable exponential slope parameter. Figure 6.3 shows a charge spectrum fit with Equation 6.1 and measured [38] spatial charge distribution of RPCs in the lab. Equation 6.3 is integrated over the area defined by R to find the fractional charge contribution from each point to nearby pads. These charge contributions are summed for every pad in the detector. If the total charge on a pad is greater than the tunable threshold, T, then a hit is recorded.

The RPCSIM program uses four tunable parameters to digitize GEANT4 points. By name these are the slope, a; threshold, T; d cut, d_{cut}; and Q_0. The following table 6.1 lists and summarizes the functions and values of tunable parameters.

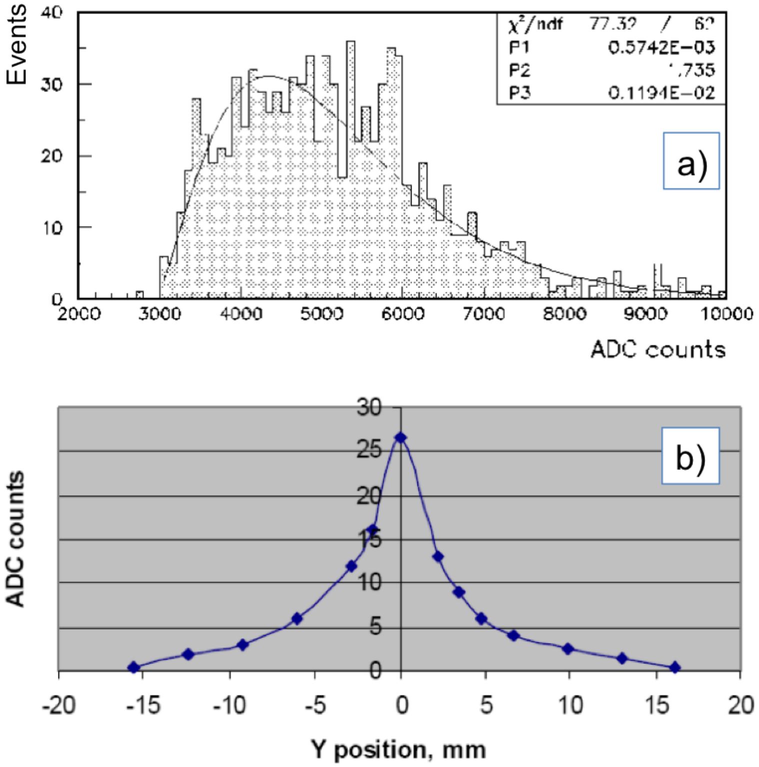

Figure 6.3. a) Charge spectrum and b) spatial distribution of charge on the plane of the readout pads (D. Underwood et. al) measured by RPCs in the laboratory.

Figure 6.4 shows the comparisons of the measured response per layer (given in hits) between a tuned RPCSIM and muon data. RPCSIM describes the data very well for typical muons as shown in the linear scale comparison. Typical muons have an average of ~ 1.5 hits per layer, where the agreement is very good. As the logarithmic comparison shows, there is a small discrepancy for muons with a large number of hits.

Table 6.1. Summary of tunable parameters in RPCSIM

parameter	function	value
d_{cut}	filters points	0.092 cm
Q_0	variations from lab	0.199 pC
Threshold, T	readout threshold	0.286 pC
Slope, a	spatial charge distribution	0.0843 cm

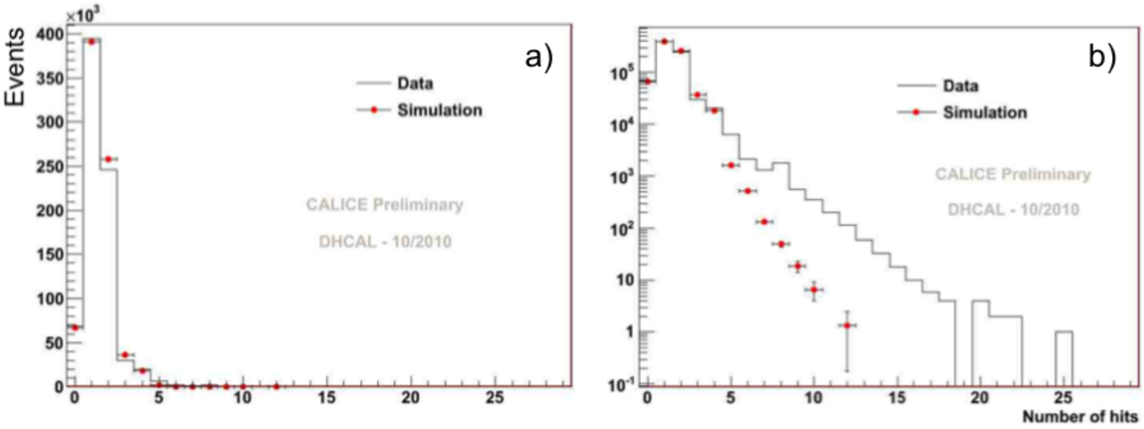

Figure 6.4. Average response over the entire detector with data (histogram) and tuned RPCSIM (red points) in a) linear and b) logarithmic scales.

CHAPTER 7

LONGITUDINAL CALIBRATION WITH WEIGHTS

The uncalibrated energy of incident particles is reconstructed in the DHCAL as the number of hits

$$E = \sum (H_i) \tag{7.1}$$

where the sum runs over all layers in the DHCAL and H_i is the number of hits in layer i.

Variations in response (efficiency, ϵ and average pad multiplicity μ) of individual layers are corrected by applying correction factors c_i to the number of hits in each layer

$$E = \sum (c_i H_i) = \sum (H_i') \tag{7.2}$$

These correction factors are a function of ϵ_i and μ_i (the efficiency and average pad multiplicity of layer i) and were taken to be

$$c_i = \frac{\epsilon_0 \mu_0}{\epsilon_i \mu_i} \tag{7.3}$$

As a next step we explore the possibility of improving the energy resolution of the DHCAL by applying individual weights to the number of hits in each layer.

$$E' = \sum (w_i H_i') \tag{7.4}$$

From this point, Equation 7.2 represents the natural response of the DHCAL and referred to as the reconstructed energy, E. Likewise, Equation 7.4 represents the weighted response and is referred to as the weighted reconstructed energy, E'. This distinction is articulated in the normalization procedure for the weights described in Section 7.3.

This approach was motivated by a similar weighting procedure, [46]. This paper reported on the analysis of data from the ATLAS electromagnetic calorimeter with its four segments in depth. A procedure to determine the weights was outlined and applied to the data of 245 GeV electrons. The method resulted in a significant improvement of the energy resolution. However, the authors failed to explore the method as a function of energy.

A similar approach, [47], to weight the energy with respect to longitudinal segmentation, but here applied to simulations of the D∅ detector. It was found that the weighting technique is able to exploit correlations between energies deposited in the different longitudinal segments of the detector and leads to a significant improvement of the energy resolution for electrons.

This thesis describes the determination of longitudinal weights for the measurement of energies of both electrons and pions in the DHCAL. The method is applied to data collected in the testbeam, as well as to simulated DHCAL data. Pions impinging on the DHCAL may travel a large distance in the calorimeter before undergoing its first hadronic interaction. The probability of a pion to survive and not undergo a hadronic interaction decreases exponentially with depth. In order to determine the longitudinal weights for pions in a similar way as for electrons, rather than considering weighting the physical layers, weights are applied to the layers after the identified (first) interaction layer IL. This way the hits in the same physical layer will be weighted differently depending on the reconstructed IL.

This chapter describes the calculation of the longitudinal weights and a procedure to normalize the weights. The uncertainties on the weights are estimated using a toy Monte-Carlo approach. The weights are determined individually for each energy point. Results are shown for three different cases:

1. The weights determined at one energy point are directly applied to the data at the same energy point.
2. The weights determined at different energy points are parameterized by a smooth function. For a given energy the weights are obtained from the parameterization.
3. The energy of a given event is estimated using the calibrated number of hits. The weights are determined from the parameterizations. This last case does not make use of the known beam energy and is the only realistic approach to the measurement of energy in theDHCAL.

7.1 Analysis Software

Data are originally written into a simple ASCII format maintaining the chronological order of events. This provides a simple machine independent and easily transferable format. ' One disadvantage is that analysis task, e.g. the formation of clusters requires many iterations of calculations for every hit, are time consuming and must be performed for every event. A strategy to minimize redundancy and simplify the analysis is to convert ASCII data into a more complicated ROOT [48] format. The software to convert ASCII data into ROOT was created and is maintained by Burak Bilki. Analysis performed in this thesis was done using this ROOT format.

Restructuring DHCAL data into the ROOT format offers many advantages. It allows the user to take advantage of native code for histogramming. Additional calculations and common analysis tasks can be stored with the original data. The ROOT structure provides a native way for pre-processing so that common analysis tasks such as constructing clusters are done only once. The use of versatile ROOT libraries is also beneficial. Finally, ROOT files are smaller in size compared to the original ASCII file when combining additional information.

Events in the ROOT data structure are stored as a DHCALEvent using customized ROOT libraries referred to as DHCALEvent libraries. This is a complicated data structure that organizes events into classes and methods. Hits are stored as C^{++} classes that contain position coordinates and a timestamp as well as cluster information.

Analysis code is written within ROOT CINT script and uses a modular format so that common analysis tasks can be grouped together. Several modules of code are used to compute the weights. A main program called *DHCALAnalyzer.C* controls the order that the analysis software modules process data and depends on user input to decide which module to run. First an application called *Histos.cc* processes ROOT DHCAL data and writes a set of initial histograms into a new root file that contains histograms used by subsequent analysis modules. Next in the analysis chain is *ErrorsCov.cc*. This application takes as input the data ROOT file and the output root file from *Histos.cc* and computes unnormalized weightsand writes them into an ASCII file, *weights.txt*. WeightsLayers.cc takes the data ROOT file and weights text file to generate the final histogram ROOT file containing the results of the weighting technique.

In addition to these DHCALEvent based ROOT analysis modules, there are smaller macros used throughout the analysis process. To fit weights, described in Section 7.5, a macro called *fitWeights.C* reads *weights.txt* and uses default ROOT libraries in order to store the fit parameters to another ASCII file called *wparsLayer.txt*. *WeightsLayers.cc* reads the parameters in *wparsLayer.txt* in order to select weights as a function of energy. Finally, *resolution.C*, reads the histograms stored in *WeightsLayers.cc* output ROOT file to generate performance plots that are used to evaluate the results of the analysis.

7.2 Calculation of Longitudinal Weights

Each layer in the shower has an associated weight such that the total weighted number of hits in the DHCAL is written as equation 7.4. The weights are later determined to minimize the energy resolution. For pions, weights are applied to shower layers, where the shower layers are those layers between the interaction layer and the last layer of the DHCAL.

Recalling Equation 7.4, longitudinal weights have only been discussed in the definition of reconstructed of energy in the DHCAL. We now derive an analytical form of the weights in such a way that the energy resolution can be minimized. First, some definitions are made.

Weights start in layer i-IL for pions so that the weight in MIP layers is 1

$$w_i = \begin{cases} 1 & \text{for } i < \text{IL} \\ w_{i-IL} & \text{for } i \geq \text{IL} \end{cases} \tag{7.5}$$

where w_{i-IL} is the weight for each shower layer. The following calculations determines the analytic form for the set of weights associated with the layers in a shower. Shower layer, i, is beyond the IL and has an associated physical layer, p, such that,

$$p = IL + i \tag{7.6}$$

for physical layers, p, greater than the IL. The maximum physical layer number is the number of layers in the Main Stack (MS) of the DHCAL; p then is an index to the location of the layer in the MS. Since the shower layer fluctuates with the interaction layer, it is associated to a different physical layer. Therefore, there is a number of events, N_i, that shower layer i is associated to a layer in the MS.

The mean number of hits in each shower layer is defined as

$$\mu_i = \frac{\sum_{k=1}^{N_i} H_i^{(k)}}{(N_i)} \tag{7.7}$$

where N_i is the number of events that shower layer i is associated to a physical layer in the MS, such that i+IL < 38.

The correlations between hits in shower layers are estimated with a sample covariance matrix.

$$Cov[i,j] = \frac{\sum_{k=1}^{N_{i,j}} \left(H_i^{(k)} - \mu_i\right)\left(H_j^{(k)} - \mu_j\right)}{N_{i,j} - 1} \tag{7.8}$$

where $N_{i,j}$ is the number of events where both shower layers have hits present. The summation runs over the entire set of events in a set of runs at a given energy. A sample covariance is an approximation to the true covariance because there is a finite number of events contributing to the calculation.

Energy resolution is defined as the standard deviation of the reconstructed energy divided by the mean of the reconstructed energy

$$\frac{\sigma\left(E'\right)}{\mu\left(E'\right)} \tag{7.9}$$

where there is one value of $\sigma\left(E'\right)$ and $\mu\left(E'\right)$ for each set of runs at a given energy but there as many μ_i, as in Equation 7.7, as layers in the shower in the same set of runs at a given energy.

Variance of reconstructed energy $\sigma^2\left(E'\right)$ is defined in terms of weights and covariance

$$\sigma^2\left(E'\right) = \sum_{i,j} w_i w_j C_{ij} \tag{7.10}$$

and the mean of reconstructed energy $\mu\left(E'\right)$ is defined in terms of weights and mean number of hits in each shower layer

$$\mu\left(E'\right) = \sum_i w_i \mu_i \tag{7.11}$$

To simplify notation for the minimization procedure later, it is now convenient to switch to vector notation. First rewriting the standard deviation $\sigma\left(E'\right)$ and mean $\mu\left(E'\right)$ of reconstructed energy respectively

$$\sigma\left(E'\right) = \sqrt{\vec{w}^T C \vec{w}} \quad \text{, and} \tag{7.12}$$

$$\mu\left(E'\right) = \vec{w}^T \vec{\mu} \tag{7.13}$$

the energy resolution in terms of weights is

$$\frac{\sigma\left(E'\right)}{\mu\left(E'\right)} = \frac{\sqrt{\vec{w}^T C \vec{w}}}{\vec{w}^T \vec{\mu}} \tag{7.14}$$

The vector, $\vec{w}$ contains the set of weights w_i, C is a matrix representing the covariance $C_{i,j}$, $\vec{\mu}$ represents the mean hits in each shower layer, and superscript T denotes a transpose.

