Performance of High Granularity Calorimeter prototypes for the CMS HL-LHC upgrade in beam test experiments at CERN

A thesis

submitted in partial fulfillment of the requirements

Of the degree of

Doctor of Philosophy

By

Shubham Pandey
ID No.: 20142022

INDIAN INSTITUTE OF SCIENCE EDUCATION
AND RESEARCH PUNE

2022

CERTIFICATE

Certified that the work incorporated in the thesis entitled *Performance of High Granularity Calorimeter prototypes for the CMS HL-LHC upgrade in beam test experiments at CERN* submitted by Shubham Pandey was carried out by the candidate, under my supervision. The work presented here or any part of it has not been included in any other thesis submitted previously for the award of any degree or diploma from any other University or institution.

Dr. Seema Sharma

Date: 3 January 2022

Declaration

I declare that this written submission represents my ideas in my own words and where others' ideas have been included, I have adequately cited and referenced the original sources. I also declare that I have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in my submission. I understand that violation of the above will be cause for disciplinary action by the Institute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when needed.

Shubham Pandey
Registration ID: 20142022

Date: 3 January 2022

If I have seen further, it is by standing on the shoulders of Giants.

Issac Newton (1642–1727)

Abstract

The CMS collaboration has opted for a High Granularity Calorimeter (HGCAL) to replace the current endcap electromagnetic and hadronic calorimeters in the view of high-luminosity phase of the LHC (HL-LHC). The HL-LHC, expected to start at the end of this decade, aims to accumulate proton-proton collision data corresponding to an integrated luminosity of 3000 fb^{-1} over a span of ten years, which is ten times more data than the ongoing phase of the LHC. High instantaneous luminosity needed to achieve these goals will result in an average of 140 to 200 proton-proton interactions per bunch crossing. These pose very stringent requirements on radiation tolerance of the endcap detectors as well as their pileup mitigation capabilities. The HGCAL is a silicon and scintillator based sampling calorimeter with unprecedented longitudinal and transverse granularity, which will facilitate efficient particle-flow reconstruction, particle identification, and pileup rejection.

The CMS HGCAL collaboration is extensively testing the detector components and its associated electronics in test bench based experiments and in beam test experiments with single particles. A prototype of the electromagnetic and hadronic section of the HGCAL was built and tested along with the CALICE Analogue Hadron Calorimeter (AHCAL) prototype in the beams of single particles at CERN's Super Proton Synchrotron (SPS) in October 2018. The combined prototype was exposed to the beams of positrons and charged pions with momenta ranging from 20 to 300 GeV/c, and muons of 200 GeV/c. This thesis presents the performance of individual prototype silicon sensor modules as well as construction, commissioning, calibration of the HGCAL prototype used in the beam test experiments. A first detailed study of performance of combined HGCAL and AHCAL prototype to hadronic showers generated by charged pions using the data collected in October 2018 is reported in this thesis. The results are also compared against simulated hadron showers modeled using GEANT4-based detector simulation framework.

Acknowledgements

Completing my integrated Ph.D. program at IISER Pune has been quite a journey with its ups and downs, as happens with every aspect of anyone's life. When I came to IISER as a freshly graduated engineer with dreams of becoming a researcher, I had no idea what research entails. As I am writing this thesis, I can confidently say that I have come a long way from where I started.

Dr. Seema Sharma has played a substantial role as my advisor in this journey. I could possibly not give her enough credit as she has practically taught me how to conduct scientific research. She instilled the sense of temperament that helped me develop as an individual both academically and personally. The projects that we undertook were challenging and fun simultaneously, be it Particle Flow Cluster Calibration during Run-2, the HGCAL beam test experiment and analyses, or setting up HEP lab at IISER Pune. I am really thankful to her for having faith in me throughout these years. Her dedication to research and her incredible intuition is something I hope to achieve someday.

As my Research Advisory Committee (RAC) members, Prof. Bhas Bapat and Dr. Satyaki Bhattacharya kept me in the right direction with their valuable feedback and suggestions. Their profound insights into physics, which were evident through their pin-point questions, have added more depths and dimensions to my overall academic growth and research experience.

I was fortunate enough to visit CERN quite a few times during Run-2 and for the HGCAL beam test experiment. During my visit, I had the pleasure of working with brilliant physicists, including Kenichi Hatakeyama, David Barney, Andre David, Katja Kruger, Clemens Lange, Arnaud Steen, Artur Lobanov, Thorben Quast. The feeling of seeing the detector in action and the beautiful event displays in real-time can not be explained in words. It was a truly fantastic experience to participate in the beam test experiment.

Our research group at IISER has had some of the most amazing people that I have known. I thank Vinay Hegde and Aditee Prabhakar for their wise advice. Bhumika Kansal and Alpana Sirohi, thank you for all the fun discussions on research and current affairs. Though I felt like an old man, but working with undergrads, Niramay, Rahul, and Nitish was fun. Thanks to Irene Dutta and Prachi Atmasiddha, who showed me the monstrosities of not caring about cyber security.

I would also like to thank our admin staff, Prabhakar Anagare and Dhanashree Sheth, who lowered the burden of bureaucracies during my entire term at IISER Pune. The IT experts at IISER, Nisha Kurkure, Goldi Mishra, and late Sachin Behra, has always been prompt to fix computing and networking related issues. I also thank Mr. Nilesh and Mr. Prashant, who have always helped me with the lab issues and instrument procurement.

Finally, I am eternally grateful to my family, who has always supported and helped me pursue my dreams. The unconditional and boundless love and care I have received from my siblings, Anurag, Saumya, and Rachna is more than I could ever ask for. My mom is the strongest woman I have known. Papa, I wish you were here. I love you all.

Contents

Chapter 1

Introduction

For eons of time, many philosophers have wondered about the fundamental constituents of the matter. Philosophers in the west such as Greek philosopher *Leucippus of Miletus* [1] and his student *Democritus of Abdera* [2] in 5^{th} century BCE, and in the east such as Indian philosopher *Kaṇāda* [3] in 4^{th} century BCE, suggested that the matter is made of indivisible basic units. The Greeks called it *atomos*, meaning indivisible, which became the root of the word *atom*.

The idea of atoms in modern physics was brought upon by English chemist John Dalton in the early 19^{th} century [4]. He performed several experiments on gases and proposed that the elements are made of atoms that are identical in weight and property. Later in the 19^{th} century, J. J. Thompson discovered electrons [5], and in the earlier 20^{th} century famous experiment performed by E. Rutherford discovered the atomic nucleus and later protons [6]. These discoveries laid the foundation of the atomic structure of matter. In parallel, theoretical work by Louise de Broglie, Schrödinger, Heisenberg, and others, established the quantum mechanical framework that describes the dynamics of particles at the atomic level. Louise de Broglie's theory of wave-particle duality of matter showed that with increasing energy, the wavelength of the *particle wave* decreases [7]. By the mid 20^{th} century, physicists started using particle accelerators delivering high energy particles as powerful microscopes to probe internal structure of the atomic nuclei. The experiments with more powerful accelerators demonstrated that nucleons themselves, i.e., protons and neutrons, are not fundamental particles. Instead, they are composed of quarks. Equipped with higher accelerating power and colliding beams, and increasingly advanced particle detection techniques, the later experiments discovered various composite and fundamental particles [8]. Thus, the branch of physics, called *particle physics*, got well established.

The field of particle physics is dedicated to the study of elementary particles. The combined effort of theoretical and experimental particle physicists of over almost 50 years has created a mathematical model of elementary particles observed in nature and their interactions, called *standard model* (SM) [9, 10]. The SM incorporates the theory of three out of four fundamental forces namely, strong force, electromagnetic force, and weak force. The basic building blocks of matter are classified as quarks and leptons, and have half integer spin (1/2), called fermions. The interactions among these particles are understood to be mediated via particles with spin 1, called bosons. The strong force is mediated by gluons (g), the weak force is mediated by $W^{\pm}$ and Z^0 bosons, and the electromagnetic force is mediated by photons (γ). The predictions made by the SM have been tested in a variety of collider and non-collider experiments, and have been found to be in agreement with high precision. For example, the measured anomalous magnetic moment of electron agrees with the SM prediction with more than per billion precision [11]. Numerous other measurements made at various experiments such as UA1, UA2, LEP, CDF, Tevatron, ZEUS have shown that the SM describes the nature remarkably well at the fundamental level [12, 13, 14, 15, 16, 17].

Although the SM has been found to describe various processes in particle physics very well, it fails to answer numerous questions. The SM does not include the description of gravity, the oldest known force of nature. It does not have any answer to the nature of the dark matter the existence of which has been proven beyond any doubt through astronomical observations. The reason behind matter dominating over anti-matter in the observable universe is still unknown. These are just a few of many questions which the SM can not answer, indicating that it is an incomplete theory. Many novel theories extend beyond the SM (BSM) physics and try to address these questions. The aim of collider-based experiments, such as experiments at the Large Hadron Collider (LHC), is to test the limits of SM and to search for signatures of new physics [18]. The LHC at CERN is the most powerful particle accelerator and collider in the world at present. It collides beams of protons at the center of mass energy ($\sqrt{s}$) of 13 TeV at four points along its 27 km long ring. At each collision point, a detector is placed to detect the particles emanating from high energy collisions. The Compact Muon Solenoid (CMS) [19] is one such detector. The CMS is a general-purpose detector designed to measure particles of momenta ranging from a few hundreds of MeV/c to a few TeV/c.

The LHC started taking physics data in 2010 and will continue for next five years colliding protons at $\sqrt{s}$ = 13 TeV. The CMS and ATLAS (another general purpose detector at LHC) discovered the Higgs boson in 2012, the last missing particle of the SM. However, physics data collected till 2018 has shown no sign of new physics. More data is required to make precision measurements of various SM processes that have smaller cross-sections such as Higgs boson self-coupling, and also to discover particles predicted by many BSM theories, e.g., supersymmetry (SUSY) [20]. The High-Luminosity operation of LHC, called HL-LHC, is planned towards this very goal [21]. The HL-LHC phase of the operation is expected to start from 2027 and will operate for at least ten years. It will deliver ten times more physics data as compared to the LHC run, bringing many physics opportunities. However, this poses various challenges for detectors. With increased collision rate of HL-LHC, the detectors will suffer from immense radiation damage, and will face a large number of particles emerging from additional proton-proton collisions at the interaction point called *pileup*. The current CMS detector is not designed to operate in such harsh conditions and will not be able to deliver adequate physics performance. For example, the current endcap electromagnetic (ECAL) and hadronic (HCAL) calorimeters will suffer irreversible radiation damage and become insensitive to traversing particles. The CMS Collaboration has proposed a comprehensive upgrade programme of its sub-detectors in the light of HL-LHC operation. As part of the upgrade programme, the inner tracker will be made more granular and its acceptance will be extended, the endcap muon systems will deploy gas electron multiplier (GEM) technology, HCAL photodetectors will be replaced by silicon photomultipliers, and trigger systems will be upgraded to handle high event rates [22]. The endcap ECAL and HCAL calorimeters will be replaced with a more radiation tolerant and highly granular calorimeter, called the HGCAL [23]. The HGCAL is a silicon and scintillator technology-based sampling calorimeter.

This thesis describes the ongoing research and development of the HGCAL detector and its components. A prototype of silicon-based electromagnetic and hadronic section of the HGCAL was built and tested in the experiments with single particle beams at CERN Super Proton Synchrotron (SPS) in October 2018 [24]. The beam test experiment was performed along with scintillator-on-SiPM based CALICE Analogue Hadron Calorimeter (AHCAL) detector prototype [25]. The AHCAL helped extending the longitudinal depth of the prototype calorimeter making it closer to the realistic depth of proposed HGCAL. The prototypes

were exposed to beams of positrons and charged pions with energies ranging from 20 to 300 GeV to evaluate electromagnetic and hadronic shower performance, and 200 GeV muons for inter-cell calibration. The topics covered in this thesis range from testing the most basic unit of the HGCAL, the silicon sensor modules, in the laboratory-based test benches to analyzing the performance of a large scale endcap calorimeter detector prototype in the beam test experiment.

The thesis begins with a brief introduction of the standard model of particle physics in chapter 2. It describes the interaction of the particles in the matter and various detection techniques that are employed in nuclear and particle physics experiments. This chapter sets the stage for calorimetric concepts and the GEANT4 simulation framework that are critical in performance studies of modern detectors including HGCAL.

Chapter 3 describes the LHC, the CMS experiment, and gives a brief summary of physics results. Chapter 4 presents the LHC road-map, technological challenges for detectors, and a brief overview of the CMS detector upgrades with a specific focus on the HGCAL.

Chapter 5 describes the system tests of silicon module, the smallest independent detector unit which is replicated to make the HGCAL active layers. The system test includes silicon module's electrical characterization, testing its readout electronics, the data acquisition system, pedestal noise measurement, and the measurement of cosmic muon signals using scintillator based cosmic test-stand.

Chapter 6 is dedicated to the description of HGCAL and AHCAL detector prototype used in October 2018 beam test experiment. This chapter describes signal reconstruction in data and simulation, detector alignments, cell-to-cell response equalization, and signal-to-noise ratio studies of silicon sensors employed in the detector.

Chapter 7 presents a comprehensive performance study of the combined HGCAL + AHCAL detector prototype to hadronic showers using beams of charged pion (π^-) of momenta ranging from 20 to 300 GeV/c. The analysis makes use of the fine longitudinal and transverse segmentation of the HGCAL to develop an algorithm that identifies the shower start location with exquisite efficiency. The calorimetric performance is studied in terms of hadronic shower energy reconstruction and its evolution in longitudinal and transverse direction. The results from October 2018 beam test experiment is compared against GEANT4 based simulation in order to validate hadronic shower simulation framework in the HGCAL detector.

This thesis presents first performance study of a large scale CMS HGCAL prototype to hadronic showers initiated by charged pions in the momenta range 20 to 300 GeV/c using the data collected in the beam test experiments at CERN in October 2018.

Chapter 2

Overview of Standard Model and experimental techniques

The standard model (SM) is the most successful theory of particle physics so far that describes nature at the most fundamental level. This chapter provides a brief overview of the SM, the primary processes through which the particles interact with the matter, various detection techniques focusing on the calorimeters to measure high energy particles. Using the knowledge of these interaction mechanism, Monte Carlo simulation frameworks have been developed that are indispensable tools in high energy physics research. Key aspects of the GEANT4 detector simulation toolkit [26] are discussed in this chapter. It is a widely used application for detailed simulation of the detector response and electronics effects and is also used in the work presented here.

2.1 Standard Model of particle physics

The SM is a quantum field theory (QFT) [27] that describes the dynamics of three out of four fundamental forces of nature, i.e. strong, weak, and electromagnetic force; and classifies all known elementary particles. In QFT, quantum mechanics is combined with the classical fields and special relativity, giving rise to quantum fields. The quantum fields are considered as the fundamental entity of nature, and the excited quantum of the field is interpreted as the *particle*. The particles carry an intrinsic angular momentum, called *spin*. Particles with half integer spin are called *fermions*, and particles with integer spin are called *bosons*.

The fermions are the basic building blocks of matter. The SM classifies the fermions as quarks and leptons in three generations with increasing mass order as listed in Figure 2.1. The quarks come in six *flavours*: up (u), down (d), charm (c), strange (s) , top (t), and bottom (b). The quarks carry either $2/3$ or $-1/3$ electric charge (e), and have a colour charge quantum number, namely Red (R), Green (G), or Blue (B). All quarks also carry another inherent quantum number, called *baryon number*. The baryon number for each quark is $1/3$. All quantum numbers are conserved quantities in the interactions, except flavour quantum number which is violated in the weak interactions.

Similar to quarks, there are six leptons that are classified into three generations with increasing mass order. Three out of six leptons carry electric charge of -1, namely electron (e) , muon (μ), and tau (τ). The other three leptons are electrically neutral, namely electron-neutrino (ν_e), muon-neutrino (ν_μ), and tau-neutrino (ν_τ). Each generation of lepton is associated with a *lepton-flavour* quantum number, i.e. L_e, L_μ and L_τ. For example, $L_e = 1$ for electron and ν_e; $L_\mu = 1$ for muon and ν_μ; $L_\tau = 1$ for tau and ν_τ. The lepton-flavour quantum number has been found to be conserved experimentally. Each fermion in the SM has a

corresponding *anti-particle*. The anti-particle is same as the particle in every aspects except for its charge and associated quantum numbers, which are opposite to that of particle.

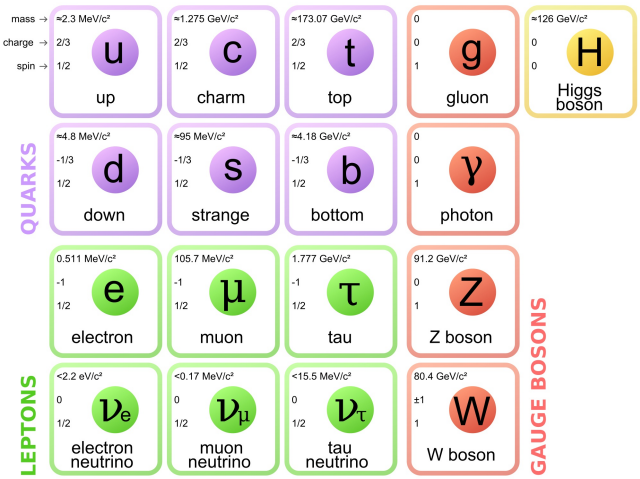

FIGURE 2.1: The table of elementary particles in the SM.

The fundamental interactions between fermions are mediated by spin-1 vector bosons, namely gluons (g) for strong interaction, $W^{\pm}$ and Z^0 for weak interaction, and photons (γ) for electromagnetic interaction. The SM is formulated with three internal gauge symmetries, $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$, where C, L and Y are three quantum numbers, viz., colour charge, weak isospin, and hypercharge, respectively. The gauge symmetries lead to conservation of these quantum numbers as a consequence of Noether's theorem [28].

The electromagnetic interaction corresponds to $U(1)$ abelian gauge symmetry group, and it is described by quantum electrodynamics (QED) [29]. The theory of QED is the result of imposing $U(1)$ gauge symmetry on Dirac's equation. Dirac's equation can be encapsulated in the form of Lagrangian density ($\mathcal{L}_{\text{Dirac}}$) as shown below:

$$\mathcal{L}_{\text{Dirac}} = \bar{\psi}(i\gamma^{\mu}\partial_{\mu} - m)\psi \tag{2.1}$$

where ψ is fermion spinor, γ^{μ} are Dirac matrices, $\bar{\psi}$ is $\psi^{\dagger}\gamma 0$ where $\psi^{\dagger}$ is the hermitian conjugate of ψ. Here, natural units, i.e. $c = \hbar = 1$, is used and Einstein's notation has been followed, i.e. summation is implied on the repeated indices. The greek letters are reserved for space-time indices. In the following text the Lagrangian density and Lagrangian is used interchangeably.

For equation 2.1 to remain invariant under $U(1)$ gauge transformation, the partial derivative (∂_{μ}) is replaced with covariant derivative (D_{μ}) which is defined as:

$$D_{\mu} = \partial_{\mu} - ig_{em}A_{\mu} \tag{2.2}$$

where g_{em} is the electromagnetic coupling constant and A_{μ} is a vector gauge field that transforms simultaneously under $U(1)$ as follows:

$$A_{\mu} \rightarrow A'_{\mu} = A_{\mu} + \frac{1}{g_{em}}\partial_{\mu}\theta \tag{2.3}$$

where θ is the generator of $U(1)$ symmetry group. From equation 2.3 and 2.2, one can verify that following $\mathcal{L}_{\text{fermion}}$ is invariant under $U(1)$ gauge transformation:

$$\begin{aligned}\mathcal{L}_{\text{fermion}} &= \bar{\psi}(i\gamma^\mu D_\mu - m)\psi \\ &= \bar{\psi}(i\gamma^\mu \partial_\mu - m)\psi + g_{em}\bar{\psi}\gamma^\mu A_\mu \psi\end{aligned} \tag{2.4}$$

Thus, while making $\mathcal{L}_{\text{Dirac}}$ gauge invariant, a vector field A_μ is introduced that couples to fermionic field ψ. The $\mathcal{L}_{\text{fermion}}$, however, does not contain the dynamics of the vector field itself. Therefore, a kinetic term $-\frac{1}{4}F_{\mu\nu}F^{\mu\nu}$, which is already invariant under $U(1)$, is added to $\mathcal{L}_{\text{fermion}}$. For electromagnetic interaction, $F_{\mu\nu}$ is the electromagnetic potential field tensor defined as:

$$F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu \tag{2.5}$$

Thus, the following Lagrangian ($\mathcal{L}_{\text{QED}}$) is obtained that describes the electromagnetic interaction.

$$\mathcal{L}_{\text{QED}} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \bar{\psi}(i\gamma^\mu \partial_\mu - m)\psi + g_{em}\bar{\psi}\gamma^\mu A_\mu \psi \tag{2.6}$$

Note that there is no mass-term for A_μ in $\mathcal{L}_{\text{QED}}$. A mass term of the form $-m^2 A_\mu A^\mu$ can be added to equation 2.6, however, the resulting Lagrangian does not remain invariant under $U(1)$ gauge transformation. Therefore, A_μ must be a massless field and by extension its associated vector boson, i.e. photon, must be massless. Since the photons are massless therefore their field of influence extends to infinity. However, the strength of electromagnetic coupling (g_{em}) diminishes rapidly as $1/r^2$. The photon couples to particles that carry a non-zero electric charge which is a conserved quantity.

The strong interaction is described by quantum chromodynamics (QCD) gauge theory. The QCD describes the dynamics of the strong force at intra-nuclear distances ($\sim$ few fm). The theory of QCD belongs to $SU(3)$ non-abelian gauge symmetry group. The corresponding field tensor for QCD, $G_a^{\mu\nu}$ is defined as follows:

$$G_a^{\mu\nu} = \partial^\mu g_a^\nu - \partial^\nu g_a^\mu + g_s f_{abc} g_b^\mu g_c^\nu \tag{2.7}$$

where g_a is the gluon gauge field with index $a = 1, 2, ..., 8$; g_s is the coupling constant for strong force, and f_{abc} is $SU(3)$ structure constant. Since there are eight independent generators of $SU(3)$ group, therefore there are eight mediator gluons. The strong force couples to particles that carry colour charge, i.e., only quarks and gluons. The QCD Lagrangian is constructed as follows:

$$\begin{aligned}\mathcal{L}_{\text{QCD}} &= -\frac{1}{4}G_{a\mu\nu}G_a^{\mu\nu} + \sum_{k=1}^{6}\bar{q_k}(i\gamma^\mu D_\mu - m_k)q_k \\ &= -\frac{1}{4}G_{a\mu\nu}G_a^{\mu\nu} + \sum_{k=1}^{6}\bar{q_k}(i\gamma^\mu \partial_\mu - m_k)q_k + \sum_{k=1}^{6} g_s \frac{\lambda_a}{2}\bar{q_k}\gamma^\mu g_\mu^a q_k\end{aligned} \tag{2.8}$$

where q_k are six quarks and λ_a are generators for $SU(3)$ group. Expanding the first term of $\mathcal{L}_{\text{QCD}}$ in equation 2.8 using the definition of $G_a^{\mu\nu}$ from equation 2.7, one obtains three as well as four point interaction vertices for gluons. It signifies that gluon can interact with itself via strong force. These interaction vertices are possible in QCD because gluons themselves carry colour charge unlike QED where photons are electrically neutral and hence can

not have self-interaction vertices.

One of the features of strong interaction is that the force strength remains constant between two quarks with increasing distance. Therefore, the potential energy of the quark system increases if the distance between the quarks is increased. At a point of about $\sim$ 1 fm, the energy between the quarks is sufficient to produce quark-antiquark pairs. This leads to a phenomenon called *color confinement* in which it is energetically favourable to produce new quark-antiquark pairs at larger distances that combine with the initial quarks and form colour-neutral composite bound states, called *hadrons*. Because of color confinement, quarks are never observed in isolation. Based on the number of quarks in the hadron, it is called either *baryon* (composite state of three quarks) or *meson* (composite state of two quarks). For example, protons and neutrons are baryons consisting of uud and udd bound states, respectively. Similarly, pions (π^+) and kaons (K^+) are mesons consisting of $u\bar{d}$ and $u\bar{s}$ bound states, respectively. The QCD also exhibits a phenomenon called *asymptotic freedom*, that is, at very small distances the strong force strength decreases rapidly, and the quarks become almost free, e.g. within hadrons.

The electromagnetic and weak interaction is unified by imposing $SU(2)_L \otimes U(1)_Y$ gauge symmetry. The corresponding gauge fields are W^i_μ and B_μ for $SU(2)_L$ and $U(1)_Y$ group, respectively. The $SU(2)_L$ field interacts with left-handed (right-handed) chiral eigenstates of fermions (anti-fermions) that carry weak-isospin charge quantum number, T_3. In the SM, the fermions can be expressed as left-handed weak-isospin doublet ($T_3 = \pm 1/2$) and right-handed weak-isospin singlet ($T_3 = 0$). For example, left-handed up quark (u_L) and electron-neutrino (ν_{eL}) have weak-isospin T_3 = +1/2, and left-handed down quark (d_L) and electron (e_L) have weak-isospin T_3 = -1/2. All right-handed quarks and charged leptons have weak-isospin $T_3 = 0$. In this prescription, right-handed neutrinos (i.e. ν_{eR}, $\nu_{\mu R}$, and $\nu_{\tau R}$) can not be accommodated. For example, the first generation leptons can be written as:

$$L_e = \begin{pmatrix} \nu_e \\ e_L \end{pmatrix} \tag{2.9}$$

$$R_e = (e_R) \tag{2.10}$$

where ν_L, e_L and e_R are defined as follows:

$$e_L = \frac{1-\gamma^5}{2} e \tag{2.11}$$

$$\nu_L = \frac{1-\gamma^5}{2} \nu \tag{2.12}$$

$$e_R = \frac{1+\gamma^5}{2} e \tag{2.13}$$

where $\gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3$. Similarly, first generation of quarks can also be expressed as left- or right-handed chiral eigenstates: $\binom{u_L}{d_L}$, (u_R), (d_R). This definition can be extended to second and third generation of leptons and quarks. Thus, electroweak Lagrangian ($\mathcal{L}_{\rm EW}$) is constructed as follows:

$$\mathcal{L}_{\rm EW} = -\frac{1}{4} W_{a\mu\nu} W_a^{\mu\nu} - \frac{1}{4} B_{\mu\nu} B^{\mu\nu} + \sum_{f=\psi_L,\psi_R} \bar{f} i\gamma^\mu D_\mu f \tag{2.14}$$

where $f = \psi_L$ (ψ_R) is left-handed (right-handed) fermion spinor as defined earlier. The field tensors $W_a^{\mu\nu}$ and $B^{\mu\nu}$ are defined as:

$$\begin{aligned} W_a^{\mu\nu} &= \partial^\mu W_a^\nu - \partial^\nu W_a^\mu + g_2 f_{abc} W_b^\mu W_c^\nu \\ B^{\mu\nu} &= \partial^\mu B^\nu - \partial^\nu B^\mu \end{aligned} \tag{2.15}$$

where W_a^μ and B_μ are gauge fields, f_{abc} is the structure constant for $SU(2)$ group and index $a = 0, 1, 2$. The covariant derivative D_μ in equation 2.14 is:

$$D_\mu = \partial_\mu - ig_1 \frac{Y}{2} B_\mu - ig_2 \frac{\tau^a}{2} W_\mu^a \tag{2.16}$$

where Y and τ^i are the generators for $U(1)$ and $SU(2)$ symmetry groups, respectively. g_1, g_2 are the coupling strength between fermions and gauge fields. The combined $SU(2)_L \otimes U(1)_Y$ symmetry imposes the conservation of so-called *weak hypercharge*, which is defined as follows:

$$Y_W = 2(Q - T_3) \tag{2.17}$$

where Y_W is the weak hypercharge, Q is the electric charge and T_3 is the weak-isospin of the particle. As mentioned earlier in the example of QED, the gauge bosons must be massless as including their mass terms break gauge invariance. Furthermore, due to $SU(2)_L$ symmetry, the addition of mass term to fermions are also forbidden, the reason of which is described as follows: adding a mass term $\mathcal{L}_{\text{mass}}$ implies:

$$\begin{aligned} \mathcal{L}_{\text{mass}} &= -m\bar{\psi}\psi \\ &= -m\bar{\psi_L}\psi_R + \text{h. c.} \end{aligned} \tag{2.18}$$

where h.c. is the hermitian conjugate of the first term. If ψ_L and ψ_R each have a distinct antiparticle, then it requires them to obey Dirac equation 2.1 such that it remains invariant under $SU(2)_L$ transformation, i.e.

$$i\gamma^\mu \partial_\mu \bar{\psi_L} = m\psi_R \tag{2.19}$$

However, left-handed and right-handed fermions carry different $SU(2)_L$ charge, i.e. $\bar{e_L}$ ($\bar{\psi_L}$) has T_3 = -1/3 and e_R (ψ_R) has T_3 = 0 as pointed out earlier. Therefore, equation 2.19 does not hold, and thus, the mass term is not allowed for fermions under $SU(2)_L$ gauge symmetry. If the SM abandons weak interaction such that it contains only $SU(3) \otimes U(1)$ then fermions are allowed to have masses, but our universe has weak interaction as a fundamental force of nature, therefore it can not be overlooked. However, as established by experiment measurements that the mediator bosons for weak interaction as well as quarks and charged leptons have non-zero mass, therefore a prescription is required for these particles to be massive. This is achieved by invoking *Spontaneous Symmetry Breaking* as explained in the next section.

2.1.1 Spontaneous symmetry breaking

The problem mentioned above is solved by introducing a scalar field in the SM [30]. The scalar field, called *Brout–Englert–Higgs* field or simply *Higgs* field, breaks the electroweak symmetry spontaneously and gives masses to bosons and fermions by interacting with them. This mechanism is called Brout–Englert–Higgs (BEH) mechanism. The process is described in the following. Introduce a complex scalar Higgs field (Φ) which is a doublet of $SU(2)$, i.e.

$$\Phi = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix} \tag{2.20}$$

where ϕ_1, ϕ_2, ϕ_3, and ϕ_4 are real scalar fields. The Lagrangian ($\mathcal{L}_{\text{Higgs}}$) of Higgs field is written as :

$$\begin{aligned} \mathcal{L}_{\text{Higgs}} &= T(\Phi) - V(\Phi) \\ &= \left(D_\mu \phi\right)^\dagger \left(D^\mu \phi\right) - \mu^2 \phi^\dagger \phi - \lambda \left(\phi^\dagger \phi\right)^2 \end{aligned} \tag{2.21}$$

where D_μ is the covariant derivative of $SU(2)_L \otimes U(1)_Y$ (equation 2.16). With $\mu^2 < 0$, the Higgs potential $V(\Phi)$ has minima for:

$$\phi^\dagger \phi = \frac{-\mu^2}{2\lambda} \equiv \frac{v^2}{2} \tag{2.22}$$

At this minima, the real component of $\phi^\dagger \phi$ satisfies following relation:

$$\Re(\phi^\dagger \phi) = \frac{\phi_1 + \phi_2 + \phi_3 + \phi_4}{2} \tag{2.23}$$

Above equation 2.23 implies that there can be infinite solutions to equation 2.22. One can choose any direction in the $SU(2)$ space (ϕ_i space) to define vacuum configuration, thus breaking the $SU(2)$ symmetry. We choose the vacuum (ϕ_0) to be:

$$\phi_0 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix} \tag{2.24}$$

where $v = \sqrt{\frac{-\mu^2}{\lambda}}$ from equation 2.22. Note that for above vacuum configuration, we have chosen $\phi_3 = v$, while $\phi_1 = \phi_2 = \phi_4 = 0$. We could have chosen any other vacuum configuration to break the symmetry. This occurs without any external agency, therefore it is termed as *spontaneous symmetry breaking*. The vacuum excitation can be analyzed by introducing a quantum fluctuation $H(x)$ in the ϕ_0, and expanding it around the minima:

$$\phi_0 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H(x) \end{pmatrix} \tag{2.25}$$

Substituting ϕ_0 from equation 2.25 into equation 2.21, the Higgs potential can now be written as:

$$V(\Phi) = -\frac{\mu^4}{4\lambda} - \mu^2 H^2 + \lambda v H^3 + \frac{\lambda}{4} H^4 \tag{2.26}$$

From above equation, one can interpret that $H(x)$, which is a real scalar field, has corresponding scalar boson (called Higgs boson, H) whose mass is $m_H = \sqrt{-2\mu^2} = \sqrt{2\lambda} v$. Moreover, the Higgs boson can interact with itself because of cubic and quartic term in $V(\Phi)$. Since λ is the free parameter of BEH mechanism, therefore precise measurement of Higgs self-coupling helps correctly determine the shape of Higgs potential.

Similarly, kinetic term of $\mathcal{L}_{\text{Higgs}}$ can be expanded by substituting ϕ_0 in the first term of equation 2.21 and using the definition of D_μ from equation 2.16, i.e.

$$\begin{aligned} T(\Phi) &= \left(D_\mu \phi\right)^\dagger \left(D^\mu \phi\right) \\ &= \frac{1}{2}(\partial_\mu H)^2 + \frac{g_1^2}{8}(W_\mu^1 - iW_\mu^2)(W^{1\mu} - iW^{2\mu})(\nu + H)^2 \\ &+ \frac{1}{8}(g_1 W_\mu^3 - g_2 B_\mu)(W^{3\mu} - g_2 B^\mu)(\nu + H)^2 \end{aligned} \tag{2.27}$$

With the linear combination of W_μ^a and B_μ, one can define following gauge fields:

$$\begin{aligned} W^+ &= \frac{-W^1 + iW^2}{\sqrt{2}} \\ W^- &= \frac{-W^1 - iW^2}{\sqrt{2}} \\ W^0 &= W^3 \end{aligned} \tag{2.28}$$

and,

$$\begin{aligned} A_\mu &= \frac{g_2 B_\mu + g_1 W_\mu^0}{\sqrt{g_1^2 + g_2^2}} \\ Z_\mu &= \frac{g_1 B_\mu + g_2 W_\mu^0}{\sqrt{g_1^2 + g_2^2}} \end{aligned} \tag{2.29}$$

where coupling constants are related as:

$$\begin{aligned} e &= \frac{g_1 g_2}{\sqrt{g_1^2 + g_2^2}} \\ \sin(\theta_W) &= \frac{g_1}{g_1^2 + g_2^2} \\ \cos(\theta_W) &= \frac{g_2}{g_1^2 + g_2^2} \end{aligned} \tag{2.30}$$

where e is the electric charge and θ_W is called *weak mixing angle*. Substituting these values into equation 2.27, we obtain:

$$\begin{aligned} T(\Phi) =& \frac{1}{2}(\partial_\mu H)^2 + m_W^2 W^{+\mu} W_\mu^- + \frac{2m_W^2}{\nu^2} W^{+\mu} W_\mu^- H \\ &+ \frac{1}{2} m_Z^2 Z^\mu Z_\mu + \frac{m_Z^2}{\nu} Z^\mu Z_\mu H + \frac{m_Z^2}{2\nu^2} Z^\mu Z_\mu H^2 \end{aligned} \tag{2.31}$$

Thus, the gauge bosons for weak interaction, i.e. charged $W^\pm$ and neutral Z^0, acquire mass m_W and m_Z, respectively which is given by:

$$m_W^2 = \frac{g_2^2 \nu^2}{4}, \text{ and } m_Z^2 = \frac{m_W^2 \nu^2}{\cos^2(\theta_W)} \tag{2.32}$$

The weak mixing angle θ_W is a free parameter, which is determined by precisely measuring the masses of W and Z boson. The gauge boson corresponding to field A_μ and

corresponding boson i.e. photon, remains massless.

The masses for fermions are also generated using Higgs field. The interaction between scalar field Φ and fermions is added in the form of so-called *Yukawa coupling* ($\mathcal{L}_{\text{Yukawa}}$). For example, for electrons the form of Yukawa coupling is as follows:

$$\mathcal{L}^{e}_{Yukawa} = g_e \bar{L}_e \Phi e_R + \text{h.c.} \tag{2.33}$$

where L_e is $SU(2)$ doublet defined in equation 2.9, and h.c. is the hermitian conjugate of first term. $\mathcal{L}^{e}_{Yukawa}$ is invariant under $SU(2)_L \otimes U(1)_Y$ transformation. Expanding Φ field according to equation 2.25, one obtains:

$$\mathcal{L}^{e}_{Yukawa} = \frac{g_e v}{\sqrt{2}}(\bar{e}_L e_R + \bar{e}_R e_L) + \frac{g_e}{\sqrt{2}}(\bar{e}_L e_R + \bar{e}_R e_L)H \tag{2.34}$$

Thus, the mass of electron can be interpreted from the first term as $m_e = \frac{g_e v}{\sqrt{2}}$. The second term corresponds to the interaction between electron and the Higgs field. This prescription of Yukawa coupling can be extended to quarks and leptons of all generations. Thus, we can write a more generalized form of Yukawa coupling as follows:

$$\mathcal{L}_{Yukawa} = g \bar{\psi}_L \Phi \psi_R + \text{h.c.} \tag{2.35}$$

Equipped with Lagrangian of all the fermions and bosons as well as their interactions, the SM can be summarized as:

$$\mathcal{L}_{SM} = \mathcal{L}_{gauge} + \mathcal{L}_{fermion} + \mathcal{L}_{Higgs} + \mathcal{L}_{Yukawa} \tag{2.36}$$

where $\mathcal{L}_{\text{gauge}}$ is the boson field dynamics for $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$ gauge symmetry, $\mathcal{L}_{\text{fermion}}$ is the fermion field dynamics corresponding to equation 2.4 with appropriate covariant derivatives, $\mathcal{L}_{\text{Higgs}}$ corresponds to equation 2.21, and $\mathcal{L}_{\text{Yukawa}}$ corresponds to equation 2.35. A schematic presented in Figure 2.2 shows all the possible interactions among the elementary particles in the SM. To summarize:

- All the charged particles, viz. quarks, charged leptons, and $W^{\pm}$ boson, can interact via electromagnetic force.
- Only quarks and gluons can interact via strong force. The gluons can interact with themselves and have self-coupling.
- All the fermions as well as $W^{\pm}$, Z^0, and Higgs boson can interact via the weak force.
- Neutrinos can only interact via weak force.
- Higgs field couples to all the massive particles as well as to itself.

2.1.2 Limitations of the SM

With the discovery of the Higgs boson in July 2012, the last missing piece of the SM was completed. Experimental measurements made so far have shown very good agreement with the SM predictions. It is, however, an incomplete theory as it fails to provide answers to many questions. A few of them are listed below:

- The SM in its description does not include the most well-known force of nature, i.e. gravity. A self-consistent and experimentally verifiable quantum theory of gravity is yet to be found.

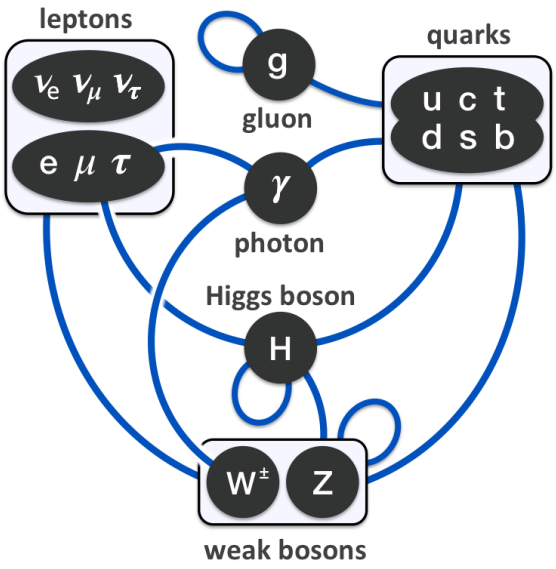

FIGURE 2.2: The schematic shows possible particle interactions in the SM. The blue line corresponds between two blocks corresponds to coupling between particles and the loop corresponds to self-coupling.

- It does not explain why the observable universe is dominated by matter over antimatter, known as *baryon asymmetry* problem.
- The observation of neutrino-oscillation (neutrinos changing from one flavour to other) requires neutrinos to be massive. In contrast, there is no prescription through which neutrinos can acquire mass in the SM.
- The SM calculations suggest that higher order corrections on the Higgs boson mass should bring its value near to Plank scale ($\sim 10^{19}$ GeV) [31]. However, the experiments have measured the mass of Higgs boson to be 125 GeV. The SM has no explanation other than postulating delicate cancellations of astronomical order for a 125 GeV Higgs boson to exist. This is known as *hierarchy problem*.
- The SM does not describe the nature of dark matter (DM), the existence of which has been proven without a doubt through astronomical and cosmological observations [32].

There are various *Beyond Standard Model (BSM)* theories such as Supersymmetry [20], extra dimensions [33], sterile neutrinos [34], etc. that aim to answer one or more of the above mentioned questions. Most of these theories predict new fundamental particles which could be produced in high energy collisions or affect predictions of the SM. Numerous experiments are being carried out to test the validity of such BSM physics models and find any deviation from SM prediction in precision measurements. The experiments at the Large Hadron Collider at CERN are dedicated to this very purpose.

2.2 Particle interactions with matter

The collider-based experiments aim to probe the nature by studying the particles produced during high-energy collisions. Heavy and unstable particles that are produced in the collision decay into lighter particles that traverse through the detector that are built around the

collision point. Based on their interaction with the detector material, these particles leave signatures in different parts of the detector system. These signatures are used to reconstruct as well as interpret each collision event that helps in new physics searches. Therefore, it is crucial to understand the interactions and processes through which the particles lose their energy in matter. A particle can interact with a given material via electromagnetic, strong, or weak interaction based on its type. For example, energetic hadrons interact with the nuclei of the material via strong interaction and develop a cascade of shower. Charged particles interact with the atomic electrons losing their energy via electromagnetic interaction as well as producing a measurable signal. Hence, various detection techniques are employed depending on the nature of particles and their interactions with materials in order to optimize the detection systems. In the following, particle interactions with matter and detection techniques has been discussed.

2.2.1 Energy loss by charged particles

Ionization energy loss

A charged particle traversing through the medium can ionize or excite the atoms. It results in energy loss of the incident-charged particle. The energy loss per unit traversed length via ionization process can be given by *"Bethe-Bloch formula"* [35] shown in the following expression:

$$-\frac{dE}{dx} = \rho N_A \frac{Z}{A} \frac{4\pi\alpha^2\hbar^2}{m_e} \frac{q^2}{\beta^2} \left[\ln\left(\frac{2m_e c^2 \beta^2 \gamma^2}{I} \right) - \beta^2 - \frac{\delta}{2} \right] \tag{2.37}$$

where, x is the traversed length in the medium, ρ is the density of the medium, N_A is the Avogadro's number, Z (A) is atomic number (weight) of the material, m_e is the mass of electron, c is the speed of light in vacuum, q is the electric charge of the incident particle, β is v/c, γ is the Lorentz factor (= $(1-\beta^2)^{-1/2}$), I ($\sim 10\times Z$ eV) is the ionization potential of the medium and δ is the correction for dielectric screening effect at highly relativistic speed of particle.

From the formula, one can infer following interesting features:

- the ionization energy loss is higher for slow particles ($1/\beta^2$ term) because the particle spends more time in the influence of the Coulomb field of the atom.
- energy loss is directly proportional to the density (ρ) of the medium.
- higher the electric charge of the incident particle (q^2 term), more will be the energy loss.
- at very high energies, where β becomes almost 1, the logarithmic term in the R.H.S. comes into effect and increases the ionization loss in the medium, called *"relativistic rise"* of energy loss.

The energy loss is usually expressed in terms of MeV g^{-1} cm^2 by dividing both sides of equation 2.37 by ρ. The transformed quantity is treated as dE/dX. The behaviour of dE/dX as a function of particle momentum is shown in Figure 2.3 for different material.

The figure shows that the energy loss increases rapidly for low momentum and decreases as the particle's momentum increases. The energy loss reaches a minimum around $\beta\gamma \sim$ 3-4. At this point, the ionization energy loss by the charged particle is smallest and the particle is called ***"minimum-ionizing-particle"*** or ***"MIP"***. After this minimum, the energy loss starts to rise due to relativistic expansion of particle's electric field, and it finally

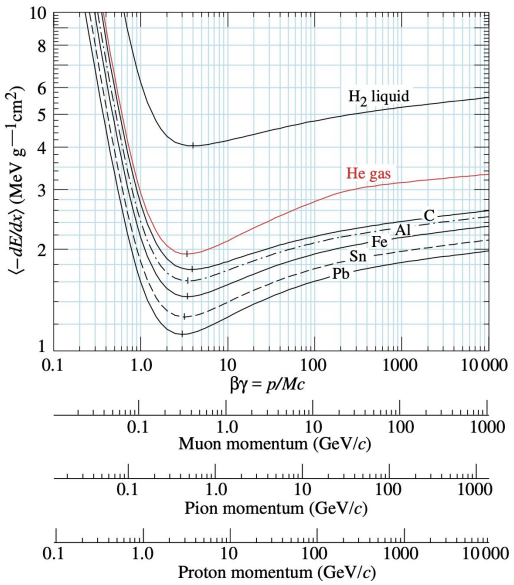

FIGURE 2.3: Mean energy loss of charged particle as a function of $\beta\gamma$ in different materials. The x-axis scale is also shown as a function of momentum for different charged particles [36].

reaches a value asymptotically.

The curve shown in the Figure 2.3 corresponds to the *average* energy loss. Since the ionization is a statistical (quantum-mechanical) process, the actual energy loss distribution has a non-zero width around the average value. The width is larger for lighter charged particles (e.g. $e^{\pm}$), and smaller for heavier charge particles (e.g. α particle).

Radiative energy loss

A charged particle in the influence of the electric field of a nucleus of an atom can accelerate or decelerate, which results in the emission of radiation in the form of a photon. This radiation is called ***bremsstrahlung***. The photon emitted from bremsstrahlung carries a part of incident charged particle's energy, thus lowering it and also changing particle's direction. Figure 2.4 shows the feynman diagram of this process.

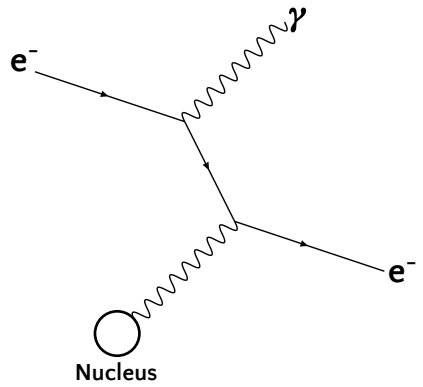

FIGURE 2.4: Feynman diagram of bremsstrahlung process.

The probability of bremsstrahlung process is dependant on the material as given below:

$$\sigma_{\text{brem}} \propto \frac{Z^2\alpha^3}{m^2c^4} \tag{2.38}$$

From the equation 2.38, it follows that the probability of the bremsstrahlung process is directly proportional to the square of atomic number (Z); therefore, for dense materials, the radiative energy loss is the dominant process. Also, it is inversely proportional to the mass (m) of the charged particle. It implies that the probability to radiate *brem* photon by a muon is 40,000 times smaller than that of electron as the muon is $\sim$ 200 times massive than electron. Hence, energetic muons ($\mathcal{O}$(10-100) GeV) produced in p-p collisions at the LHC mostly behave as MIPs in the detector.

At this point, we can define two quantities, namely *"radiation length"* (X_0) and *"critical energy"* (ϵ) as characteristics of a material which are important parameters in detector design as discussed later in the chapter.

- **Radiation length (X_0):** The radiation length is defined as the length at which an electron loses, on an average, $\sim$ 63.2% of its incident energy via radiative losses.
- **Critical energy (ϵ):** It is the energy at which ionization loss and radiation loss becomes equal for electron and can be approximated as $\epsilon \sim 560/Z$ (MeV).

The charged leptons, i.e., electrons and muons lose their energy via above two discussed processes, with ionization being the dominant process for muons.

2.2.2 Energy loss by photons

Photons lose their energy via other EM processes which are listed as follows.

- **Photoelectric effect:** In this process, a photon is absorbed by the atom of the material, and an electron is released, called *photoelectron*. The photon has to have a minimum energy, called *work function* (W_p), for this process to occur. The typical value of W_p is of the order of a few eV. The rest of photon energy is carried away by photoelectrons in the form of kinetic energy. The probability of undergoing photoelectric process depends on the energy of incident photon and the material as shown below:

$$\sigma_{\text{photoelectric}} \propto Z^5\alpha^4 \left(\frac{m_ec^2}{E_\gamma}\right)^n \tag{2.39}$$

  where n = 7/2 for $E_\gamma << m_ec^2$ and n= 1 for $E_\gamma >> m_ec^2$. The most striking feature of equation 2.39 is that the cross-section is directly proportional to the fifth power of Z. It implies that in highly dense materials such as lead (Pb) or uranium (U), low energy photons ($\sigma_{\text{photoelectric}} \propto 1/E_\gamma^n$) are absorbed mostly via photoelectric effect.

- **Compton scattering:** In this process, a photon scatters off of an electron, and its wavelength, hence energy, changes. This is the dominant process of energy loss by photons carrying energies $\sim$ 0.5 to 3.5 MeV with following cross-section dependence:

$$\sigma_{\text{compton}} \propto \frac{\ln E_\gamma}{E_\gamma} \tag{2.40}$$

- **Pair production:** If the energy of incident photon is more than $2m_ec^2$ (i.e. $>$ 1.1 MeV), then it may convert to electron-positron pair in the electric field of a charged particle, called *pair-production*. Figure 2.5 shows feynman diagram of pair production process.

It is to be noted that a photon disappearing and producing electron-positron pairs in a vacuum is not possible as it violates the law of momentum conversation, therefore a recoil object is required for this process to occur.

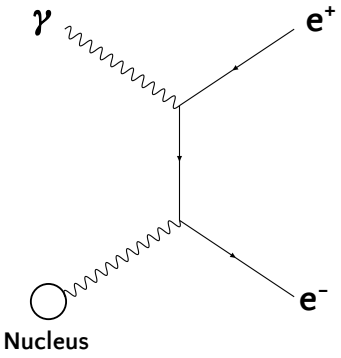

FIGURE 2.5: Feynman diagram of pair-production process.

The cross-section of this process is dependent on the material as shown below:

$$\sigma_{pair} \propto \frac{7}{9}\frac{A}{N_A}\frac{1}{X_0} \tag{2.41}$$

where A is the atomic weight, N_A is the Avogadro's number and X_0 is the radiation length of the material. Equation 2.41 indicates that the mean free path of a photon in the material before converting into electron-positron pair is:

$$L_{mean} \sim \frac{9}{7} \times X_0 \tag{2.42}$$

Therefore, a photon encounters $9/7$ times longer radiation length for the same traversed length in the material as compared to an electron. The energy ranges at which these processes dominate are different for electrons and photons, and it has an important implications in designing the electromagnetic calorimeters as discussed later in this chapter. Figure 2.6 shows the region of dominating processes as a function of incident energy of photons (left column) and electrons (right column) in different materials.

For electrons:

- At the lower energies, ionization loss is the major process through which an electron loses its energy.
- For the materials with higher Z, the critical energy (ϵ) is smaller. For example, for uranium $\epsilon \sim 9$ MeV, whereas for carbon $\epsilon \sim 95$ MeV. The radiation loss starts to dominate at an earlier stage for higher Z material which follows from equation 2.38 as σ_{brem} is directly proportional to Z^2.

For photons:

- In high Z materials, e.g., uranium, the photoelectric effect plays a dominant role at lower photon energies.
- As the A of the material increases, the pair production occurs at a lower threshold and is the dominant process through which photon losses its energy by converting into e^+-e^- pair.

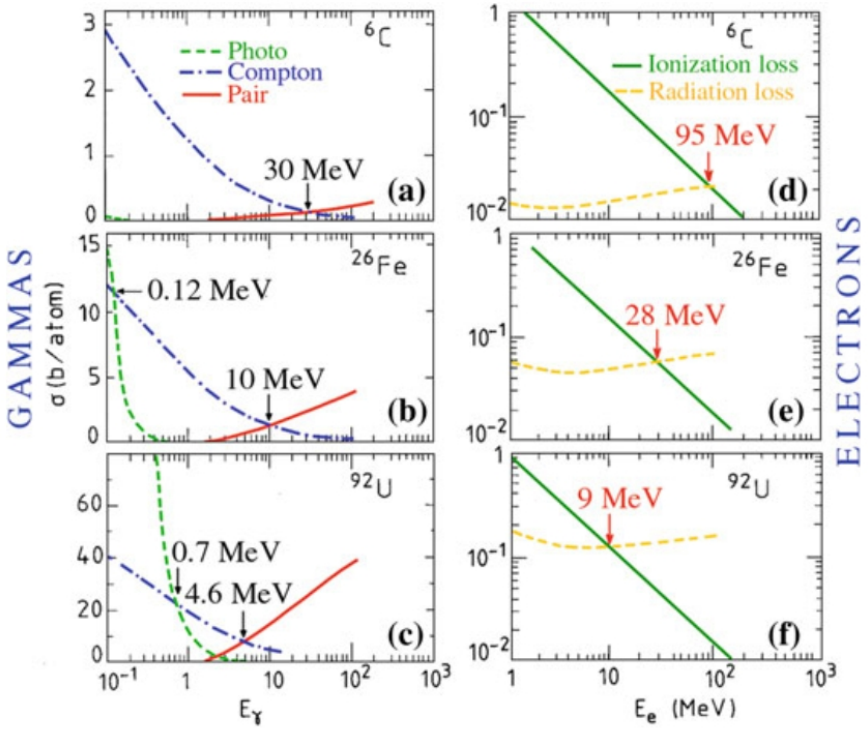

FIGURE 2.6: Figure shows the cross-sections for different processes for photons (left colum) and electrons (right column) in carbon (top row), iron (middle row), and uranium (top row). The critical energy, indicated in each plot, decreases as the Z of material increases [37].

2.2.3 Energy loss by hadrons

The hadrons, apart from losing their energies via ionization (if charged), can interact with the nuclei of the material via strong force and initiate nuclear processes, irrespective of their electric charge. The interaction can either be elastic or inelastic. In elastic strong interaction, the incident hadron scatters from the nucleus and loses some part of its initial kinetic energy, and continues down its path at a different angle. The angle of deflection depends on the energy transfer and impact parameter, thus making it a stochastic process. The inelastic strong interaction, however, is a far more complicated process in which secondary hadrons are produced, as shown in Figure 2.7. This process takes place in two steps: fast intra-nuclear cascade and nuclear evaporation, which are described below:

- **Fast intra-nuclear cascade:** The incoming projectile hadron collides with the nucleus of the atom of material. It interacts with the nucleon and transfers some part of its energy. The nucleon then gains kinetic energy and collides with other nucleons, and this process follows very rapidly. The nucleons break out of the nucleus and can also produce other hadrons. The nucleons that break out of the nucleus need to overcome the nuclear binding energy; therefore, they must have sufficient energy. The energy spent to overcome binding energy is thus lost and not available for generating any signal in the detector.

- **Evaporation:** Other nucleons that do not have enough energy to break out of the nucleus bounce back and excite the nucleus. While coming to the ground state, the nucleus emits other hadrons (mostly neutrons), photons, alpha radiation, etc. This is a relatively slow process. If the nucleus is very heavy, such as uranium, it can also undergo nuclear fission.

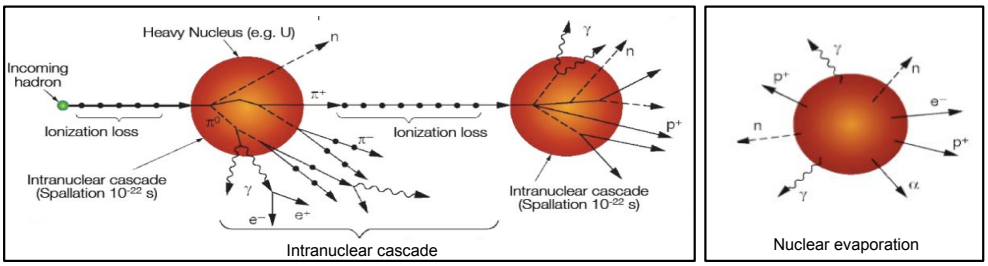

FIGURE 2.7: Cartoon representation of intra-nuclear cascade (left) and evaporation (right). The original image is created by H. C. Schoultz Coulon.

Figure 2.8 shows the elastic and total cross-section for π^+-p and p-p interaction as a function of center of mass energy. For energies above 1 GeV, the cross-section for inelastic hadronic interaction (the difference in height between σ_{total} and $\sigma_{elastic}$ in the figure) is higher than elastic collision. It increases with the increasing center of mass energy. It should also be noted that the total cross-section for p-p is higher than π^+-p interaction.

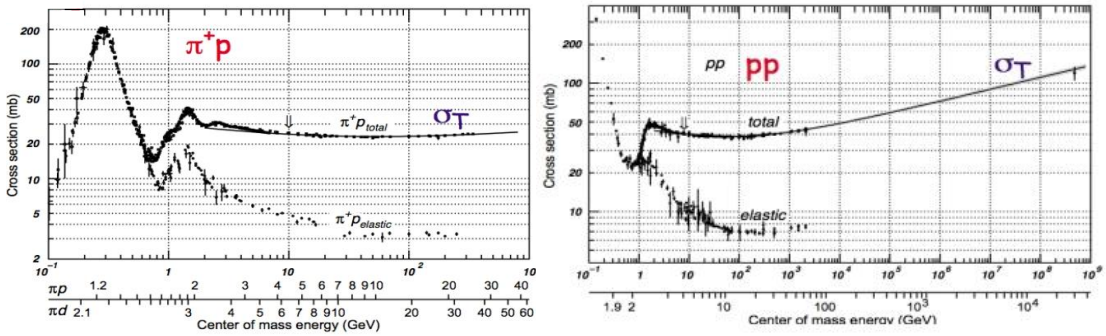

FIGURE 2.8: Cross-section for π^+-p (left) and p-p (right) interaction as a function of center of mass energy [36].

$$\sigma^{pp}_{total} = 38 \text{ mb, at center of mass energy 100 GeV}$$
$$\sigma^{\pi p}_{total} = 24 \text{ mb, at center of mass energy 100 GeV}$$

This difference can be explained by the fact that the physical size of the pion is smaller than the proton, which results in a higher probability of p-p interaction. Here, another characteristic property of material can be defined in regards to hadronic processes, which is also a very useful quantity (similar to X_0 in EM processes), called *"nuclear interaction length"*:

Nuclear interaction length (λ_{int}): It is the mean traversed length in the material at which the probability of nuclear interaction is 63.2%. The nuclear interaction length (λ_{int}) can be calculated using following expression:

$$\lambda_{int} = \frac{A}{N_A \sigma_{total}} \sim 35 A^{1/3} \text{ g-cm}^{-2} \tag{2.43}$$

where A is the atomic weight, N_A is the Avogadro's number, and σ_{total} is the total cross-section of hadronic interaction. Usually λ_{int} is quoted for protons at 200 GeV [36], however it has an implication for the pions. Since σ_{total} for protons is higher as compared to σ_{total} of pions, therefore for a given traversed length, proton *sees* more material as compared to pions in terms of λ_{int}. Table 2.1 lists various properties of different materials. Depending

on the requirements of the experiment, different materials are opted as explained in the following section.

TABLE 2.1: Properties and characteristic lengths of different materials widely used in particle physics experiments.

Material	Z	A	ρ [g/cm^2]	X_0 [cm]	R_M [cm]	λ_{int} [cm]
Beryllium	4	9.01	1.85	35.28	6.58	38.83
Carbon	6	12.01	2.21	19.32	5.01	38.83
Silicon	14	28.08	2.33	9.37	4.94	46.52
Iron	26	55.84	7.87	1.76	1.72	16.77
Copper	29	63.55	8.96	1.44	1.57	15.32
Zinc	30	65.38	7.13	1.74	1.95	19.42
Tungsten	74	183.84	19.30	0.35	0.93	9.95
Lead	82	207.2	11.35	0.56	1.6	17.59
Uranium	92	238.03	18.95	0.32	1.0	6.258

2.3 Cascade of particles in matter

The particles produced in the high energy collisions, e.g., at LHC, are highly energetic and do not lose their energy all at once. Instead, they develop a cascade of shower particles in the material. These shower particles, in turn, produce low energy secondary particles that deposit their energy via different processes, as discussed in the previous section. The shower initiated by the particles is broadly categorized into two types based on the interactions involved, namely electromagnetic (EM) shower and hadronic (Had) shower. These two types of showers have different features and characteristics that are important to understand in order to make the right choice for detectors.

2.3.1 Electromagnetic shower

A high energy electron (photon) develops a cascade of particles via bremsstrahlung (pair production). Figure 2.9 shows a representation of EM shower development in the material. In the figure, the photon enters the material from the left and converts into a pair of electron and positron, which propagate further while emitting *brem* photons. The particle multiplicity doubles at each *generation*. The process continues until energy of the secondary particles down the generation reaches critical energy (ϵ). After which, the particles get absorbed via ionization, photoelectric effect, or Compton scattering process depending on the particle type.

In such cascades, the shower of particles propagates in both longitudinal as well as lateral direction. Figure 2.10 shows simulated EM shower development in longitudinal direction as a function of depth (in X_0) for 10 GeV electron in different block of material. From the figure, one observes that the energy deposition first increases and reaches a maximum and then decreases exponentially while penetrating deeper into the material. The depth at which the energy deposition is maximum is called *shower maxima* (t_{max}). The figure shows that the t_{max} is greater for high Z material, e.g., lead (Pb), as compared to low Z material, e.g., aluminium (Al). Also, the tail of the shower is observed to be longer for Pb as compared to Al. Both of these observations can be understood by the fact that high Z materials have low critical energy (Figure 2.6). Therefore, particle multiplication continues at relatively lower energy and, hence, have a higher generation-order of particle cascade. However, it

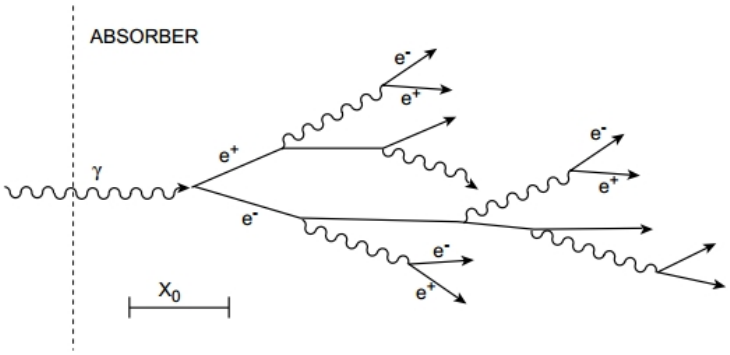

FIGURE 2.9: Schematic diagram of electromagnetic shower cascade development in the block of absorber material [38].

is to be noted that even though the shower spreads deeper into the calorimeter in a high Z absorber, the physical length (i.e., depth in meters) is considerably smaller (table 2.1).

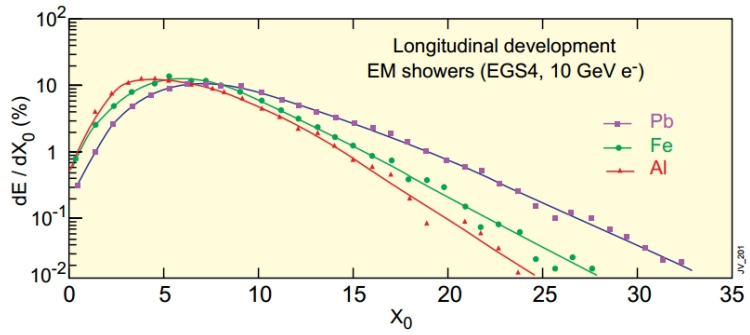

FIGURE 2.10: Longitudinal shower development of 10 GeV electron in different material (shown in different colours). The shower spreads further into the absorber for high Z material, e.g. Pb (shown in purple colour) [38].

The development of EM shower in longitudinal direction can be well modelled by Longo's parametrization as shown below:

$$\left\langle \frac{1}{E}\frac{dE(t)}{dt} \right\rangle = \frac{(\beta t)^{\alpha-1}\beta \exp(-\beta t)}{\Gamma(\alpha)} \tag{2.44}$$

where t is the material depth in X_0, Γ is the Gamma function, α and β are the shape and scaling parameter that depend on the material and incident energy of $e^{\pm}/\gamma$.

Figure 2.11 (left) shows longitudinal shower development for electrons of energies ranging from 1 GeV to 1 TeV in the block of copper. Figure 2.11 (right) shows the longitudinal depth (in X_0) required to contain fraction of total shower energy for different incident electron energies. One of the most interesting and useful feature that can be inferred from this figure is that the shower maxima (t_{max}) scales as log(E). It means we do not require twice as much material to contain shower of $e^{\pm}/\gamma$ that has twice as much energy. About $\sim 28\ X_0$ of material is sufficient to contain more than 99% of the shower initiated by 1 TeV e^{-}.

The shower also spreads in the lateral direction because of scattering of secondary $e^{\pm}$ or emission of γ away from the shower axis. If one imagines an infinite cylinder with its axis aligned with the shower axis and starts increasing the radius (R) from zero, then the energy contained in this imaginary cylinder can be plotted as a function of R which shows shower development in the lateral direction. The radius at which fractional contained energy is

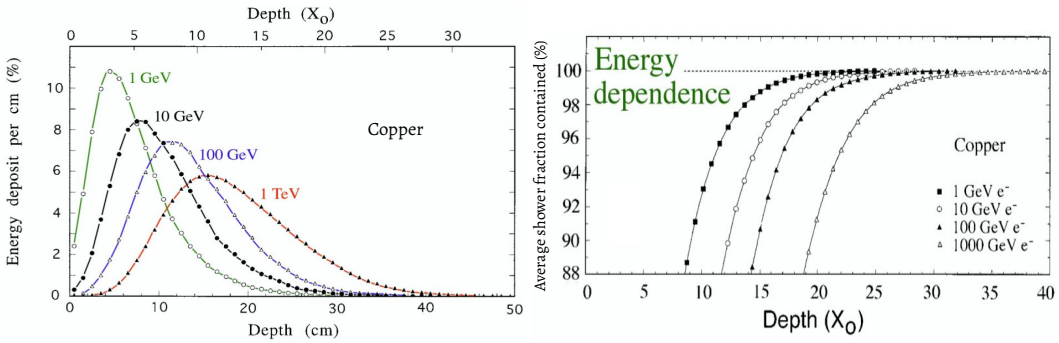

FIGURE 2.11: Longitudinal shower development (left) and average shower containment (right) in the block of copper for different energies of incident electrons.

90% is called *Moliére radius* (R_M), which is a property of the material. Figure 2.12 shows fractional energy contained as a function of radius in units of R_M for the shower initiated by 50 GeV e^- in $PbWO_4$. The Moliére radius for $PbWO_4$ is ~ 2.2 cm, which is also the lateral size of scintillator crystal used in the CMS electromagnetic calorimeter (discussed in next chapter).

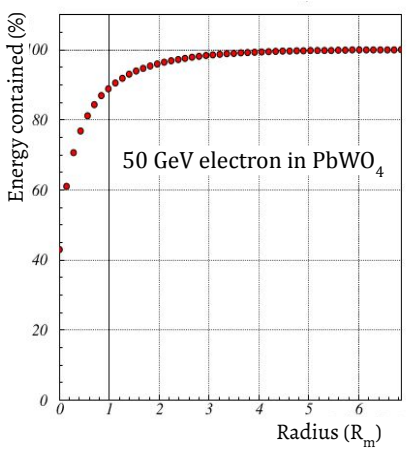

FIGURE 2.12: Transverse shower development for 50 GeV electron in $PbWO_4$. The x-axis is the radius or distance from the shower axis in terms of Moliere radius, of an infinite cylinder, and y-axis is the fractional energy contained in the the cylinder of radius R. By definition, 90% energy is contained at 1 R_m as indicated by the vertical black line.

2.3.2 Hadronic shower

As discussed in section 2.2.3, a hadron undergoes nuclear interaction while traversing through the matter, which produces various secondary hadrons. The secondary hadrons may further undergo nuclear interaction. The process of particle multiplication continues as a cascade, similar to electromagnetic shower until the secondary hadrons reach a threshold for hadron production. Hadronic showers are substantially more complicated than EM showers because of the involvement of strong interaction and a large variety of possible nuclear processes. The hadronic showers have two components: electromagnetic and hadronic, as shown in Figure 2.13.

The EM component is the result of neutral pion (π^0) decaying into a pair of photons almost instantly. The photons then develop the shower, which is purely electromagnetic

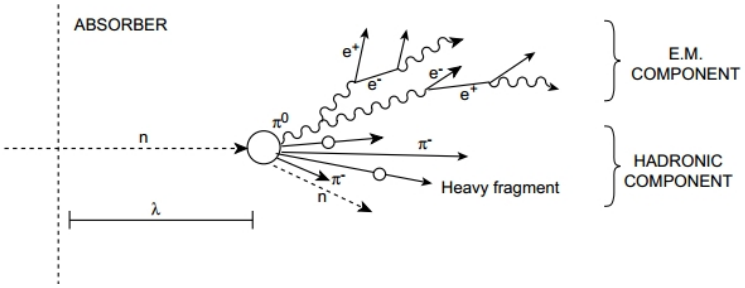

FIGURE 2.13: A cartoon representation of hadronic shower. The hadron interacts with the nucleus of material and produces secondary hadrons out of which π^0 decays almost instantly into a pair of photon and constitutes pure electromagnetic part of the hadronic shower [38].

in nature. In comparison, the hadronic component comprises spallation hadrons, evaporation neutrons, etc. A part of the energy of incident hadron is *lost* while breaking up the nuclei, hence, it is not available for any signal generation in the detector unlike EM shower, where all energy is deposited in the material. This has a far-reaching consequence on the energy response of the detector to hadronic showers, as discussed later in this chapter (section 2.6.2).

In hadronic showers, the generation of particle multiplication continues until the energy of the shower particles goes below the energy threshold to produce secondary hadrons. The typical value of this threshold is about 1 GeV for pions [39]. Whereas the corresponding threshold for the electromagnetic shower, i.e., critical energy (Figure 2.6), is significantly smaller than hadronic showers. Therefore, for the same incident energy of electron and pion, the electron shower has a higher particle multiplicity than the pion shower.

The longitudinal shower development of hadronic showers is somewhat similar to electromagnetic showers except that they penetrate deeper into the material. Figure 2.14 shows the longitudinal shower profile for different incident energies of charged pions. The figure on the right shows the average fractional energy contained in the material with varying depths. Comparing similar numbers for EM showers in Figure 2.11 (right), one infers that it takes more material to contain hadronic showers as compared to EM showers. For example, it requires $\sim 8.2\ \lambda_{int}$ (~ 137 cm) of iron to contain 99% of the shower of 100 GeV π^- whereas it requires 25 X_0 ($\sim$ 44cm) of iron to contain 99% of shower initiated by 100 GeV e^-.

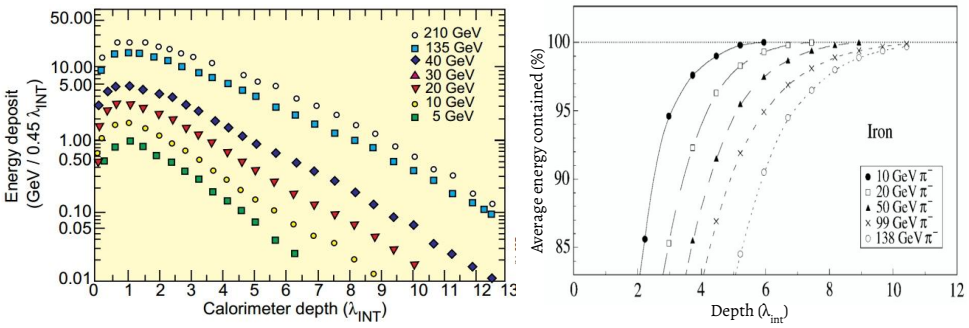

FIGURE 2.14: The longitudinal shower development (left) and average fractional shower containment (right) as a function of material depth by pions of different energy.