With the resolution defined in terms of the weights, it is possible to analytically minimize the resolution with respect to the weights. So differentiating equation 7.14 with respect to the weights and setting the results to zero

$$\begin{aligned} 0 = \frac{\partial}{\partial \vec{w}^T}\left(\frac{\sqrt{\vec{w}^T C \vec{w}}}{\vec{w}^T \vec{\mu}}\right) &= \frac{1}{(\vec{w}^T \vec{\mu})^2}\left[\frac{(C\vec{w})(\vec{w}^T \vec{\mu})}{\sqrt{\vec{w}^T C \vec{w}}} - \vec{\mu}\sqrt{\vec{w}^T C \vec{w}}\right] \\ &= \frac{\left[C\vec{w}\vec{w}^T\vec{\mu} - \vec{\mu}\vec{w}^T C \vec{w}\right]}{(\vec{w}^T \vec{\mu})^2 \sqrt{\vec{w}^T C \vec{w}}} \end{aligned} \tag{7.15}$$

is equivalently written as

$$(\vec{w}^T \vec{\mu})\vec{w} = (\vec{w}^T C \vec{w}) C^{-1} \vec{\mu}. \tag{7.16}$$

The quantities in parenthesis on both sides of the equation are scalars while the remaining are vector quantities. Therefore, a set of weights $\vec{w}$ that satisfies this equation must have the same direction as the vector $C^{-1}\vec{\mu}$ and the weights are written in vector form as

$$\vec{w} = \lambda(C^{-1}\vec{\mu}) \tag{7.17}$$

where λ is a dimensionless normalization constant to be determined. An analytic form of weights, in summation notation, follows as

$$\tilde{w}_i = \lambda \sum_{k=1} C_{ik}^{-1} \mu_k \tag{7.18}$$

With $\lambda = 1$, this is the analytic form used to calculate unnormalized weights. Normalization of weights is discussed in the next section.

7.3 Normalization of Weights

Normalizing weights sets the energy scale of the weighted hits. For the DHCAL, the energy scale is defined as the total number of hits as a function of beam energy, given in units of $hits/GeV$. Applying unnormalized weights results in degraded linearity response as shown with pions in Figure 7.1. The blue points are weighted hits and the black points are unweighted hits. In both cases, the points are fit with a power law function.

As the weights are originally calculated at each energy points, they are not in the same scale to beam energy as the total hits. This means that the weights must be normalized. There are many options to normalize the weights since this can be done arbitrarily. The method taken by [46], was to normalize the weights in all layers to the weights to the weight in the layer with most energy deposited. The analogy to the DHCAL would be to normalize weights to the shower layer with the most number of hits. The problem with this approach for the DHCAL is that the layers are finely segmented and the shower max layer fluctuates between events.

One method is to normalize to the weight in the shower max layer.

$$\tilde{w}_i = \frac{w_i}{w_M} = \frac{\sum\limits_{k=1} C_{ik}^{-1} \mu_k}{\sum\limits_{k=1} C_{Mk}^{-1} \mu_k} \tag{7.19}$$

where i is the layer number, M is the shower max layer and the weights are defined as before in Section 7.2. This method assumes that the weight in the shower max layer is a good representation of the energy scale, which is a rather poor approximation. This method doesn't work because the energy scale doesn't match the total unweighted hits and the energy linearity is degraded. As a result of the degraded energy linearity, these normalized weights do not have a clear dependence on the beam energy. Parameterization with the beam energy is essential in order to apply this longitudinal calibration without knowledge of the beam energy. So this normalization, suggested by [46] was abandoned. A preferred normalization is one that scales linearly with beam energy, and has a weak dependence on the event to event fluctuations.

In order to normalize the weights to produce a linear response of weighted reconstructed energy, E', an iterative procedure is applied. First the unnormalized weights are applied to all energies and the response of the total hits to beam energy is plotted. The unnormalized total hits (blue) have a highly non-linear response as shown in figure 7.1. This also serves to illustrate the role of normalization in setting the energy scale of the weights.

A normalization factor is calculated for each of the beam energy points as follows

$$s_i^{(k)} = \frac{E'}{E_i \cdot 14.74 hits/GeV} \tag{7.20}$$

where 14.74 is the desired linear response of the DHCAL, E_i and s_i are the beam energy and normalization factor of the i^{th} energy point, and the superscript k refers to the k^{th} iteration. E' is the weighted reconstructed energy given by Equation 7.4. The normalization factor is the ratio of the weighted reconstructed energy (blue points) in figure 7.1 to a straight line described by 14.74 hits/GeV.

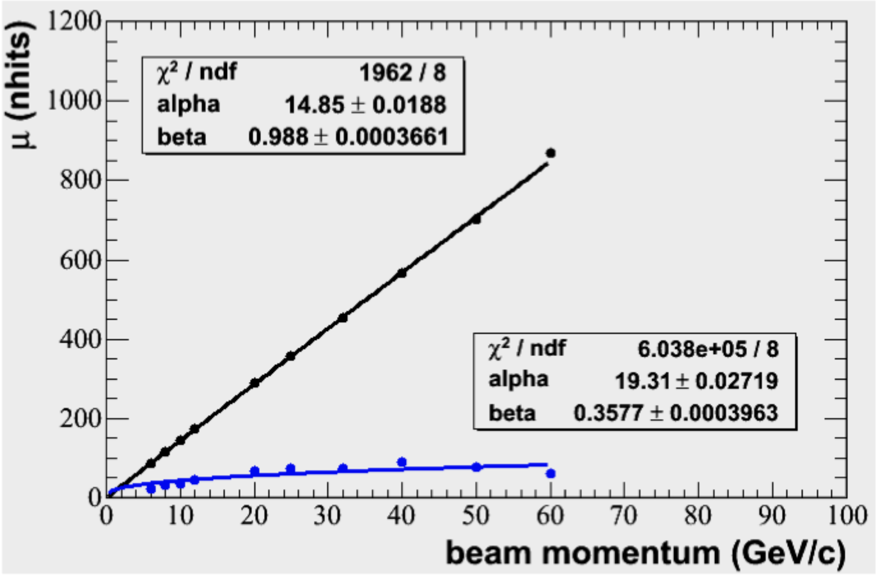

Figure 7.1. Raw weights applied to pion data.

The value acquired for each energy point is then use to normalize the weights so that the normalized weights are

$$w^k = \frac{w^0}{s_i^k} \tag{7.21}$$

where the superscript refers to the sequence of iterations, 0 refers to the unnormalized weights and i refers to each energy point. A single normalization value is applied to every layer's weight value. The weights are applied to data using this first set of normalized weights. Then another normalization value is calculated to correct for the previous normalization's nonlinearity. This iteration is repeated until the resulting linearity is within 5%. Iterations are needed because the initial unnormalized weights are so small that reconstructed energies at 6 and 8 GeV have many events with 0-10 hits, which is unphysical behavior or pions at this energy. This iterative procedure is beneficial in that it results in a linear response after applying weights and the energy resolution values do not need a large correction due to nonlinearity. The final normalization is a product of all the normalization constants from each iteration as follows

$$s_i^k = s_i^{k-1} s_i^{k-n} ... s_i^0 \tag{7.22}$$

so that the final set of normalized weights is

$$w^k = \frac{w^0}{s_i^{k-1} s_i^{k-n} ... s_i^0} \tag{7.23}$$

Pions require only two iterations of normalization to achieve a linear response and positrons require a single iteration. The response is considered linear when the new normalization ratio is greater than 0.9 implying that each point is within 5% of a linear response. The final weights are then just a result of the un-normalized weights and the product of normalization constants.

7.4 Estimation of Uncertainties

In this section we explore various contributions to the uncertainties of weights. Statistical uncertainties are those arising from sample sizes, i.e. the number of events, fluctuations in the physical behavior of interactions in the DHCAL and the propagation of uncertainties through calculations. The number of events is maximized by including as much data as possible but the efficiency of particle ID reduces this number further. Fluctuations in the physics behavior of interactions in the DHCAL includes such things as the fluctuation in the number of hits in a given layer from event to event.

Uncertainties on the quantities, C^{-1} and $\vec{w}$ propagate through the calculation of the weights. A procedure to estimate the propagation of these uncertainties is discussed in Section 7.4.1. Various sources of systematic of uncertainties are discussed in Section 7.4.2 and their values are estimated.

7.4.1 Statistical Uncertainties on Weights

If the uncertainties on the average number of hits in each layer and the corresponding set of covariance values were not correlated then the uncertainties on the weights would follow the typical analytical uncertainty propagation. This is, however, not the case, so the uncertainties on the weights need to be estimated using a different approach. In this we follow [46] and use a simple Monte-Carlo technique. In this technique the covariant matrix elements and average number of hits per layer are smeared such that 10,000 weights for each layer are generated. These generated weights are distributed according to a gaussian function. The standard deviation of the resulting gaussian distribution of MC generated weights is then used as an estimation of the uncertainty on the weights. This smearing is done for the weight of each shower layer.

With this procedure it is necessary to transform the weighted mean hits in each layer with a matrix, U, such that the uncertainties of the transformed hits are uncorrelated. Such a matrix is unitary, $U^{-1} = U^T$ and diagonalizes the original covariant matrix.

This transformation diagonalizes the covariance matrix

$$\hat{C}_{rs} = \sum_{i,j} U_{ri} U_{sj} C_{ij} \tag{7.24}$$

where r,s refer to layers in the uncorrelated space and the only non-zero elements are for r=s. The hat on C associates it with the transformed space. The mean number of hits is also transformed using the same U

$$\hat{\mu}_r = \sum_i U_{ri} \mu_i \tag{7.25}$$

Now the quantities $\hat{C}_{rs}$ and $\hat{\mu}_r$ are uncorrelated. The standard deviations of mean hits per layer are directly given by the square root of the diagonal elements of the transformed covariance matrix

$$\hat{\sigma}_r = \sqrt{\hat{C}_{rr}} \tag{7.26}$$

The statistical uncertainty on the average number of hits per layer is the standard deviation for that layer divided by the square root of the number of events that contain that shower layer. It is written as

$$std.dev.[\hat{\mu}_r] = \frac{\hat{\sigma}_r}{\sqrt{N}}. \tag{7.27}$$

Fluctuations in a sample covariance matrix follow a Wishart distribution. It is the multivariate analogue to the χ^2 distribution. With no correlation between layers' mean hits, the covariance, C, uncertainties follow a χ^2 distribution. Uncertainties in means of hits in each layer do not follow a gaussian distribution as evident from histograms from data and simulation. Nevertheless, gaussian behavior is assumed for the fluctuations of mean hits, μ_i.

The number of events is not constant for the mean of hits in each layer or the covariance. For this reason, a vector containing these values should also be transformed but this was not done at the time of analysis, therefore it is discussed in the section on systematic uncertainties, 7.4.2.

A toy monte-carlo smears $\hat{C}_{rs}$ and $\hat{\mu}_r$ according to their associated probability functions, transforms these values back to the correlated space and generates a smeared weight. This procedure is performed 10,000 times and a distribution of weights for each layer is generated.

First, an average mean is randomly sampled from a gaussian distribution whose mean is equal to the average mean and standard deviation equal to the statistical uncertainty on the average mean as shown in equation 7.27.

A covariance element is then randomly sampled from a χ^2 distribution. The number of degrees of freedom is $N_r - 1$, where N_r is the number of events that layer r has hits. The result of randomly sampling this χ^2 distribution is a smeared number of events. A random sample of this χ^2 distribution produces the quantity $(\hat{\sigma}_r^*/\hat{\sigma}_r)^2$ where the $\hat{\sigma}_r^*$ is the desired smeared variable and the asterisk denotes a smeared quantity. Then the set of $\hat{\sigma}_r^*$ and $\hat{\mu}_r^2$ for all the layers is transformed back to correlated space and a smeared weight is calculated.

$$w_i^* = \lambda \sum_{j=1} {C^*}_{ij}^{-1} \mu_j^* \tag{7.28}$$

where i and k are used to denote the original correlated space and λ is the normalization constant. This smearing process is repeated 10,000 times to produce a distribution of weights w_i^* for each shower layer. The standard deviation of each distribution is an approximation to the uncertainty of the weight in the corresponding shower layer. An example of these distribution is represented in figure 7.2.

A summary plot of all the uncertainty values for normalized weights as a function of layer after the interaction layer for data pions at each energy is shown in figure 7.3. A typical behavior occurs at each energy that the minimum uncertainty is generally near the shower max. This is because there are more hits in these layers providing a more precise calculation of the weights values. The uncertainties are larger at the back of the shower because there are less events contributing to the calculation of those weights. The magnitude of the uncertainties represented in this graph do not reflect any particular normalization of weights in the final analysis.

7.4.2 Estimate of Systematic Uncertainties

Systematic uncertainties reflect common deviations inherent in the calculation or measurement of various quantities and are not affected by statistical sample size.

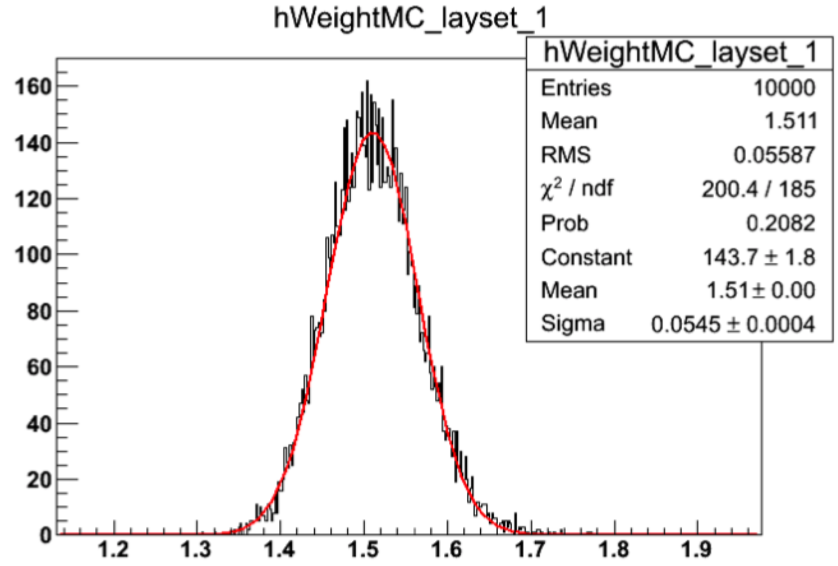

Figure 7.2. Sample distribution of smeared weights for the first layer after the interaction layer, generated by toy Monte-Carlo to estimate uncertainties on weights at 10 GeV..

Statistical uncertainties may be exactly, or nearly, calculated but systematics can only be estimated. For the use of longitudinal weights, sources of systematic uncertainty may affect the value of weights, the quality of various fits throughout the procedure, and the mechanisms used to apply a set of weights to a set of runs. A detailed estimation of systematics due to PID is presented, followed by a discussion of other possible sources.