The shower maxima for hadronic showers are rather broad compared to a pure EM

shower. The production of π^0s at the first hadronic interaction contributes significantly to the shower maxima. The longitudinal shower development can be parametrized by two constituents: a function peaking at the initial stage of shower development corresponding to the EM component and an exponentially falling function corresponding to the hadronic component. A choice of such parametrization is discussed in detail in [40], and there are more such discussions available in literature.

The hadronic showers also have a substantial event-to-event fluctuation due to the production of neutral pions as the shower develops. The neutral pions may be generated anywhere along the shower and carry different energies based on when and where they are produced. Figure 2.15 shows such a scenario for four individual events of hadronic showers in the block of matter. The π^0s deposit their energy in a small depth of material in the form of a pure EM shower. π^0 produced earlier in the shower have a high probability of carrying large fractional energy of the incident hadron. These event-to-event fluctuations contribute to the precision with which one can measure the energy of hadronic showers.

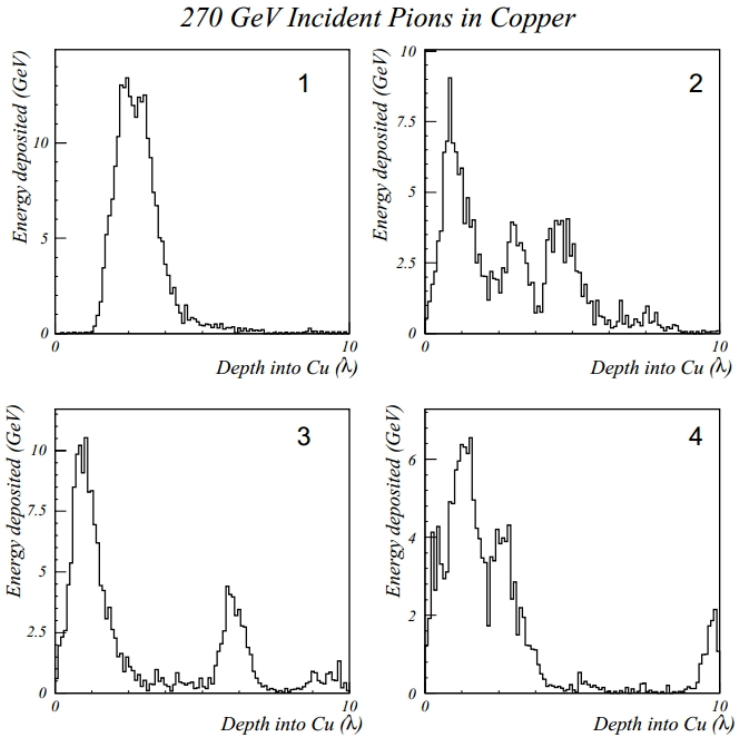

FIGURE 2.15: The four images show longitudinal shower development for four individual events where shower is initiated by 270 GeV pion in copper. The large energy deposit peak corresponds to π^0 production in the shower and shows large event-to-event fluctuation [38].

The lateral spread of hadronic showers is also wider than the EM showers. It is due to the fact that the secondary hadrons undergo multiple scattering quite often with a large angle of deflection from the shower axis. Figure 2.16 shows energy deposited in the towers of detectors as a function of distance from the shower axis. The lateral shower profile can also be parameterized with two components: a dense EM core near the shower axis and a non-EM beam halo that falls exponentially. More details can be found in [37].

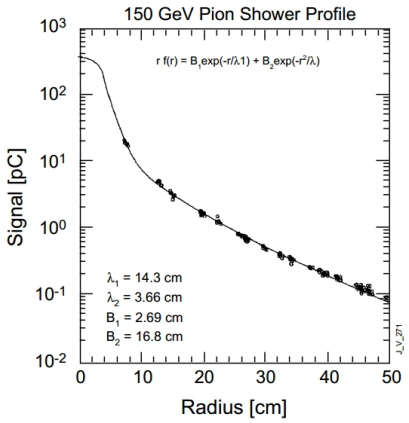

FIGURE 2.16: The signal corresponding to the energy deposited by the 150 Gev pion shower as a function of distance from the shower axis. The lateral profile is fitted with a function with two parts corresponding to electromagnetic (sharply falling near the shower axis) and hadronic component (exponential decay away from shower axis) [38].

2.4 Detection techniques

The energy deposited in a detector material needs to be converted into a measurable signal to infer the properties of the particle that traversed through the material, such as its energy, position and type. Few major detection techniques are discussed in the following.

2.4.1 Gaseous detectors

Gaseous detectors are filled with noble gas, such as argon. A charged particle while traversing the detector ionizes the gas atoms and produces electron-ion pair along its ionization path. The electron and ions move towards the electrodes under the influence of an electric field and produce a pulse that can be amplified using electronics. Typically, about 30 to 40 eV is required to produce an electron-ion pair in gaseous detectors. Based on the voltage applied across the electrodes, it can be used in different modes. Figure 2.17 shows the average total number of electron-ion pairs produced in terms of charge collected as a function of applied voltage.

In the *"proportional counting region"* as indicated in Figure 2.17, the pulse generated at the output is proportional to the energy deposited in the detector. This configuration is used in many types of gaseous detectors such as ionization chambers, wire chambers, drift chambers etc., where it is mainly used to reconstruct the track of the particle by measuring position and timing of pulse. For example, the muon stations in the CMS detector (described in next chapter), viz. drift tubes, cathode strip chamber, resistive plate chambers, are based on these detection principles.

If the applied voltage is very high ($\geq$ 750 Volts for argon), then the primary electron and ion acquire enough energy to ionize the gas and generate a secondary electron-ion pair. This process continues and results in an avalanche. The output pulse becomes independent of the energy deposited by the charged particle. This operating region is called Geiger-Muller region and is used as the counter.

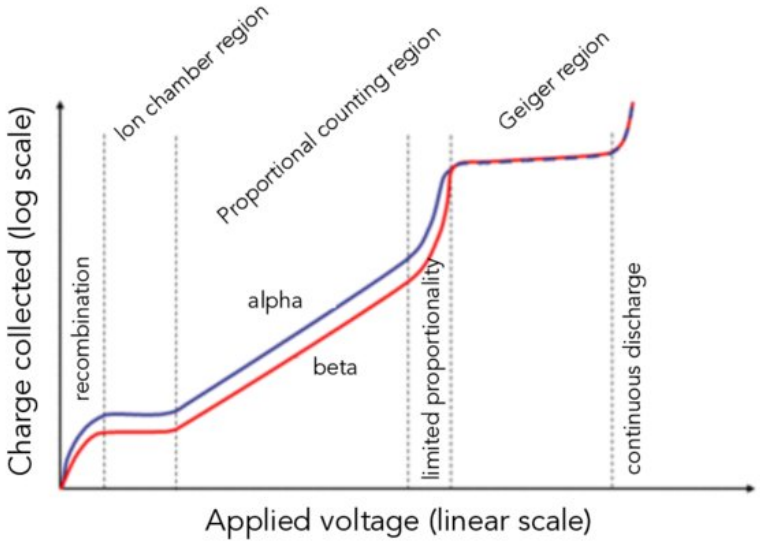

FIGURE 2.17: Figure shows a typical plot of charge collected on the electrodes due to generated electron-ion pairs as a function of applied voltage in the gaseous detectors.

2.4.2 Scintillator detector

In some materials, when a charged particle passes through it, then the material produces flashes of visible (or UV) light, called *scintillation* photons [41]. Such materials are called *scintillators*. This phenomenon occurs because of the energy band structure of the material. Apart from the ground and excited state, the scintillators have a quasi-stable state of their constituent molecules. The incident particle excites the material, and its molecules transition from the ground to the excited state. The molecules, however, do not come down to the ground state directly. Instead, they first transition to the quasi-stable state by dissipating their energy to the lattice in the form of vibrations and then jump to the ground state by emitting a photon whose wavelength typically lies in the visible or UV range. The whole process takes about a few tens of nanoseconds. Depending on the material, the average energy required to produce a photon ranges from tens of eV to a few hundreds of eV. Table 2.2 lists few scintillating materials and their properties. The light output and peak emission wavelength are important aspects required to choose appropriate photodetectors. The decay time is also an important characteristic of scintillator detector when event rates are high, e.g. at LHC.

The scintillators are coupled to photo-detectors that produce electrical pulse corresponding to scintillating photons. The electrical pulse can be processed electronically and analyzed to give information about the incident particle. The photo-detectors are photo-multiplier devices such as photo-multiplier tubes (PMT), as shown in Figure 2.18. The main components of a PMT are photo-cathode, focusing electrodes, accelerating electrodes, and dynodes. The photo-cathode produces photo-electrons corresponding to incident scintillating photons. The photo-electrons are focused and accelerated towards dynode arrangement. Each dynode generates multiple electrons, about ten electrons per incident electron, which are accelerated towards the next dynode under the influence of the electric field. Typically, a PMT has about six dynodes that provide a combined amplification of 10^6 per photo-electron, thus providing an amplified signal in the form of a pulse. The pulse is directly proportional to the energy deposited in the scintillator, which can be used for spectroscopy or counting of incident radiation. In nuclear and particle physics, scintillators are

used to detect the particles' energy and, because of their fast response time ($\sim$ few ns), to generate the triggers that decide whether to keep or throw the detector information corresponding to an event. Apart from PMTs, there are many other photodetectors which are commercially available. For example, hybrid photodiodes (HPDs) and silicon photomultiplier (SiPMs) are used in the CMS HCAL, and avalanche photodiode (APDs) and vacuum phototubes (VPTs) are used in the CMS ECAL detector, as discussed in section 3.2.

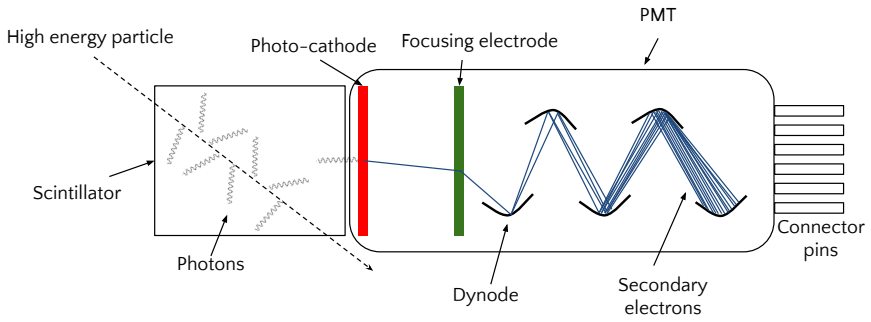

FIGURE 2.18: Scintillator + PMT setup representation. A high energy particle produces scintillating photons some of which enters PMT. The photo-cathode at the window aperture produces a photo-electron that gets multiplied by the multiple dynodes and generates electrical pulse that is read out via connector pins.

TABLE 2.2: List of scintillator material and their properties

Scintillators	NaI(Tl)	CsI(Tl)	BGO	PbWO4	LAr	Plastic
ρ [g/cm^3]	3.67	4.51	7.13	8.28	1.4	1.03
X_0 [cm]	2.59	1.85	1.12	0.89	13.5	42.4
R_M [cm]	4.5	3.8	3.4	2.2	10	9.1
λ_{int} [cm]	41.4	36.5	22.0	22.4	65.0	78.9
Decay time [ns]	250	1000	300	15	6	1-5
Peak emission [nm]	410	565	480	420	128	370-430
Rel. light yield [%]	100	45	9	0.7	N/A	28-34

2.4.3 Cerenkov detector

When a charged particle traverses a dielectric medium, it polarizes the molecules. The polarized molecules emit radiation (photons) while coming to the ground state. If the velocity of the charged particle is greater than the velocity of radiation in the medium, then the wave-front of the radiation overlaps and forms a shock-wave emanating at an angle (θ) along the path of the particle as shown in Figure 2.19. This effect is called *Cerenkov radiation* [42]. The angle (θ) of radiation is given as:

$$\cos\theta = \frac{c}{vn} \tag{2.45}$$

where c is the velocity of light in vacuum, v is the velocity of particle in the medium and n is the refractive index of the medium.

The photons from Cerenkov radiation are fed to photo-multiplier devices to produce a measurable electrical signal. The Cerenkov detectors have a fast response time (almost

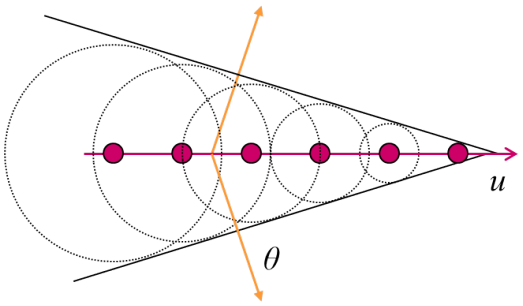

FIGURE 2.19: Cerenkov radiation generated by creates by a charged particle passing through the dielectric medium with a velocity u which is greater than velocity of light in the same medium. The Cerenkov photons radiate at an angle θ from particle trajectory.

instantaneous) and are also radiation hard. However, they have a very low light yield as compared to scintillators. Therefore, this detection principle is used for specific purposes such as in the very forward region (HF) of the CMS detector.

2.4.4 Semiconductor detectors

The semiconductors are a special type of material whose band gap between conduction and valence band lies in between insulators and conductors. The conduction band is the energy level where electrons are free to move, whereas, in the valence band, electrons are bound to the atom and can not move freely. The most common semiconductors are silicon (Si) and germanium (Ge). The pure crystals of such materials are called *intrinsic* semiconductors. The intrinsic semiconductors, however, are not suitable for particle detection at temperatures above few hundred kelvin because at these temperatures the signal produced by traversing particle is indistinguishable from inherent noise due to thermal fluctuation[1]. Instead, a combination of semiconductors with externally added impurities (called dopants) is used, as described below.

The conductivity of the intrinsic semiconductor can be increased by introducing specific impurities with elements whose atomic number differs by ± 1 with respect to the intrinsic semiconductor, called donor or acceptor atoms. For example, an impurity of phosphorus (aluminium) is introduced to silicon, such that they have one extra (one less) electron. In the case of phosphorus impurity, a donor, the lattice has an extra one electron loosely bound to the phosphorus atom that can jump to the conduction band with thermal energy alone and is called *n-type* semiconductor. Conversely, in the case of aluminium, an acceptor, the lattice lacks one electron. It borrows an electron from neighbouring atoms creating a "hole" in the valence band and is called *p-type* semiconductor. This process of introducing impurities is called *doping*. The p- and n-type semiconductors have holes and electrons as majority charge carriers, respectively, with its conductivity greater than intrinsic semiconductors.

The p-type and n-type semiconductors can be joined together using special techniques to form a p-n junction diode, as shown in Figure 2.20. At the junction, the majority charge carriers diffuse to the opposite side and recombine. Since both types of semiconductors are neutral to begin with, the transport of the majority charge carriers to the opposite side leaves a net charge at the junction, i.e., positive charge in the n-type and negative charge in the p-type region. It creates a net electric field and potential difference across the junction

[1]Experiments such as EDELWEISS [43] and SuperCDMS [44] employ ultra-pure intrinsic semiconductor detectors at cryogenic temperatures, to suppress thermal noise, for dark matter searches.

and the majority charge carriers stop from diffusing further into opposite side, and create a region depleted of charge carriers. This region is called *depletion zone*. In the depletion zone, electric field and the potential difference can further be increased by subjecting the p-n junction diode to an external voltage supply, with the negative contact made at the p-type side and positive contact made at the n-type side. Any electron-hole pair produced in the depletion zone gets swept away from the zone. This configuration is called *reverse biasing* the diode. The default current induced in this configuration is negligible which is of the order of nanoAmperes. If any ionizing radiation passes through the depletion zone of the p-n junction semiconductor, then it generates electron-hole pair along its path which are swept away under the influence of electric field and induces a non-negligible current that can be measured as a signal which is proportional to the energy deposited by the ionizing particle. The average energy required to produce electron-hole pair is $\sim$ 3 eV for silicons, which is very small compared to gaseous detectors ($\sim$ 30-40 eV). Thus, semiconductor detectors produce a high multiplicity of electron-hole pairs in a small volume, resulting in low statistical fluctuation and excellent signal resolution.

Silicon sensors being radiation hard and compact in size (thus providing good spacial resolution) as compared to other detectors, are used in tracker [45] near the interaction point in fixed target and collider experiments.

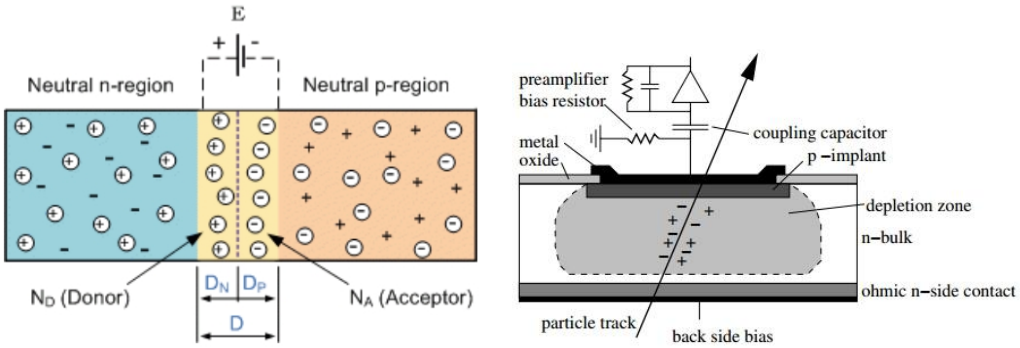

FIGURE 2.20: Schematics of a typical p-n junction diode (left) and a transverse view of ionizing particle traversing through silicon sensor (right) [45].

Modern high energy physics experiments are very sophisticated to carry out numerous measurements in order to make discoveries. The experiments such as the CMS [19] employ different types of these detection technologies simultaneously in order to achieve measurement precision of the highest order. The detectors that are used to measure the energy of the particles are called *calorimeters* and is described in detail in the following section.

2.5 Calorimeters

Calorimeters are used to measure the energy and the position of high energy charged and neutral particles by completely absorbing them. A valuable property of calorimeters is that the energy resolution improves with increasing energy of the particle (discussed in section 2.6.3) in contrast with trackers where track momentum resolution degrades with increasing energy. It makes calorimeters attractive for energy measurement of ultra-relativistic particles using the particle-flow method (see section 3.3).

Calorimeter detector systems are designed to fully contain the cascades of showers produced by $e^{\pm}$, photons and hadrons. The medium in which the incident particle develops

the shower is called *absorber* medium, and the medium in which the energy of the shower particles are converted into the measurable signal is called *active* medium. Based on the type of absorber and active medium construction, the calorimeters can broadly be segregated into two types: homogeneous and sampling calorimeter, described in the following section:

1. **Homogeneous calorimeter:** A single medium serves both as the absorber as well as the active medium. In practice scintillator detectors, such as liquefied Argon (LAr), dense crystal scintillator (BGO, $PbWO_4$) are used as homogeneous calorimeter. The shower generated by electromagnetic particles ($e^{\pm}/\gamma$) is relatively compact as compared to hadrons; therefore these calorimeters are mainly used to measure the energy of $e^{\pm}/\gamma$, and the calorimeter is called electromagnetic (EM) calorimeter. Figure 2.21 shows a simulated event display of 20 GeV electron shower in liquid-Ne in 17.4 kGauss magnetic field.

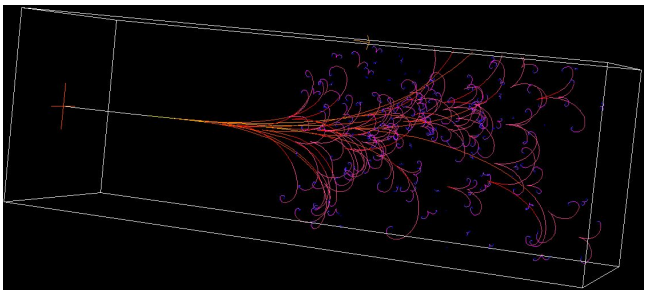

FIGURE 2.21: A simulated view of 20 GeV electron shower in a homogeneous calorimeter [46].

2. **Sampling calorimeter:** In the sampling calorimeters, block of absorbers, called passive layers, alternate with the active layers. The incident particle develops a cascade of shower in the absorber material, and the shower is sampled at various stages of development in the longitudinal direction as shown in Figure 2.22. Essentially, each active layer gives a snapshot of the shower development, which is put together to make a complete picture of overall shower development. Usually, the absorber of high Z (such as Fe, Cu, Pb, U, etc.) is used in order to achieve maximum characteristic lengths (X_0 and λ_{int}) with minimum physical depth (in meters) of the detector. These calorimeters are primarily used where a larger depth of calorimeter is required to contain the shower or the energy measurement precision is not of the highest priority. For this reason, sampling calorimeters are mostly used in the energy measurement of hadronic showers as they have considerably larger longitudinal spread as compared to electromagnetic showers, and such calorimeters are called hadronic calorimeters.

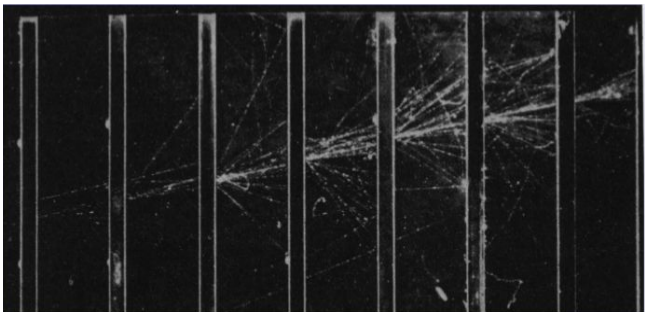

FIGURE 2.22: An image of photon shower in a bubble chamber with lead absorbers [47].

Unlike homogeneous calorimeters, only a part of shower energy deposited in the active layers contributes to signal generation. Hence, they have worse resolution than their counterpart (discussed in next section). For sampling calorimeters, the fraction of energy deposited in the active layers can be quantified in terms of *sampling fraction* (S.F.) as defined below:

$$\text{S.F.} = \frac{E_{\text{active}}}{E_{\text{active}} + E_{\text{passive}}} \tag{2.46}$$

here, E_{active} is the energy deposited in the active layers and E_{passive} is the energy deposited in the passive layers. Thus, for a sampling calorimeter with S.F. = 1%, only 1 GeV of 100 GeV incident particle contributes in signal generation. Typically, the S.F. is less than a couple of percent.

2.6 Energy response and resolution of calorimeters

Having discussed the calorimeters and its types, it is imperative to discuss the figure-of-merits of the calorimeter and the factors dictating them. There are two figure-of-merits for calorimeters, namely energy response and energy resolution. The *energy response* is defined as the total signal generated by the calorimeter in the response of energy deposited by the incident particle. The *energy resolution* is the fluctuation in the calorimeter energy response. The energy response of the calorimeter depends on whether the particle develops an electromagnetic shower or hadronic shower as discussed below.

2.6.1 Response to electrons and photons

The calorimetric response to electromagnetic shower scales linearly with the incident energy of $e^{\pm}/\gamma$. This is because of the fact that, in electromagnetic showers, all of the shower energy is deposited in the calorimeter and contributes to signal generation. Thus, the energy response to electromagnetic shower is inherently linear for any type of calorimeter. However, the response can deviate from the linear trend because of instrumental effects such as saturation of photodetectors at higher energies or shower leakage. Such instrumental effects can be avoided with appropriate precautions, e.g., lowering the PMT gain, choosing sufficient calorimeter depth, etc. Figure 2.23 shows energy response of Si-W ECAL calorimeter of CALICE [48] as a function of electron incident energy. The plot on the left shows measured signal in terms of MIP equivalent of energy deposits as a function of electron beam energy, fitted with a straight line. The plot on the right shows the residual of each point with respect to the fitted straight line. The residuals are also shown for simulation (red open square). The energy response shows a linear trend (within uncertainties) as expected for electromagnetic showers.

2.6.2 Response to hadrons

As described in section 2.3.2, hadronic showers have two components, viz. electromagnetic and hadronic. While the electromagnetic part (predominately from the di-photonic decay of π^0) deposits all of its shower energy in the calorimeter, the hadronic component has an invisible part resulting from breaking up atomic nuclei of the calorimeter medium. The invisible energy does not contribute to the calorimeter signal. Hence the calorimetric response to the hadronic shower is different as compared to a pure electromagnetic shower. In other words, if the calorimeter shows energy response "e" to 50 GeV electron and "π" to 50 GeV pion then:

$$\frac{e}{\pi} \neq 1$$

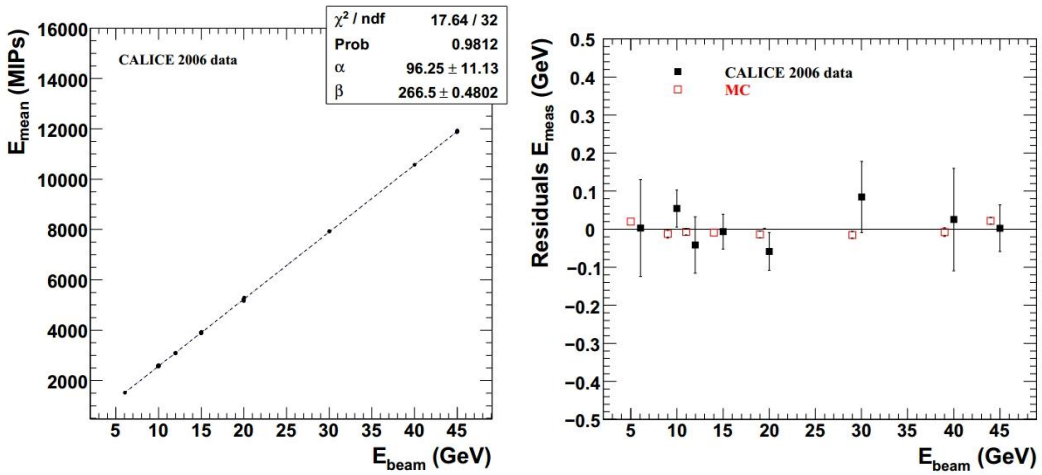

FIGURE 2.23: Left plot shows measured calorimeter energy as a function of electron beam energy fitted with a straight line and the right plot shows energy linearity in terms of residuals [48].

The energy response to hadrons can be defined as the combination of the electromagnetic (em) and non-electromagnetic (non-em) part of the hadronic shower as shown below:

$$\pi = f_{em} \cdot e + (1 - f_{em}) \cdot h \tag{2.47}$$

where f_{em} is the em fraction in the shower, e is the calorimeter response to em part and h is the calorimeter response to hadronic (non-em) part. Making algebraic manipulations to equation 2.47, one arrives at following two equations:

$$\frac{e}{\pi} = \frac{e/h}{1 - f_{em} \cdot (1 - e/h)} \tag{2.48}$$

and

$$\frac{e}{h} = \frac{1 - f_{em}}{\pi/e - f_{em}} \tag{2.49}$$

Equation 2.48 shows that based on the e/h ratio, calorimeter can exhibit either higher, lower, or equal response to hadronic showers with respect to electromagnetic shower. Thus, the calorimeter can be segregated into following three categories:

- **Compensating calorimeter:** for $e/h = 1$
- **Under-compensating calorimeter:** for $e/h > 1$
- **Over-compensating calorimeter:** for $e/h < 1$

The e/h ratio of the calorimeter is a fundamental property of the detector that affects how the calorimeter responds to hadronic showers. The value of the e/h ratio can not be measured directly. However, it can be extracted by measuring the response ratio e/π of an electromagnetic shower (e.g. the beam of electrons) and hadronic shower (e.g. charged pions) along with equation 2.49.

One of the other and equally important feature of hadronic shower response is the energy-dependence of fraction of em component (f_{em}). This fraction is not a constant, rather it varies non-linearly with respect to the incident energy of the hadron. The average em component can be parametrized [39] as follows:

$$\langle f_{em} \rangle = 1 - \left[\frac{E}{E_0} \right]^{k-1} \tag{2.50}$$

where E is incident energy of the hadron, E_0 is average energy needed for pion production ($E_0 \sim 0.7$ and 1.3 GeV for Cu and Pb, respectively), and $k-1$ is related to average pion multiplicity in the shower ($k \sim 0.82$ for both Cu and Pb). This introduces a non-linearity in the calorimeter response to hadronic showers. Figure 2.24 (left) shows the non-linear trend of em fraction (with fitted parametrization 2.50) as a function of pion energy as measured by SPACAL and QFCAL experiment [37]. Figure 2.24 (right) shows energy response of CMS HCAL to hadronic showers. The energy response is non-linear, especially at lower energies. The CMS HCAL calorimeter is calibrated using fixed energy pions. Therefore the response goes above one after $\sim$ 50 GeV. This non-linearity in the energy response is fixed by applying offline calibration, as discussed in appendix B. From equation 2.48, one can note that if e/h is one then the non-linear contribution from f_{em} vanishes and the response becomes linear.

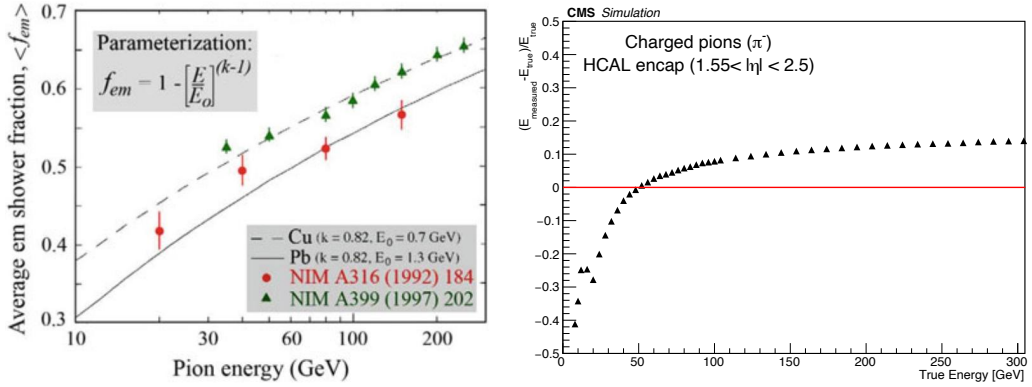

FIGURE 2.24: Energy dependence of em fraction in hadronic showers is shown on the left and right plot shows non-linear energy response of CMS HCAL calorimeter to charged pions.

2.6.3 Energy resolution of calorimeters

The energy resolution of a calorimeter can be attributed to the following three major contributions:

- **Stochastic fluctuation:** This is the most fundamental fluctuation in the energy response and originates from the statistical nature of particle multiplicity in the shower. The particle multiplicity (n) in the shower follows a Poisson distribution with a standard deviation of $\sqrt{n}$. Therefore the relative resolution (standard deviation normalized by mean) scales as $1/\sqrt{n}$. Because of comparatively lower shower particle multiplicity and large event-to-event fluctuation, hadronic showers have significantly worse resolution than pure electromagnetic showers. This is an irreducible fluctuation and is most dominant in the intermediate energy range, i.e., ten to hundreds of GeV.

- **Noise:** This fluctuation is due to the electronics noise introduced by the various components instrumented in the detector. The particles emanating from the pileup also contribute to the energy distribution fluctuation. This fluctuation is most dominant at the lower energies ($\leq$ few GeV), and the contribution vanishes at higher energies.

- **Fluctuation due to imperfect detector:** The imperfection in the detector also contributes to the fluctuation in the energy response. These imperfections include shallow detectors resulting in shower leakage, imperfect cell-to-cell calibration, etc. These fluctuations dominate at very high energies, i.e., hundreds of GeV and above.

All of the above factors contribute to the overall energy resolution of a calorimeter independently. The relative energy resolution of a calorimeter can be parametrized using the following expression:

$$\frac{\sigma}{E} = \frac{S}{\sqrt{E}} \oplus \frac{N}{E} \oplus C \tag{2.51}$$

where S is the stochastic term corresponding to stochastic fluctuation, N is the noise term corresponding to noise fluctuation, and C is the constant term corresponding to fluctuation due to imperfect detector. All these terms are added in quadrature as they are independent of each other.

In sampling calorimeters, only a part of shower energy contributes to a recordable signal. Therefore it suffers from a further fluctuation in the sampled energy. Thus, the sampling calorimeters always have a worse resolution as compared to homogeneous calorimeters. To first order, fluctuation in the energy response due to sampling fluctuation can be absorbed in the stochastic term.

2.7 GEANT4 simulation

In particle and nuclear physics, the discovery of new physics relies on careful analysis of experimental data collected during its operation. In practice, the data analysis, however, can not be done on an event-by-event basis as billions of collision events are collected. Instead, a statistical analysis is done in the kinematic phase-space and is compared against physical theories of interest. Based on the knowledge of the theories, the events are simulated in the experimental conditions using *MONTE CARLO* methods [49]. Simulating the events accurately is a multi-faceted task that starts with, for example in LHC experiments, p-p collision, particle production, decay modes, the interaction of decay products with the detector, and finally, the detector response to these interactions. Each of these steps is done for billions of events repeatedly to generate a *simulated data-set*. The reliability of simulation prediction depends on: (a) robustness of physics models that generates event dynamics at the *"truth"* level (called event generators), (b) the realistic description of detector geometry and construction, and (c) accurate modelling of particles interaction with detector material and the detector response.

For a detector of the size of CMS, the last part is the most complex, challenging, and resource-intensive task. The CMS experiment uses the GEANT4 simulation toolkit [26] for this very purpose. The GEANT4 is an object-oriented software framework that allows to construct detector geometry of any scale and at the same offers various (and ever-improving) physics models to simulate particle interactions with the material.

2.7.1 GEANT4 detector construction

The first step towards detector simulation in GEANT4 is detector construction. The GEANT4 toolkit provides modular geometry building in which detector geometry is built out of several volumes. The largest *volume* is called the *world volume* which contains all other volumes.

The *volume* is a solid which can be of any geometrical shape, dimension, and material. The material is an important input as that dictates the characteristic lengths (X_0, λ_{int}) of the detector. GEANT4 has a PDG [36] based database of the elements, material, isotopes and also allows to instantiate user-defined material compositions, such as alloys.

The modular geometry of GEANT4 allows to create multiple instances of a volume which are placed in the world volume. Based on the experiment, a part of the detector volume is made as *"sensitive"* material, i.e., the material of the detector. Although the particle deposits energy in all of the detector material by various interactions, energy deposited in the sensitive material is retained for energy reconstruction. Using all these details, the final detector geometry is constructed. For example, Figure 2.25 shows a simulated CMS detector geometry constructed using the GEANT4 toolkit.

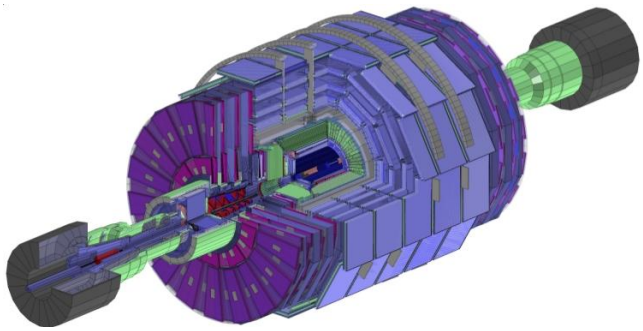

FIGURE 2.25: Simulated geometry of the CMS experiment using GEANT4 toolkit

2.7.2 Particle interaction simulation

GEANT4 provides an exhaustive list of particles and models for their interactions with the matter that the user can instantiate according to their requirement. Based on the detector geometry description, GEANT4 propagates each instance of particle through the detector material for each event. In this process, the particle can undergo different interactions based on particle type and its energy. Each particle is propagated and tracked until it reaches a minimum energy threshold below which the particle is *"killed"*, and the corresponding instance is destroyed. For GEANT4, various physics models have been developed based on the knowledge of all the possible particles interactions with the material, as discussed in this chapter. It offers separate models for simulating electromagnetic and hadronic shower cascade described briefly in the following.

Electromagnetic shower cascade

Electromagnetic physics includes interactions of charged particles and photons such as multiple Coulomb scattering, ionization loss, radiation loss, pair production, etc. The theoretical models for these interactions are well developed and are easy to implement in the GEANT4 framework. Numerous validation for GEANT4 electromagnetic shower physics have been made with the experimental data and have been found to be in good agreement within a few percent [50].