7.4.2.1 Systematics Due to PID

The main source of systematic uncertainty for a longitudinal calibration is the misidentification of particles due to PID. Positrons and pions have drastically different behavior with respect to

- correlations of energy deposition between layers
- mean number of hits in each layer

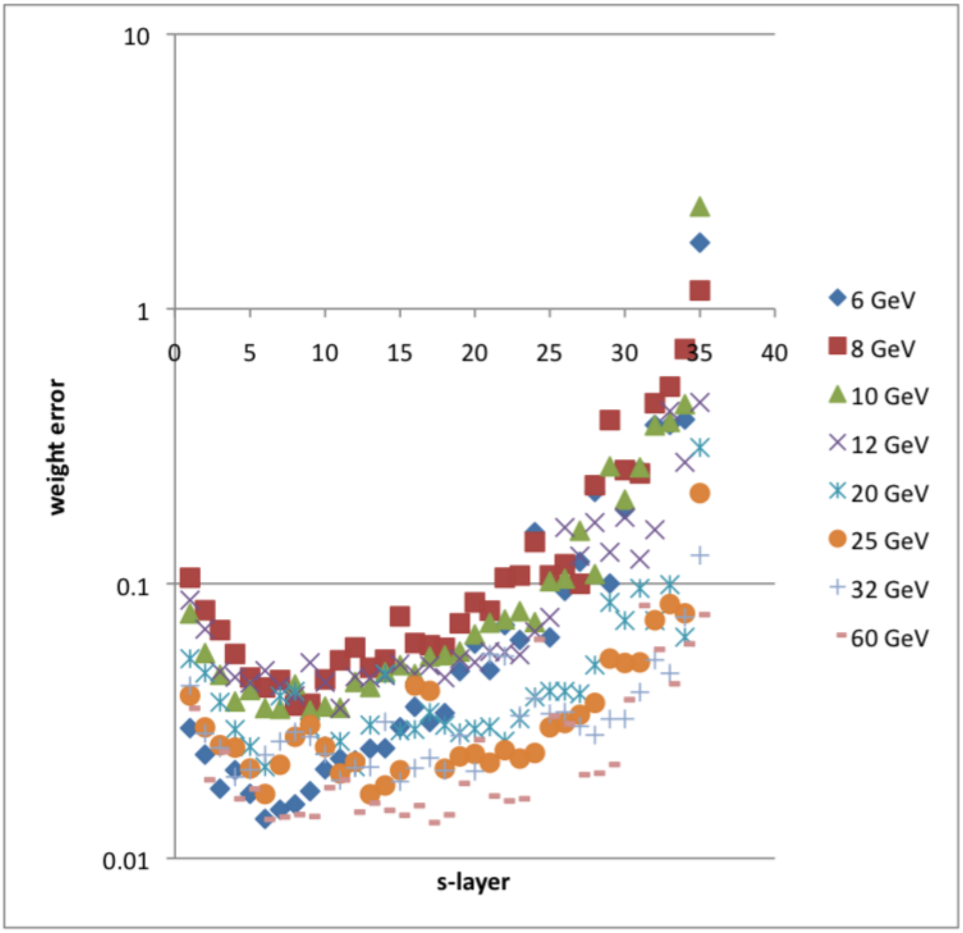

Figure 7.3. The uncertainty values, (RMS of smeared distributions), for all layers in all energies of data pions..

- the fraction of leakage
- energy response in linearity and resolution

Weights depend directly on two of these quantities. The effect that leakage has on weights is shown in Section 7.5.5. Positrons are highly non-linear whereas pions become increasingly non-linear starting at 25 GeV. Therefore, a large impurity of positrons in a pion sample may have adverse effects on the resulting weights' values. In the interest of time, systematic uncertainties due to PID are estimated from pion data. The items previously listed may have their own inherent systematics but an estimation of those, and other possible, values are left for future work.

The systematic uncertainty of PID, specifically the affect of having an impure sample of pions due to mis-identification of positrons is estimated for pions. Positrons are injected into a sample of data pions using positron ID, that is estimated to have 96% purity (see Section 5.2.2). The original pion sample has an unknown amount of positrons already in it. The systematic uncertainty estimation relies then on the affect of adding positrons. In the interest of time, simulation is not used to estimate this uncertainty and is also left for future work. Positrons are injected into the pion sample so that the increased number of events, that is positrons, is 10%.

Figure 7.4 compares the uncertainties of different sets of weights at 10 GeV. A set of weights are generated at 10 GeV using the regular pion sample and using the pions with 10% positrons. The uncertainties calculated from the MC statistical calculation are shown along with the systematic uncertain given as the difference between the weight values at each layer. It appears that there is something wrong with the PID since the systematic uncertainty is large compared to the statistical calculation. Actually, the statistical uncertainties are too small. A possible mistake was found in the Monte-Carlo procedure to estimate statistical uncertainties. During the MC procedure, a vector containing mean hits in each layer and a matrix of the covariance elements is transformed into an abstract mathematical space. This allows the MC technique to smear values. This is the correct procedure when the number of events for each layer is constant but this is not the case for pions since weights float with the IL. Therefore, a vector containing the number of events for each layer should

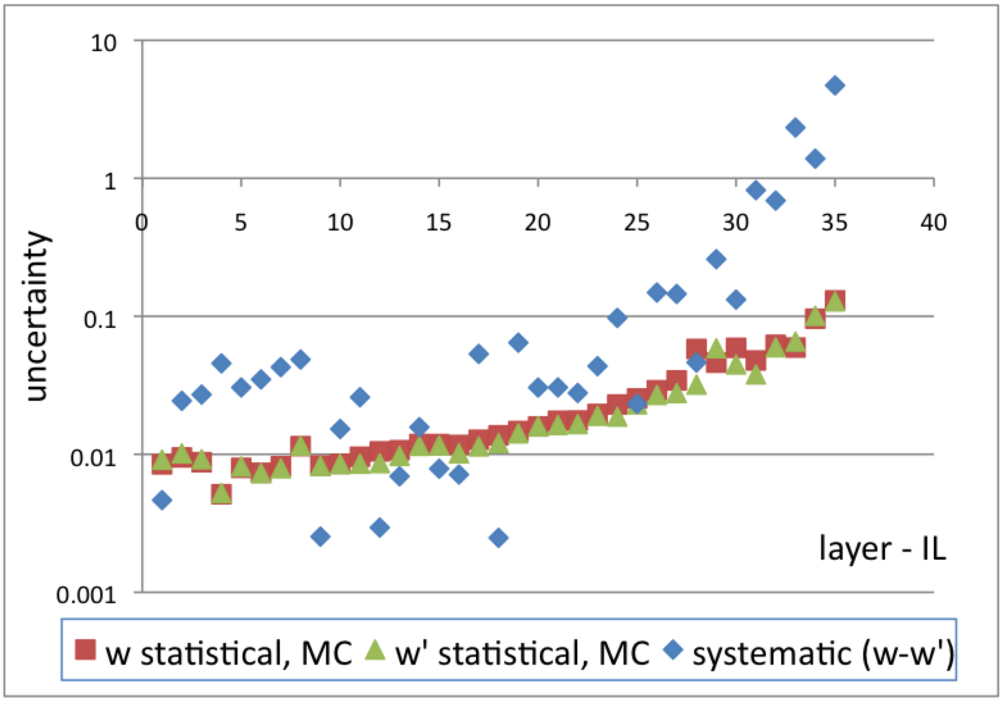

Figure 7.4. Statistical and systematic uncertainties at 10 GeV due to PID. MC refers to Monte-Carlo.

also be transformed. Additionally, when the systematic uncertainty is included with the statistical uncertainty according to

$$\sigma_{TOT} = \sigma_{stat} \oplus \sigma_{sys} \tag{7.29}$$

where $\oplus$ represents addition in quadrature, The total uncertainties reflect the expected statistical uncertainty given the spread of the weight values, shown in Figure 7.5 The size of the systematics appears to reflect the expected size of the statistical uncertainties on weights due to error propagation. The large uncertainties for the back layers are dominated by systematics. In this case, the statistical uncertainties cannot be used to compare estimations of the systematic uncertainties. Therefore, the mis-calculation of the statistical uncertainty is a source of systematic uncertainty in that statistical uncertainties are incorrectly calculated for every set of weights.

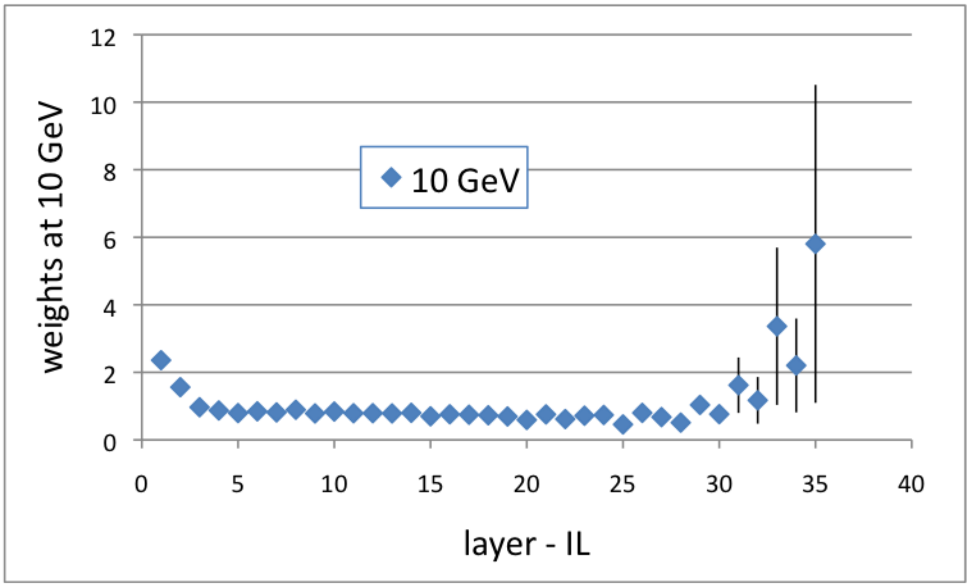

Figure 7.5. Weights at 10 GeV with statistical and systematic uncertainties combined.

Additional possible sources of systematic uncertainty on the weights are as follow

- identification of the interaction layer, IL
- momentum spread of the Fermilab test beam
- propagation of uncertainties in weights' calculation

Each of these sources is discussed, but not calculated in detail, in the following sub sections.

7.4.2.2 Identification of the Interaction Layer

As the weights are determined as a linear function of shower layer, an uncertainty in identification of the interaction layer will propagate into the calculation of the weights. If the interaction layer is misidentified then the weights are applied to the wrong layer.

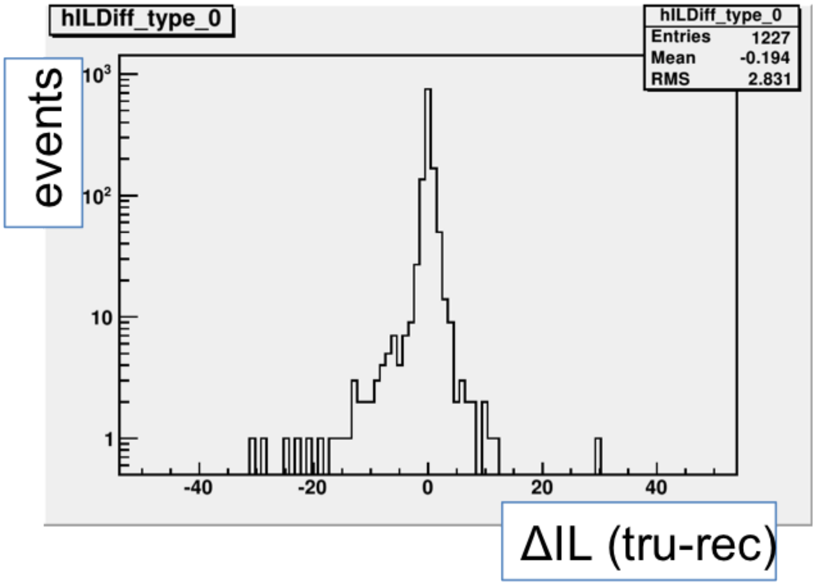

Figure 7.6. Histogram of true IL - reconstructed IL for 20 GeV pions. IL deviates by about 1.5 layers about the true IL.

It is possible to estimate interaction layer mis-identification as a source of systematic uncertainty by using simulation. The IL is known in simulation so that the difference between the true IL and the reconstructed IL can be directly calculated. Figure 7.6 shows a histogram of the difference between the true IL and the reconstructed IL for a small sample of pions at 20 GeV. The RMS is 2.831 which means that on average the reconstructed IL will fluctuate about the true IL by approximately 1.5 layers. Since the weights for a single energy are relatively constant, except the first few layers, IL is not expected to have a large contribution to weights' systematic uncertainties.

The interaction layer is suspected to bias the weight value of the first two layers since there is a strict requirement on the number of hits. The IL must be the first of two consecutive layers with at least 4 hits. This requirement can possibly affect the calculation of the covariance matrix. In a test of the effect of weight values for the back layers, the weights for the last 10 layers of pions are taken as the average of the weights from the middle 10 layers. The different between using an average weight and using the directly calculated weights for each layer resulted in no change to the energy resolution. Likewise the first 2 layers have fewer hits relative to the hits in the middle 10 layers. Therefore, IL is expected to have minimal affect as a source of systematic uncertainty.

7.4.2.3 Momentum Spread of the Fermilab Test Beam

According to the website of the Fermi Test Beam Facility (FTBF) website the momentum spread of the test beam is estimated at 2.0-2.5%. The weights calculated for a given layer depends on the average number of hits in that layer. In turn, the average number of hits per shower layer, or the longitudinal shower profile, however varies with the energy of the incident particle. Therefore the spread in momentum spread of the test beam affects the value of the weights at each energy setting.

7.4.2.4 Propagation of Uncertainties in Weights' Calculation

Part of the calculation to estimate the statistical uncertainty on weights required transforming the covariance matrix and mean hits vector to an uncorrelated space in order to smear weights. It was found that the number of events for each layer is also required to transform into uncorrelated space but this was not done. To estimate this uncertainty, a vector containing the number of events for each layer is also transformed to complete the calculation. The resulting uncertainties on weights are compared to the uncertainties used in this analysis as an estimate of the systematic uncertainty from this source.

7.5 Application of Weights

The previous section has been a description of the computation of the weights. In practice, the beam energy should not be used to select weights since the energy of incident particles in a real experiment is not known *a priori*. In order to apply a set of weights without knowledge of the beam energy, the generated weights are fit as a function of beam energy and nhits from each event is used as a first estimate of the shower energy to select the appropriate set of weights. A separate fit is used for each shower layer. The uncertainties on the fit parameters follow through to the uncertainty on the reconstructed energy and resulting energy resolution with the normal error propagation techniques.

Two ways of applying the weights are studied to test various aspects of the weighting procedure. In order to establish the ultimate performance limit, all events at a single energy are weighted using the weights calculated from that precise beam energy. If the energy resolution is not improved at this stage, then a practical application of weights that does not use knowledge of the beam energy will also not improve the energy resolution. In order to verify the fitting of weights as a function of beam energy a second method to apply weights is tested. In this test, weights are calculated for each event using the fitted function at the precise beam energy of the events. If the resulting energy resolution is very different from the first test results, then this implies an incorrect fitting procedure. Once these parts of the weighting procedure are verified, events can be weighted by calculated weights from the fitted function using the number of hits in the event. This is the desired method to select and apply weights so that the beam energy is not used at the application stage. Normalized weights that are generated at each energy are also referred to as the direct weights or directly calculated weights. Weights calculated from the fits are referred to as the weights from fits or from fitted functions. One set of weights can be calculated from fitted functions using the beam energy and other can be calculated using the number of hits from each event.

A fourth method to apply weights is explored in order to apply weights in the energy range in which the third method improved energy resolution but use unweighted hits in the energy range that it degraded energy resolution.

The two methods that use the beam energy give essentially the same result. The energy resolution is improved for the entire range. There is a larger improvement for low energy than high energy in both cases. Although this is not a valid procedure to apply weights in a real experiment it serves as a validation of the method and set a theoretical limit for any real or practical implementation of the weighting method. If resolution is not improved in either of these cases then any procedure to weight without knowledge of the beam energy will not work either.

7.5.1 Pions in Data and Simulation

Linearity and resolution of weighted (using beam energy) hits for pions is shown is figures 7.7 and 7.8 respectively. Since at each energy point, weights are normalized directly to a straight line the resulting weighted hits is almost perfectly linear. As a result the correction to resolution is nearly negligible. This provides the necessary benchmark to compare various applications of energy independent weighting algorithms.

The next step is to apply weights calculated from a fit function of the weights. The first part of this step is to calculate weights from the fitted functions directly at the beam energy of the runs. This will validate the quality of the fits if no degradation of energy resolution is observed. The following figures, 7.9 and 7.10 is the linearity and resolution response, respectively, of the pions using weights from fitted function at the beam energy.

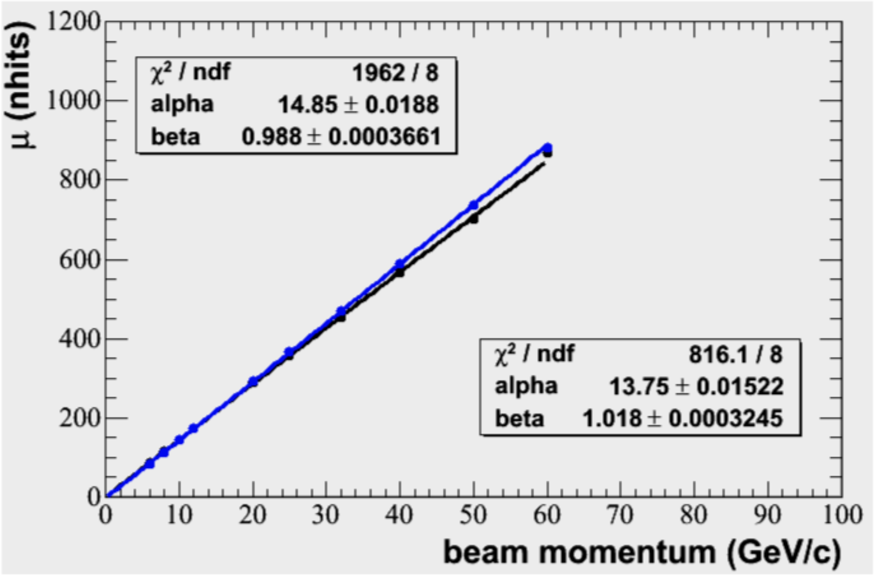

Figure 7.7. Linearity of weighted hits applying weights directly from run set at each beam energy.

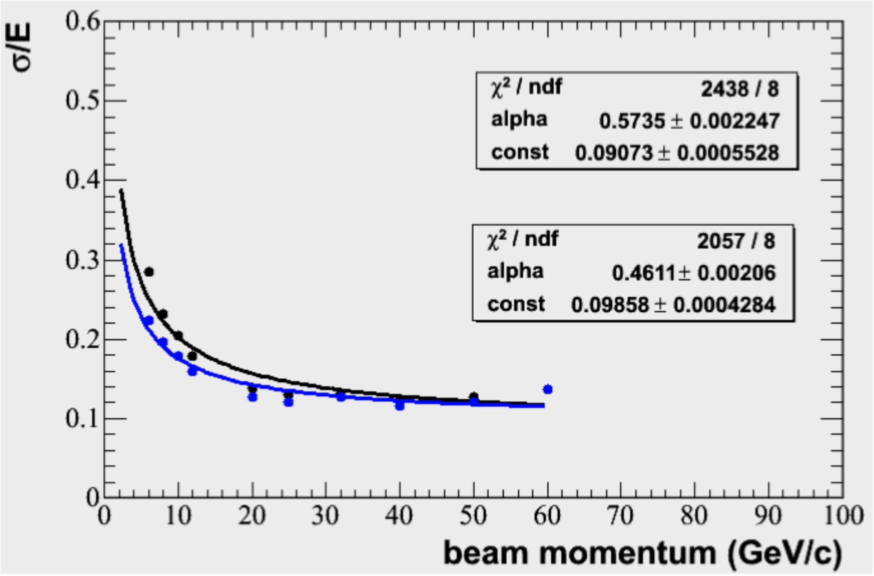

Figure 7.8. Resolution of weighted hits applying weights directly from run set at each beam energy.

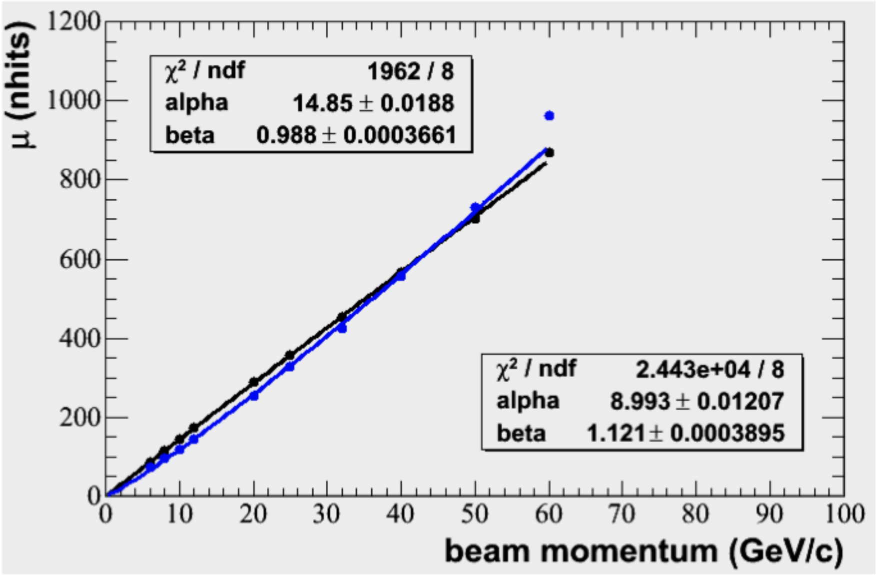

Figure 7.9. Linearity of weighted hits selecting weights from fits at the beam energy.

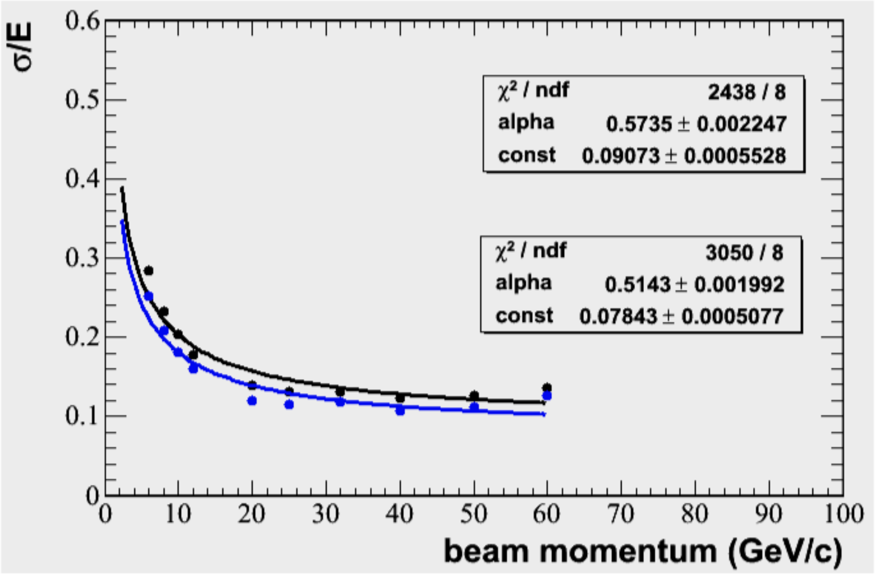

Figure 7.10. Resolution of weighted hits selecting weights from fits at the beam energy.

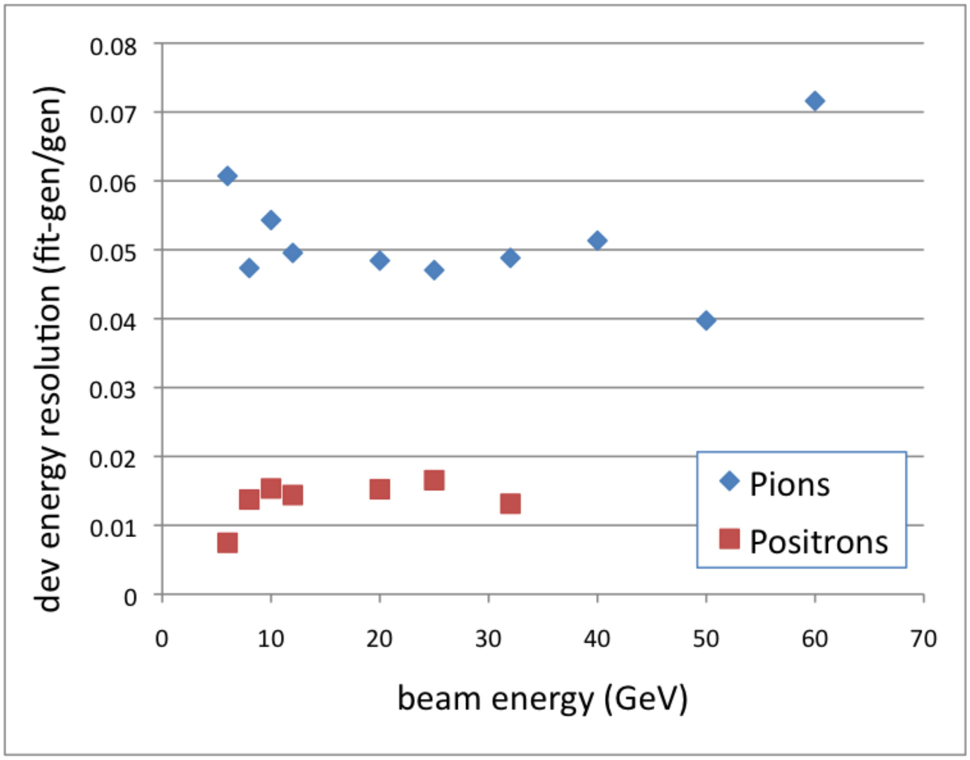

Figure 7.11. Graph showing the deviation of the energy resolution in calculating weights from fitted functions compared to directly generated weights. Simulated pions are in blue, positrons are in red.

There are some differences to notice in switching from using the directly generated weights or from their fits at the beam energy. First for the linearity response, weights applied from fits shows a degradation of the linearity from a straight line. This is because the fitted weights are forced to a straight line. Forcing the weights to have a linear dependence on energy improves the energy resolution by approximately 5% over using the directly generated weights. This improvement is shown if Figure 7.11. Deviations are positive for simulated pions and positrons, showing that forcing weights to a linear dependence on beam energy raises the theoretical upper limit for a practical application of weighting.

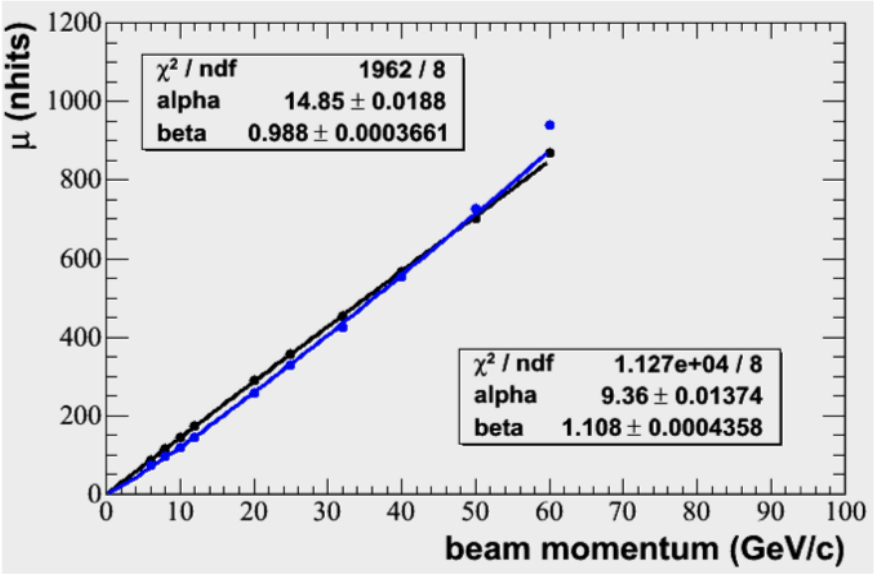

Figure 7.12. Linearity of Pions in Data using weights selected from fits at each event's energy.

Since the fitting of weights as a function of beam energy is validated, the fits provide a way of associating the energy of each event with an appropriate weight. This is the approach taken here. The energy used to select weights originates from the number of hits in each event. Hits are converted to energy by using the inverse of the power law fit to unweighted hits, i.e. the black line in figures 7.1 - 7.10, explicitly written as

$$E_i = \left(\frac{h_i}{\alpha}\right)^{1/\beta} \tag{7.30}$$

where $\alpha = 14.85$ and $\beta = 0.988$ and i runs over the events.

Figure 7.12 shows the linearity of the DHCAL response for weights selected using each event's reconstructed energy. Linearity is not affected much by using the calculated weights from fits as shown by comparing Figure 7.12 to Figure 7.9. The energy resolution when using weights from fits calculated from each event nhits is shown in figure 7.13.

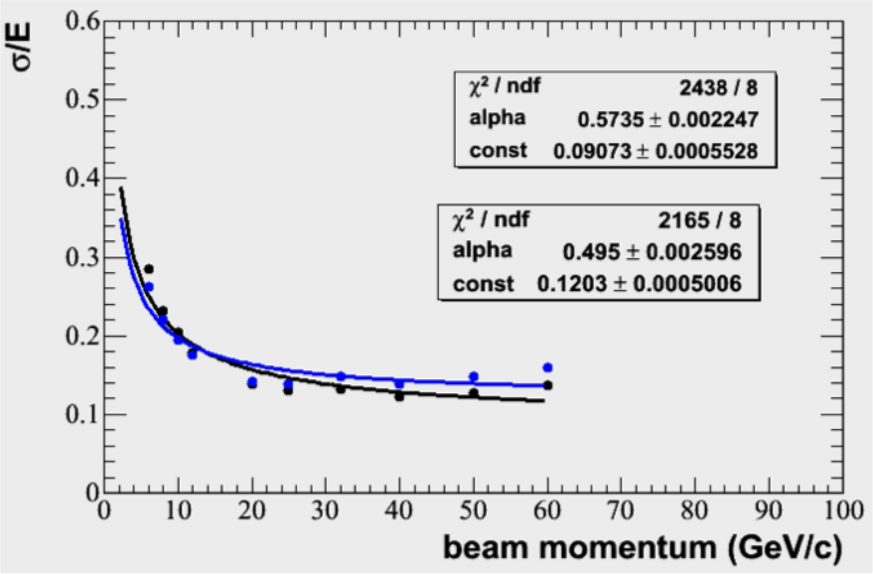

Figure 7.13. Resolution of Pions in Data using weights selected from fits at each event's energy.