Hadronic shower cascade

Hadronic showers, on the other hand, are more complex than electromagnetic showers and are non-trivial to simulate. In hadronic shower cascade, the strong interactions include elastic scattering, inelastic interaction, capture, fission, evaporation for the incident hadron as well as for the target atomic nuclei. Calculating the final states analytically is an extremely difficult task. Therefore, the hadronic physics models make use of various approximations and parametrizations that describe the hadronic processes in different energy regimes. Based on the energy and DeBroglie wavelength of the projectile hadron, two different types of models attempt to simulate hadron-nucleon interaction: intra-nuclear cascade and parton string model. More details can be found in the GEANT4 physics manual [51]. A few of such models are described below:

- **Bertini cascade:** In the energy range of 0 to 10 GeV (maximum limit), the Bertini (BERT) cascade model approximates the nuclear medium as three concentric constant density shells. The nucleon momentum is sampled according to the Fermi gas model. The point of entry for the projectile in the nucleus is sampled over its projected area. The projectile is transported inside the nucleus in the straight line, and the mean free path is determined according to the free hadron-nucleon total cross-section. The secondaries produced in the reaction are also transported in a similar manner. They may reflect back or penetrate through the concentric shells of the nuclear medium, depending on their energy. Finally, the nucleus itself may undergo nuclear evaporation if it has sufficient energy to emit neutron or alpha particles. The photons are emitted for energies below 0.1 MeV.

- **Binary cascade:** Binary cascade is an alternative to the Bertini cascade model. The energy region of validity corresponds to 0 to 3 GeV for protons/neutrons, 0 to 1.5 GeV for pions, and 0 to 3 GeV/A for light ions. In binary cascade, first a 3D model of the nucleus is built. Then the interaction cross-section of the hadron-nucleon is evaluated according to the momentum of projectile and nucleon momentum (sampled according to the Fermi gas model). The primary and secondary hadrons are propagated in a curved path through the nuclear potential. Thus, a cascade of secondary hadrons is developed along with the left-over excited nucleus. This model also implements the Coulomb barrier for charged hadrons which is not implemented in the BERT model. The de-excitation of the nucleus is handled by GEANT4's inbuilt pre-compound model [51]. The Binary cascade model has been found to work well in the energy regimes up to 10 GeV.

- **Quark-Gluon string model:** If the energy of projectile hadron is large enough, the interaction among individual quarks of hadron and nucleon becomes possible. In such cases, parton string models are used. The Quark-Gluon String model (QGS) handles the interactions with proton, neutron, pions, and kaons in the energy range of approximately > 20 GeV. The QGS first builds a 3D model of the nucleus and is collapsed to a 2-D disk for a highly boosted projectile. Based on the impact parameter, the collision probability is calculated using quasi-eikonal model [51] and Gaussian density distributions of hadron and nucleon. The strings of gluons are formed between a single quark of the projectile hadron and the nucleon mediated by n sampled Pomerons. The string is fragmented and is proceeded by hadronization, creating secondary hadrons. It is then followed by intra-nuclear cascade by one of the above two mentioned models. The de-excitation of the nucleus is handled by GEANT4's inbuilt pre-compound model [51].

- **Fritiof model:** The Fritiof (FTF) model is also a parton string model used to describe hadronic interaction for large energy. Unlike QGS, the string formation in FTF is dictated by the momentum exchange in diffractive cross-section between hadron and nucleon in binary reactions [52]. The string fragmentation proceeds via the LUND-string fragmentation model [53]. The rest of the hadronization and de-excitation processes are common for FTF and QGS models.

In practice, a single model may not necessarily describe the true hadronic interactions at all energy regimes. Therefore, the physics models are combined together to form a *"physics list"*. The different models are used in the physics lists with a fixed region of validity with a small overlapping region between two models. The physics model is chosen randomly with equal probability in the overlap region to make a smooth transition. In this thesis, the following two physics lists have been used to compare and validate GEANT4 based simulation against experimental data.

- **FTFP_BERT_EMN:** It uses the Bertini cascade model for energies less than 12 GeV and Fritiof model for energies greater than 3 GeV. The EMN corresponds to the electromagnetic physics model that gives extra care on handling multiple scattering, which is critical for sampling calorimeter.
- **QGSP_FTFP_BERT_EMN:** It employs the Bertini cascade model for energies up to 8 GeV, the Fritiof model between 3 to 25 GeV, and the QGS model for energies above 12 GeV.

2.7.3 Event simulation in GEANT4

Having constructed the detector geometry and the list of particle interaction models in GENAT4, the required number of events can be generated. The GEANT4 provides two manager classes to handle the simulation, namely `G4RunManager` and `G4EventManager` class. The `G4RunManager` encapsulates all the necessary information needed to simulate the events such as geometry and physics. The `G4EventManager` encapsulates the representation and structure of the event. A hierarchy of processes, called `RunAction`, `EventAction` and `StepAction`, are invoked while simulating the events. The process of simulation is described briefly in the following.

For each event, a particle (particles) is (are) generated using either GEANT4 inbuilt event-generator or an external event-generator interfaced with GEANT4 such as `PYTHIA` [54]. Each particle is tracked along its path, called `G4Track`. The `G4Track` is propagated in steps, called `G4Step`. At each `G4Step`, a stepping action is invoked where the step is propagated to the next step. The length of the step can be adjusted according to the user's requirement based on resource consumption and required precision. Based on the current state of the step, such as kinetic energy, particle type, the stepping action invokes all the available physical processes that the step (the particle) can undergo. At this stage, the secondary particles and their `G4Track` instances are created in the process are stored along with corresponding information, i.e., kinetic energy, particle type, position etc. The change in the kinetic energy of the primary particle is also noted and updated. If the particle traverses through the *sensitive* material of the detector, then a *hit* is created and following associated `G4Step` information is saved:

- energy deposited by the step
- the position of the step
- the time-stamp of the step

- the geometrical information of the detector

The hit is called a *simHit* in GEANT4 terminology. The associated geometrical information of the detector and the energy deposited is crucial for event reconstruction. For example, based on geometrical information, the simHit belongs to a particular sensor of the detector, and the deposited energy is converted into signal generation by the sensor. With the help of GEANT4's inbuilt method called `simWatcher`, various other information of `G4Step` (be it in sensitive material or otherwise) can be stored according to the user's requirement, such as particle type, energy, generated number of secondaries etc.

Thus, for each event, each `G4Track` is propagated and tracked inside the detector until it comes to rest or it goes beyond the boundaries of the detector geometry. The *simHit* information is stored for each event and is used for event reconstruction. The events can be repeated as many times as required for a run. The simulation process is completed at the end of the run. More details about the GEANT4 application can be found here [55].

Chapter 3

The CMS experiment at the LHC and recent physics results

The SM has been put to stringent tests in a number of particle physics experiments such as proton-proton collisions using the CMS and other experiments at the Large Hadron Collider (LHC). The CMS detector comprises various sub-systems optimized for measuring different particles information with the help of a dedicated event reconstruction algorithm called Particle-Flow algorithm. Using the CMS detector, cross-sections of various SM processes as well as properties of the Higgs boson have been measured. The data from a very rich program of physics searches have been found to be consistent with the SM expectations even in the most tight kinematics phase spaces including multi-particle final states. This chapter provides a brief overview of the LHC and the CMS experiment along with a very brief glimpse to the current status of physics measurements with it. In the following, I describe the key features of the CMS detector and how we go about making these studies in the proton-proton collision data delivered by the LHC.

3.1 Large Hadron Collider

The Large Hadron Collider or the LHC [18] is the world's largest and most powerful particle accelerator and collider built which is operated by *European Organization for Nuclear Research (CERN)*. The LHC is a 27 km long accelerator ring, located $\sim$ 100 m underground at the border of Switzerland and France as shown in Figure 3.1. The LHC operates in three collision modes: (a) proton-proton collision, (b) proton-lead collision, and (c) lead-lead collision. Different collision modes serve different experimental purposes. There are four major detectors located around each collision point, namely *ATLAS* [56], *CMS* [19], *ALICE* [57], and *LHCb* [58]. The ATLAS and the CMS are general purpose detectors designed to cover a broad physics program, whereas ALICE and LHCb are dedicated for heavy-ion physics and b-physics, respectively. A discussion on heavy-ion program is beyond the scope of this thesis. In the proton-proton (p-p) collision mode, the protons are collided at the center of mass energy of 13 TeV in the current LHC operation.

Accelerating protons to a very high energy requires a series of accelerating units. The performance of each unit is optimized for a particular energy range, therefore protons pass through various stages of CERN's accelerator complex, shown in Figure 3.2, before getting injected into the LHC ring [59]. The proton beams can also be extracted from intermediate stages, when required, to perform fixed target or beam test experiments. For example, Figure 3.2 shows *North Area*, where the HGCAL beam test experiments were performed, which is the topic of this thesis work. The steps involved in the proton beam acceleration are described below.

FIGURE 3.1: An aerial view the LHC ring with four collision points and corresponding experiment.

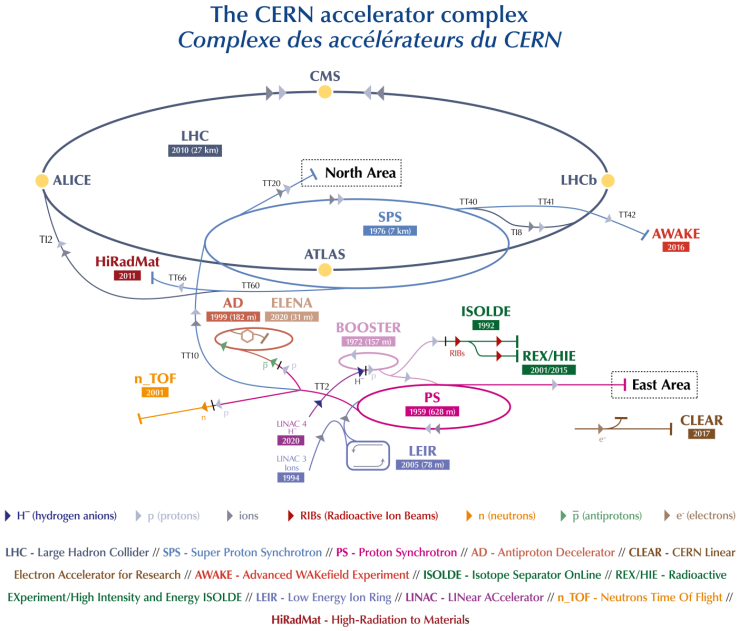

FIGURE 3.2: A sketch of the CERN accelerator complex [60].

The source of protons in the LHC ring is the hydrogen gas. The protons are created by stripping electrons from the hydrogen gas source. The protons are sent to radiofrequency quadrupole (RFQ), where they are focused and bunched together. The bunches of protons are injected into a linear accelerator called *LINAC2*, where the protons are accelerated to 50 MeV. The proton beam from LINAC2 passes through a set of synchrotron accelerators, namely Proton Synchrotron Booster (PSB), Proton Synchrotron (PS), and Super Proton Synchrotron (SPS). The PSB accelerates the proton to 1.4 GeV and injects the beam into PS. The PS increases the proton beam energy to 26 GeV, and SPS increases it to 450 GeV before injecting it into the LHC ring. The synchrotron accelerators also help reduce the transverse beam size and increase beam brightness. Finally, the LHC ring accelerates the proton beam to 6.5 TeV. The protons circulate in the LHC ring as *bunches*. Each bunch contains about $\sim 1.15\times10^{11}$ protons and are made to collide every 25 ns. With a bunch spacing of 25 ns, the LHC can accommodate 3560 bunches in its 27 km long ring. However, only 2808 bunches are filled at a time in order to introduce a sufficient time window for: (a) proton bunch injection to maintain collision rate at the interaction point (IP), and (b) safe dumping of the proton beam in case of emergency.

The event rate for a particular process can be estimated with the help of cross-section of the process and instantaneous luminosity (L). The instantaneous luminosity is a characteristic of the collider machine and is related to cross-section and event rate in the following manner:

$$\frac{dN}{dt} = L \cdot \sigma \tag{3.1}$$

where $\frac{dN}{dt}$ is the event rate of the process, σ is the cross-section of the process, and L is the instantaneous luminosity. From the relation 3.1, the unit of instantaneous luminosity is expressed in terms of $\text{cm}^{-2}\text{s}^{-1}$. The total number of events generated can thus be obtained by integrating equation 3.1 over the time during which the collider is operational:

$$N_{total} = \int \sigma \cdot L dt = L_{int} \cdot \sigma \tag{3.2}$$

where L_{int} ($= \int L dt$) is the integrated luminosity. In order to accumulate sufficient statistics for rare processes, such as Higgs boson production (σ_H = 50 $\times10^{-37}$ cm^2 or 50 *pb* or 5×10^6 *outhouse*), a high instantaneous luminosity is required. In particle and nuclear physics, the cross-section is expressed in the units of barns, outhouse, or shed and are defined as follows: 10^{-24} cm^2 = 1 barn = 10^6 outhouse = 10^{24} shed.

The instantaneous luminosity is a combination of various accelerator parameters as shown below:

$$L = \frac{N_b^2 n_b f_{rev} \gamma_r}{4\pi\epsilon_n \beta^*} F \tag{3.3}$$

where N_b is the number of particles in a bunch, n_b is the number of bunches circulating in the ring, f_{rev} is the frequency of the bunch revolution, γ_r is the Lorentz boost factor, ϵ_n is the normalized transverse emittence, β^* is the beta function at the collision point, and F is the geometric luminosity reduction factor. The factor $\epsilon_n\beta^*$ defines the beam spot size at the IP and the factor F is defined as follows:

$$F = \frac{1}{\sqrt{1 + \left(\frac{\theta_c \sigma_z}{\sigma_{xy}}\right)^2}} \tag{3.4}$$

where θ_c is the crossing angle, σ_z is the root-mean-square (rms) longitudinal size of bunch length, and σ_{xy} is the rms size of bunch in the transverse direction.

The LHC collides beams of protons at the instantaneous luminosity of 2×10^{34} cm^{-2}s^{-1}. The collimator magnets of the LHC will be upgraded in the coming years that will reduce β^* and crossing angle θ_c. It will result in increased instantaneous luminosity as a direct consequence of equation 3.3 and 3.4.

At the LHC, multiple p-p interactions occur during the collision. Most of the these interactions are soft interactions, i.e. momentum transfer is not large, producing soft particles. A few are hard interactions with a large momentum transfer such that massive SM or BSM particles can be produced, and these interactions are of interest. The soft interactions are called *pileup interactions* (μ). The pileup poses many challenges for the detector and is discussed in section 4.3. The instantaneous luminosity dictates the number of p-p interactions during a bunch crossing. With the total p-p inelastic cross-section σ^{pp}_{inel}, L_{inst}, and time between bunch crossing (ΔT), one can calculate the average pileup ($\langle\mu\rangle$) using equation 3.2 as follows:

$$\langle\mu\rangle = L_{inst} \times \Delta T \times \sigma^{pp}_{inel} \tag{3.5}$$

Substituting the corresponding values for LHC, i.e., $L_{inst} = 2\times10^{34}$ cm^{-2}s^{-1}, $\Delta T = 25$ ns, and $\sigma^{pp}_{inel} = 80$ mb at 13 TeV center of mass energy, we obtain $\langle\mu\rangle$ to be 40. The calculated value of $\langle\mu\rangle = 40$ is very close to LHC operation during 2018 as shown in Figure 3.3. Figure 3.3 (left) shows peak instantaneous luminosity for different years of LHC operation since its beginning in the year 2010. Figure 3.3 (right) shows pileup event distribution for corresponding years. The average pileup increases with increasing instantaneous luminosity.

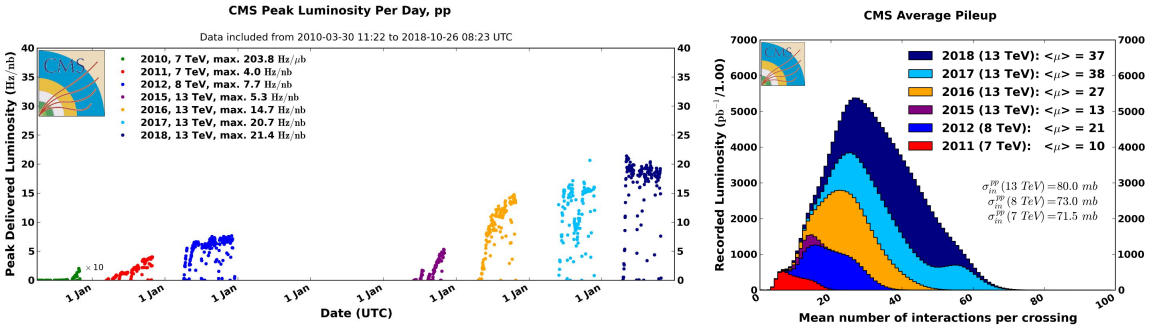

FIGURE 3.3: Left figure shows the peak instantaneous luminosity recorded by the CMS during different years of operation, and corresponding pileup distribution is shown on the right [61].

3.2 The CMS experiment

The primary motivation of the CMS experiment is to shed light on the electroweak symmetry breaking. This, however, is not the only physics goal behind the conception of the CMS detector. The CMS experiment is also aimed to probe Beyond Standard Model (BSM) physics in the direct searches of new particles in GeV to TeV mass scale as predicted by the various novel theories, such as Supersymmetry (SUSY) [20].

To cover such a vast landscape of physics program, the detector requirement for the CMS is such that it can provide:

- excellent muon identification, and good resolution of muons of very high momentum ($\sim$ TeV/c)
- excellent energy measurement of $e^{\pm}/\gamma$
- excellent tracking of charged particles to facilitate event tagging such as τ leptons and jets originating from the b quark
- hermetic coverage to ensure precise estimation of missing transverse energy (E_T^{miss})

All these requirements ensure that the CMS detector can yield a good trigger efficiency, good di-lepton and di-photon mass resolution to help in Higgs searches and at the same time provide good estimation of E_T^{miss} and track reconstruction for probing new physics. The final commissioned CMS detector design is shown in Figure 3.4. The CMS is built in an onion-like structure where different sub-detectors are placed in concentric cylinders around the collision points. The innermost part is the *Tracker* that is used for reconstructing the trajectories of charged particles traversing through it. The electromagnetic calorimeter (*ECAL*) is placed around the Tracker, and it is used for energy measurement of electrons and photons as they deposit all of their energy in the ECAL. However, the hadronic shower extends beyond ECAL, therefore a hadron calorimeter (*HCAL*) is placed around ECAL to measure the energy of charged and neutral hadrons. A large superconducting solenoid surrounding the HCAL provides a magnetic field of 3.8 T in its large bore of diameter $\sim$ 6.3 m. The trajectory of the charged particles bends under the influence of this magnetic field, which helps to determine the momentum of the charged particles from the curvature of their tracks. A large cylindrical yoke of iron is used for returning the magnetic field. Alongside the return iron yoke, *Muon stations* are integrated that helps in identification and momentum measurement of muons traversing the detector. The details of each sub-detector are described briefly in the following.

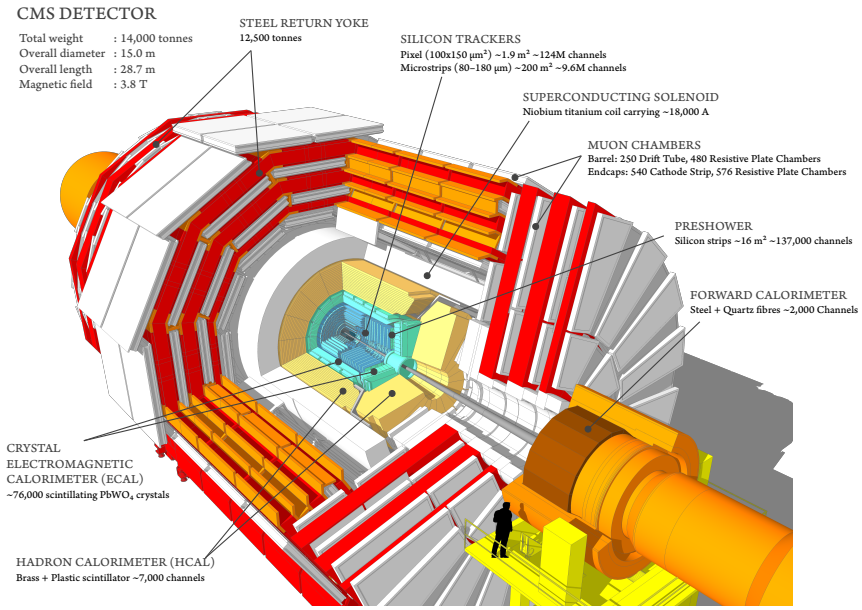

FIGURE 3.4: A schematic view of the CMS detector with each part indicated in the image.

3.2.1 Tracker

The inner tracker [62] is a combination of silicon-based pixel and strip layers. The tracker is located 4.4 cm away from the p-p collision point, hence it is subjected to a large amount of radiation. Therefore, the detector material is chosen to be silicon as it is more radiation hard as well as compact as compared to other detector materials. The tracker provides a pseudo-rapidity (η) coverage of $|\eta| < 2.5$ with the central barrel covering $|\eta| < 1.2$ and two endcaps disks in either side of the barrel covering $1.2 < |\eta| < 2.5$. The tracker has four (three) pixel and ten (twelve) strip layers in the barrel (endcap) region. The area of each cell in the pixel layer is about 100×150 μm^2 with 300 μm thickness and provides a position resolution of $\sim$ 10 μm. The pitch of strip sensors varies from 80 μm to 183 μm depending upon the distance from the beam spot. Each pixel and strip layer registers a hit when a charged particle passes through it. With the combination of all the hits in the tracker layers, the track of the charged particle is reconstructed. The resolution of the transverse momentum (p_T) as estimated from the curvature of the track is given below:

$$\frac{\sigma}{p_T} = 1.5 \times 10^{-4} \cdot p_T \oplus 0.005$$

3.2.2 Electromagnetic calorimeter (ECAL)

The ECAL [63] is a homogeneous calorimeter made of $PbWO_4$ scintillator crystals. The ECAL is divided into a barrel and two endcap disks, providing coverage of $|\eta| < 1.479$ and $1.479 < |\eta| < 3.0$, respectively. The photodetectors in the ECAL uses avalanche photodiodes (APDs) in the barrel and vacuum phototriodes (VPTs) in the endcaps. The VPTs, although have lower quantum efficiency than APDs, are used in the endcaps because of their relatively higher radiation tolerant capability, and have larger area to cover the back face of the crystal (thus increasing light collection). There are a total of 61200 (7324) $PbWO_4$ crystals used in the ECAL barrel (each endcap). The physical dimension of each crystal is $2.2 \times 2.2 \times 23$ cm^3 ($2.9 \times 2.9 \times 22$ cm^3) in the barrel (endcap) region. The lateral area of crystals correspond to $\sim 0.0174 \times 0.0174$ in $\eta - \phi$ plane. The lateral dimension is equivalent to one Moliére radius and the longitudinal length is $\sim$ 25 radiation lengths (X_0), thus ensuring $\sim$ 98% electromagnetic shower containment in the longitudinal direction. The energy resolution of ECAL to electromagnetic shower is as follows:

$$\frac{\sigma}{E} = \frac{2.8\%}{\sqrt{E}} \oplus \frac{12\%}{E} \oplus 0.3\%$$

A thin detector, called *preshower*, is placed in front of each endcap disk. The preshower is made up of two lead (Pb) absorbers of ~ 2 X_0 and 1 X_0, respectively, each followed by a silicon micro-strip layer. The silicon strip layer is placed perpendicular to each other and has a pitch of 1.9 mm. The fine transverse granularity of preshower helps differentiate between prompt photons (i.e., generated at the interaction vertex) and photons originating from neutral pion decay ($\pi^0 \to 2\gamma$).

3.2.3 Hadron calorimeter (HCAL)

The HCAL [64] is a sampling calorimeter made of plastic scintillators as the active layers interspersed between brass absorbers. The HCAL is divided into barrel (HB) and two endcap disks (HE), providing a coverage of $|\eta| < 1.3$ and $1.3 < |\eta| < 3.0$, respectively. The HCAL employs 17 layers of scintillator tiles of 3.7 mm thickness for all layers except first layer which is of 9 mm thickness. The photons produced by the traversing particles in scintillator tiles are fed to the wavelength-shifting (WLS) fibers and are carried to the photo-detectors, viz. hybrid photodiodes (HPDs) in the barrel and silicon photomultipliers

(SiPMs) in the endcaps [65], to match the spectral response of photo-detectors. The light is optically summed up from multiple fibers resulting in loss of longitudinal segmentation of HB sampling layers. The HB (HE) is read out as one (seven) longitudinal segment(s). During Phase-1 upgrade, HPDs will be replaced by SiPMs in HB, and more longitudinal segmentation will be possible (discussed in section 4.2).

The lateral dimension of the scintillators corresponds to a rather coarse granularity in the $\eta - \phi$ plane $\sim 0.087 \times 0.087$ in the barrel , and $\sim 0.17 \times 0.17$ in the endcap. The total permissible longitudinal length of HCAL was limited to the left-over space between ECAL and solenoid which is about 6 to 10 interaction lengths (λ_{int}) depending on $|\eta|$. A tail-catcher, called Outer Hadron calorimeter (HO), is placed outside the solenoid in the barrel region to collect signal from late starting showers. The energy resolution of HCAL to hadronic showers is given as follows:

$$\frac{\sigma}{E} = \frac{110\%}{\sqrt{E}} \oplus 9\%$$

In the very forward part, a steel-quartz fiber-based calorimeter (HF) is placed to extend the pseudo-rapidity coverage to $|\eta| < 5.0$. The HF is installed $\sim$ 11 m downstream of the interaction point to intercept the particles flying in the very forward region. The radiation-hardened quartz-fiber are inserted into specially drilled grooves in the steel absorbers in the longitudinal direction. The fibers are of two types: one of which covers the full longitudinal range of HF, called long-fibers, and the other starts from 22 cm downstream of the front face of HF, called short-fiber. The two fiber types help reconstruct the electromagnetic and hadronic components in the hadronic showers. The details of the hadronic shower are discussed in section 2.3.2. The HCAL has a total 9072 number of channels with 2592 in barrel and endcap, and 3888 in the HF. The hermetic coverage of HCAL is crucial in jet reconstruction and estimation of missing transverse energy (E_T^{miss}) in the event.

3.2.4 Superconducting solenoid

One of the main features of the CMS detector that stands out among other particle physics experiments is the magnetic field of 3.8 T supplied by its super-conducting solenoid in the bore of such a large diameter of about $\sim$ 6 m and length of 12.5 m. The superconductor used in the solenoid is an alloy of Niobium and Titanium (NbTi) that can conduct a nominal current of $\sim$ 19.14 kA. At its peak, the solenoid can generate a magnetic field of 4 T and has a stored energy of 2.6 GJ. Its stored energy and the *cold mass* ratio is about $\sim$ 11.6 kJ/kg, which is the highest among other solenoid-based magnets. The cold mass in a superconducting magnetic is the total mass of the magnet that operates at a few kelvin temperature which is required for superconductors. It mainly includes magnet windings, and its housing. The magnetic field is limited to 3.8 T instead of 4 T because at 4 T the magnetic field lines extends until the CMS supporting structure resulting in power losses through *Eddy current heating*. The magnetic field is returned with the help of 10,000 tonnes of iron yoke consisting of three layers that surround the solenoid.

3.2.5 Muon stations

The primary task of muons stations is to identify and reconstruct muons traversing through the detector and provide the trigger on muons. The muon stations are built with three technologies: drift tube (DT) covering the barrel region $|\eta|$ <1.2, cathode strip chambers (CSC) covering endcap region $1.0 < |\eta| < 2.4$ complemented with resistive plate chambers (RPC). The DTs and CSCs help in muon identification, momentum measurement and provide triggers on muons. With its excellent timing capability, the RPC provides a redundant trigger

on muons with relatively worse position measurement. There are 12 (6) measuring planes per station in the barrel (endcap) region. Each station has a position resolution of about $\sim$100 μm. A global track is reconstructed with the combination of hits in the inner tracker and muon stations. The momentum is estimated through the *saggita* of its trajectory. The complementary information provided by the inner tracker and the muon stations results in excellent momentum resolution for muons in a wide-angle coverage. The combined system of inner tracker + muon stations provides a momentum resolution of 1% (10%) at 50 GeV (1 TeV).

3.2.6 Triggers and data acquisition system

At the LHC, with the collision rate of 40 MHz, it is not possible to store and process the data of each collision event because of limited capacity of offline storage and CPU power. Therefore, the event rate is reduced substantially with the help of *trigger* system. The CMS incorporates two levels of trigger, namely Level-1 (L1) Trigger and High-Level Trigger (HLT). The L1 Trigger is an on-detector customised hardware-based trigger. The L1 Trigger has been designed such that the triggered events are of interest as much as possible. This is done by using coarsely segmented data from calorimeters and muon stations to look for a predefined pattern, such as high E_T or high p_T physics objects. The current technologies allow L1 Trigger a time window of 3.4 μs to make a *positive* trigger decision, during which the data from all the subdetectors are kept in a pipeline. If the event fails L1 Trigger, the collision data for the corresponding event is discarded permanently. Otherwise the data from the pipeline is passed to CPU farms located on the surface. The L1 Trigger reduces the data volume by limiting the rate from 40 MHz to 100 kHz. The HLT [66], on the other hand, is a software based trigger that runs on the aforementioned CPU farms. It has access to all the sub-detector information, therefore it can a perform a better analysis on the events . A fast event reconstruction is performed at the HLT level, and is subjected to a set of predefined criteria. If at least one criteria is satisfied, then the event is stored permanently along with a `True` flag for all other satisfied conditions. Otherwise the event is discarded. The HLT reduces the data rate from 100 kHz to 500 Hz.

In the CMS, the average event data size is about $\sim$ 1 MB per event, and scales linearly with the instantaneous luminosity. An event rate of 100 kHz at L1 Trigger implies that about 100 GB of physics data is generated per second. The data acquisition (DAQ) system in the CMS is responsible for transmitting the data volume generated by different sub-detectors to the CPU farms for further analysis. For this purpose, the CMS employs optical links, and VME or μ-TCA based telecommunication architecture. It provides a large bandwidth and fast data transfer over its network. More details of the CMS DAQ system can be found here [19].

3.3 Particle flow algorithm

The particles originating from p-p collisions deposit energy in different sub-detectors while traversing the CMS. Most heavy and unstable particles, such as Z^0 boson and top quark, decay into lighter and relatively stable particles before reaching the detector. The signature left by these lighter particles in different sub-detectors is used to reconstruct the particle identity and their four-momentum vector. The CMS employs the concept of *particle flow* for particle identification and event reconstruction [67].

In the particle flow (PF) method, the information from different sub-detectors is combined in order to identify individual physics objects and assign appropriate energy. The

implementation of this approach in the CMS detector owes credit to the detector design that provides a powerful magnetic field, an excellent spatial-granularity tracker, highly segmented ECAL, hermetic HCAL, and its excellent muon stations. The redundant measurements allow physicists to choose the best available measurement in the event reconstruction, leading to better efficiency and resolution. Figure 3.5 shows a sketch of the transverse slice of the CMS detector and shows the signature left by different types of particles in different sub-detectors.

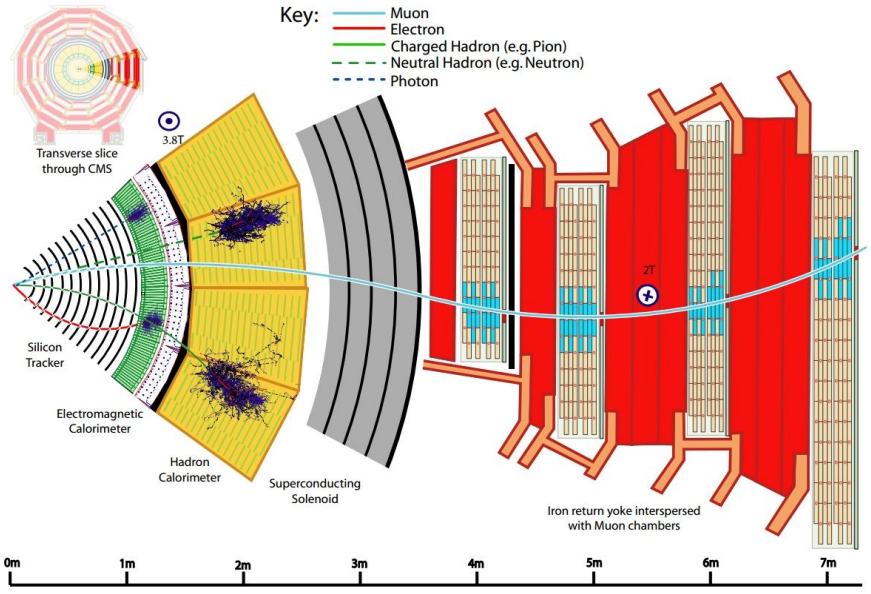

FIGURE 3.5: Schematic view of transverse slice the CMS detector showing energy deposited by different particles in different sub-detectors.

The PF algorithm is a set of various reconstruction steps and has numerous interconnected parts. A detailed description of the algorithm can be found in [67]. In summary, for each event, the PF algorithm begins by reconstructing the tracks from all available hits in the tracker. The track reconstruction is performed in an iterative manner in which the best quality tracks from consecutive hits are first reconstructed and is removed from further iterations. In subsequent iterations, the next best quality tracks are reconstructed, and the iteration continues until all the hits are exhausted.

The tracks reconstructed by the PF algorithm is used in the p-p interaction vertex reconstruction and primary vertex identification [68]. The vertex reconstruction is performed by a dedicated *deterministic annealing* algorithm [69]. This algorithm clusters the tracks together on the basis of their z coordinate of their points of closest approach to the beam spot. An *adaptive vertex fit* [70] determines the position of each vertex. With at least 50 reconstructed tracks, the resolution of reconstructed primary vertex is about $\sim$ 20 (30) μm in the transverse (longitudinal) direction of proton beam axis. The precise identification of primary vertex is crucial to mitigate pileup contribution by rejecting the particles emanating from secondary vertices in the event. Also, it is used in the identification of *displaced vertices* originating from on-flight decay of particles away from the beam spot, such as τ-leptons and B hadrons, and identify jets originating from b-quarks. These displaced vertices are used to reconstruct the original particle. For example, τ-leptons are reconstructed with the so-called *Hadron-Plus-Strip* (HPS) algorithm [71], integrated in the PF algorithm. The performance of vertex reconstruction algorithm, hence pileup mitigation, is entirely dependent

on the precise track reconstruction by the PF algorithm which in turn is dependent on the spacial resolution of the tracker.

As mentioned earlier, the electromagnetic objects, i.e., $e^{\pm}$ and γ, deposit all of their shower energy in the ECAL, whereas the shower of hadrons extends into the HCAL. Based on the clusters of energies in ECAL and HCAL $e^{\pm}$, γ and hadrons are identified. The clusters of energy deposited in the ECAL and HCAL are reconstructed separately. The particles are identified by linking the clusters and tracks together. A dedicated linking algorithm performs this task in the PF algorithm. The links are made if they satisfy a predefined criteria based on proximity in η-ϕ plane. The electrons or positrons, being electrically charged, leave a track in the silicon tracker and deposit their energy in the ECAL. Therefore, if a link between a track and ECAL cluster is found, the object is identified as electron or positron. The electric charge on the object is determined by the bending direction of associated track in the magnetic field. The energy of the object is assigned based on the track momentum. Similarly, charged hadrons are reconstructed by linking the tracks together with ECAL (if available) and HCAL clusters.

The muons are reconstructed with the hits available in the inner tracker and associated hits in muon stations. A global fit is performed to all the hits, and the momentum of the muon is estimated from the track.

The neutral particles, such as photons and neutral hadrons, do not leave any track in the tracker and do not bend under the influence of the magnetic field. Therefore, neutral particles are reconstructed based on isolated energy clusters in ECAL and HCAL. A photon is created if an ECAL cluster with no associated HCAL cluster is found. Neutral hadrons are created from the HCAL cluster with (or without) associated ECAL cluster.