The same procedure of normalization is also applied to simulation in the same energy range. Weights applied to simulated pions produce comparable results to data but simulated pions only require a single iteration to normalize weights to produce a linear response.

Figure 7.14 shows the initial response of un-normalized weights applied to simulated pions. No significant difference is observed between data and simulation at this stage.

Normalized weights are applied directly to the set of runs from which they were calculated. There is a noticeable difference here between simulation and data that the improvement to energy resolution remains constant throughout the entire energy range. Figure 7.15 shows the linearity and figure 7.16 the resolution for simulated pions response.

Continuing the procedure used for data, the weights are fit with a straight line as a function of beam energy. Figures 7.17 and 7.18 show the linearity and resolution of

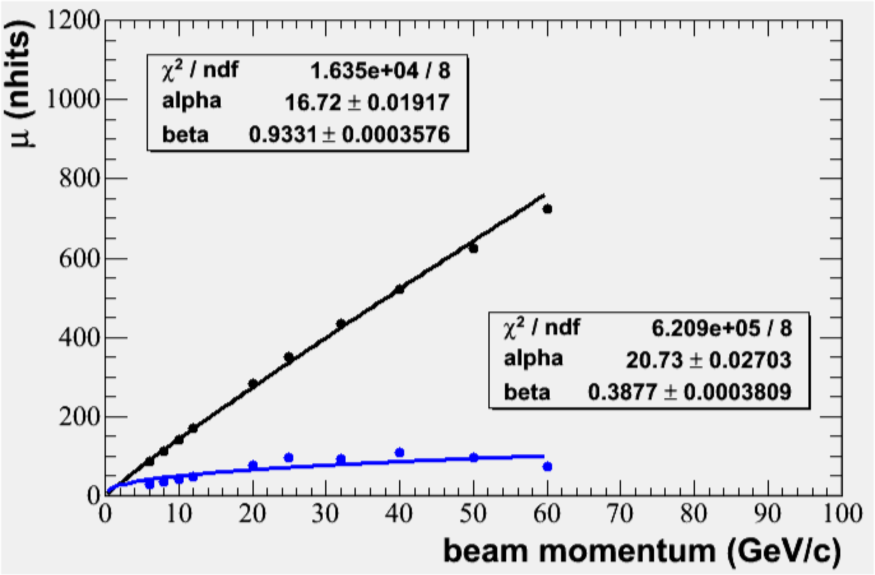

Figure 7.14. Raw weights applied to simulated pions.

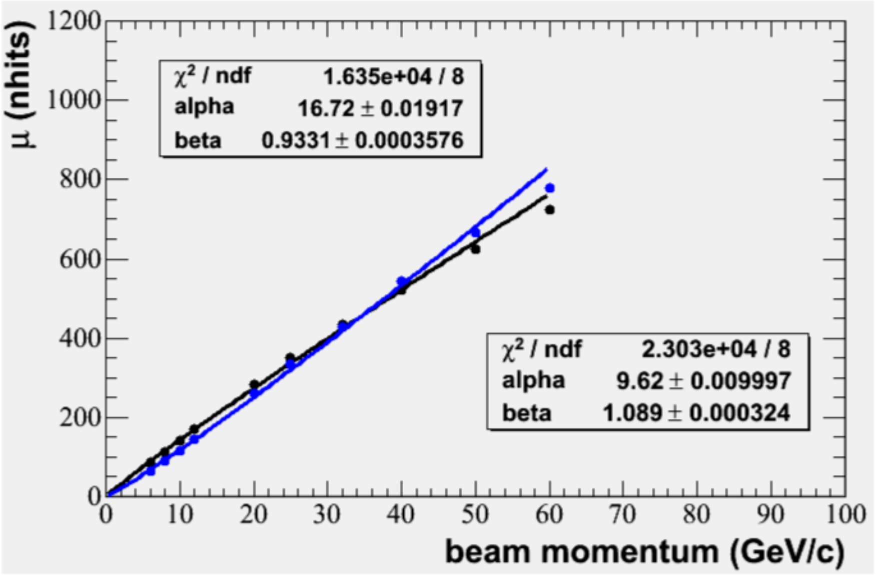

Figure 7.15. Linearity of weighted hits applying weights directly from run set at each beam energy for simulation.

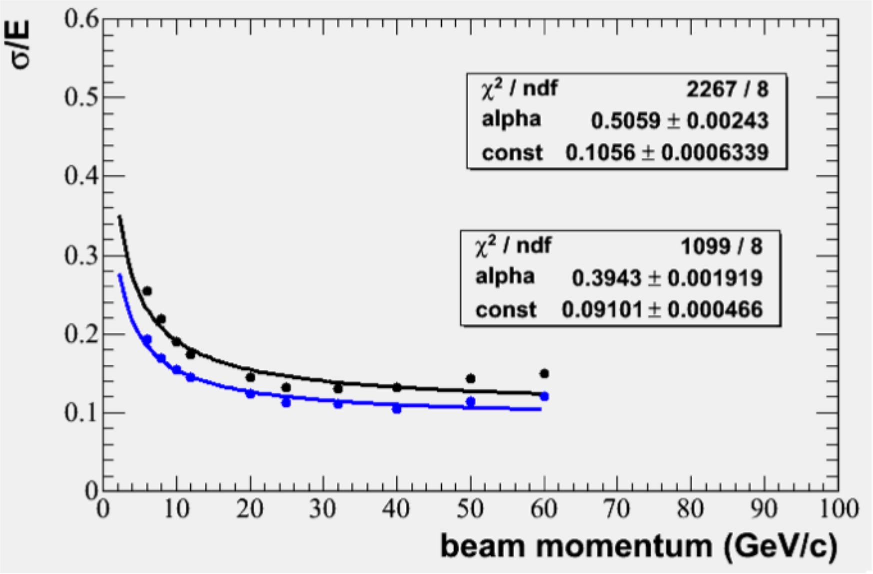

Figure 7.16. Resolution of weighted hits applying weights directly from run set at each beam energy for simulation.

weighted hits when the weights are selected and applied on an event basis according to the energy calculated from the event hits. Linearity is again maintained and a similar behavior for the resolution is observed. Weights improve resolution through 12 GeV runs but resolution increasingly degrades with beam energy.

Weights have been applied to pions in data and simulation. The effect of weighting on energy resolution is the same for both data and simulation showing a limited range of improvement. The same procedure is applied to positrons as a check to verify the method and test algorithms. It is clear from the case of calculating weights without the beam energy, i.e. using weights parameterized by the beam energy, that there is a limitation to the use of this longitudinal calibration. A more thorough study to the cause of the degradation of energy resolution is described in Section 7.5.6.

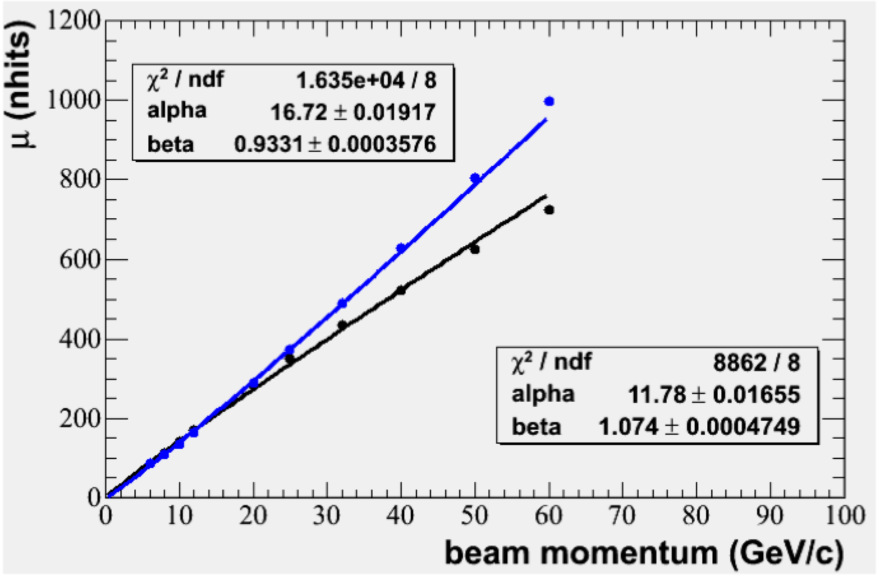

Figure 7.17. Linearity of weighted hits selecting weights from fits at each event's energy.

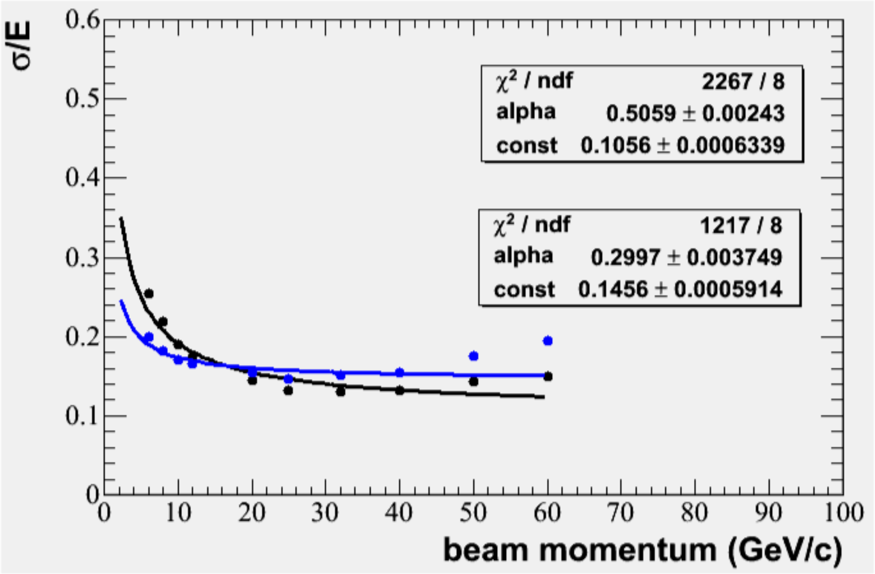

Figure 7.18. Resolution of weighted hits selecting weights from fits at each event's energy.

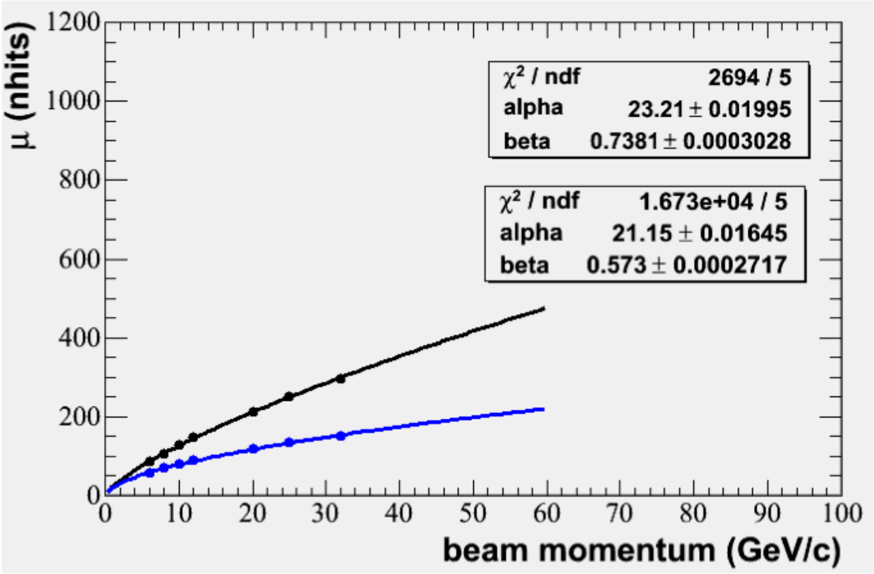

Figure 7.19. Raw weights applied to simulated positrons..

7.5.2 Positrons

No modification to the weighting procedure is required for positrons. Weights were generated, normalized, fitted and applied using simulated positrons in the energy range 6 - 32 GeV. Positrons serve as a validation of the code and the method. Weighting is expected to result in maximum improvement of energy resolution in positrons since similar methods have been applied and shown to be successful [46, 47]. From the derivation of weights using the correlations in energy deposition between the layers it is also assumed that the improvement of energy resolution increases with the strength of correlations.

Again, the first step is to weight the positrons with the raw weights and iterate normalizations until a linear response is achieved. The DHCAL does not have a linear response to positrons but forcing linearity removes the need to correct the resolution for non-linearity. Linearity of positrons using un-normalized weights are shown in figure 7.19.

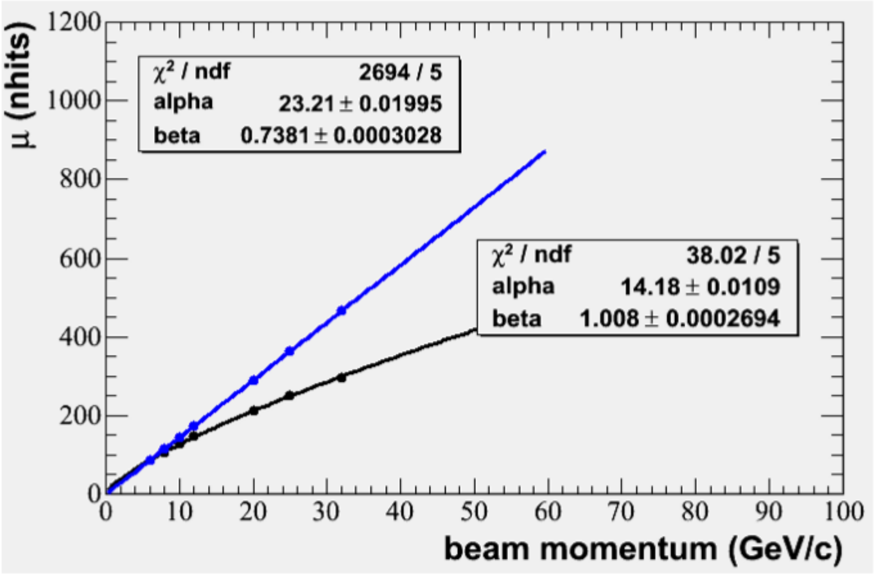

Figure 7.20. Linearity of simulated positrons with weights applied directly.

As is evident from the beta parameter in the power law fit to positron linearity in Figure 7.20 , only one iteration of normalization is required to produce a linear response over the entire energy range. Weights used for Figure 7.20 and 7.21 are the directly generated weights at each beam energy. Linearity is a result of normalization rather than weights.

Weighting improves the energy resolution by about 35% at each energy point, as shown in figure 7.21. As a comparison, pions showed a decreasing improvement in energy resolution with increasing beam energy.

Figures 7.22 and 7.23 are the linearity and resolution response of, respectively, when calculating weights by using the beam energy as input into the weights' fitted functions. As shown previously in Figure 7.11 there is only a 1% difference between using weights calculated from fits at the beam energy compared to using directly generated weights. This serves to validate the fits and also provides a best case

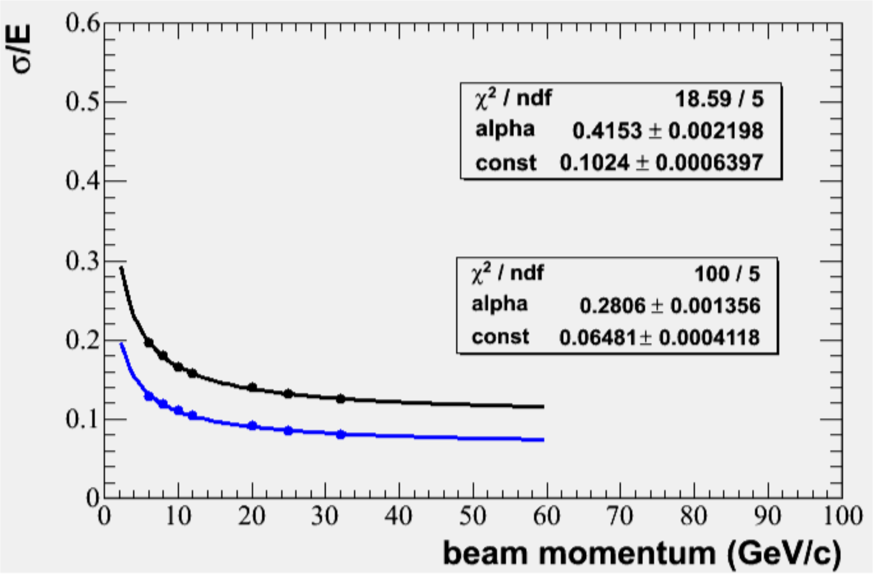

Figure 7.21. Resolution of simulated positrons with normalized weights applied directly.

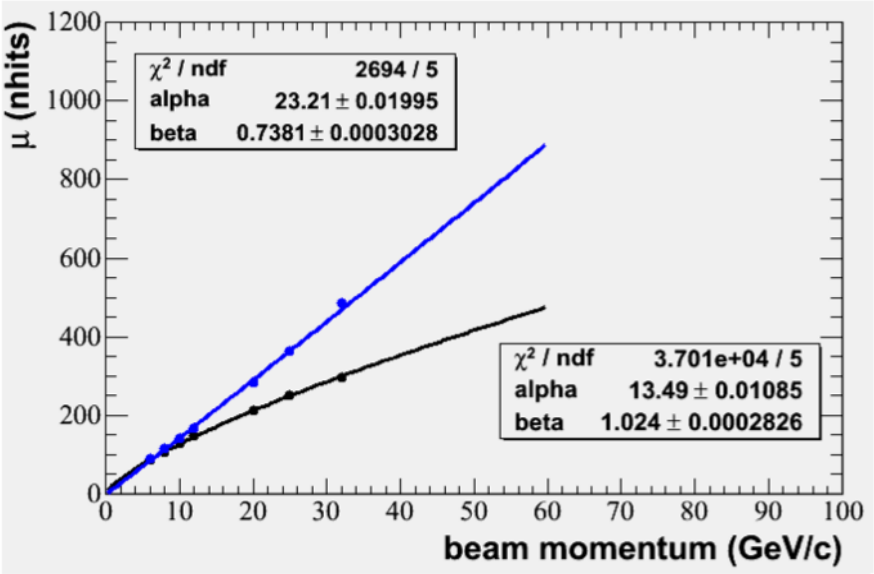

Figure 7.22. Linearity of simulated positrons with weights selected from fits at each beam energy.

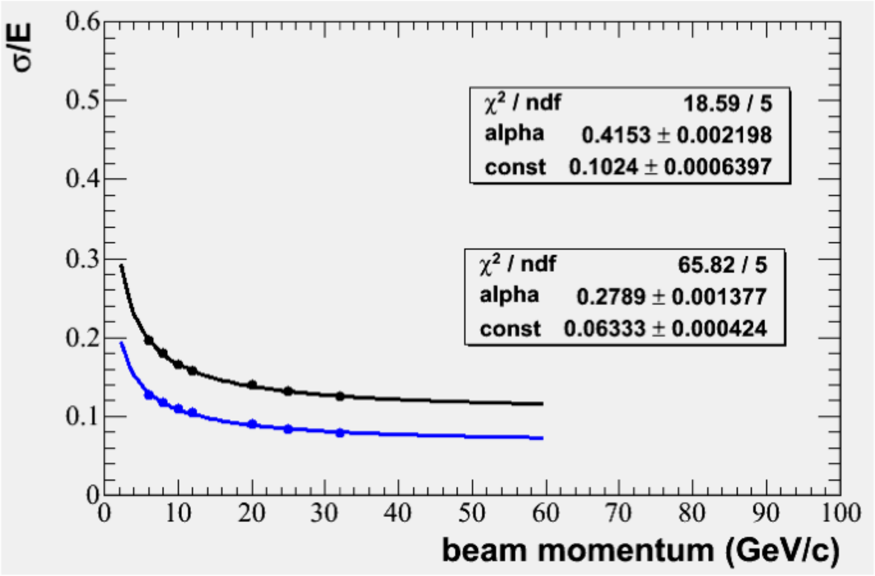

Figure 7.23. Resolution of simulated positrons with weights selected from fits at each beam energy.

scenario for using each event's energy to calculate weights from fits since the latter scenario fluctuates with the number of hits event to event.

With the fits validated it is possible to use the fits to calculate weights on an event by event basis using the energy calculated from nhits, the total number of hits in an event. As in the pion case, nhits in each event is converted to energy using equation 7.30 but with $\alpha = 23.21$ and $\beta = 0.7381$ that come from the power law fit to unweighted positrons nhits distributions. Linearity and resolution using each events' reconstructed energy (not beam energy) is shown in figures 7.24 and 7.25 , respectively.

7.5.3 Summary of Weighting Pions and Positrons

A summary of the relative improvement in energy resolution is shown in figure 7.26. Each point is the relative difference between weighted and un-weighted energy resolutions. The different colors represent the application of weights to data pions

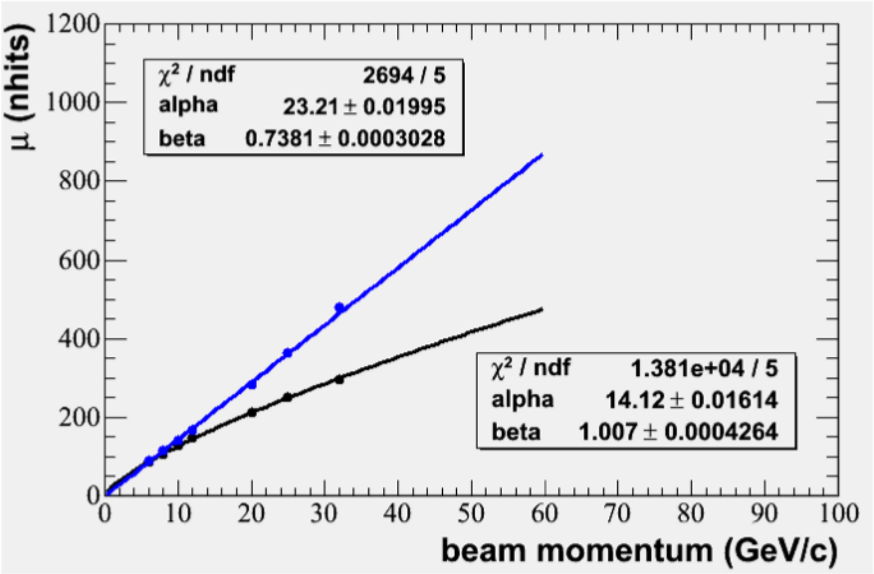

Figure 7.24. Linearity of simulated positrons weights selected from fits at each events' energy.

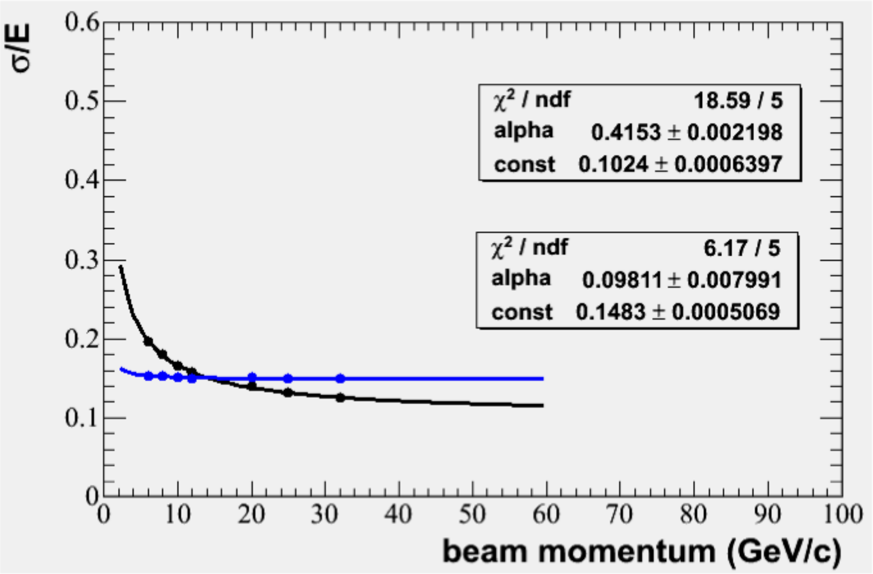

Figure 7.25. Resolution of simulated positrons with weights selected from fits at each events' energy.

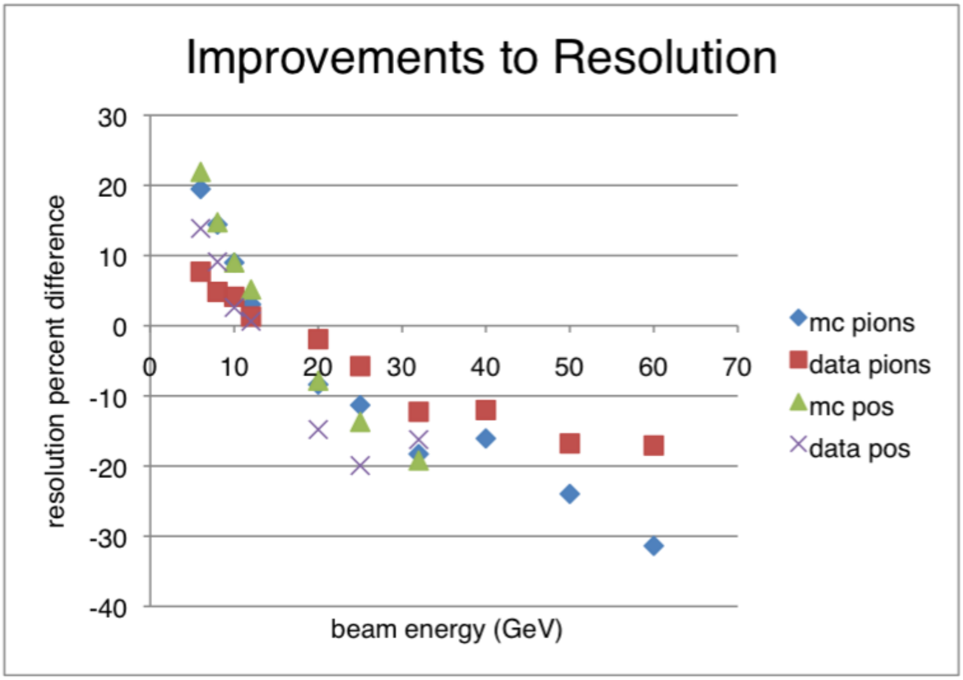

Figure 7.26. Summary graph of the effects of weighting on the various particle types from data and simulation..

and data positrons as well as simulated pions and simulated positrons. A line at zero represents the unweighted case. These are the cases of applying weights based on nhits from each event, independent of knowledge of the beam energy. As it shows, weighting improves the energy resolution for all types below 20 GeV. The 32 GeV point for data positrons is higher due to the low purity of positron ID, see Section 5.2.2. The dependence of the energy resolution improvement on events' nhits is described in Section, 7.5.6.

7.5.4 Energy-Cutoff Weights Application

The longitudinal weighting technique only improves the energy resolution up to 12 GeV. A method is developed to turn off the application of weights for events with energies outside this range. The decision to apply weights is calculated based on the energy calculated on an event by event basis. If the energy falls below the cut then weights are selected from fits as usual, otherwise weights are not applied.

Since resolution is slightly improved for 12 GeV and slightly degraded for 20 GeV it is likely that 16 GeV is a neutral region. Assuming that the spread of 16 GeV pions is the average of that from 12 and 20 GeV runs, the cutoff energy is set to 2σ lower than 16 GeV, or 11.104 GeV. Since test beam provides mono-energetic beams, the energies of events are grouped by beam energy. At the transition, 11.104 GeV, many events in 10 GeV and 12 GeV samples are not weighted. Therefore, the energy scale of weighted reconstructed energy and un-weighted reconstructed energy need to match. Figure 7.27 shows the reconstructed energy (expressed in hits) at 10 GeV when un-weighted events are not matched to the same energy scale as weighted events. The peak on the left is weighted using the event's hits while the peak on the right is un-weighted. This is due to the fact that energy is calculated from nhits using a power law fit to the un-weighted hits as shown by the black line in figure 7.28 but the fully weighted linearity follows a different power law. Therefore the un-weighted events should be scaled to match the energy scale of the weighted events. Its best not to scale weighted events due to the normalization of weights.

In order to find the correct scaling of un-weighted events, two power law functions are used. In the case events are weighted in the entire energy range, as in Section 7.5.1, then the power law fit in the linearity plots provides a conversion of hits to energy, denoted 'wenergy'. If hits are converted to energy using the power law fit of linearity when no weighting is applied at any energy, then this is just called energy. In the case that events in only part of the entire energy range are weighted,

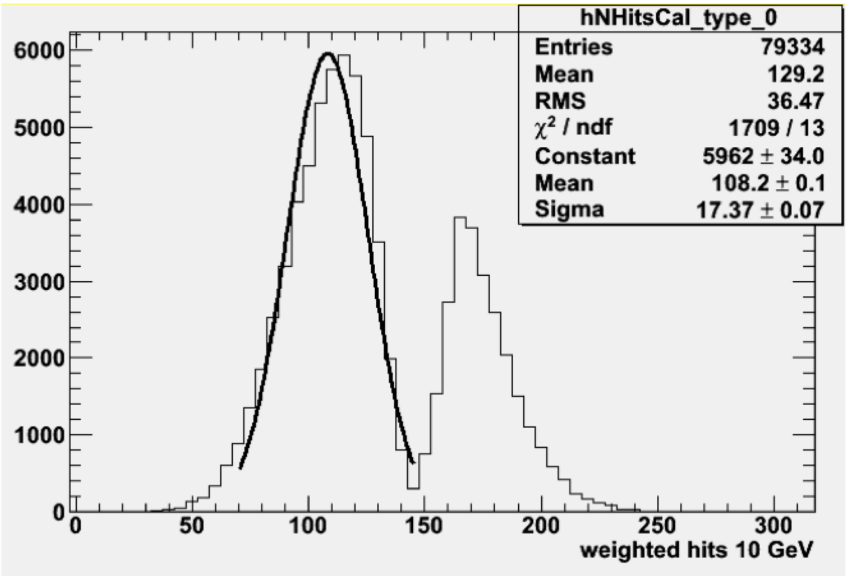

Figure 7.27. Distribution of nhits from 10 GeV pions partially weighted.