The energy scale of ECAL and HCAL clusters, which are input to the PF algorithm, are set by their respective calorimeter level calibrations. For example, ECAL is calibrated using *isolated* electrons, and HCAL is calibrated using *isolated* charged hadron with track-momenta ranging from 40 to 60 GeV that do not undergo hadronic interaction in ECAL. With different energy scales of calorimeters, the combination of ECAL and HCAL clusters does not yield the correct energy of incident hadrons. Also, since hadrons exhibit non-linear energy response (see section 2.6.2), therefore, cluster calibration for hadrons becomes a necessity to correctly assign their energy before proceeding further ahead. The PF hadron cluster calibration is described in appendix B, and the concept makes the basis of the pion energy reconstruction in beam test data of HGCAL prototypes presented in chapter 7.

Thus, for each event, the PF algorithm creates a *global event description* containing a list of all the particle candidates with the most precise energy or momentum measurement by taking advantage of different sub-detectors. The particle candidates are then used as input in the jet-clustering algorithm to create *jets* by combining all the available information of particle candidates. The jets are the experimental signature of quarks and gluons in the detector. The CMS uses a specialized anti-k_T jet-clustering algorithm [72] with the distance parameter (dR) of 0.4, called AK4 jets for abbreviation. Figure 3.6 shows the jet energy response and resolution comparison between two AK4 jets in barrel region when clustered using calorimeter information, called calo-jets, and using PF information, called PF-jets. It is observed that PF-jet shows superior performance in the terms of improved response and resolution as compared to calo-jets.

This state-of-art detector system and PF based global event description makes it possible to make precision measurements of the SM predictions as well as search for new physics in a variety of final states as very briefly discussed in the following section.

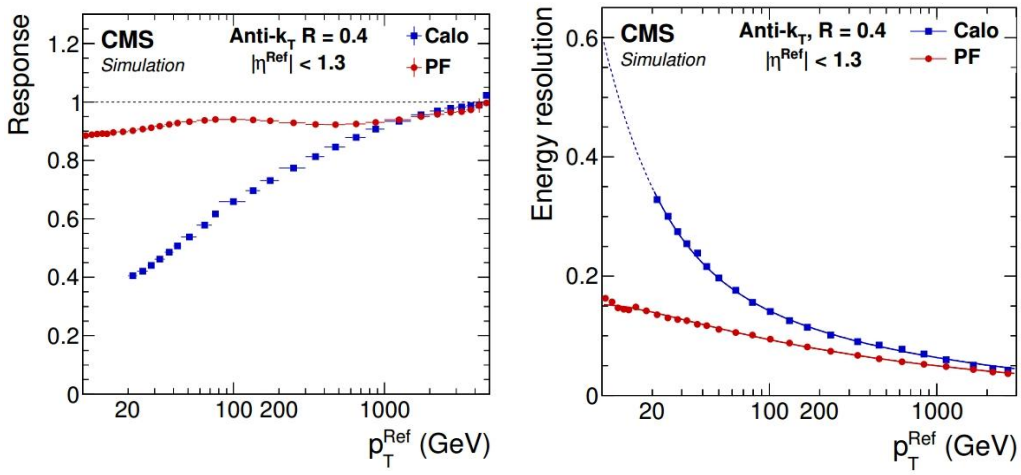

FIGURE 3.6: The response (left) and resolution (right) comparison between PF-jet (red marker) and calo-jets (blue markers) [67].

3.4 A glimpse of the CMS physics results and way forward

Since 2010, the LHC has collided p-p bunches at $\sqrt{s}$ = 7 TeV (2010-11), 8 TeV (2012) and 13 TeV (2015-2018). The Run-1 (physics data collected during 2010 to 2012) and Run-2 (physics data collected during 2015 to 2018) proved to be a milestone in particle physics with the discovery of the Higgs boson in 2012, which was the highlight of the decade. The production cross-section for various SM processes has been measured with high precision, and have been found to be consistent with the SM prediction across ten orders of magnitude, as shown in Figure 3.7. The processes covered in the figure correspond to vector boson production, vector boson pair production, top quark pair production, associated production of top quark production with W, Z and H, and several Higgs processes. The measurements are performed not only inclusive in a production mode but also in differential channels, such as number of jets, and in many kinematic variables any deviation in which can give hints of new physics effects. Most of the measurements are found to be consistent with the theory predictions within the current experimental and theoretical uncertainties.

The SM makes very definite and precise predictions for the Higgs boson decays into various channels, and hence, its coupling strengths to various SM particles. The Higgs coupling measurements to massive SM fermions and bosons show good agreement with the SM as shown in Figure 3.8 (left). Using full Run-2 data, a direct evidence of Higgs boson decaying to second generation fermions (H $\rightarrow \mu\mu$) have been observed for the first time with 3 σ statistical significance [73]. With combined Run-1 and 2016 data, the CMS experiment has further reduced uncertainty in the Higgs boson mass measurement , viz. m_H = 125.38 $\pm$ 0.18 GeV as shown in Figure 3.8 (right), the most precise measurement of Higgs mass till date [74].

The data from Run-1 and 2 has shown no sign of new physics. Nevertheless, it has provided valuable insights by constraining the parameter space of various BSM physics

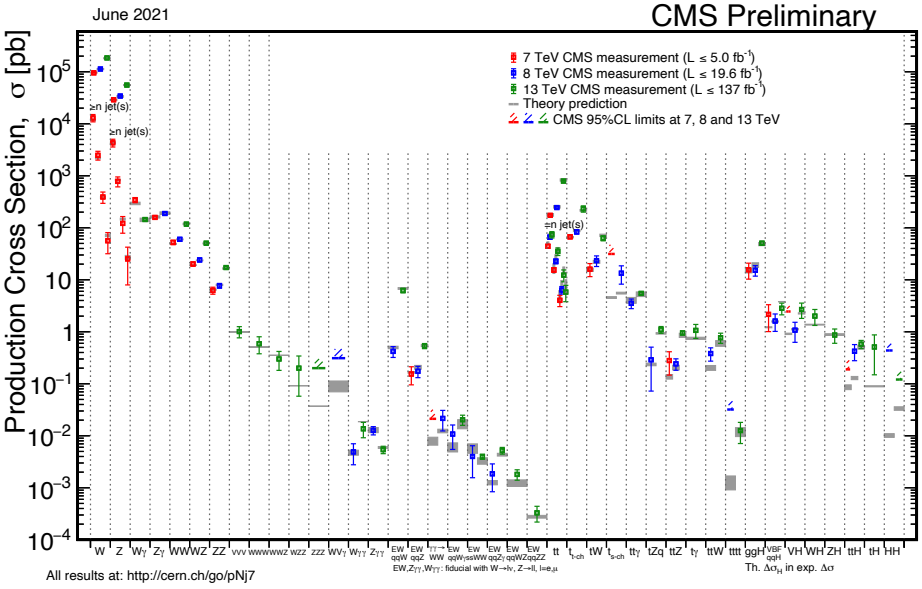

FIGURE 3.7: Production cross-sections of different SM processes.

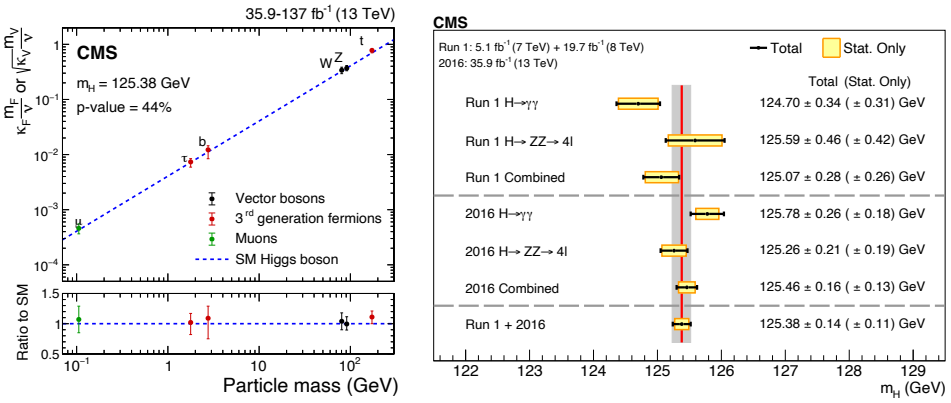

FIGURE 3.8: Left plot shows measurement of Higgs coupling to different SM particles using Run-2 data, and dotted blue line shows SM prediction [73]. The plot on the right shows Higgs mass measurement using different data sets and decay channels [74].

models. Extensively reviewing the CMS physics program for BSM searches is beyond the scope of this thesis. However, a few examples have been presented in the following. Figure 3.9 shows the mass reach for various squarks ($\widetilde{q}$) in different decay modes. The squarks, e.g. $\widetilde{t}$ and $\widetilde{b}$, are supersymmetric partners of SM quarks t and b, respectively. From the Figure 3.9, one can see that possibilities of up to TeV range particles have been explored at the CMS experiment predicted under various hypotheses like SUSY, extra dimensions etc. The experiment has also probed several scenarios that predict long lived particles. Please refer to [75] for a complete set of results.

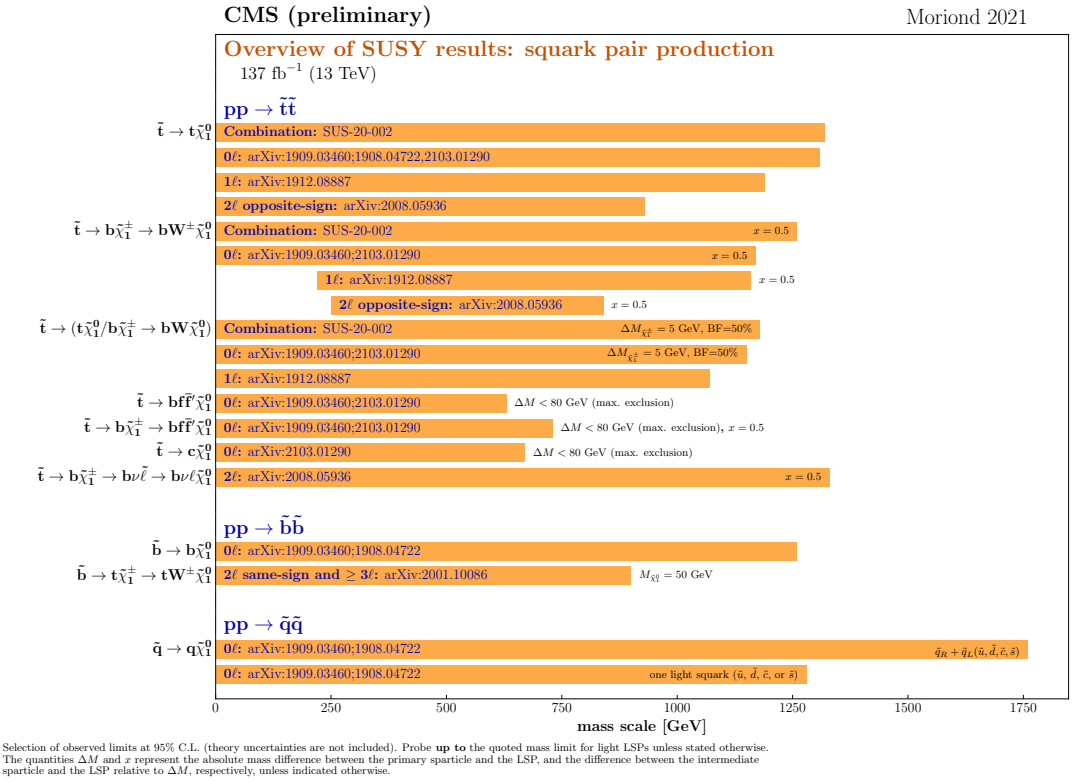

FIGURE 3.9: Excluded mass scale for squarks i.e. SUSY particle [76].

All the measurements have shown good agreement with the SM. However, a high level of statistical precision is yet to be achieved on various SM measurements. Many BSM models predict the existence of multiple Higgs bosons, therefore a precise measurement of the Higgs coupling is a robust tool to test the validity of the SM. Any deviation from the SM prediction in the Higgs property would bring us closer to finding new physics. Also, the existence of new particles as predicted by BSM theories is expected to have extremely small production cross-sections. Therefore, more collision data is required to perform resonance searches with high statistical significance.

The preparation for next run of physics data taking (Run-3) is on-going and will start from mid 2022. The Run-3 will continue till 2025 accumulating a combined (Run 1 + Run 2 + Run 3) data set of 300 fb^{-1}. Immediately after Run-3, the upgrade of the LHC will begin that will increase its instantaneous luminosity. The next phase of data-taking with increased luminosity is called Phase-2 or HL-LHC. The detectors will also be upgraded and improved. The HL-LHC will be operational for another ten years accumulating a total of 3000 fb^{-1} physics data which brings many physics opportunities. For example, Figure 3.10

(left) shows the projections of uncertainties in the measurement of Higgs coupling to different SM fermions and bosons. The projection shows that with high statistics, the uncertainty will be below 4% for most of the Higgs coupling measurements. The di- and tri-Higgs production measurement will become accessible with full Phase-2 data, which is the key to measure Higgs self-coupling and its scalar potential [77]. Figure 3.10 (right) shows projected uncertainties in different decay channels of di-higgs production mode. The combination of all the channels will allow to reduce the uncertainties even further.

With higher statistics, physicists will be able to cover more phase space of BSM physics. Figure 3.11 (left) shows the excluded phase space for SUSY particle $\widetilde{\chi}_1^{\pm}$ corresponding to 80 fb^{-1} Run-2 data, and the projection for full 3000 fb^{-1} data shows that mass reach will increase by a factor of more than two. Also, various detector upgrades will increase sensitivity to many BSM searches. For example, introduction of a new timing detector (MTD) in the CMS, discussed in next chapter, will provide measurements of timing and β (i.e., $\frac{v}{c}$) of the decay particles. It will help in the direct searches of Long-Lived-Particles (LLP) allowed in many BSM extensions. The precise timing information can be used to discriminate between signal and background processes, improving the search sensitivity. Figure 3.11 (right) shows the mass resolution of an LLP, the stau ($\widetilde{\tau}$) of mass 432 GeV, with and without MTD detector.

Thus, the HL-LHC brings exciting opportunities that will either establish in favor of or rule out various potential BSM physics models. The details of the CMS upgrades for the HL-LHC phase are discussed in the next chapter.

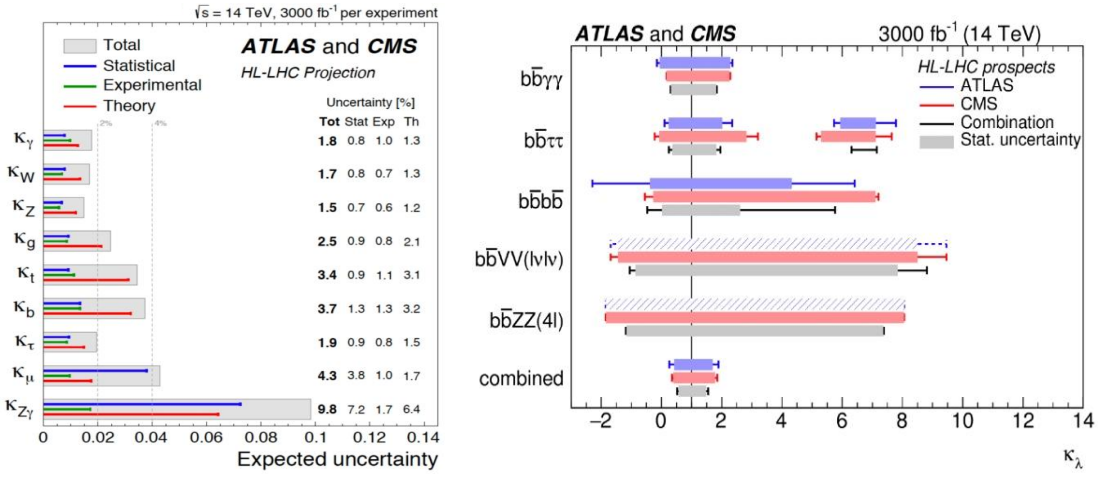

FIGURE 3.10: The HL-LHC projections of uncertainties in the measurements of Higgs coupling to different SM fermions and bosons (left), and uncertainties in higgs-self coupling measurement in different di-higgs decay channels (right) [78, 79, 80, 81].

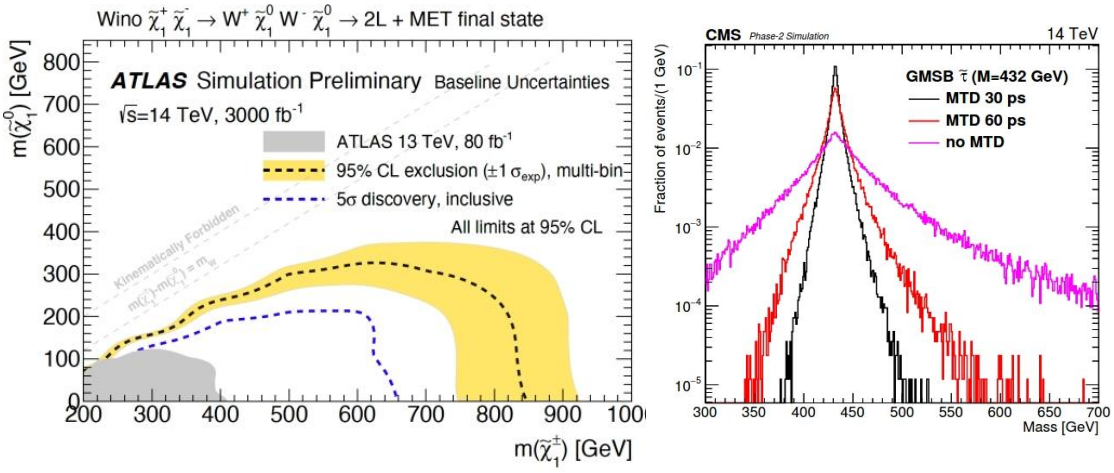

FIGURE 3.11: The figure on the left shows excluded phase space of charged neutralino ($\tilde{\chi}_1^{\pm}$) with 80 fb^{-1} Run-2 data set and full HL-LHC data set. The right figure shows reconstructed mass of an LLP $\tilde{\tau}$ with different timing resolution of MTD [82, 83].

Chapter 4

The CMS detector upgrades for HL-LHC and the High Granularity Calorimeter

As discussed at the end of chapter 2, more collision data is required to further constraint the SM and BSM scenarios. The HL-LHC will provide such opportunity by increasing the collision rate substantially. However, it also brings many challenges for the detectors, such as increased radiation damage and pileup. Therefore, the detectors need to be upgraded in order to maintain adequate physics performance in the harsh run conditions of the HL-LHC. This chapter gives a brief summary of the LHC timeline. It presents a review of various parts of detector upgrades from the CMS point of view. The last section of the chapter describes the HGCAL detector and its components.

4.1 The LHC road-map

The LHC has been designed to explore physics at TeV energy scale during its operation span of about 30 years. The duration of LHC operation has been roughly equally divided into two phases [84], namely Phase-1 and Phase-2. Figure 4.1 shows a road-map of LHC operation. In each phase, the physics data is taken in a series of a long period of time, referred to as *runs*. The runs are interleaved between shutdowns that allow physicists and engineers to access the LHC tunnel and caverns for accelerator and detector upgrades. The Phase-1 was interspersed by two long shutdowns LS-1 and LS-2, each of two years, during which the accelerator instrumentation such as superconducting magnets and collimators are upgraded [65]. This allows the LHC to ramp up its peak luminosity to 2×10^{34} cm^{-2} s^{-1}, more than twice as compared to the beginning of the LHC operation. At the time of writing this thesis, we are at the end of LS-2. The Run-3, spanning three (or four) years, is scheduled to begin in May 2022, and marks the last part of the Phase-1 of LHC operations. At the end of the Phase-1, the CMS and ATLAS experiments are expected to accumulate total integrated luminosity corresponding to more than 300 fb^{-1} at $\sqrt{s} = 13$ TeV.

The Phase-2 [84] will begin with a three years long shutdown, called LS-3. During LS-3, the LHC accelerator ring will receive a major upgrade which will allow its peak instantaneous luminosity to reach as high as 7×10^{34} cm^{-2} s^{-1}. However, the luminosity at the CMS experiment is proposed to be *levelled* (limited) at 5×10^{34} cm^{-2} s^{-1}, and to accumulate ~ 250 fb^{-1} per year for the next ten years. This will mark the beginning of the High-Luminosity LHC (HL-LHC) operation. At the end HL-LHC, the accumulated physics data will correspond to a total integrated luminosity of ~ 3000 fb^{-1}, ten times more than the LHC operation. With such high statistics, it will open up the opportunities [85] for new physics searches and, at the same time, achieve a high level of precision in the Higgs

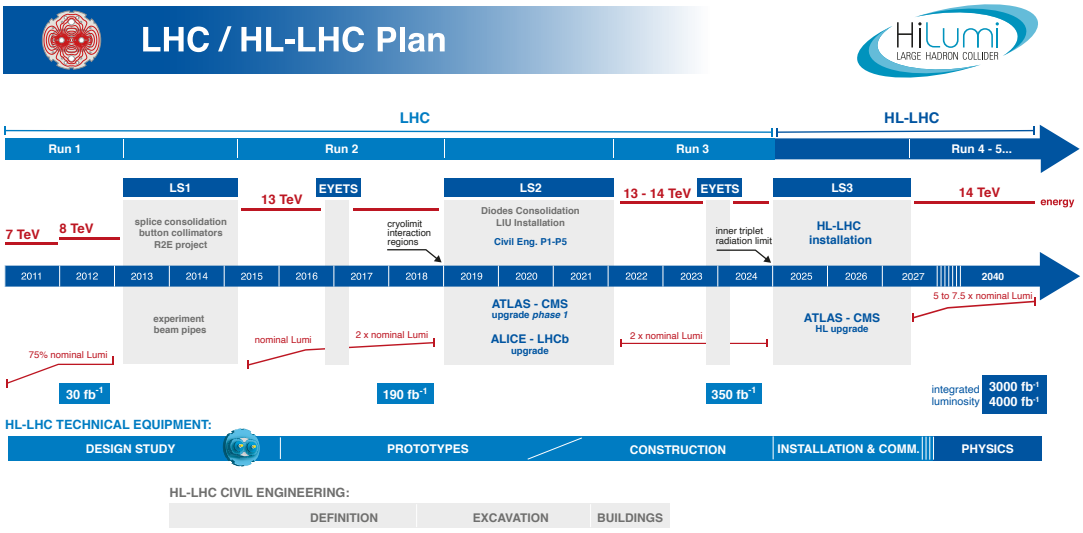

FIGURE 4.1: The timeline for LHC and HL-LHC operations along with physics runs and shutdown schedule.

measurements as discussed in the last chapter.

While the accelerator upgrades bring opportunities for new physics discoveries, it also bring challenges for the detectors such as (a) maintaining optimal physics performance in the hostile radiation environment, (b) effective mitigation of pileup, and (c) maintaining efficiency of trigger with increased detector occupancy. In order to fully utilize collision data delivered by the LHC, the detectors should be able to overcome all of these challenges. The CMS collaboration has planned various improvements and upgrades for Phase-1 and Phase-2 operation of LHC to optimize the detector performance. A summary of detector upgrades is discussed in the following.

4.2 Phase-1 CMS detector upgrades

In the light of accelerator upgrade in Phase-1 operation of LHC, various parts of the detector are replaced or upgraded during LS-1 and LS-2 shutdowns. The major upgrade [65] of the CMS detector during LS-1 and LS-2 comprises the pixel detector, hadron calorimeter, muon system, triggers, and data-acquisition (DAQ) system which are summarized below.

Tracker: The pixel detector and its front-end electronics were designed for an instantaneous luminosity of 1×10^{34} cm^{-2} s^{-1}. The read-out chips (ROC) will, therefore, suffer a heavy data loss during Phase-1 operation because of increased luminosity. The pixel detector upgrade [86] during LS-1 includes an additional fourth pixel layer in the barrel region, improved ROCs, overall low in material budget by minimizing the electronics, cabling, and cooling. Figure 4.2 shows a schematic diagram of the original and upgraded pixel detector. The improved ROCs preserves the data rate with minimum loss of efficiency. The addition of the fourth pixel layer helps improve the pixel-hit-based tracking, thus increasing the efficiency and resolution of the tracks and decreasing the fake rates or track misidentification. It also allows for more accurate primary and secondary vertex identification. The low material budget will minimise the multiple scattering and photon-conversions, thus improving track efficiency and resolution.

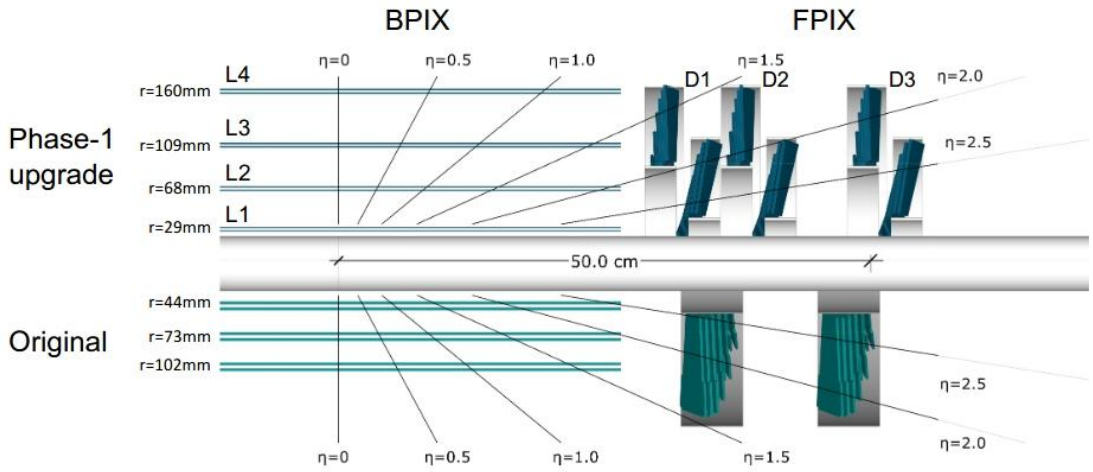

FIGURE 4.2: Schematics of pixel detector in the original design and after Phase-1 upgrade. The *BPIX* and *FPIX* in the image correspond to pixel layers in barrel and endcap region, respectively.

HCAL: The initial choice for photo-detectors in HCAL was Hybrid Photo Diodes (HPDs) because of their radiation hardness, ability to work in high magnetic fields, and cost. However, the performance of HPDs was found to be degrading over time during the operation of physics data-taking. The HPDs have low gain, low signal-to-noise ratio, and they require a high operating voltage ($\sim$ 10 kV). The recent development in the solid-state photo-detectors has made it possible to use Silicon Photo-multipliers (SiPMs) in numerous applications. The SiPMs have many advantages over HPDs, such as high signal-to-noise ratio, compactness, multiple order of magnitude higher gain, and very low operating voltage ($\sim$ 100V). The SiPMs are not affected by the magnetic field and are also cost-effective. Therefore, the HPDs are replaced by SiPMs in the Phase-1 upgrade [87] in the barrel (HB), endcap (HE) and outer (HO) part of HCAL. The HPDs in the HO and HE are replaced with SiPMs during LS-1, and are replaced in HB during LS-2. Also, since SiPMs have significantly larger gain than HPDs, it allows for longitudinal segmentation of HB (from one to three segmentation) and HE (from three to seven segmentation). Hence, the hadronic shower can be probed in a finer detail resulting in improved HCAL performance.

Muon stations: With increased luminosity, the current muon endcap is expected to suffer trigger efficiency loss for muons with transverse momenta (p_T) $<$ 25 GeV after LS-2. To compensate for this loss, a new set of detector based on Gas Electron Multiplier (GEM) technology [88] , are installed, during LS-2, in the endcap in one muon station covering 1.6 $< |\eta| <$ 2.2. The GEM, being very thin and able to withstand flux of high particle multiplicity, is the primary choice for Phase-1 muon upgrade.

Trigger: The occupancy in the detector will increase significantly with the increased instantaneous luminosity, resulting in high background rates and lower efficiency of Level-1 Trigger. At the time of trigger construction, the full granularity of *calorimeter towers*, i.e., ECAL+HCAL energy in 0.087 $\times$ 0.087 in $\eta \times \phi$ direction, could not be exploited because of limited bandwidth and processing technologies [65]. Developments in the data storage and communication technology have made it possible to make use of full ECAL and HCAL trigger tower granularity. Similarly, the information from GEM is integrated to generate triggers for muons with high efficiency in a high particle flux environment. More efficient trigger algorithms, implemented in customised hardware, are being developed [89] to increase the overall trigger efficiency while maintaining the data rate limited to 100 kHz.

DAQ: With increased luminosity, the event rate is expected to go higher; therefore, the DAQ system needs to be fast, reliable and should be able to transfer data over limited bandwidth. Towards this requirement, the DAQ system is being upgraded from older commercial VME technology to new μ-TCA based technology. The μ-TCA technology is already being used in various telecommunication and military applications for fast data transfer over the network.

4.3 Phase-2 CMS detector upgrades

As mentioned earlier, the Phase-2 operation of the LHC is also termed as HL-LHC [21] because of substantial increase in the instantaneous luminosity. It will pose even harsher operating conditions for detectors than Phase-1. The major challenges for the detectors are briefly discussed below.

Radiation damage: A more than two fold increase in instantaneous luminosity compared to Phase-1 will increase particles emanating from the collisions during each bunch crossing. Majority of these particles are pions both charged as well neutral. The charged pions ionize the detector volume while traversing and also produce a cascade of shower particles that further interact with the detector material. The neutral pions decay into a pair of photons, which further produce a cascade of electromagnetic shower particles. The ionization radiation dose received each year by the detector during Phase-2 will be equivalent to the total radiation dose received throughout the entire span of Phase-1 [22]. Figure 4.3 (left) shows the projected absorbed ionization dose (in terms of *Gray* unit) in the CMS detector by the end of HL-LHC operation. In the forward part of the detector, the radiation dose will exceed 1 MGy. The projected flux of neutral and charged particles seen by the CMS detector is shown in Figure 4.3 (right) in terms of 1 MeV neutron fluence per cm^2 in the CMS detector by the end of HL-LHC.

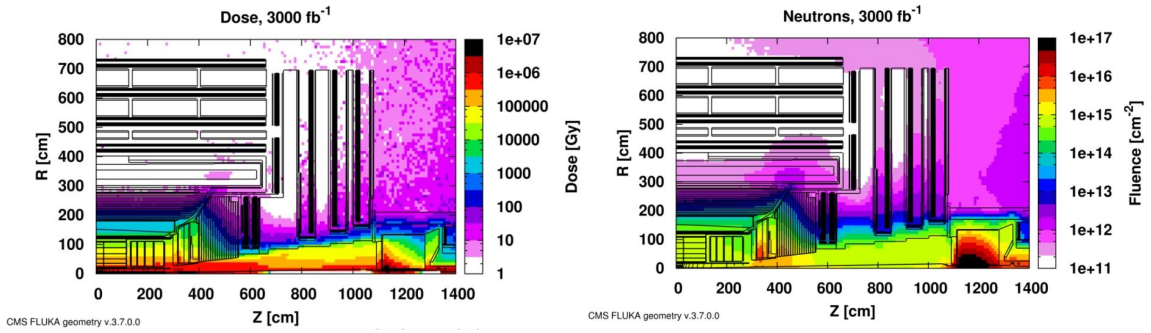

FIGURE 4.3: The projected total absorbed dose (left) and neutron fluence per cm^2 (right) in the CMS detector by the end of HL-LHC operation [23].

Such conditions will cause severe radiation damage to the detectors and on-detector electronics. Constant irradiation to particles results in defects in semiconductor devices [90]. The defects are mainly produced by non-ionizing radiation such as neutrons as they recoil off of silicon atoms and displace them from their sites in the lattice. The displacement of silicon atom creates interstitial and vacancy pair sites, known as *Frenkel pair* defects. It changes the electrical properties of the silicon, e.g., increases leakage current and creates trap centers within the lattice that traps the charge carriers, thereby decreasing the charge collection efficiency. Also, the damage in the semiconductor bulk requires a high reverse

bias voltage to achieve full depletion width resulting in higher power consumption. Eventually, the damage becomes so much that the silicon sensors can no longer produce a desirable signal, and the optimal physics performance is lost.

In scintillators, such as $PbWO_4$ in ECAL and plastic scintillator in HCAL, the radiation damage does not have much effect on the scintillation property. However, the light propagation within the scintillator suffers tremendously [91]. The radiation damage creates absorption bands in the scintillator material that results in *colour center* formation. These colour centers attenuate the light propagating through the scintillator, resulting in total light output yield degradation. Thus, most of the scintillating photons get absorbed, making the scintillator less *transparent*. In more severe cases, the transparency can degrade by more than 90%, as shown in Figure 4.4 for ECAL crystals. The figure shows the loss in crystal transparency in different η regions of ECAL as a function of time. For the pseudo-rapidity $|\eta| > 2.7$, the crystal lost almost all of its transparency by the end of Run-2. Furthermore, the scintillator also suffers from phosphorescence i.e., glowing after irradiation, which causes an increase in noise in the detector.

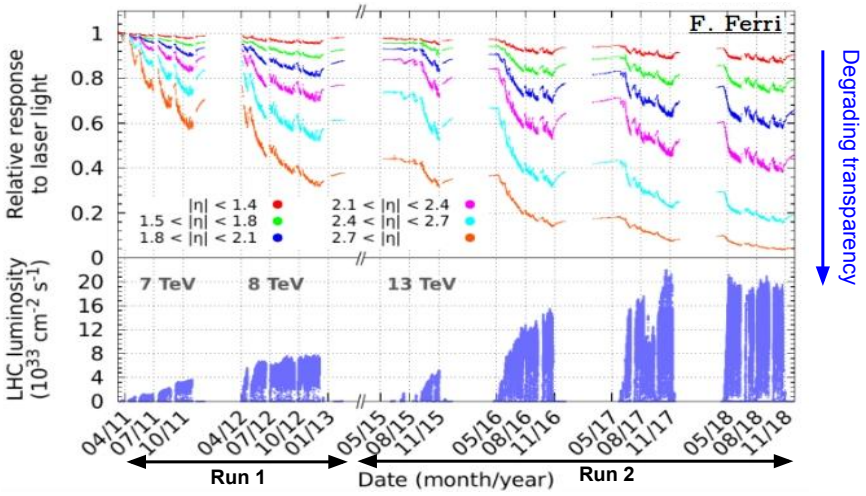

FIGURE 4.4: Figure shows total light output yield of ECAL crystals in different $|\eta|$ regions as a function of LHC data taking [92].

Increased Pile-up: As mentioned in section 3.1, multiple p-p interactions occur during each bunch crossing (BX), called pileup. The average pileup scales with instantaneous luminosity delivered by the LHC as shown in equation 3.5 and Figure 3.3. During HL-LHC operation, the average pileup will be $\sim$ 140 to 200. There are two types of pileup: In-Time (IT) pileup and Out-Of-Time (OOT) pileup. The IT-pileup corresponds to energy deposits in the detector from the current BX. In contrast, OOT-pileup corresponds to the remaining energy left from the previous BX or energy added up from later BX riding on decaying signal from current BX. The OOT pileup is relatively easy to mitigate as their contribution lies away from the *signal time-bucket* of 25 ns in the detector. The majority of the pileup in the event is IT-pileup, in which the particles, although soft, contribute to the additional hits in the tracker as well as deposit energy in the calorimeter along with the particles from hard interaction. It has many undesirable consequences on the event reconstruction and interpretation. For example, Figure 4.5 shows a simulated event display of p-p interactions in HL-LHC condition, and the tracks emanating from it. The additional hits in the tracker increase the complexity with increased combinatorics of hit collection for track reconstruction. This results in loss of tracking efficiency due to *confusion* in the tracking algorithm.

The extra energy deposit in the calorimeter results in mismeasurement of jet energy reconstruction and missing energy E_T^{miss} estimation in the event. Furthermore, it also affects the *isolation*, the activity around the object, of physics objects such as $e^{\pm}/\gamma$ resulting in loss of either object reconstruction or loss of data object quality. The combined effect of these issues results in the degradation of the physics performance of the detector.