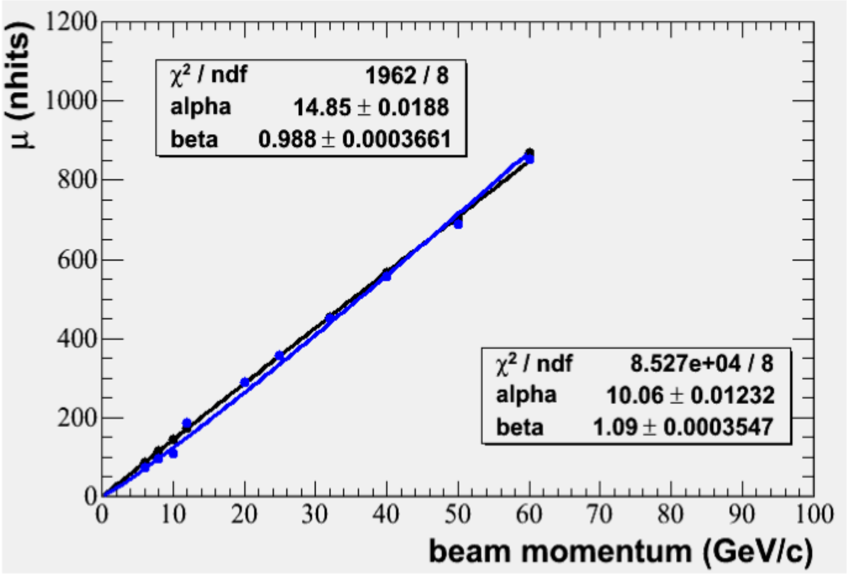

Figure 7.28. Linearity of un-weighted and paritally weighted events..

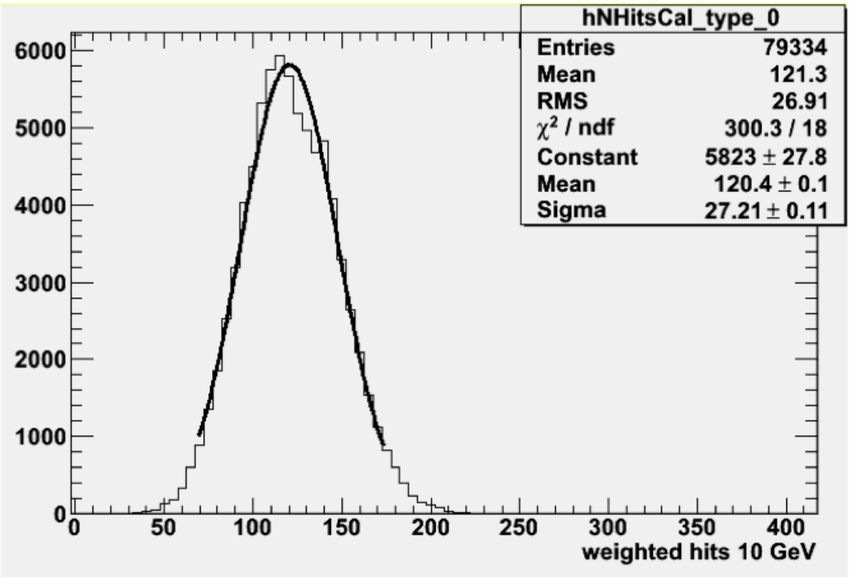

Figure 7.29. Distribution of nhits from 10 GeV pions using a scaled partially weighted.

it is possible to scale unweighted events to match the energy scale of weighted events by applying the ratio,

$$ediff = (wenergy - energy)/energy. \tag{7.31}$$

Then the events above 11.104 GeV are scaled by multiplying by $1 - ediff$. This scales the un-weighted events to the power law fit of weighted events. Comparing the results of this scaled partial weighting at 10 GeV, 7.29, to the previously unscaled partial weighting of 7.27 shows that the double peaks are fixed. The linearity of partially weighted events without scalling and completely un-weighted events is shown in Figure 7.28. Conversely, Figure 7.29 shows the linearity when the partially weighted events are scaled by $1 - ediff$. This procedure of scaling unweighted events is called scaled partial weighting.

The energy resolution of scaled partially weighted events should match the fully weighted resolutions in the transition region, if the scaling is implemented properly. The scaled partially weighted energy resolution is within 2% of the fully weighted

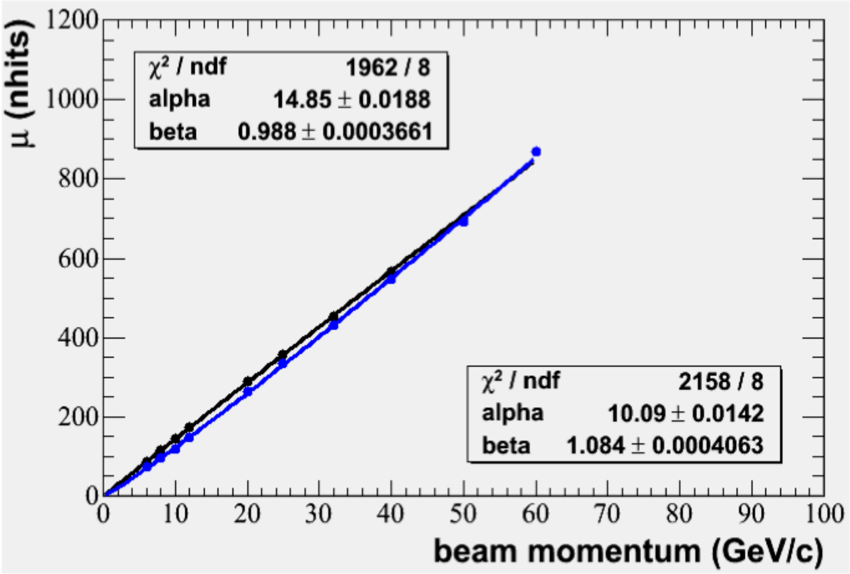

Figure 7.30. Linearity of un-weighted and scaled partial weighted events.

case for both 10 and 12 GeV pions. The linearity and resolutions for scaled partial weighting is shown in figures 7.30 and 7.31, respectively.

Overall, the scaled partial weighting procedure has no end affect on the energy resolution. This may be due to unknown systematics or other uncertainties added during the procedure. The effect of scaling and linearity correction doesn't cancel exactly. At this point no claim about energy resolution improvement or degradation can be made about weighting events below an energy-cutoff. If a longitudinal calibration using the hits in each layer is going to be pursued then a partially weighting technique should be explored further.

7.5.5 Strength of Correlations

One possible source of limitation to improving the energy resolution with longitudinal weights could be that the correlations are too small. This may be suggested by the fact that energy resolution is most improved in positrons which have a known

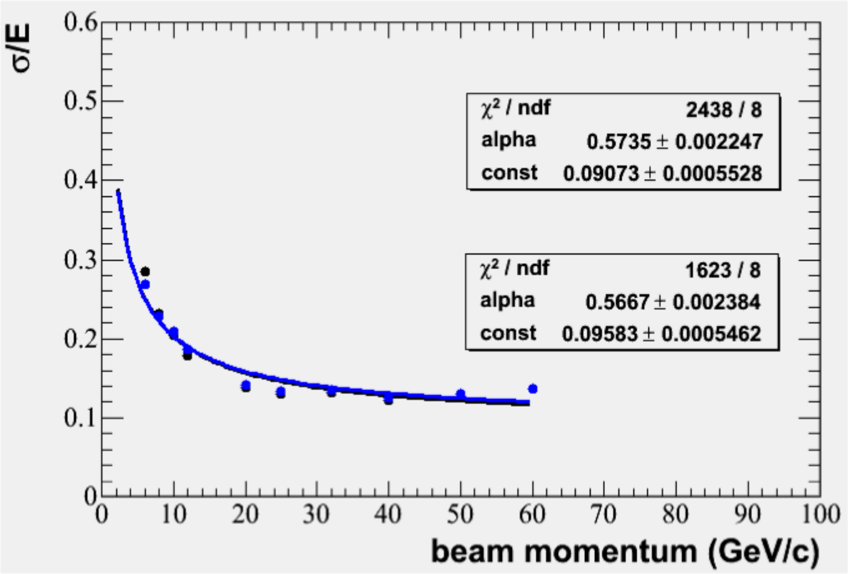

Figure 7.31. Resolution of un-weighted and scaled partial weighted events.

correlation in the energy deposition between layers. Conversely, resolution was not improved by using weights calculated from the linear fitted functions. Figure 7.32 shows the covariance matrix elements for the number of hits in each layer for 50 GeV pions. Each box represents a number of entries so that big boxes contain many entries. If the hits in each layer are perfectly un-correlated to each other then only the diagonal elements are filled, but this is not the case. A definition for the strength of correlations, S, has the form

$$S = \frac{\sum R_{ij} N_{ij}}{\sum N_{ij}} \tag{7.32}$$

where R_{ij} is the distance that each element is away from the diagonal and N_{ij} is the number of entries in each bin. Equation 7.32 gives a measure of the amount of correlation so that large values of S represent large correlations between the hits in each layer.

Correlation strength is plotted in Figure 7.33 for data pion, simulated pion, data positrons, simulated positrons and simulated pions with the correct leakage

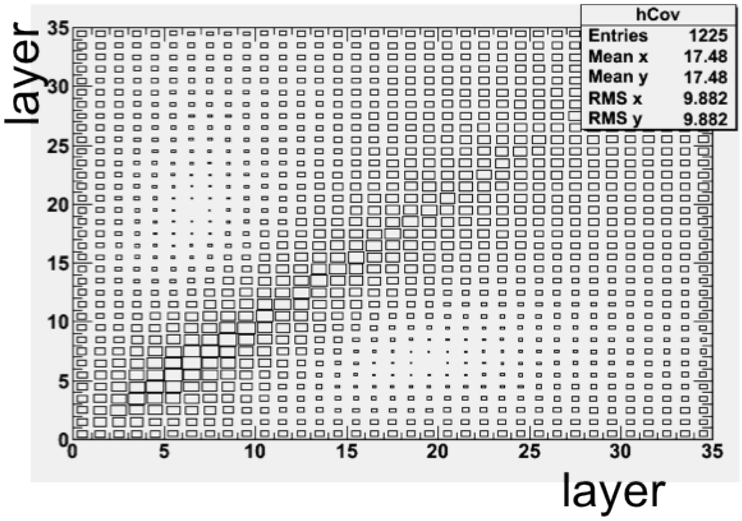

Figure 7.32. Covariance matrix at 50 GeV.

requirements. Correlation strength also showed that a previous definition of leakage was biased. The bias therefore affected weights so it was abandoned for the current one, shown by the light blue crosses. The other pion points are only present to illustrate the affect of poor leakage requirements.

The main point of Figure 7.33 is the dependence of correlation on beam energy and the difference between pions (blue crosses only) and positrons. Positrons have a stronger correlation than pions and are rather constant between 6 and 32 GeV. Pions have a slight dependence on beam energy. There is no clear evidence of an effect between 12 and 20 GeV where the transition between improvement and degradation occur for energy resolution using longitudinal weights. Therefore, the lack of correlations in the deposited energy between layers does not limit the improvement to the energy resolution.

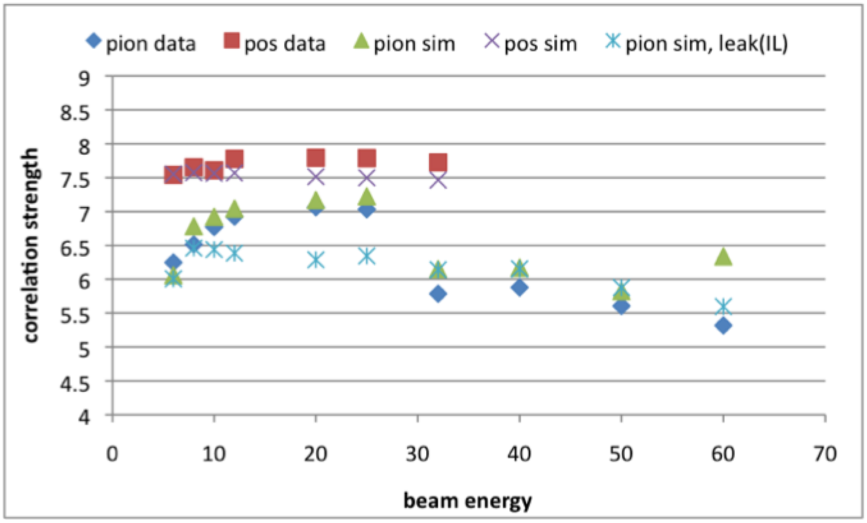

Figure 7.33. Correlation strength for data pion, simulated pion, data positrons, simulated positrons and simulated pions with the correct leakage requirements.

7.5.6 Dependence of Weighting on Initial Energy Resolution

Up through 12 GeV, longitudinal calibration (from this point forward the energy range is assumed to be from 6 - 60 GeV) improves energy resolution for both pions and positrons when calculating weights based on the number of hits from events. In this section, evidence is shown to suggest that the fluctuations of nhits, at a particular beam energy, limits the improvement of energy resolution. Weight selection for each event fluctuate proportionately to nhits so its possible to test the dependence of weight calculation to nhits fluctuations.

Weight calculations from fitted functions are done on an event by event basis where the total number of hits is converted to an energy using the inverse power function coming from the unweighted linearity response and has the form

$$E' = \left(\frac{n'}{\alpha}\right)^{1/\beta} \tag{7.33}$$

. This converted energy is the input energy for the linear functions of weights in each shower layer with the form

$$weight = a + bE' \tag{7.34}$$

. The initial, unweighted, energy resolution determines how much the nhits at a particular beam energy, fluctuate from event to event. In turn, fluctuations of nhits results in a fluctuation in the calculation of weights.

To study the source of the degradation of energy resolution at higher energies a test was devised to artificially reduce the resolution, used to calculate weights, of a high energy pion run, 50 GeV, to each of the lower energy runs' resolution. The number of hits n' , as in Equation 7.33 used to calculate weights is scaled with f according to

$$f(n - n_0) = (n' - n_0) \tag{7.35}$$

where, n_0 is the mean number of total hits at a lower (relative to the mean total hits at 50 GeV) beam energy and n is the unweighted number of hits from an event. Taking n' from Equation 7.35 as the input to Equation 7.33 a scaled weight is calculated and used to weight the event. Essentially the scale factor f allows 50 GeV pions to calculate weights using the fluctuations from an energy with a smaller fluctuation. The resulting energy resolutions for each scale factor f is shown in Figure 7.34. Each point is the energy resolution at 50 GeV when weights were calculated assuming the energy resolution at a lower energy. The x-axis is given in units of scale factor, f, with the lowest coming from 6 GeV and the highest at 25 GeV.

The energy resolutions in Figure 7.34 use the linear fitted functions for weights that are parameterized by energy. The energy resolution scales linearly with decreasing resolution. A scale factor of 0 represents using the weights' fitted functions at the precise beam energy. The energy resolution at 50 GeV using directly generated

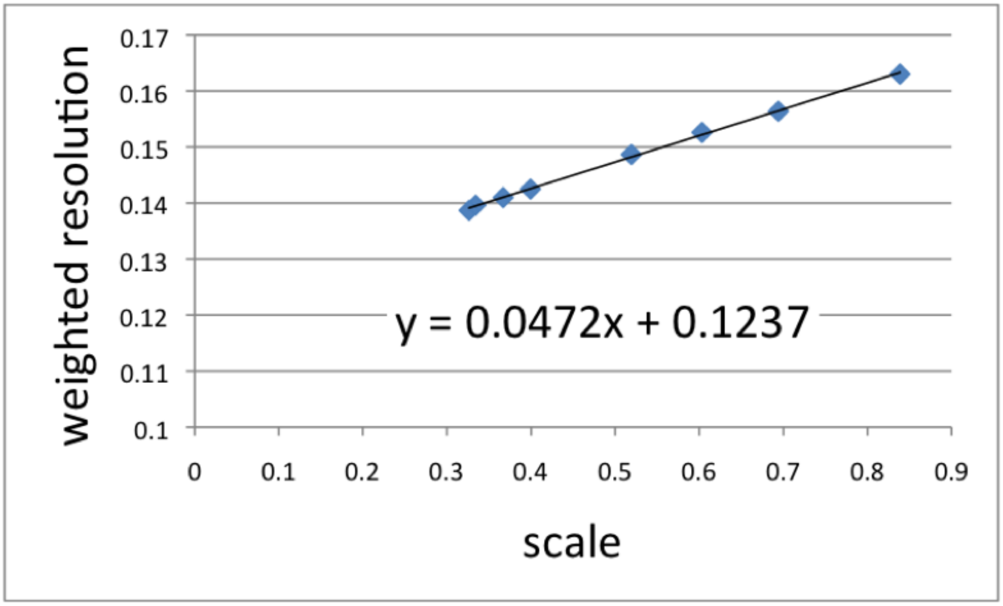

Figure 7.34. Energy resolution at 50 GeV when weights are calculated by artificially assuming the energy resolution of a lower energy according to the scale factor f.

weights is 0.124. So if the energy resolution at 50 GeV is artificially reduced to 0 then the energy resolution can be improved. But this is not possible if the number of hits is to be used to calculate the proper weights for each event. Therefore, the improvement to energy resolution using longitudinal weights is limited by the initial resolution.

CHAPTER 8

CONCLUSIONS

8.1 Results

The concept of digital hadron calorimetry is well on its way to be validated. Many test beam runs were performed taking tens of millions of events. A deeper understand of RPC technology and construction has been developed. Assumptions on the calibration of a digital calorimeter are being tested with various strategies as well as the use of shower variables to perform offline analysis such as the weighting procedure discussed in this thesis.

A very complex electronics readout system has been developed and extensively tested. The DHCAL presents a record number of channels in a hadron calorimeter. Experience in testing electronics in test beam environments has also lead to many insights about RPC performance in a calorimeter. These tests present the first digital pictures with very high granularity of hadron showers. As a result, this data is expected to provide new insight into the passage of hadrons through matter in addition to improvements on particle flow algorithms.

Weights are generated and applied to pions and positrons in data and simulation using the longitudinal behavior of showers to minimize energy resolution. Effects on the longitudinal behavior of pion showers due to fluctuations in the interaction layer have been minimized by basing weights on the start of the shower. It was found that longitudinal weights depend on the beam energy. Therefore the weight in each layer is parameterized to energy using a first order polynomial. Given this parameterization, events are calibrated using the number of hits in each event.

Single particle pion and positron energy resolutions are improved when using knowledge of beam energy (more so for positrons) but limited up through 12 GeV when weights are selected on an event by event basis using the energy calculated from each event.

Improvement to energy resolution is proportional to fluctuations of energy from event to event. The absolute value of these fluctuations, with respect beam energy, is a limiting factor to the energy range for which longitudinal calibration improves the energy resolution. The amount of fluctuations limiting resolution improvements is different between pions and positrons. Weighting events longitudinally has a fundamental limitation due to the fluctuation of nhits between events and the dependence of weights on energy. Although these weights improve single particle energy resolution when using the beam energy they do not improve resolutions when knowledge of the beam energy is removed. Some other parameterization of weights is most likely required.

Estimates of the systematic uncertainty on the weights due to particle ID have been calculated. It was found that the statistical uncertainty on the weights due to error propagation are highly underestimated and require additional work. Additional sources for systematic uncertainty, such as leakage and mis-identification of the true interaction layer, have been identified and explored. Rough estimates for systematics due to interaction layer s

8.2 Future Work

The technique presented in this thesis calculates weights according to layer number, hence longitudinal calibration. Specifically, it takes advantage of the correlations of hits, or energy deposited by pions and positrons, between layers. The binning of hits by layer is common to other analysis such as measurement of efficiency, multiplicity, and various calibration strategies. Correlations may also exist between hits that are grouped in other ways. The calculation of weights presented in this thesis only uses properties of hits in a group, rather than properties of the group itself. Therefore, further work will explore how correlations arise in alternate methods of grouping hits.

For example, density is used in this thesis for particle identification. Assigning a density value to hits suggests that grouping hits into bins of density may reveal a new correlation. A definition for the density of a hit is already provided. The next step is to partition hits into bins of density and identify possible correlations between density bins. Then it is possible to use the framework developed in this thesis to calculate a set of weights.

The energy dependence of weights, along with the initial energy resolution, seem to have the greatest impact on limiting the energy range that longitudinal weights can improve the energy resolution. An exact relationship between initial energy resolution and expected improvement has yet to be quantified.

Many possible sources of systematic uncertainty can still be explored. More exact methods for estimating systematics are possible with validated simulations. Further study can be performed on the Monte-Carlo technique to estimate the statistical uncertainties on weights.

REFERENCES

[1] D. Perkins, *Introduction to High Energy Physics.* Cambridge, UK: Cambridge University Press, 2000.

[2] J. Brau *et al.*, “International Linear Collider reference design report. 1: Executive summary. 2: Physics at the ILC. 3: Accelerator. 4: Detectors,” 2007.

[3] S. Glashow, “Partial Symmetries of Weak Interactions,” *Nucl.Phys.*, vol. 22, pp. 579–588, 1961.

[4] S. Weinberg, “A model of leptons,” *Phys. Rev. Lett.*, vol. 19, pp. 1264–1266, Nov 1967. [Online]. Available: http://link.aps.org/doi/10.1103/PhysRevLett.19.1264

[5] A. Salam, “Weak and Electromagnetic Interactions,” *Conf.Proc.*, vol. C680519, pp. 367–377, 1968.

[6] (2012) A generic fitter project for hep model testing. [Online]. Available: http://project-gfitter.web.cern.ch/project-gfitter/

[7] M. Baak *et al.*, “Updated status of the global electroweak fit and constraints on new physics,” *The European Physical Journal C*, vol. 72, pp. 1–35, 2012. [Online]. Available: http://dx.doi.org/10.1140/epjc/s10052-012-2003-4

[8] R. Kogler, “The electroweak sm fit,” presented at the Recontres de Moriond, La Thuile, March 2013.

[9] H. Aihara *et al.*, “SiD Letter of Intent,” 2009.

[10] T. Abe *et al.*, “The International Large Detector: Letter of Intent,” 2010.

[11] D. Buskulic *et al.*, “Performance of the ALEPH detector at LEP,” *Nucl.Instrum.Meth.*, vol. A360, pp. 481–506, 1995.

[12] B. E. Cox, F. Loebinger, and A. Pilkington, "Detecting higgs bosons in the b$\bar{b}$ decay channel using forward proton tagging at the lhc," *Journal of High Energy Physics*, vol. 2007, no. 10, p. 090, 2007. [Online]. Available: http://stacks.iop.org/1126-6708/2007/i=10/a=090

[13] S. Heinemeyer, V. Khoze, M. Ryskin, W. Stirling, M. Tasevsky, and G. Weiglein, "Studying the mssm higgs sector by forward proton tagging at the lhc," *The European Physical Journal C*, vol. 53, pp. 231–256, 2008. [Online]. Available: http://dx.doi.org/10.1140/epjc/s10052-007-0449-6

[14] L. Xia, "Development of a particle flow algorithm at argonne," *AIP Conference Proceedings*, vol. 867, no. 1, pp. 531–537, 2006. [Online]. Available: http://link.aip.org/link/?APC/867/531/1

[15] M. Thomson, "Particle flow calorimetry and the pandorapfa algorithm," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 611, no. 1, pp. 25 – 40, 2009. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0168900209017264

[16] D. Notz, "ZEUS: A Detector for HERA," 1990.

[17] S. Abachi *et al.*, "The D0 Detector," *Nucl.Instrum.Meth.*, vol. A338, pp. 185–253, 1994.

[18] Bruckmann *et al.*, "ON THE THEORETICAL UNDERSTANDING AND CALCULATION OF SAMPLING CALORIMETERS," 1987.

[19] C. Adams, A. Anastassov, and et. al., "Design, construction and commissioning of the digital hadron calorimeter - dhcal," in preparation.

[20] R. Santonico and R. Cardarelli, "Development of resistive plate counters," *Nuclear Instruments and Methods in Physics Research*,

vol. 187, no. 23, pp. 377 – 380, 1981. [Online]. Available: http://www.sciencedirect.com/science/article/pii/0029554X81903633

[21] A. M., "The simulation of resistive plate chambers in avalanche mode," in *IV International Workshop on Resistive Plate Chambers and Related Detectors*, Castel dell'Ovo, Napoli, Oct. 1997, pp. 197–219.

[22] A. Akindinov, A. Alici, and et. al., "Study of gas mixtures and ageing of the multigap resistive plate chamber used for the alice tof," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 533, no. 12, pp. 93 – 97, 2004, ¡ce:title¿Proceedings of the Seventh International Workshop on Resistive Plate Chambers and Related Detectors¡/ce:title¿. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0168900204014111

[23] B. Bilki, J. Butler, and et. al., "Measurement of the rate capability of resistive plate chambers," *Journal of Instrumentation*, vol. 4, no. 06, p. P06003, 2009. [Online]. Available: http://stacks.iop.org/1748-0221/4/i=06/a=P06003

[24] P. Camarri, R. Cardarelli, A. Ciaccio, and R. Santonico, "Streamer suppression with sf6 in rpcs operated in avalanche mode," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 414, no. 23, pp. 317 – 324, 1998. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0168900298005762

[25] (2011) Calice software. [Online]. Available: https://twiki.cern.ch/twiki/bin/view/CALICE/SoftwareMain

[26] M. Brandel. (1999, July) 1963, the debut of ascii. Internet article. [Online]. Available: http://edition.cnn.com/TECH/computing/9907/06/1963.idg/

[27] B. Bilki, J. Butler, and et. al., "Calibration of a digital hadron calorimeter with muons," *Journal of Instrumentation*, vol. 3, no. 05, p. P05001, 2008. [Online]. Available: http://stacks.iop.org/1748-0221/3/i=05/a=P05001

[28] (2012) Badger air-brush co. air-opaque paints. [Online]. Available: http://www.badgerairbrush.com/air-opaque.asp

[29] (2013) Beam details. [Online]. Available: http://www-ppd.fnal.gov/MTBF-w/beam/energy.html

[30] (2013) Mtest beam details. [Online]. Available: http://www-ppd.fnal.gov/MTBF-w/beam/purity.html

[31] J. Repond, private communication, 2010.

[32] B. Bilki, private communication, 2013.

[33] P. CMS. (2006) Mokka. [Online]. Available: http://polzope.in2p3.fr:8081/MOKKA

[34] S. Agostinelli, J. Allison, and et. al., "Geant4: a simulation toolkit," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 506, no. 3, pp. 250 – 303, 2003. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0168900203013688

[35] J. Allison *et al.*, "Geant4 developments and applications," *Nuclear Science, IEEE Transactions on*, vol. 53, no. 1, pp. 270–278, Feb.

[36] B. Bilki, J. Butler, and et. al., "Measurement of positron showers with a digital hadron calorimeter," *Journal of Instrumentation*, vol. 4, no. 04, p. P04006, 2009. [Online]. Available: http://stacks.iop.org/1748-0221/4/i=04/a=P04006

[37] G. Drake, J. Repond, *et al.*, "Resistive plate chambers for hadron calorimetry: Tests with analog readout," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated*

Equipment, vol. 578, no. 1, pp. 88 – 97, 2007. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0168900207007577

[38] D. Underwood *et al.*, “Rpc investigation using finely spaced 2-d strip readout,” in *Nuclear Science Symposium Conference Record, 2007. NSS ’07. IEEE*, vol. 1, 26 2007-Nov. 3, pp. 618–622.

[39] (2008) Reference physics lists. [Online]. Available: http://geant4.cern.ch/support/proc_mod_catalog/physics_lists/referencePL.shtml

[40] C. Adloff *et al.*, “Validation of geant4 monte carlo models with a highly granular hadron calorimeter,” in preparation.

[41] G. I. Lykasov, G. H. Arakelyan, and M. N. Sergeenko, “The quark–gluon string model: soft and semihard hadronic processes,” *Physics of Particles and Nuclei*, vol. 30, no. 4, pp. 343–368, 1999. [Online]. Available: http://link.aip.org/link/?PAN/30/343/1

[42] *Physics Reference Manual*, geant4 9.6.0, GEANT4, 2012.

[43] M. Guthrie, R. Alsmiller, and H. Bertini, “Calculation of the capture of negative pions in light elements and comparison with experiments pertaining to cancer radiotherapy,” *Nucl.Instrum.Meth.*, vol. 66, pp. 29–36, 1968.

[44] H. Bertini and M. Guthrie, “News item results from medium-energy intranuclear-cascade calculation,” *Nucl.Phys.*, vol. A169, pp. 670–672, 1971.

[45] C. Adloff *et al.*, “Construction and Commissioning of the CALICE Analog Hadron Calorimeter Prototype,” *JINST*, vol. 5, p. P05004, 2010.

[46] M. Delmastro and M. Fanti, “Energy resolution optimization through layers measurements weighting: analytical solutions and numerical strategies.” 2002.

[47] A. Peryshkin and R. Raja, “On sampling fractions and electron shower shapes,” 2011.

[48] (2013) Root. [Online]. Available: http://root.cern.ch/drupal/

BIOGRAPHICAL STATEMENT

Jacob Russell Smith was born in Austin, TX in 1982. He received his B.S. in 2006, M.S. in 2009 and Ph.D. degrees from The University of Texas at Arlington in 2013, all in Physics. From the middle of his undergraduate years to the end of his M.S. degree in 2006 he worked on R&D for GEM detectors for future HEP experiments. He continued detector R&D at Argonne National Laboratory starting the summer 2009 through the end of his graduate career. While there, he lead many aspects of construction, testing and analysis for the world's first large-scale hadron calorimeter with digital readout and embedded front-end electronics, providing never before seen detailed images of hadronic showers. His current research interest is in High Energy Physics experiments analysis and detector development. He is a member of the American Physical Society and its Forum on International Physics.