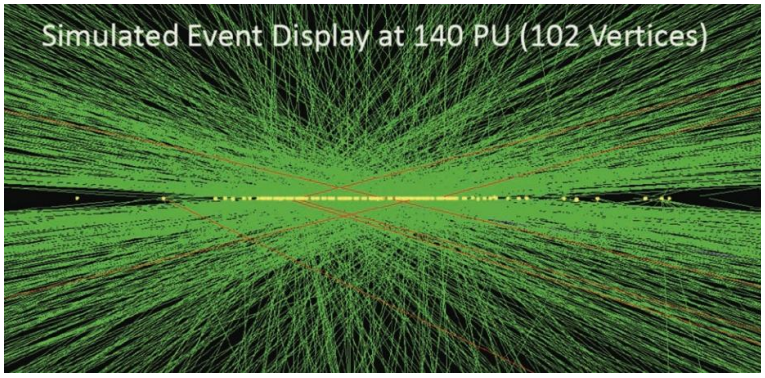

FIGURE 4.5: A simulated view of event display of all the p-p interaction vertices (yellow points) during HL-LHC run. Each green line correspond to soft particles coming out of bunch-crossing [22].

Thus, the Phase-2 upgrade of LHC demands a detector that is radiation tolerant and can perform effective pileup mitigation, an impossible task for the current CMS detector in these conditions. The CMS collaboration is planning a comprehensive upgrade for all of its detector components including tracker, endcap calorimeters, muon system, trigger, and DAQ, in order to maintain optimum physics performance throughout its operation during HL-LHC [22]. A summary of the CMS Phase-2 upgrade program is presented in the following.

Tracker: The Phase-2 upgrade of the CMS revolves around the above two mentioned challenges during HL-LHC, i.e., radiation tolerance and pileup mitigation. By the end of Run-3, the radiation damage in the tracker will be beyond recovery. Therefore, the tracker will be replaced entirely [93]. The *new* tracker will employ silicon sensors of smaller thickness i.e. 100 to 150 μm as opposed to Phase-1 silicon sensors that have a thickness of 300 μm as mentioned in section 3.2.1. The smaller thickness silicon chosen for Phase-2 tracker upgrade because of the fact that the performance degradation of silicon sensors with smaller thickness is smaller as compared to silicon sensors with larger thickness (discussed in more detail in the context of endcap calorimeter in section 4.4.1). Also, studies [94] have shown that silicon with high oxygen concentration requires lower bias voltage to maintain full depletion as compared to low oxygen concentration silicon after being imposed to a very high radiation dose, thus increasing its radiation tolerance. Therefore, oxygen rich silicon sensors will be fabricated for the new tracker. Furthermore, the CMS relies on particle-flow reconstruction for efficient pileup mitigation, therefore, the new tracker will be more granular, about four times more than the Phase-1 tracker. This will help in efficient track separation and accurate reconstruction of interaction vertices, thereby helping in pileup mitigation. Many changes in the tracker design will lead to a low material budget tracker. The tracker coverage will extend till $|\eta| < 4$ and contribute to L1 trigger formation helping in online pileup background rejection.

Endcap calorimeter: The current endcap ECAL and HCAL calorimeters will suffer from

heavy radiation damage by the end of Run-3 and lose their optimal physics performance. Therefore, the current ECAL and HCAL endcap will be replaced by a more radiation-hard detector. There were two competing designs [95] that were initially considered for this replacement:

- **Shashlik design with scintillating crystal for endcap ECAL (EE) and rebuilt of endcap HCAL (HE):** In this design, the EE is proposed to be replaced with a sampling calorimeter that are built in modules. Each module consists of 28 plates of 2.5 mm thick tungsten (W) used as the absorber and 29 plates of 1.5 mm LYSO crystals as the active medium [95, 96]. The transverse size of the front face of the module corresponds to 14×14 mm^2, and the longitudinal length corresponds to 114 mm, about $\sim 22.3\ X_0$. The tungsten is chosen as the absorber because of its small Moliére radius (~ 0.9 cm), resulting in compact showers. This helps in better shower separation and pileup mitigation. In this design, the modules are constructed in *shashlik* configuration and to be read out with the help of WLS fibers. The fibers are coupled with either GaInP photo-sensors or SiPMs of area 3×3 mm^2. The only longitudinal segmentation is proposed to be near the shower maxima. This design is more inline with the Phase-1 ECAL design with $PbWO_4$ crystals. For the HE, the transverse granularity is twice as compared to Phase-1 HE and employs a more radiation tolerant scintillators.

- **High Granularity Calorimeter (HGCAL):** In this design, the EE and HE are replaced by a highly granular sampling calorimeter based on silicon and scintillator+SiPM technologies [23]. The silicon-based detectors, such as the CMS tracker, have demonstrated an excellent performance even after being exposed to a large amount of radiation. Therefore, silicon sensors are employed in the high radiation area in this design. The individual silicon cell has an area of 0.5 to 1.1 cm^2, thus providing a fine transverse segmentation. It uses Pb, CuW, Cu as the absorber in the ECAL section with 28 sampling layers and steel in the HCAL section with 22 sampling layers, offering a fine segmentation in the longitudinal direction. This detector is inspired by the calorimeter design by CALICE collaboration for ILC [97].

The HGCAL is a suitable detector for particle-flow reconstruction because of its high longitudinal and transverse segmentation. The performance of such design has already been established in the beam test experiments conducted by CALICE collaboration [48]. The shashlik-based design, on the other hand, faces many challenges such as technological advancement in radiation hard crystals, WLS fibers and its proposed photo-sensors. Whereas, the CMS tracker group has already been studying and shown that the silicon sensors can retain adequate charge collection efficiency and signal-to-noise ratio even after being subjected to an intense amount of irradiation [93]. After much deliberation within the CMS collaboration, the shashlik-based design has been dropped in favor of the HGCAL design. The details of HGCAL are discussed in section 4.4.

MIP timing detector: The beam spot (p-p interactions in each BX) has a spatial spread of ~ 5 cm (rms) in longitudinal direction which corresponds to a spread of 180 ps in time. Therefore, having precise timing information of each interaction vertex introduces another dimension in pileup mitigation by associating time to each vertex. Therefore, the CMS collaboration has decided to add a MIP timing detector (MTD) between the tracker and the ECAL in both barrel and endcap regions [83]. The MTD is instrumented with LYSO:Ce crystals of size 12×12 mm^2 coupled with 4×4 mm^2 SiPMs in the barrel region, called Barrel Timing Layer (BTL), of MTD. In the endcaps of MTD, called Endcap Timing Layer (ETL), due to radiation tolerance limitations, silicon based Low-Gain-Avalanche-Diodes (LGAD) [98] of size 3 mm^2 are instrumented [99]. The MTD will provide an exquisite time measurement

with a resolution of $\sim$ 30 ps.

Muon systems: The GEM detectors will be fully operational in the first two endcap muon stations as opposed to Phase-1, where only one endcap muon station is instrumented with GEM detector. The later two stations will be complemented with RPCs to provide trigger and positional measurement, while GEMs will provide fast and redundant triggers resulting in efficient muon trigger. The endcap will be extended till $|\eta| \leq 3$.

Trigger: To increase the physics acceptance during HL-LHC operation, the L1 trigger rate will be increased to 750 kHz by upgrading the front-end electronics in all of its sub-detector systems. The trigger latency will be about 12.5 μs (as opposed to 3.4 μs in Phase-1) which will provide sufficient time for online track reconstruction and its inclusion in the decision-making logic along with calorimetric and muon system information.

DAQ: The staggering amount of read-out channels and high trigger rate demands a fast and high bandwidth as well as high computing power to handle large data volume. The projected technological development in the telecommunication sector is expected to bear all the requirements towards the CMS data traffic networks during HL-LHC operation [100].

4.4 The High Granularity Calorimeter (HGCAL)

As discussed in the previous section, the CMS detector will receive high amount of radiation dose during HL-LHC operation, especially in the endcap regions (Figure 4.3). The endcap ECAL and HCAL calorimeters are not designed to operate beyond an integrated luminosity of 500 fb^{-1}. Therefore, it will be replaced with a more radiation tolerant and highly granular calorimeter, the HGCAL [23]. The HGCAL is a sampling calorimeter with unprecedented longitudinal and transverse granularity. Figure 4.6 shows a longitudinal schematic view of the HGCAL detector.

The electromagnetic section of HGCAL (CE-E) has 28 sampling layers with silicon as the active material interspersed between Cu, CuW, Pb absorbers. The CE-E is $\sim$ 34 cm deep and the total depth corresponds to $\sim$ 25 X_0 and $\sim$ 1.3 λ_{int}. The dense absorbers in CE-E with a small Moliére radius ($\sim$ 2.8 cm) ensure the compactness of electromagnetic showers. With its fine longitudinal and transverse segmentation, the CE-E is capable of excellent shower separation, pileup rejection and has a good electromagnetic energy resolution.

The hadronic section of HGCAL (CE-H) has 22 sampling layers. It is divided into two parts: front hadronic section with silicon as the active medium in high radiation region (CE-H-Si) and back hadronic section with scintillator as the active medium in low radiation region (CE-H-Sci). The active layers are interspersed between layers of steel absorbers. The CE-H is $\sim$ 157 cm deep and the total depth corresponds to $\sim$ 8.5 λ_{int}.

4.4.1 Active materials in HGCAL

Due to its radiation hard property, the silicon is chosen as the active material for the bulk of the HGCAL detector where the radiation level is too high. In the relatively low radiation region, scintillator tiles directly readout by SiPM are used as the active material.

The silicon sensors are fabricated in 8-inch hexagonal modules to fully utilize the base silicon wafer. Each sensor module is further divided into smaller hexagonal silicon cells of area either $\sim$ 0.5 cm^2 or $\sim$ 1.1 cm^2. The silicon sensors of three different active thicknesses

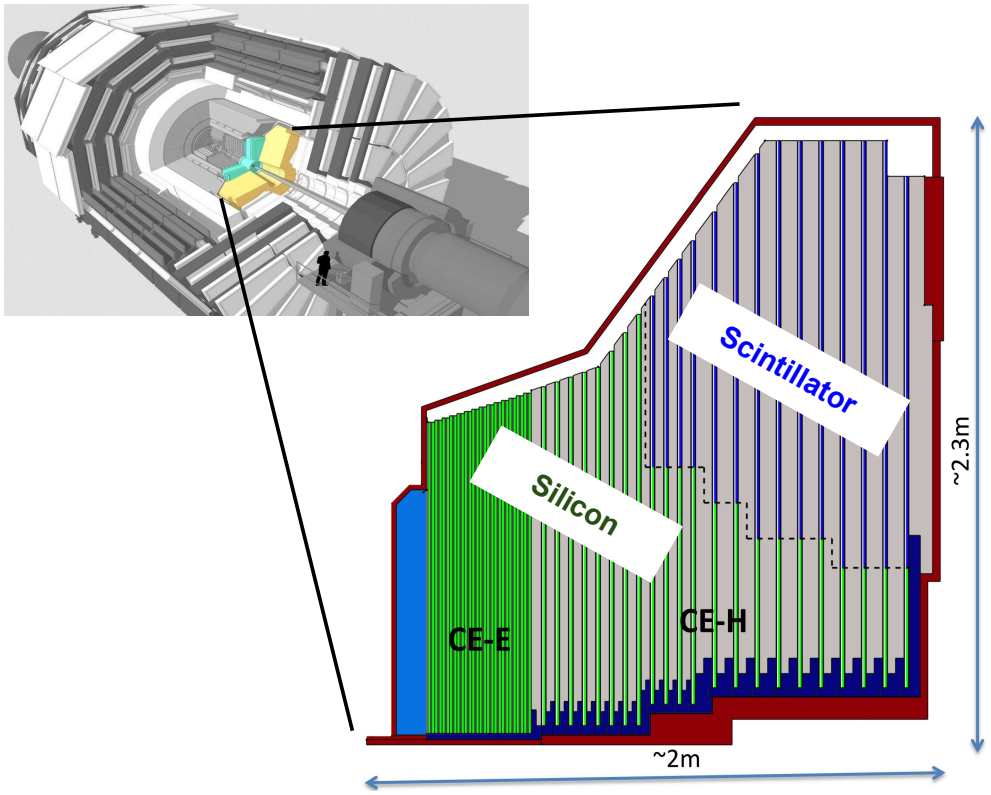

FIGURE 4.6: The longitudinal cross-sectional view of proposed HGCAL calorimeter.

120 μm, 200 μm, and 300 μm will be instrumented at different distances from the beam-pipe. To maintain fine transverse granularity with a good signal-to-noise ratio, silicon sensors of 120 μm active thickness and 0.5 cm^2 area will be instrumented at the inner-most radii (high pseudo-rapidity) of the HGCAL layers where the fluence and radiation dose will be largest. Figure 4.8 (left) shows the schematic of one CE-E disk indicating location of different silicon sensors. The different dimensions of the silicon sensors have been chosen after considering the impact on its electrical properties with different amounts of radiation dose. Studies have shown that with increasing irradiation, the decrease in the charge collection is lesser for thinner silicon sensors as compared to thicker silicon sensors [23]. Figure 4.7 shows the signal produced by silicon sensors of different thicknesses as a function of radiation dose in terms of 1 MeV neutron equivalence per cm^2. It is observed that the charge collection efficiency degrades slowly for silicon sensors of active thickness 120 μm as compared to 200 μm or 300 μm active thickness sensors. Also, by the end of HL-LHC, the bias voltage required for maintaining depletion width is smaller for smaller active thicknesses silicon sensor, thereby consuming lower power. The silicon will be fabricated using crystal growth techniques that introduce high oxygen concentration in the silicon crystal in order to increase its radiation tolerance, as mentioned for phase-2 tracker silicon sensors (section 4.3).

Furthermore, the irradiation studies point to the fact that the radiation dose causes *acceptor-like* defects in the silicon lattice [101]. This makes n-type silicon susceptible to *type-inversion*, i.e. changing from n- to p-type. In contrast, it adds to the majority charge carrier in p-type silicon. Hence, the n-on-p type silicon sensors, i.e., n-type deposition in p-type bulk, show much higher radiation tolerance than p-on-n type silicon sensors [102]. In addition, the p-on-n type sensors show non-Gaussian noise behaviour after being irradiated,

which leads to a worse signal-to-noise ratio. Whereas, n-on-p type sensors do not show such behaviour. Therefore, the CMS HGCAL collaboration has opted to use n-on-p type silicon sensors. The various parameters of the HGCAL silicon sensors are listed in table 4.1. The silicon cells are expected to retain sufficient signal-to-noise ratio (S/N $\sim$ 2.2 as shown in table 4.1) end of HL-LHC lifetime that can discern signal from minimum-ionizing-particles, thus allowing for acceptable level of inter-cell calibration and hence will maintain adequate physics performance.

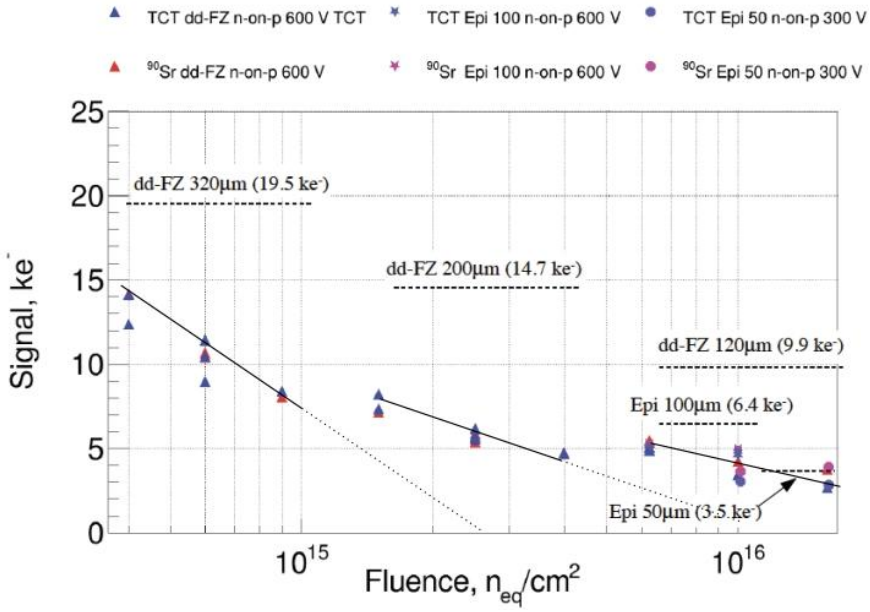

FIGURE 4.7: The signal produced in terms of charge collected for different silicon thicknesses as a function of amount of irradiation in terms of 1 MeV neutron fluence per cm^2. The horizontal lines show the silicon cells of different active thicknesses [23].

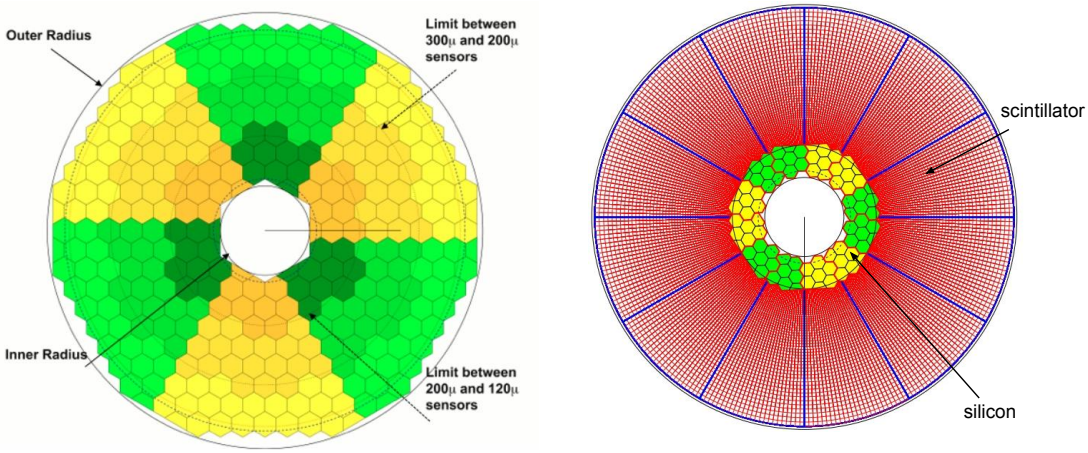

FIGURE 4.8: A schematic diagram of 9^{th} layer of CE-E (left) and 12^{th} layer of CE-H (right). The image also indicates the location of silicon sensors of different thicknesses as well as location of scintillators [23].

The plastic scintillators of approximately square shape with area $\sim$ 5 cm^2 (30 cm^2) and 3 mm thickness will be instrumented in inner (outer) radii of the backside of CE-H disks,

TABLE 4.1: Various parameters of silicon sensors used in HGCAL.

Active thickness (μm)	**300**	**200**	**120**
Area [m^2]	245	181	72
Largest lifetime dose [Mrad]	3	20	100
Largest lifetime fluence [n_{eq}/cm^2]	0.5×10^{15}	2.5×10^{15}	7×10^{15}
Largest outer radius [cm]	$\sim$ 180	$\sim$ 100	$\sim$ 70
Smallest inner radius [cm]	$\sim$ 100	$\sim$ 70	$\sim$ 35
Cell size [cm^2]	1.18	1.18	0.52
Initial S/N for MIP	11	6	4.5
Smallest S/N for MIP after 3000 fb^{-1}	4.7	2.3	2.2

shown in Figure 4.8 (right). The scintillator tiles will be coupled with SiPM photo-detectors integrated on printed-circuit-board (PCB), called *tile-boards* shown in Figure 4.9. It will require about $\sim$ 600 m^2 of silicon in the form of $\sim$ 28000 modules and $\sim$ 500 m^2 of plastic scintillator material in $\sim$ 3800 modules to construct both endcap disks. In order to minimize the electronics noise in silicon sensors and SiPMs, the whole HGCAL will be operated at -30^{o}C with the help of CO_2 cooling system.

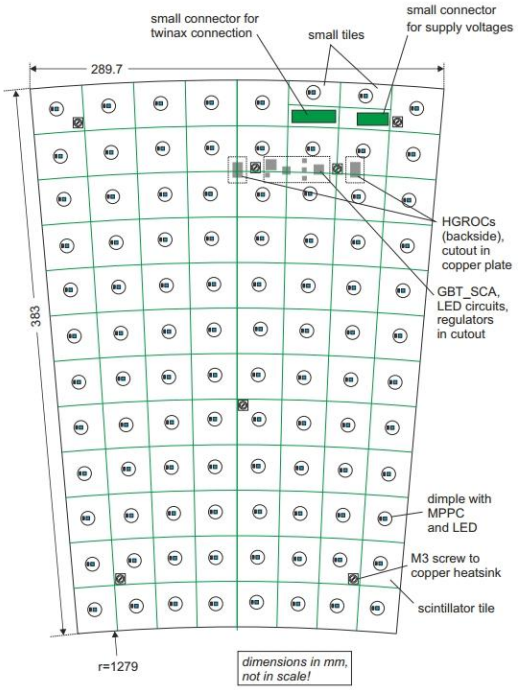

FIGURE 4.9: A schematic diagram of tile board with scitillator tiles with SiPMs, and readout electronics [23].

4.4.2 Front-end electronics and DAQ

The readout chip used in the very front-end (FE) electronics for signal digitization in HGCAL is called ***HGCROC*** [103] . The HGCROC is radiation hard, low noise, low power, with a large dynamic range and robust *application specific integrated circuit* (ASIC) designed specifically for HGCAL operation. Figure 4.10 shows a simplified block diagram of HGCROC ASIC. It offers a preamplification stage, followed by a fast and slow shaper, a 10 bit SAR

analog-to-digital converter (ADC), and a time-to-digital converter (TDC) of 25 ps bin. The analog pulse from the shaper is sampled with a 40 MHz clock and is digitized by the ADC. It provides a buffer of 12.5 μs to store data for L1 trigger decision latency. The HGCROC also provides two timing information: Time-over-Threshold (ToT) and Time-of-Arrival (ToA). The ToT is the measurement of time during which the pulse from the preamplifier stays above a certain threshold, thus making it directly proportional to the pulse height (see Figure 6.3). The ADC value and ToT data are used for energy measurement corresponding to a collected charge of 0.2 fC-150 fC and 100 fC-10 pC, respectively. Thus, the HGCROC can provide energy measurement in a wide range from 1 to 3000 MIP equivalent of energy deposited. Typically, 1 MIP energy deposit corresponds to $\sim$ 3 fC in 300 μm silicon cell. In order to maintain the linearity in energy measurement, intergain calibration is necessary. For this purpose, the HGCROC provides an 11 bit digital-to-analog converter (DAC) with two injection capacitors (0.5 pF and 8 pF) to perform in-situ intergain calibration. The ToA provides precise timing information for collected charges above $>$ 12 fC with a timing resolution of $\sim$ 50 ps. The power consumption of HGCROC is less than 15 mW per channel.

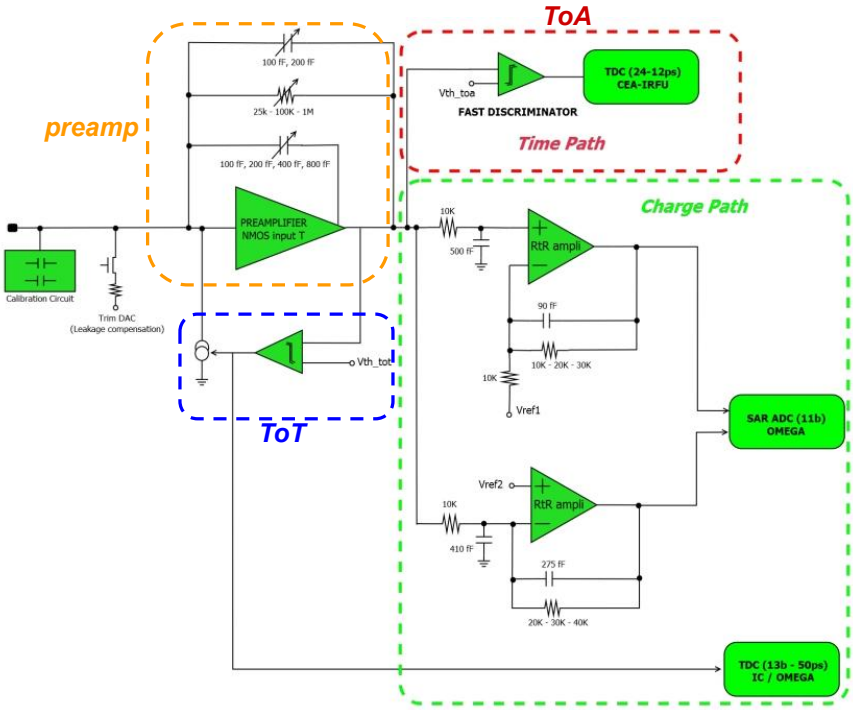

FIGURE 4.10: A simplified block diagram of HGCROC.

The HGCAL will have a staggering amount of readout channels, $\sim$ 6M for silicon and $\sim$ 240k for scintillator. The simulation-based studies show that the occupancy varies from 60% to 1% depending on proximity to the beam-pipe during HL-LHC operation. The data rate in the high occupancy region will be as high as 0.8 Gb/s per HGCROC. The event size of the data produced by all of the HGCAL FE is about $\sim$ 2.5 MB/event and scales linearly with the average pileup. With the L1 trigger rate of 750 kHz, the average data volume generated by HGCAL FE electronics will be $\sim$ 12 Tb/s. The transmission of such a huge data volume in itself is a challenge. The HGCAL DAQ will employ high speed, high bandwidth optical links such as Low Power Gigabit Transceivers (LpGBT) and VRTX+ links in the back end electronics. The data will be transferred to the central DAQ system using DAQ boards

implemented in *advanced telecommunication computing architecture* (ATCA) [100]. Further research and studies are being done to optimize the data compression, path, and bandwidth.

4.4.3 Detector design and mechanics

Each layer of HGCAL is constructed with multiple 60^o wedges, called *cassettes*. Each cassette in CE-E layers comprises silicon sensors modules placed on either side of a 6 mm thick copper cooling plate with onboard electronics as shown in Figure 4.11 (left). Six such cassettes are assembled to make one complete layer of CE-E. The supporting structure of each silicon sensor module, called *baseplate*, is made of CuW. The steel-clad Pb absorber of about $\sim$ 4.9 mm is placed in between each cassette in CE-E. Thus, CE-E has an alternating absorber of Pb and Cu/CuW with an active silicon layer sandwiched between the absorbers. One entire endcap disk of CE-E is assembled by stacking 14 such cassettes and Pb absorber layers.

The layers of CE-H are also constructed in cassettes, as shown in Figure 4.11 (right). The silicon modules and scintillator tile-boards are placed on a 6 mm thick copper cooling plate only on one side, facing the p-p collision point. Multiple cassettes are joined together to make a complete CE-H layer. Stainless steel of thickness 35 mm and 68 mm is used as the absorber for the first 12 and last 12 CE-H layers, respectively. One full endcap of CE-H is assembled by stacking up 22 cassettes and absorber layers. The CE-E and CE-H layers are then assembled to make a full HGCAL endcap disk as shown in Figure 4.12.

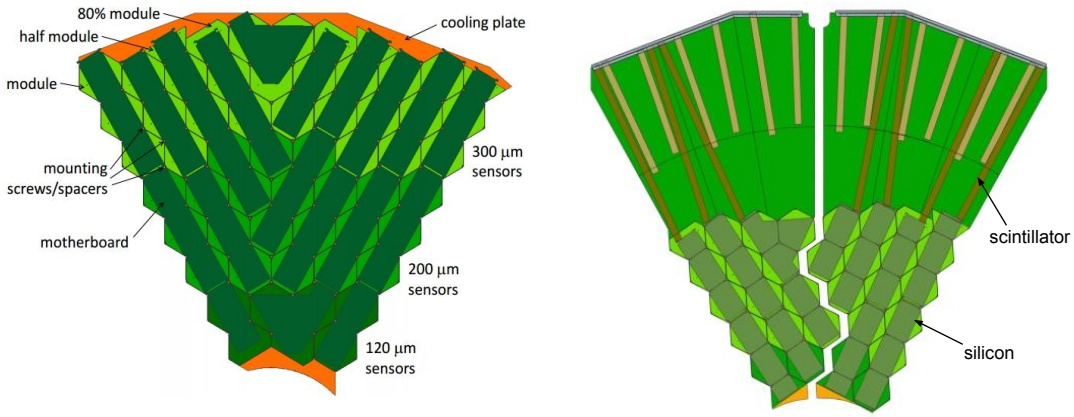

FIGURE 4.11: Schematic diagram of one of the cassette in CE-E (left) and CE-H (right).

4.4.4 Projected performance of HGCAL

The HGCAL will be operated in a dense radiation environment with high pileup conditions. The HGCAL has been designed to do 5D calorimetry, i.e., three-position measurements (x, y, z), energy, and time measurement. Thus, it can produce fine images of shower development, making it an imaging calorimeter and ideal for particle flow reconstruction. Various performance studies [85] has been made using GEANT4 [26] simulation framework. For example, Figure 4.13 shows the radius of circles needed to contain 68% and 90% of electromagnetic shower energy in each layer of CE-E. The study shows that the Moliére

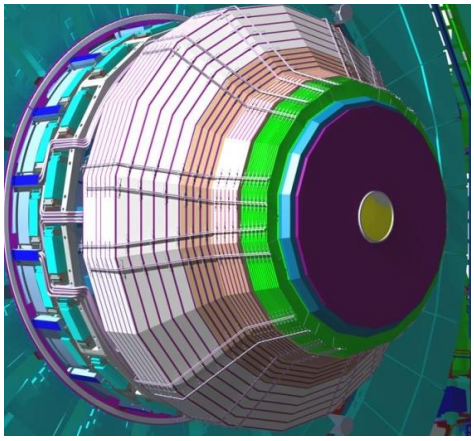

FIGURE 4.12: A computer-aided design of one of the HGCAL endcap wheel ready to be inserted in the CMS cavern.

radius of the CE-E is $\sim$ 2.8 cm, larger than the silicon sensor cells. Thus, it ensures compactness of shower development. Also, the fine granularity helps in excellent shower separation of boosted or merged jets and helps distinguish between pileup jet and the jet originating from hard interaction. For example, Figure 4.14 shows shower development in the radial or transverse direction (radial shower profile) for a jet originating from an energetic quark with transverse momentum (p_T) = 50 GeV and a pileup jet. The pileup jet contains softer particles, generally from multiple PU interactions, compared to the quark jet, thus having a wider radial profile whereas the narrowness of quark jet stands out. The radial profile along with the fine longitudinal sampling has been used to study discriminators to tag VBF and PU jets. Simulation studies have shown that by tuning the discriminant parameters, an efficiency of 80% for PU jet tagging can be obtained against jets originating from vector-boson-fusion processes (described below), at an L1 trigger rate of 10 kHz. More details can be found in [23].

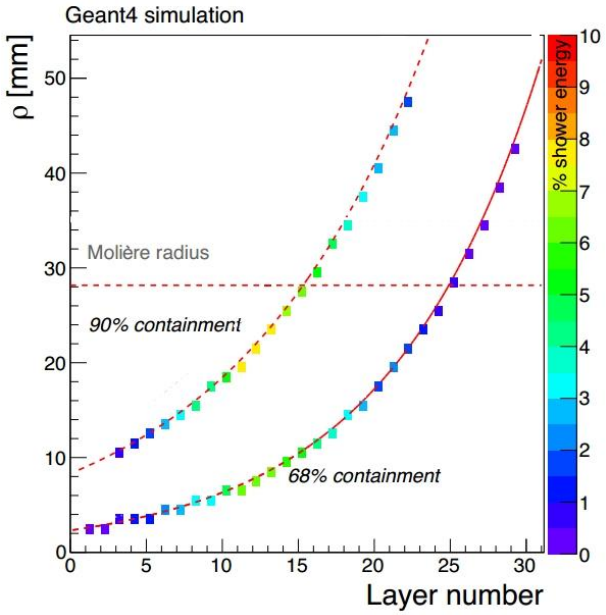

FIGURE 4.13: The plot shows 68% and 90% radial containment of electron shower at each layer of CE-E with increasing radii (ρ).

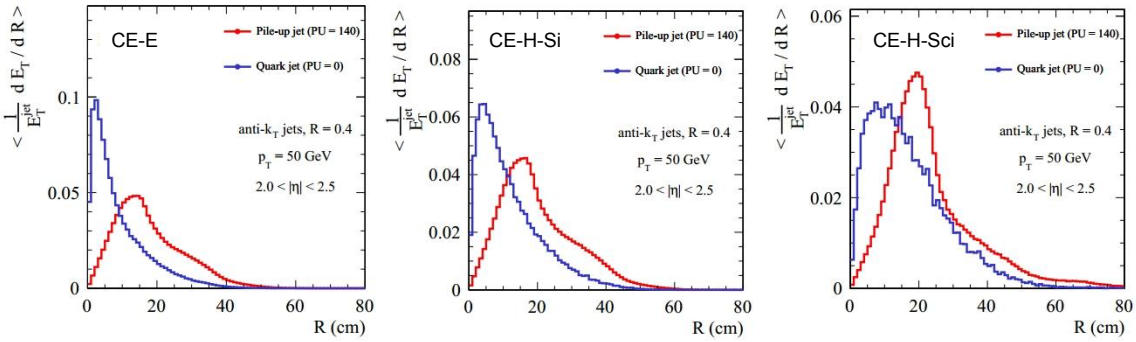

FIGURE 4.14: The radial profile of quark jet of 50 GeV p_T (blue) and pileup jet (red) in CE-E (left), silicon part of CE-H (middle) and scintillator part of CE-H (right) [23].

The excellent jet tagging and measurement in HGCAL are particularly useful for Higgs property studies by triggering Higgs events in vector-boson-fusion processes (VBF). The VBF is the second dominant mode of Higgs production with cross-section σ_{VBF} = 3.8 pb, after gluon-gluon fusion (ggF) with cross-section σ_{ggF} = 48 pb at LHC ($\sqrt{s}$ = 13 TeV). Figure 4.15 shows Feynamn diagrams of Higgs productions via ggF and VBF process. The VBF process can be identified easily with its two forward energetic jets with large gap in pseudorapidity in the event. The HGCAL can trigger on these jets, and its inbuilt precise timing capability will assist in the pileup mitigation of soft jets, and isolate high energy clusters from the VBF jets. For example, Figure 4.16 shows a VBF event (qqH; H $\rightarrow \gamma\gamma$) with one γ and VBF jet incident in the same quadrant. The figure shows event display with and without imposing timing cut. One observes that after applying the timing cut, the event is practically free of the pileup, and the energy clusters from γ and VBF jet stands out.

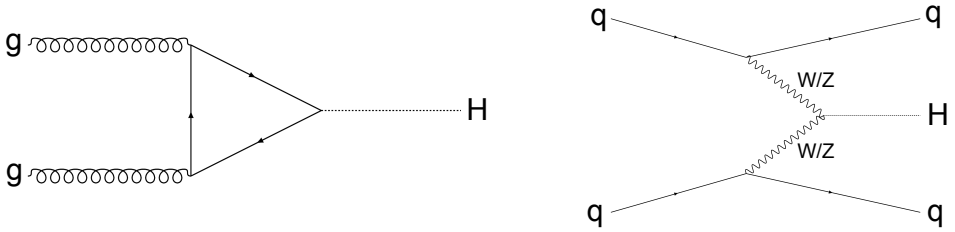

FIGURE 4.15: Feynman diagram of Higgs production via gluon-gluon-fusion process (left) and vector-boson-fusion process (right).

Thus, the HGCAL will maintain adequate physics performance in harsh radiation and pileup environment at the HL-LHC. The construction and operation of such a large scale and complex detector is a technological challenge in itself. With its 6M readout channels and data transmission of more than one TB per second requires ingenuity in the hardware and software development. Also, building the detector and its self-supporting structures, and finally its installation 100 m underground in the CMS cavern is a humongous task for civil engineering. The CMS HGCAL collaboration has laid out a detailed plan for sensor manufacturing, quality assurance, performance studies, commissioning and installation of the detector. The research & development for its design choices and detector units are being carried out across the world. This thesis presents a detailed physics performance studies of the HGCAL detector prototype and its electronics and DAQ components.

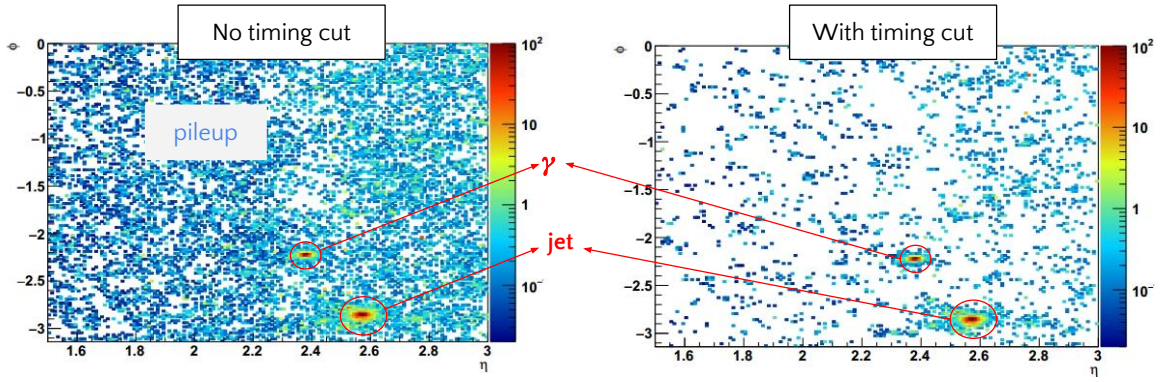

FIGURE 4.16: VBF ($H \to \gamma\gamma$) event with one γ and VBF jet in the same quadrant (shown in red circles). The image is projected onto the front face of HGCAL. The left figure corresponds to without requiring any timing cut and the right figure corresponds to after applying a time window cut of $|\Delta t| < 90$ ps on cells with charge $q > 12$ fC [23].

Chapter 5

Commissioning silicon sensor modules for the CMS HGCAL

As described in chapter 3, silicon will be used as the active material in the high particle fluence and high radiation region in the HGCAL detector. The ionizing particles passing through them will produce electron-hole pairs inducing charge across the silicon which will be collected and digitized by on-board readout electronics. The fabrication of these sensors will be based on 8-inch silicon wafers cut into hexagonal shape to order utilize maximum area of circular wafer as compared to other geometrical shapes. Each wafer is further divided into hexagonal cells of 1 cm^2 or 0.5 cm^2. The bare sensor along with its associated electronics is assembled into a silicon sensor module which is the basic detector unit comprising HGCAL sectors and arranged into layers. A total of approximately 28000 such hexagonal modules, resulting in $\sim$ 6M readout channels, will be used in the CMS endcap calorimeter during HL-LHC.

Given the complexity of the system, with fine electronics elements directly connected to the silicon, the active sensing material, small scale tests are performed for validating and improving the design choices, and benchmarking the various aspects of its technical performance. These investigations are based on test-bench setups in various institutes across the globe or in dedicated beam test experiments at the facilities like CERN, DESY and Fermilab. Based on the lessons learned and experiences gained, each of the final modules will undergo quality-assurance tests to be installed in the final detector. The quality-assurance tests for silicon sensor modules include:

- current-voltage (IV) and capacitance-voltage (CV) characterisation to check electrical properties of silicon sensor
- testing various electronic connections and readout circuitry
- characterize pedestal noise and its stability over time
- charge injection tests for inter-gain calibration
- response of silicon sensors to cosmic muons passing through it using a scintillator based coincidence setup to tag them

For the work presented in this thesis, the silicon modules employed in the beam test experiments performed in 2016 and 2018 are used, which are referred to as v2016 and v2018 respectively. The silicon sensor module prototype v2018 is similar to v2016 in many aspects apart from module assembly details and readout chip which is described in the next chapter. I had performed a detailed system test for v2016 in the laboratory at CERN, and in the IISER Pune (DAQ setup and pedestal noise measurement). In this chapter, system tests for v2016 silicon sensor prototype and the set-up of cosmic test stand in CERN laboratory are described.

5.1 Sensor module prototype v2016

Silicon sensor module version v2016, incorporates dual-PCB (printed-circuit-board) based design, i.e. the silicon sensors and the read-out electronics are constructed as separate boards and then connected together to make a full module as shown in Figure 5.1. The module prototype is a combination of various layers that are glued together as shown in Figure 5.2 (left). The bottom-most layer is the baseplate, made up of CuW (W:75%, Cu:25%) material to provide mechanical support and also acts as a heat transfer interface between the module and the cooling plate. A gold-plated Kapton^{TM} layer is glued on top of the baseplate to provide high-voltage (HV) bias contact to the silicon cells. The silicon layer is carefully glued on top of the gold-plated Kapton^{TM} layer. A PCB is glued on the silicon layer that connects with the silicon cells via wire bonds through the holes (visible in Figure 5.1 left). All the connections are routed to two connectors on the top of this first PCB. Another PCB that holds all the readout electronics, called "readout PCB" Figure 5.1 (right), is connected to the silicon sensors via these two connectors on the first PCB.

5.1.1 Silicon sensor

The sensors used for the prototype is p-on-n type silicon diodes operated in reverse-bias mode. The silicon sensor module, along with its associated electronics, is structured in the shape of hexagon, and the sensor itself is divided into small hexagonal cells. The cells at the edge of the silicon wafer, however, could not be made a complete hexagons therefore these cells are of different area. The size of a full hexagonal silicon cell is $\sim$ 1.1 cm^2 as shown in Figure 5.2 (right). The cells which are at the edge of modules are merged together to have a comparatively larger surface area such as merged *mousebites*. There are a total of 134 silicon cells which can be read out independently. The total thickness of the silicon sensor is 320 μm with the depletion (active) width of $\sim$ 200 μm. In 200 μm depletion width, a minimum-ionizing-particle will deposit about 57 keV energy [36], thus creating $\sim$ 16000 electron-hole (e-h) pairs in the bulk [1]. These e-h pairs are swept away under the influence of electric field induced by the reverse bias voltage and are collected at the electrodes.

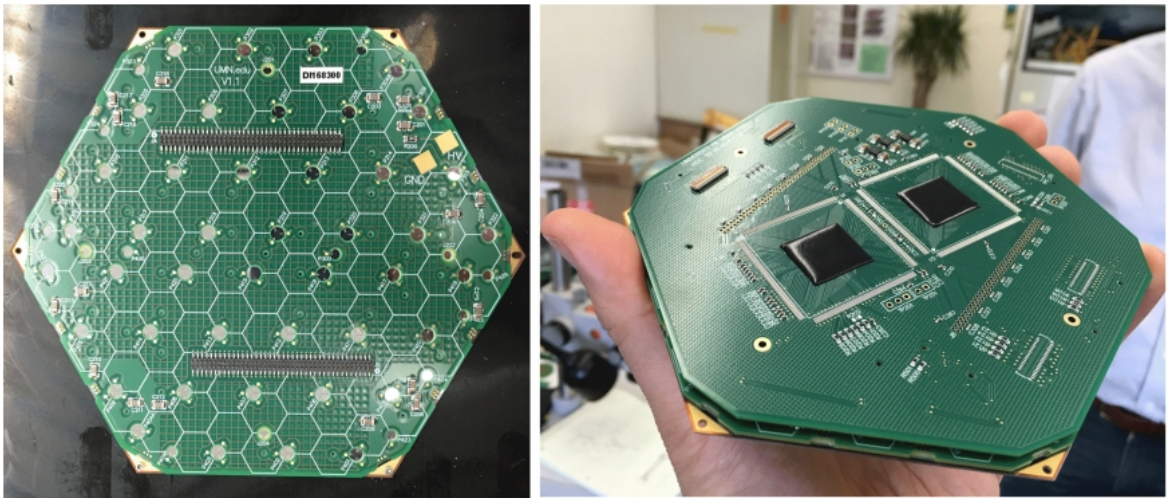

FIGURE 5.1: The dual-PCB structure of sensor module version v2016. Left image shows bottom-PCB that has silicon sensor layer and two connectors visible on top. Right image shows top-PCB with two SKIROC2 readout ASICs

[1] Mean energy required to create e-h pairs is 3.62 eV for silicon sensors at 300 K.

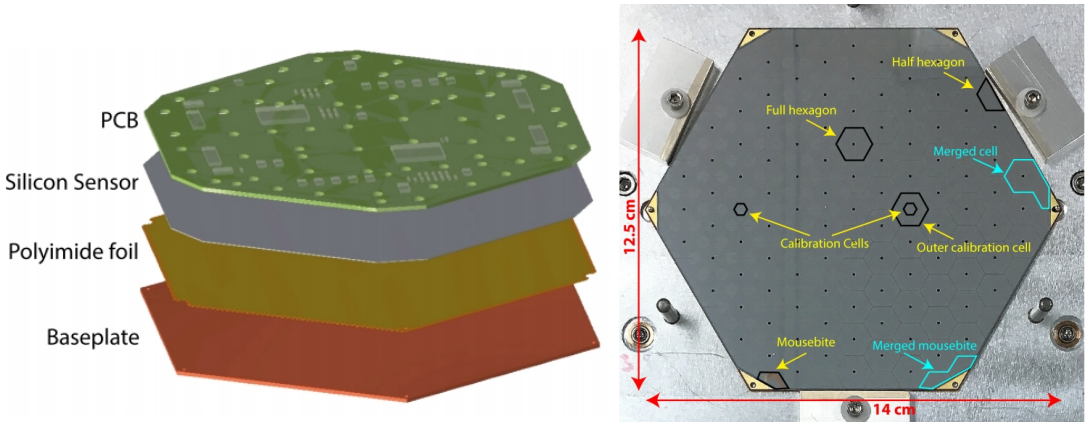

FIGURE 5.2: Left image shows different layers of hexaboard module which are glued together to form the silicon sensor prototype. Image on the right shows silicon sensor layer with cells of different sizes marked on the layer.

5.1.2 Readout PCB

The top PCB houses two SKIROC2 read-out ASICs visible in the Figure 5.1 (right) as two black squares [104]. The SKIROC2 ASIC was originally designed for CALICE collaboration [97] and was adopted for the CMS digitization requirement of 40 MHz corresponding to the p-p bunch crossing every 25 ns at the LHC in contrast to the expected 200 ms at ILC. Each SKIROC2 chip provides various functionalities such as a pre-amplification stage, pulse-shaper with two different gain setting to cover a large dynamic range of energy deposit, 15-deep switched capacitor array (*SCA*) to integrate charge for each 25 ns time sample and a 12-bit analogue-to-digital converter (*ADC*). Figure 5.3 shows the block-diagram of SKIROC2 ASIC. The signal produced by a traversing particle first passes through the pre-amplification stage. The preampliefier in SKIROC2 is a charge-sensitive low-noise amplifier. The amplification gain can be changed by adjusting its feedback capacitor (C_f) value ranging from 400 fF to 3.2 pF via slow-control parameters. The amplified signal is then shaped with CR-RC based two pulse shaping network with a time-constant of 180 ns. The two different pulse shapers have different gain that differs by a factor of 10, namely low gain and high gain, and provide a large dynamic range for energy measurement. The shaped pulse from the two shapers are sampled in the time window of 25 ns with the help of 13 switched capacitor arrays. These 13 time samples are then digitized by a 12-bit Wilkinson ADC.

Each SKIROC2 ASIC has 64 channels that can read and digitize the analog signals from silicon sensor and consumes 25 μW per channel. Two SKIROC2 ASICs are connected to 128 out of 134 silicon cells. The remaining six silicon cells are not connected and are left open. The top and the bottom PCB together make up one silicon hexaboard module prototype as shown in Figure 5.1.

5.2 Data acquisition system for silicon sensor module

In order to read and store the data from the silicon sensors, a full data-acquisition (DAQ) chain is employed. The DAQ system has been developed by engineers and physicists in the HGCAL group. It has been been designed in a way such that it can be easily scaled-up to acquire data from multiple sensor modules [105]. Figure 5.4 shows various components of the DAQ chain. The core of the DAQ system is the ZedBoard which is a commercially available development board with ZynQ^{TM} SoC (System-on-Chip) device. The ZynQ^{TM}

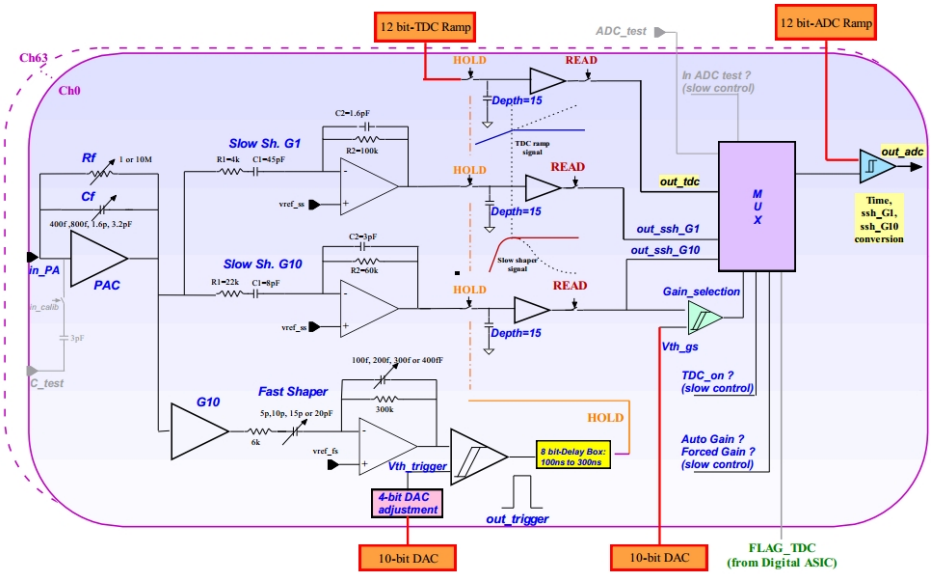

FIGURE 5.3: Block diagram of SKIROC2 ASIC channel.

device integrates an Xilinix Artix-7^{TM} FPGA for programmable logic (PL) and a dual core ARM Cortex-9^{TM} CPU for processing system (PS) on a single chip. The ZedBoard also provides 512 MB DDR3 RAM on-board memory, a 4 GB SD card, Ethernet, FMC interface etc. In essence ZedBoard is a fully functional Linux system that can communicate with the programmable logic sector of ZynQ^{TM} chip.

Such rich functionalities of ZedBoard facilitates to design software and firmware that can be used to issue triggers, send control signals and communicate with SKIROC2 ASICs. The ZedBoard is connected to a custom-designed card, called ZED-IO, with 14 HDMI ports via FMC interface. The ZED-IO board broadcasts the signals from ZedBoard to all 14 HDMI ports and funnels the incoming signals from HDMI to ZedBoard. The ZED-IO board is connected to a custom-designed card, called Dual-Daughter-Carrier (DDC) card, which houses two control FPGAs, called FMC-IO. The DDC card is a simple PCB with the only job of hosting FMC-IOs and routing the signals from ZED-IO board. Each FMC-IO board has Xilinix Artix^{TM} based FPGA and can communicate with up-to two SKIROC2 ASICs (i.e. one sensor module) via another board, called elbow board. The elbow board provides a communication path between readout chips and FMC-IO, as well as provides the high-voltage bias connection between sensor module and the HV power supply. Thus, with this DAQ system one can acquire data from up to (14 HDMI $\times$ 2 FMC-IOs $\times$ 2 SKIROC2 $\times$ 64 channels =) 3584 channels of silicon sensor modules.

The software on the ZedBoard contains instructions for PS as well as a firmware for PL. One of the major tasks of the ZedBoard is to issue a *trigger* to take the readout from silicon sensor module. In experimental particle physics, a *trigger* is defined as the signal that is used to commence the digitization and transfer the signal of interest induced in the detector for permanent storage. The software on PS performs following tasks: it detects the number of FMC-IOs connected via ZED-IO and their addresses, it provides software generated triggers and can toggle to external trigger. It sets the run number and the number of events to be recorded and finally initiates data taking. The communication between PS and FMC-IOs occur via PL on the ZedBoard. The firmware on the ZedBoard facilitates the communication between PS and FMC-IOs via FMC interface on ZedBoard and it also ensures that ZedBoard talks to one FMC-IO at a time by implementing multiplexer on the FPGA.

The addresses of each FMC-IOs is set via DIP switches present on the DDC card. The ZedBoard sends all the data with this unique address on its header so that it can talk to only one FMC-IO at a time. When ZedBoard issues a trigger, FMC-IO reads the data from SKIROC2 ASICs and stores on its on-board buffer memory. During this process, all the further triggers are rejected by ZedBoard until all the FMC-IOs complete their read-out cycle. The ZedBoard then gets ready to receive another trigger. To reduce the time during which ZedBoard remains in "*trigger veto*" mode, the data is temporarily buffered on the FMC-IO FPGA memory. Thus, data is not transferred at each trigger, rather for about 150 events the data is stored on the FMC-IO, and it is then transferred to ZedBoard for permanent storage. The reason to choose 150 event buffer is based on the fact that the DAQ system was designed for beam test experiment at CERN where the beam is delivered every 30 seconds and lasts for about 5 seconds called a "spill". In each spill, a train of about 150 single particles are delivered to the experimental hall area that impinge the detector one-by-one sequentially. Therefore, 150 events are kept in the buffer memory and transferred between two spills for permanent storage such that there is no loss of events.

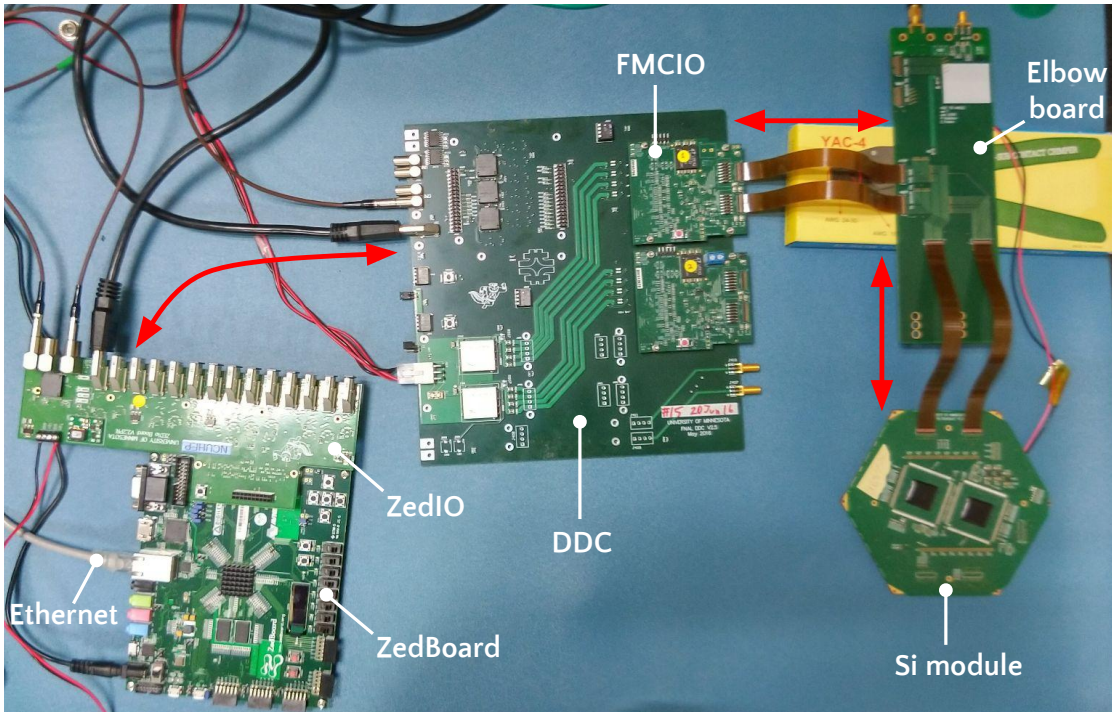

FIGURE 5.4: Data-acquisition chain with all components for sensor module version v2016. The red arrows indicate the flow of data.

Once a trigger is provided to SKIROC2 chips, which can be either internally generated through ZedBoard or externally generated through, say coincidence logic unit, it transfers the digitized signal which is referred to as "*raw data*" for the event. The raw data contains a header and ADC counts for each channel for two gain settings. The header has information of time-stamp, trigger number and event number. The ADC data is packed in a 32-bit word which are "Gray encoded" [106] and is stored in two columns (two-dimensional array) corresponding to two SKIROC2 chips. The reason to use Gray encoding is because in this scheme two consecutive numbers differ only by one bit as opposed to standard binary encoding in which more than one bit can differ between two consecutive numbers. Therefore, Gray encoded digital data is more immune to error for analog to digital conversion as well as in data transmission. More details about Gray codes can be in [106]. Each 32-bit word encodes ADC data from two channels for a given gain setting and a given SKIROC2 chip.

The ADC data for a particular gain channel occupies 16-bit memory in the data-frame. The data is stored from the high address (i.e. starting from channel 63) and low gain setting first. The first 32 rows of the raw data corresponds to low gain channels and next 32 rows correspond to high gain channels. For example: element corresponding to first row, first column has ADC counts from low gain channel 63 in the first 16-bits, the Most Significant Bits (*MSB*), and channel 62 in the second 16-bits, the Least Significant Bits (*LSB*) of SKIROC2 chip 1. Similarly, first row, second column has ADC counts from low gain channel 63 in the first 16-bits (MSB) and channel 62 in the second 16-bits (LSB) of SKIROC2 chip 2, and so on.

The raw data is first saved onto the ZedBoard memory temporarily and is then manually transferred to hard-drive for permanent storage for further analysis. The unpacking of the raw data is done offline in the CMSSW framework on an event-by-event basis. The raw data is first read from the text file where header is separated out and is used for timing and event indexing. For each event, 128 elements (from 64 rows and two columns) are read out. Each 32-bit element is then unpacked into two 16-bit values corresponding to ADC data from each gain channel. With the help of *electronic map* of the module, each 16 ADC data is mapped to corresponding channel and gain settings. The electronic map contains the location of each cell center with the IDs of corresponding chip and channel. Then the ADC data is decoded from Gray code to binary or decimal and is stored for each event, called *DIGI* format. Thus, for each event, the ADC counts from high and low gain channels for each silicon cell are retrieved.

5.3 IV characterization and pedestal noise

The IV characterization of the silicon sensor module is the first and important step in the system tests. The silicon sensor module is a collection of p-n junction diodes which operates in the reverse bias mode. An ideal p-n junction diode acts as a non-conducting element in the reverse bias mode, and hence not expected to conduct any current. However, in practice there is a small current across the diode, called *leakage current*, because of the thermal fluctuations in the depletion region that creates electron-hole pair at room temperature. If there are defects in the semiconductor lattice, such as trap centers then it increases the electron-hole pair generation due to thermal fluctuation and thus increases the leakage current and noise in the sensor. The IV characterization gives us valuable information about the electrical properties of the sensing bulk of the module. Furthermore, it helps us decide the operating voltage at which the sensor is almost fully depleted and the voltage at which the silicon enters into avalanche breakdown [2]. Similarly, CV characterization is also an important measurement to estimate the capacitance of the detector. I had performed such electrical tests, i.e. IV characterization and pedestal noise measurement of silicon sensor modules v2016 in the laboratory at CERN (and also at IISER HEP lab), and has been described below.

To obtain IV characterization of the module, Keithley source-meter 2410^{TM} is used which can go up to 1.1 kV and can measure the current as low as 1 fA. The measurements are taken at the room temperature. The connection between the source-meter and the elbow board is made using SMA cable as shown in the Figure 5.5. The bias voltage can be easily controlled from the front panel of the source-meter and its display shows the current drawn by the module. The module is housed in a wooden box and covered by a black cloth to reduce the current induced in the silicon cells by ambient light. The IV curve of multiple modules are recorded. Figure 5.6 shows reverse current (in μA) as a function of input bias voltage

[2]Increasing reverse bias voltage (V_{bias}) increases the electric-field across the diode. At a certain V_{bias}, electric field is so intense that a single electron can knockout electron for Si atom which further knocks electron out from another Si atom, resulting in a avalanche and creates a large reverse currents.

(in volts) for two example sensor modules. The module corresponding to Figure 5.6 (left) shows that reverse current first increases as a function of bias voltage and then settles at a plateau of $I_{reverse} \sim 14\ \mu A$ for $V_{bias} > 40$ volts. The plateau region corresponds to the fully depleted silicon cells. For such modules, the typical operating bias voltage was set to V_{bias} = 120 V. However, IV curve for another module, Figure 5.6 (right), shows that reverse current increases with increasing bias voltage and never settles to any plateau. The reverse current is also considerably larger as compared to the other module (approximately three times at V_{bias} = 200 V) as shown in Figure 5.6 (right). Given the high leakage current, the faulty module is not suitable for further measurements, and it is discarded. This testing procedure is performed for all the modules before these could be used for further investigations or be exposed to particle beams.

FIGURE 5.5: Setup for IV characterization. Image on the left shows front panel of Keithley Source-meter 2410 (instrument on top). Right image shows the module covered with a wooden box and black cloth.

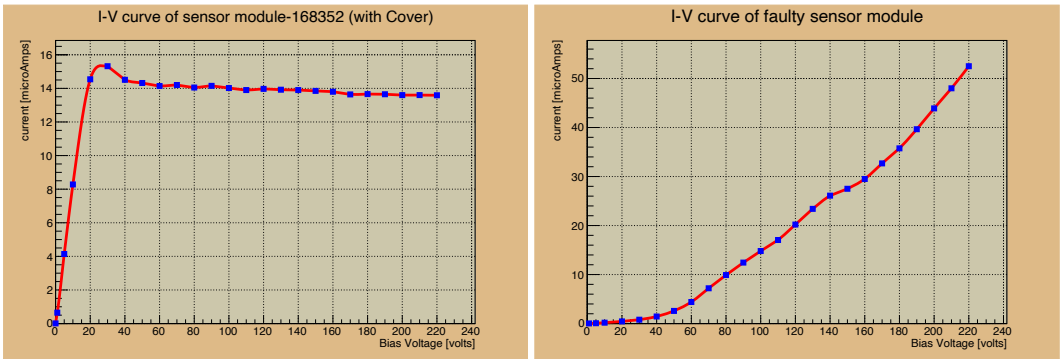

FIGURE 5.6: IV curve for two example silicon sensor modules, the points are connected by a smooth curve. Plot on the left shows IV curve of a working module and on the right shows a faulty sensor module.

After determining the IV characteristics, pedestal noise level is measured. The pedestal noise measurement requires reading the state of silicon sensors in the absence of any traversing ionizing particle. For this purpose, the silicon sensor module is connected to the *DAQ* chain and the SKIROC2 ASICs are configured appropriately. The trigger signal is supplied via a function generator as shown in Figure 5.7, to record the events randomly. The image on the top-left of Figure 5.7 shows the front panel of the function generator which is set to generate a square pulse of frequency 10 Hz with peak-to-peak voltage set to V_{pp} = 4.0 V, similar to TTL pulse as accepted by the DAQ. The function generator pulse is fed to ZEDIO board via LEMO cables and from ZEDIO to DDC board which finally passes the trigger to SKIROC2 chips. Around 6000 events are recorded using this setup following the same DAQ chain as discussed in section 5.2.

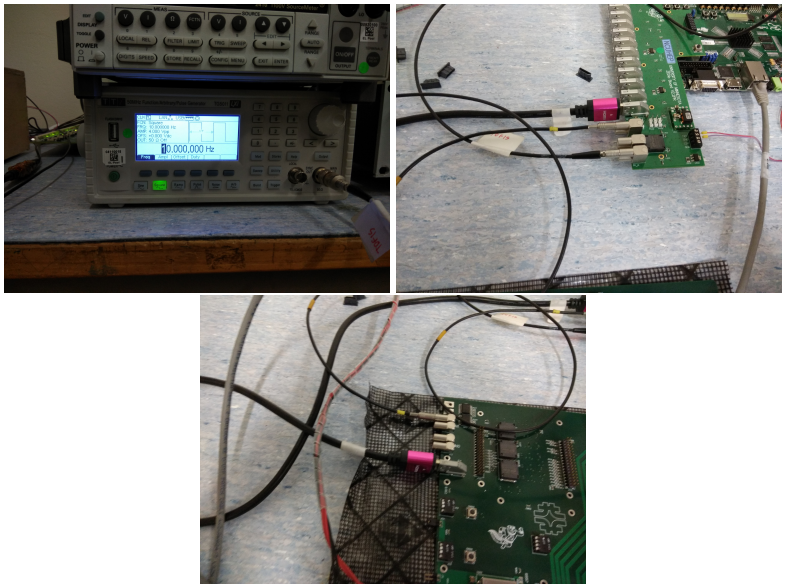

FIGURE 5.7: The setup for external triggering to v2016 DAQ setup. Top-left figure shows the front panel of function generator set to generate a square pulse 10 Hz with V_{pp} = 4.0 volts. Top right and bottom figure shows the LEMO cable connection from function generator to ZEDIO and ZEDIO to DDC board respectively.

Figure 5.8 shows distributions of ADC counts for two representative silicon cells. The distributions are close to Gaussian with non-zero value of mean and width. The non-zero mean in the ADC distribution is an artefact of read-out electronics that introduces a DC baseline in the signal, called *mean pedestal*. The fluctuation around this DC baseline, called *pedestal noise*, is because of thermal noise in the silicon as well as picked up noise from the stray signals e.g. power sources. The ADC distributions of all silicon cells exhibit similar behaviour. The ”mean pedestal value” is referred to as ”pedestal” and ”pedestal noise” as noise in the following. The ADC distributions are fitted with a Gaussian function as shown in Figure 5.8. The mean and width of the fitted Gaussian function are extracted as pedestal and noise, respectively. Figure 5.9 (left) shows the pedestal and Figure 5.9 (right) shows noise for all the channels. It is observed that the average pedestal value is $\sim$ 270 high gain ADC counts and average noise value is about $\sim$ 15 high gain ADC counts and it is fairly stable for all channels of the given module.

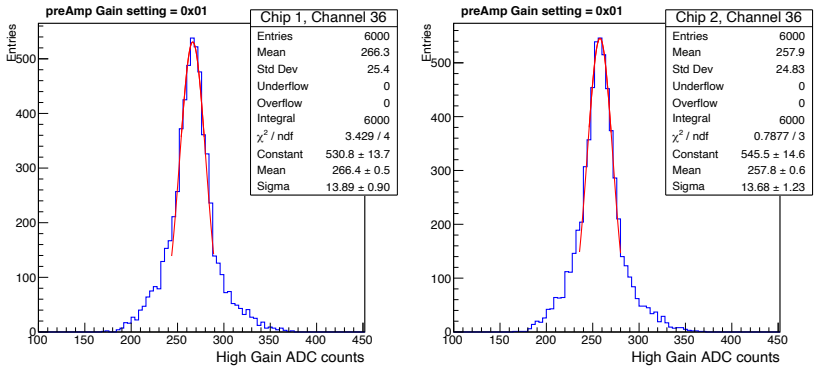

FIGURE 5.8: High gain ADC distribution of two example silicon cells in the absence of ionizing particle.

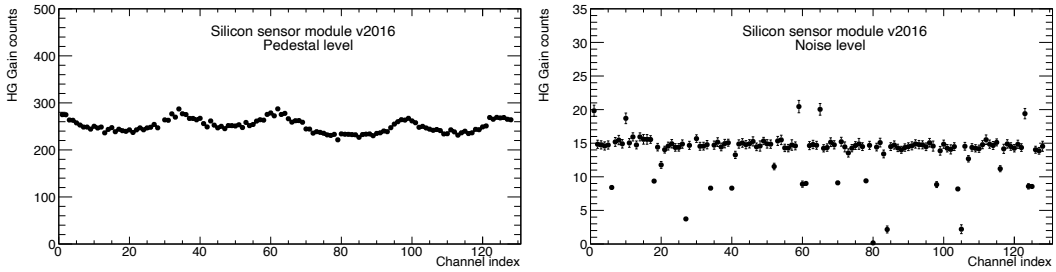

FIGURE 5.9: Left figure shows DC shift (pedestal) in the high gain ADC data, and the right plot shows fluctuation (noise) about the DC baseline, for all silicon channels

The pedestal values are subtracted from all the channels to bring the DC baseline to zero. Also, while the noise originating from the thermal fluctuation in silicon sensors, the *intrinsic noise*, can not be removed but the noise contribution from the external sources such as power source or ambient electromagnetic interference can be estimated, called *common-mode noise*, and subtracted from the ADC data on an event-by-event basis. The method for pedestal and common-mode noise estimation is discussed in section 6.5.1 for v2018 module and the procedure is similar for v2016. Figure 5.10 shows pedestal and noise after subtracting DC baseline and common-mode noise. As expected, the pedestal is brought to zero, and the intrinsic noise is $\sim$ 2.2 high gain ADC counts.

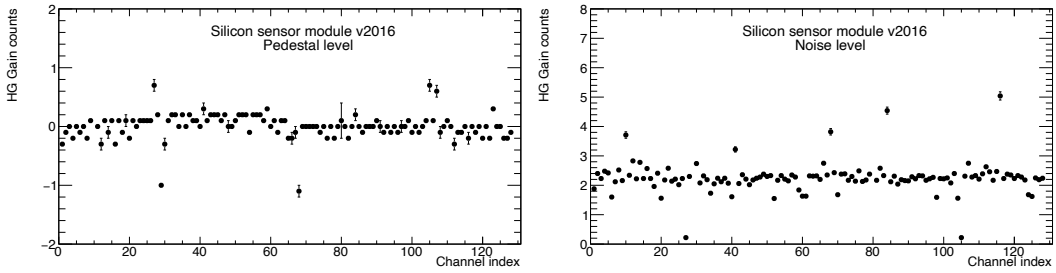

FIGURE 5.10: Pedestal and common-mode noise subtracted DC level (left) and pedestal noise (right) for all silicon channels

At the time of taking the measurements, there was no provision available for CV characterization and charge-injection for v2016, therefore it is not presented in this thesis. However, system tests performed earlier have shown very small and stable capacitance of module v2016 [107]. Method of using charge-injection to perform inter-gain calibration makes an important step for physics performance studies, and is discussed in section 6.5.1 for v2018 module used in October 2018 beam test experiments.

5.4 Measuring cosmic muon signals using a cosmic stand

After assuring the general status of the silicon sensor module with the IV characterization and stability of pedestal noise, the next step is to check if the module can discern the signal, produced by ionizing particles traversing it, from the noise. High energy muons rarely initiate shower in the detector, and hence act as minimum ionizing particles or MIPs (section 2.2.1). Therefore, the signal produced by muons is the smallest signal that the sensor can produce corresponding to a real particle and is used as the benchmark for determining its signal-to-noise ratio. Also, high energy muons are crucial tools for the SM measurements as well as BSM searches. Ideally, this is achieved by subjecting the sensors to the beam of

muons of known energy but in general, in the laboratories we do not have access to the particle beams.

Fortunately, this can be achieved with the help of naturally occurring cosmic muons. These muons are the by-product of ultra-relativistic cosmic particles, mostly protons originating from different sources such as supernova, active galactic nuclei etc, interacting with the upper atmosphere of the earth. The cosmic particles interact with the molecules of air and create a shower of particles such as charged pions ($\pi^{\pm}$), kaons etc., as shown in Figure 5.11. This phenomena was first observed by Victor Hess that won him the Nobel prize of year 1936. The shower particles decay into lighter particles. The pions, for example, decay into muons. These interactions and decay usually occurs tens of kilometers above the sea-level. However, the muons being highly boosted (highly relativistic) reach the earth surface before decaying into electron and two neutrinos. The average energy of these muons are of the order of $\sim$ 4 GeV [36]. Measurements have shown that the average flux of cosmic muons in vertical direction is about $\sim$ 1 muon per cm^2 per minute at sea-level.

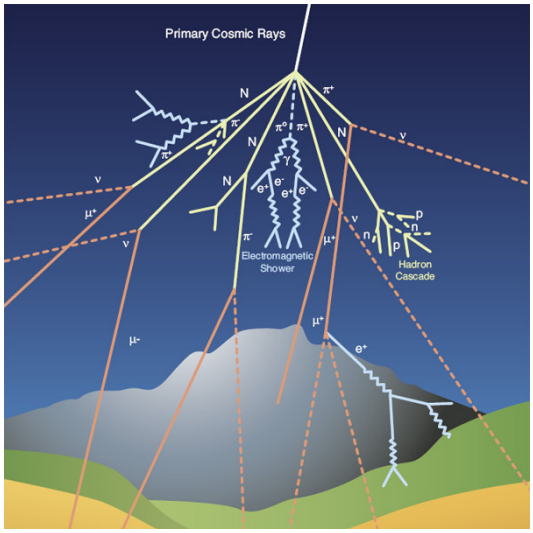

FIGURE 5.11: A representation of cosmic ray shower where primary cosmic ray (mostly protons) interact with earth's atmosphere and produce a cascade of shower particles some of which reach the earth surface.

The cosmic muons can pass through the materials, such as meters of concrete wall, before completely stopping by losing their energy via ionization process. In order to obtain the spectrum of energy deposited by cosmic muons (in terms of ADC counts), it is important to trigger the data-acquisition when the muon passes through the silicon cell in a timely manner. The slow shaper in the SKIROC2 ASIC takes about $\sim$ 180 ns to peak corresponding to charge induced by a MIP in the silicon cell [104]. The peak is rather broad and a reasonable signal can be obtained in the time window of $\pm$25 ns ($\pm$ 1 time-sample). Too early or too late triggering might result in the loss of signal.

The trigger is generated with the help of supplying a two-fold coincidence signal to the SKIROC2 ASICs using scintillator tiles coupled to photomultiplier-tube (*PMT*). The two-fold coincidence reduces the probability of trigger generation by the spurious noise in a single scintillator+PMT setup and at the same time ensures that the cosmic muon is approximately normal to the detector plane. Figure 5.12 shows the cosmic muon stand setup I used in the HGCAL system test laboratory at CERN. The sensor module is placed in between two scintillator tiles of different cross-sectional area. Scintillator on the top has a

cross-sectional area of 4 cm x 4 cm and the one at the bottom has an area of 12 cm x 10 cm. The distance between top and bottom scintillators is about $\sim$ 10 cm, hence time difference between muon (traveling at $\sim c$) reaching two scintillators is less than 1 ns. Since the response time of a typical scintillator (section 2.4.2) is of the order of few ns, therefore we can safely assume that the signal is generated simultaneously by the two scintillators when a cosmic muon passing through them.

The signal from the top and bottom scintillator+PMT setup is passed through a discriminator, level adaptor and a coincidence logic unit housed in NIM module crate. Discriminator converts analog signal to a discriminated NIM signal, level adaptor converts NIM signal to TTL signal and the coincidence logic unit performs an AND logic operation on two input TTL signals originating from two PMTs, which is fed to ZEDIO board as the external trigger. Thus, when a muon passes through both top and bottom scintillators, with sensor module placed in between, the coincidence logic unit generates a trigger which is supplied to SKIROC2 ASICs to initiate the data-acquisition.

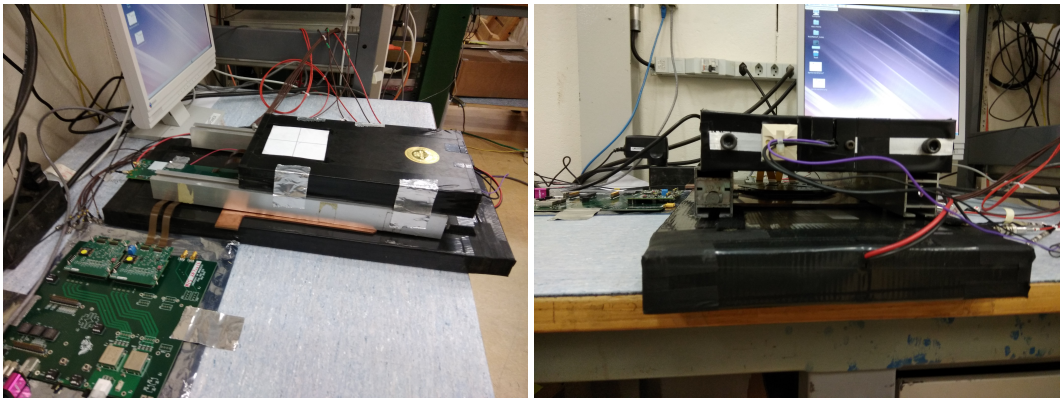

FIGURE 5.12: Two images show the cosmic muon setup from two angles. Sensor module is places in between two scintillator tiles (right).

Figure 5.13 (top) shows a block diagram of the trigger-flow for the cosmic muon stand. With the help of top scintillator with smaller cross-sectional area, we could control the location of the cells exposed to cosmic muon triggering. At a time only $\sim$ 4 cells could be correctly triggered using 4 cm x 4 cm scintillator tile. The wires and pulse-processing units introduce a delay or *latency* between trigger generation and reception. The major latency is caused by the wire length whereas the pulse-processing units cause latency of few to tens of ns depending on the instrument. In the run of data taking (18k events), no MIP signal was observed which led to a conclusion that an earlier trigger was being supplied. An additional cable of 24 ns equivalent delay was introduced as shown in the block diagram of trigger flow in Figure 5.13 (bottom), and $\sim$ 8k events was acquired with modified setup.

The data collected with additional 24 ns delay is reconstructed to obtain ADC distribution with pedestal and common-mode noise subtraction. Figure 5.14 shows the ADC distributions from the high gain channel of one of the exposed silicon cell for 8k events, it shows a distribution with a mean of zero corresponding to intrinsic noise as well as a MIP energy distribution of cosmic muons.

The intrinsic noise can be modelled as a Gaussian distribution due to its random nature. Whereas, the energy distribution of an ionizing particle traversing through a thin detector can be described by Landau distribution [108]. The Landau is an asymmetric distribution with a narrow width around its peak and has a long tail. In a thin material, the traversing

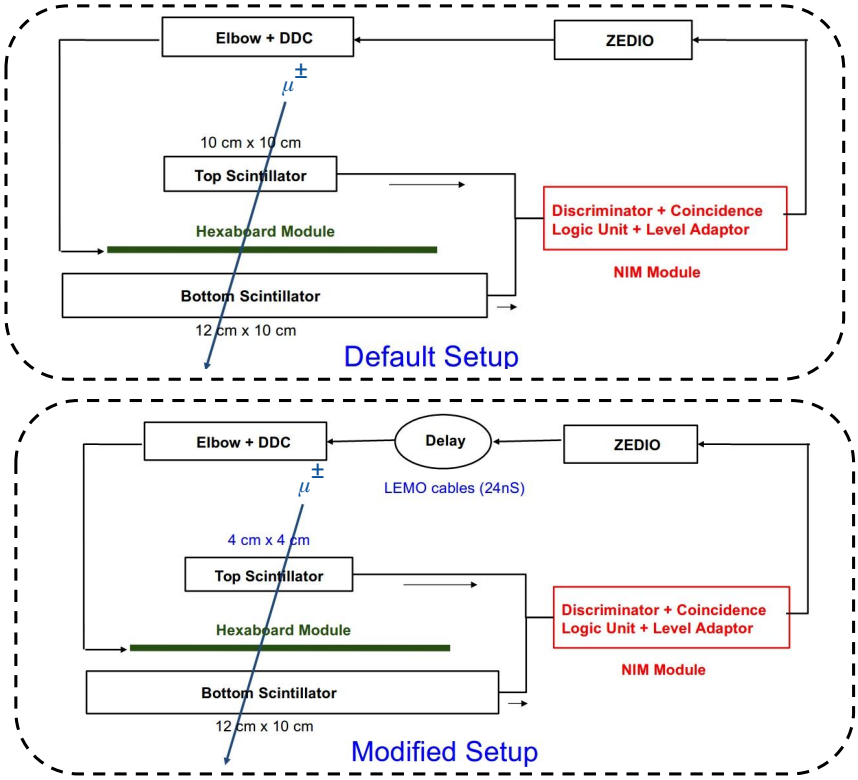

FIGURE 5.13: Figure shows the block diagram of the trigger flow from the scintillator tiles to sensor module with the default setup (top) and with additional delay of 24 ns (bottom).

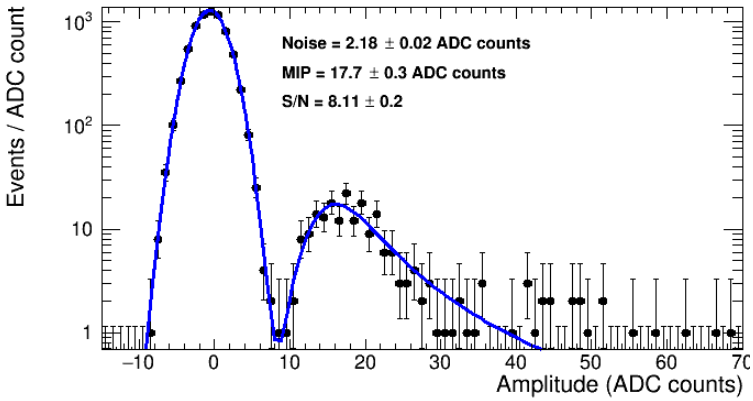

FIGURE 5.14: High gain ADC distribution for one of the channels. The left distribution centered around zero corresponds to noise and the right distribution corresponds to muon energy spectrum. The pedestal noise is fitted with a Gaussian function and MIP distribution is with a fitted with a Landau convoluted with a Gaussian function.

particle can ionize only a small number of atoms (or molecules), hence depositing a tiny amount of energy. The most probable value (MPV) of energy deposited is reflected in the Landau distribution as its peak (i.e., the *mode* of the distribution), and the fluctuation in the deposited energy corresponds to its width. However, the traversing particle has a small but non-zero probability to transfer relatively large amounts of energy in the ionization process. It results in a large amount of energy deposits in the detector and shows up as the tails of Landau distribution. Hence, the ADC distribution is fitted with following function:

$$\begin{aligned} f(\text{MPV}, \sigma_{noise}, \sigma_{MIP}, a_1, a_2) = a_1 \times \text{Gaussian}(0, \sigma_{noise}) \\ + a_2 \times \text{Landau}_1(\text{MPV}, \sigma_{MIP}) * \text{Gaussian}(0, \sigma_{noise}) \end{aligned} \tag{5.1}$$

The first term of the function corresponds to Gaussian distribution with $\mu = 0$, second term of the function is a convolution of a Landau and Gaussian distribution. The weights a_1, a_2; MPV (most probable value or mode) of Landau distribution; and width of Gaussian and Landau distribution are let to float during the fitting. The fitting gives us the MPV MIP energy deposit and intrinsic noise in terms of high gain ADC counts.

Multiple cosmic muon runs are taken by moving the top scintillator around in order to expose different cells of the sensor module. Thus, energy distribution of MIPs (cosmic muons) is obtained for multiple silicon cells in terms of high gain ADC counts. The MIP signal is also used as a "standard candle" to equalize channel-to-channel response with the help of ADC-to-MIP conversion factors, discussed in detail in section 6.8. The plot shown in Figure 5.14, 17.7 $\pm$ 0.3 high gain ADC counts corresponds to 1 MIP equivalent energy deposit.

Furthermore, this experiment also facilitates the quantification of *signal-to-noise ratio* for the silicon sensors for v2016 module. Signal-to-noise ratio is an essential figure-of-merit for any detector. It helps to quantify the level of separation between the signal and pedestal noise (signal to noise ratio or SNR). The knowledge of SNR helps physicist to decide the noise-threshold to be applied on the read-out channels in order to minimize noise-like hits without losing the signal. This is crucial to manage the bandwidth bottleneck, by minimizing the channel occupancy in the systems like the CMS experiment with millions of readout channels. At the same time, the efficient rejection of noise improves the resolution and response of energy measurements of the physics objects such as electrons or hadrons.

For silicon sensor modules, signal refers to ADC data corresponding to a MIP passing through it and noise refers to ADC data corresponding to the intrinsic noise of the silicon cell as described in the last section. Using these values, we can determine the signal-to-noise ratio of the silicon sensor. The SNR is determined by dividing MPV of the fitted Landau function by width of Gaussian function. For the 2016 module version, we obtain SNR $\sim$ 8.1 $\pm$ 0.2. Similar tests were performed for silicon modules employed in October 2018 beam test experiment, and results are documented in the next chapter along with the description of beam test experiments.

Chapter 6

Beam test experiment - prototype commissioning, calibration and event reconstruction

In parallel to improving, validating and bench-marking the most basic unit of the detector (a single sensor module for example) with the laboratory test-bench setups, the detectors have to be prepared for two important challenges: (a) performance of these basic units in the presence of realistic signals expected in the collision experiment, and (b) to scale up these units to work in synchronization for data acquisition in the presence of collision particles passing through, and understanding their response to particles of various types. A larger set of prototype modules are manufactured, qualified using laboratory system tests, and assembled into a prototype of a section of expected final endcap calorimeter detector. Such detector prototypes are then tested in a controlled environment with single particle beams of known energy and identity. These are called beam test (or test beam) experiments.

As part of the detector development, the CMS HGCAL collaboration has tested prototype setups based on silicon modules in the beam test experiments at Fermilab and CERN using electron beams of momenta in range 20-250 GeV/c [107]. The main focus of the thesis work presented here are the experiments carried out at the beam test facilities at CERN during October 2018. A complete longitudinal section of the CMS endcap calorimeter is exposed to single particle beams of e^+, π^- and μ^- of momenta varying from 20 to 300 GeV/c derived from the interactions of protons, accelerated by the Super-Proton-Synchrotron (SPS) to 400 GeV, with a Beryllium target.

The goals of the experiments performed in 2018 were to demonstrate that on-going research and development efforts are in-line with the desirable performance of the sensor prototypes and readout electronics for the purpose of the final detector and physics data taking. This requires a working data-acquisition system for multiple units of the detector to be run in synchronization, and pedestal noise being within tolerable level. Furthermore, it is also of utmost importance to demonstrate the performance of the prototype for the measurements of electromagnetic and hadronic showers in real experiments, and have a well understood dataset to compare with GEANT4 based detailed detector simulations. The task of building the prototype, conducting the beam test experiment, and the analysis of data was undertaken by the CMS HGCAL team, of which I was an active member.

In the first part of this chapter, the details of 6-inch silicon sensor prototype used in the beam test experiment , the experimental setup, secondary particle beam generation and transport at H2 beamline, are discussed. In the second part, the event reconstruction in data and simulation, detector alignment, channel-to-channel response equalization and signal-to-noise ratio of silicon sensors, are described.

6.1 Silicon sensor module prototype version v2018

Prototype silicon sensor modules used in the 2018 beam test experiments, referred to as v2018, are fabricated on 6" silicon wafers similar to sensor prototype version v2016. A representative module is shown in Figure 6.1. These prototype modules are also fabricated in similar fashion as described in section 5.1, i.e. by gluing together the baseplate, gold-plated Kapton^{TM}, silicon layer and PCB layer on top. The silicon module version v2018, being similar to version v2016 in many aspects, has many differences to its predecessor which are listed as follows:

- Many modules are built with Cu baseplate along with CuW baseplate. The CuW baseplates being dense are employed in the electromagnetic section and Cu baseplates are employed in the hadronic section of HGCAL prototype.
- Most of the silicon sensors are produced with 300 μm active thickness and a few sensors with 200 μm active thickness to test the performance of different types of silicon cells for real particle showers.
- The module design moved from double PCB design in v2016 to single PCB design in v2018 to make less bulky and easier to handle.
- The readout chip used for v2018 is SKIROC2-CMS [109], an upgraded version of SKIROC2 [104], and is closer to HGCROC design (section 4.4.2). The differences are described in section 6.2.
- Four SKIROC2-CMS ASICs are used per module to minimize the path length between individual cells and readout chips.
- The PCB has an Max10^{TM} FPGA on-board to issue the control commands and to frame the digitized data coming from SKIROC2-CMS ASIC in a proper format to hand it over to the DAQ system.

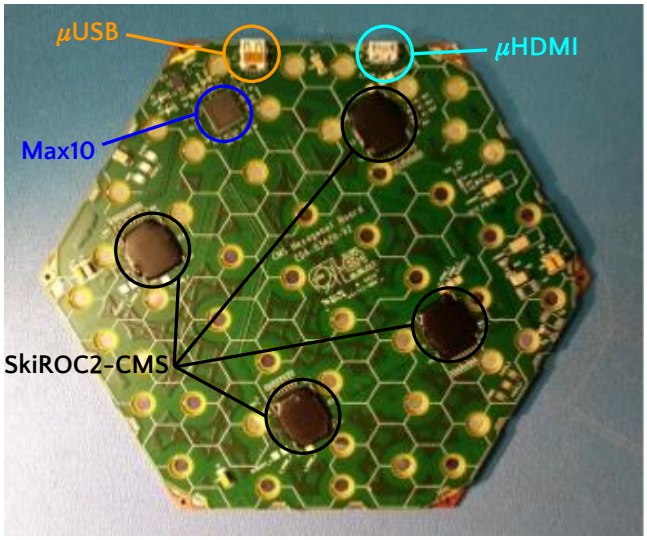

FIGURE 6.1: Silicon sensor module version v2018. The different components on PCB are shown in the image.

The PCB (Figure 6.1) contains all the electronics such as four SKIROC2-CMS ASICs readout chip, a Max10^{TM} FPGA, μHDMI and μUSB port. Max10^{TM} FPGA is used to send commands, communicate data stream and distribute clocks between SKIROC2-CMS and DAQ

system via μHDMI port and the firmware for the FPGA is loaded via μUSB port [110]. The readout and digitization of analog pulses from silicon cells is done via SKIROC2-CMS ASICs.

6.2 SKIROC2-CMS readout chip

The SKIROC2-CMS is an upgraded version of SKIROC2 ASIC used in v2016 modules. It incorporates a few additional features that were not present in SKIROC2 which are discussed in the following. Figure 6.2 shows a simplified internal block diagram of SKIROC2-CMS. Each SKIROC2-CMS ASIC can read up-to 64 channels. Each channel in SKIROC2-CMS offers a low-noise charge sensitive preamplifier with adjustable gain via slow-control registers. It has two pulse shapers of different gains, called High-Gain (HG) and Low-Gain (LG). The HG can measure very small energy deposition in the silicon, corresponding to $\sim$ 2 fC to 150 fC, whereas LG can measure high energy deposition in the silicon, corresponding to $\sim$ 100 fC to 900 fC. The shaped pulse from HG and LG stage is sampled in 13 deep switched-capacitor-array (SCA). Each capacitor integrates the charge over 25 ns. When the trigger is received, the charge stored on 13 SCAs from two gain stages is passed to 12-bit Wilkinson ADC for digitization. The ADC information from all 13 time-samples of HG and LG shaper is stored, in contrast to SKIROC2 where ADC information of only one time sample was stored.

In addition, each channel also measures the timing information, namely Time-over-Threshold (ToT) and Time-of-Arrival (ToA). To be able to maintain a large dynamic range for large signal, as foreseen in the CMS HGCAL, a ToT circuit is introduced in the readout chip that gets activated only when signal crosses a certain threshold and its output is proportional to the signal amplitude. Its working principle can be understood with the help of a cartoon representation shown in Figure 6.3. For large signals, a timing measurement is made for the time duration for which the pulse from preamplifier stays above a predefined threshold. This is achieved by feeding the pulse from the preamplifier to a discriminator with adjustable threshold (V_{thres}). The ToT can measure large energy deposit in silicon, beyond 600 fC of induced charge. A Time-to-Digital converter (TDC) ramp (with 25 ps time bin) is triggered during non-zero output state of the discriminator, and thus a digitized time measurement is performed. This time measurement is proportional to the pulse height of the preamplifier, hence input signal. Generally, the ToT is triggered for signals above few hundreds fC and can measure signals $\sim$ 10 pC. Thus, it provides a low power and a large dynamic range. Similarly, the ToA measurement is performed with a fast discriminator and a TDC with time bins of 25 ps for precise timing measurement of the particle traversing through the cell.

In this way, SKIROC2-CMS provides a large dynamic range to measure signal induced in the silicon sensor by traversing particles (a few fC to 10 pC) with low power consumption per channel (10 mW) and is closer to final design of HGCAL readout chip, the HGCROC. In v2018, there are four SKIROC2-CMS ASICs per module therefore each SKIROC2-CMS connects to 32 silicon cells and the rest 32 channels are kept open. The connection between silicon cells and the SKIROC2-CMS chip is made via a thin wire-bonding through the circular holes of the PCB drilled for this purpose. Figure 6.4 (left) shows a cross-sectional schematic diagram of the connection and the right image shows the wire-bonding through the hole. The connection is done such that "even" channels of ASIC are connected to silicon cells whereas "odd" channels are left open. The wire-bonding is an extremely delicate process and is done using specialized automated work-stations, called *Gantries*, as shown in Figure 6.5.

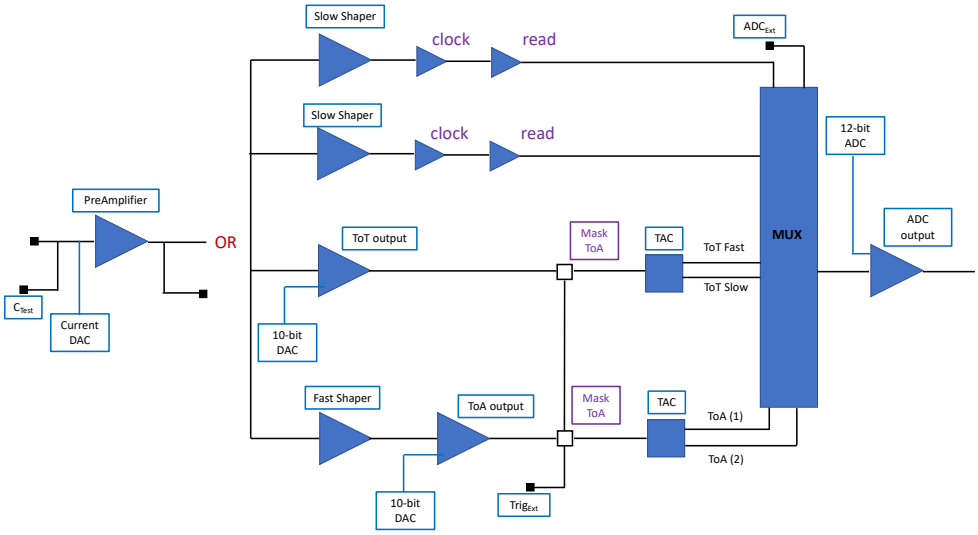

FIGURE 6.2: A simplified block diagram of SKIROC2-CMS ASIC [111] .

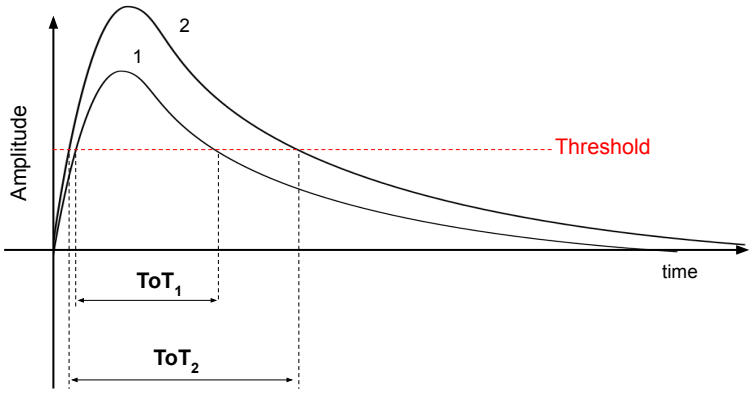

FIGURE 6.3: Cartoon representation of ToT where the ToT_1 corresponds to the pulse-1 (with smaller height) and ToT_2 corresponds to the pulse-2 (with larger height). Pulse-2 stays for a longer duration above the threshold as compared to pulse-1 thus ToT_2 will be larger than ToT_1 and will be proportional to time spent above the threshold by the incoming pulse.

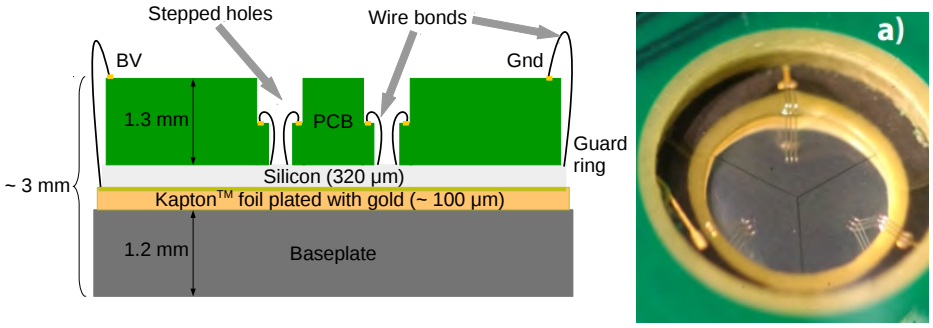

FIGURE 6.4: The connections from the silicon cells to the PCB is shown in the schematics on the left and a close-up image is shown on the right.

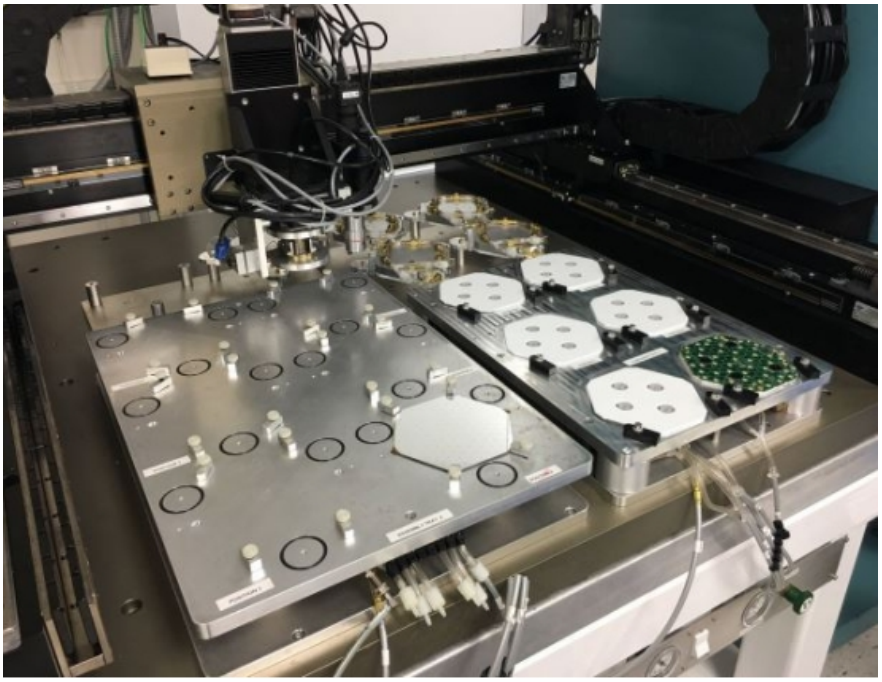

FIGURE 6.5: The image shows the gantry setup at University of California Santa Barbara. The gluing of different layers are done automatically with high precision using such gantries.

Detailed system-tests were carried out for silicon sensor modules version v2018 in the laboratory at CERN. The set of modules which qualified the features such as IV and CV characterization, pedestal noise measurement were used to build the HGCAL detector prototype used in the beam test experiment, as explained in the next section.

6.3 Beam test experimental setup

We built the prototypes of electromagnetic and hadronic sections of HGCAL, referred to as CE-E and CE-H respectively, using the silicon senor modules. More than 90 silicon sensor modules are employed resulting in a first large scale test of HGCAL prototype using high energy beams of single particles. A prototype of Analogue Hadronic Calorimeter (AHCAL) from CALICE collaboration is placed downstream of CE-H [112]. The combined calorimeter prototypes are placed on a concrete platform in the H2 beamline of SPS CERN which delivered beams of single particles. Figure 6.6 shows the image of experimental setup with different sections duly marked.

The whole experiment was a combined effort of HGCAL and CALICE collaboration where physics data taking took place during a span of three weeks of October 2018. The DAQ for HGCAL and AHCAL work separately on a common trigger and are synchronized offline based on time-stamp. The DAQ system for HGCAL and AHCAL has been described in detail in [110] and [112], respectively.

During the span of three weeks of data taking, we assembled the silicon modules in CE-E and CE-H in three different configurations, namely configuration-1, configuration-2 and configuration-3, to fulfil different experimental goals such as focusing on electromagnetic

or hadronic shower performance. One week of data-taking was allocated for each configuration. This thesis focuses on the results using data collected in configuration-1 which is described in detail in the following section.

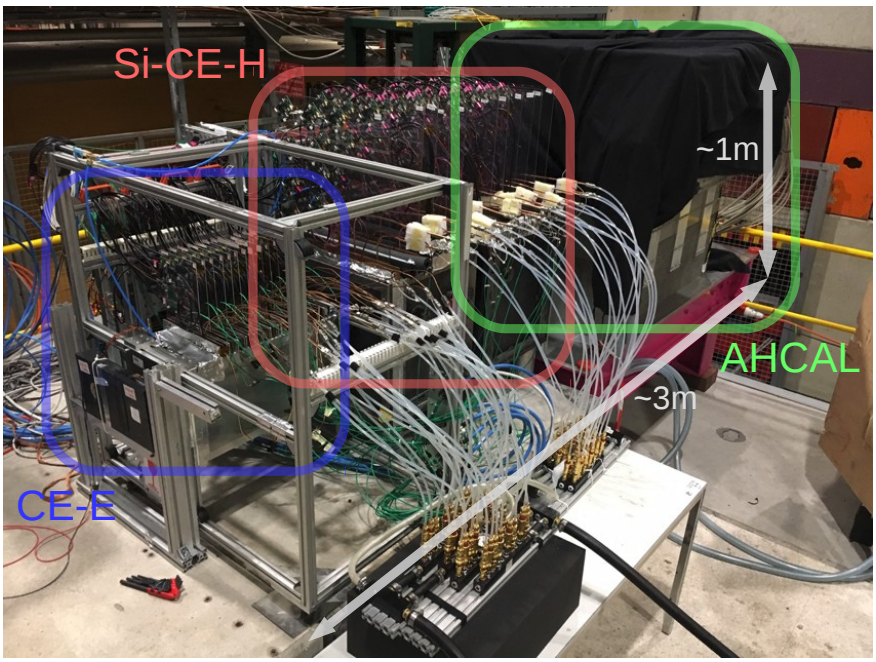

FIGURE 6.6: The detector setup for beam test experiment. Different compartments have been indicated with the different colour boxes and the beam enters the detector setup from the left.

6.3.1 CE-E prototype

The CE-E prototype is built in double-sided mini-cassettes, as shown in Figure 6.7 (left). Each mini-cassette has two sensor modules mounted on the either side of a 6 mm thick copper cooling plate. One of the sides of the mini-cassette is connected to a 4.9 mm thick lead (Pb) absorber cladded with 300 μm steel. Figure 6.7 (right) shows the side view sketch of the mini-cassette indicating the dimensions of various components. There are 14 such mini-cassettes built and placed in CE-E prototype in a hanging-file structure, thus making a total of 28 sampling layers with each active layer containing one silicon sensor module with Pb or Cu+CuW as the absorber material. The total depth of CE-E prototype is $\sim 27\ X_0$ and $\sim 1.4\ \lambda_{int}$. The transverse coverage of CE-E prototype is limited by the area of a single sensor module, i.e. 12 cm $\times$ 14 cm.

The CE-E prototype is the first section seen by the incident particle. The incident particle traversing CE-E prototype sees an alternating absorber layer of Pb and Cu+CuW. The physical length of Pb and Cu+CuW are chosen such that the absorber depth between two active layer is approximately 1 X_0. The differences in the absorber material resulted in the variation in X_0 as shown in Figure 6.8. The variation in the absorber depth is within $\sim$ 5%. Also, layer 21 and 23 are built with extra 1.2 mm Cu, and layer 25 and 27 are built with extra 0.6 mm CuW layer on their baseplate, thus causing extra ΔX_0 for the immediate downstream active layer. All the layers in CE-E except the last two, are instrumented with silicon modules with active thickness 300 μm. The last two layers are instrumented with silicon modules with active thickness 200 μm.

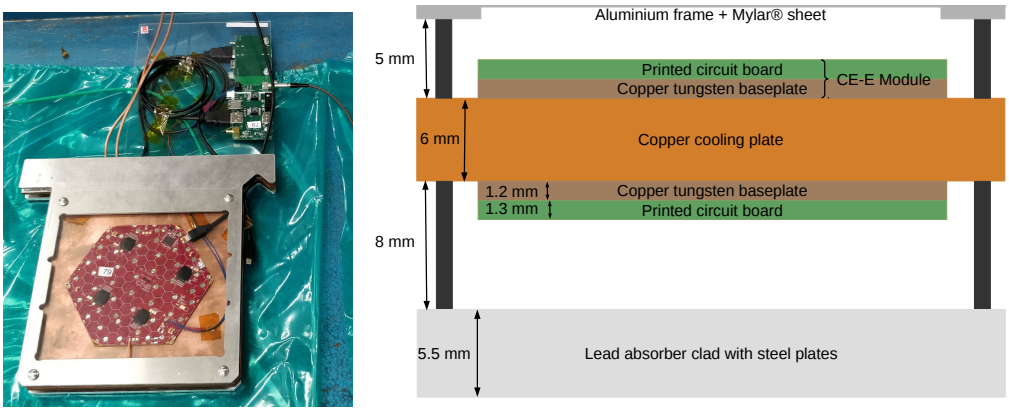

FIGURE 6.7: Figure on the left shows one mini-cassette built with single silicon sensor module each side of the cassette (only one is visible in this image). Right image shows schematic diagram of cross-sectional view for a mini-cassette with physical lengths shown in the image.

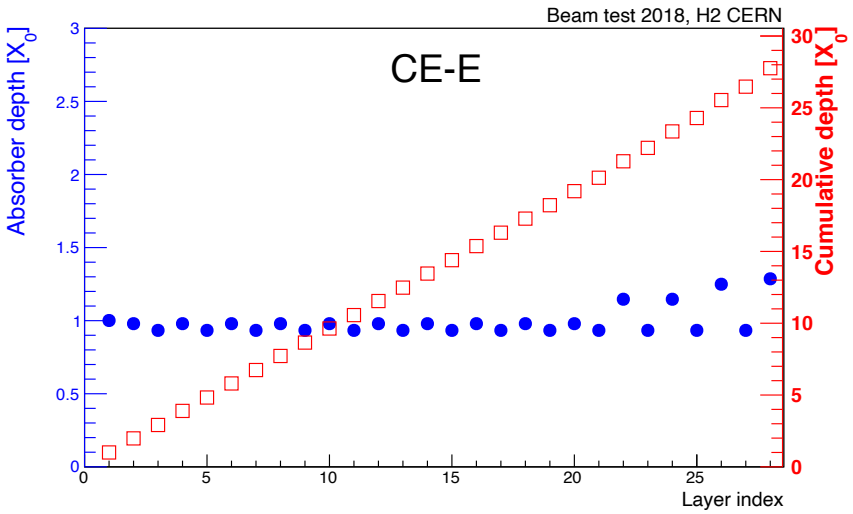

FIGURE 6.8: Absorber and cumulative depth in terms of radiation length (X_0) for CE-E layers. The values are obtained from simulated beam test geometry.

6.3.2 CE-H and AHCAL prototypes

The CE-H prototype is placed downstream of CE-E. The CE-H has 12 sampling layers interleaved between 4 cm thick steel absorber. The first 9 active layers of CE-H is instrumented with 7 silicon sensor modules arranged in a daisy-like structure to increase the transverse coverage of CE-H. The modules are mounted on 6 mm Cu cooling plate as shown in Figure 6.9. The last 3 active layers of CE-H are instrumented with only one silicon sensor module (due to lack of module availability) placed on the central location of the layer. Therefore, the lateral coverage of first nine and last three CE-H layers corresponds to 37×35 cm^2 and 12×14 cm^2, respectively. The active thicknesses of all the silicon sensors are 300 μm except two modules which have 200 μm and are located at off-center location in layer 5 and 6 of CE-H. The layers are built in a hanging file structure and placed in two separate iron boxes (with wall of 9 mm thickness), each containing 6 CE-H layers. The total depth of CE-H is approximately 3.4 λ_{int}.

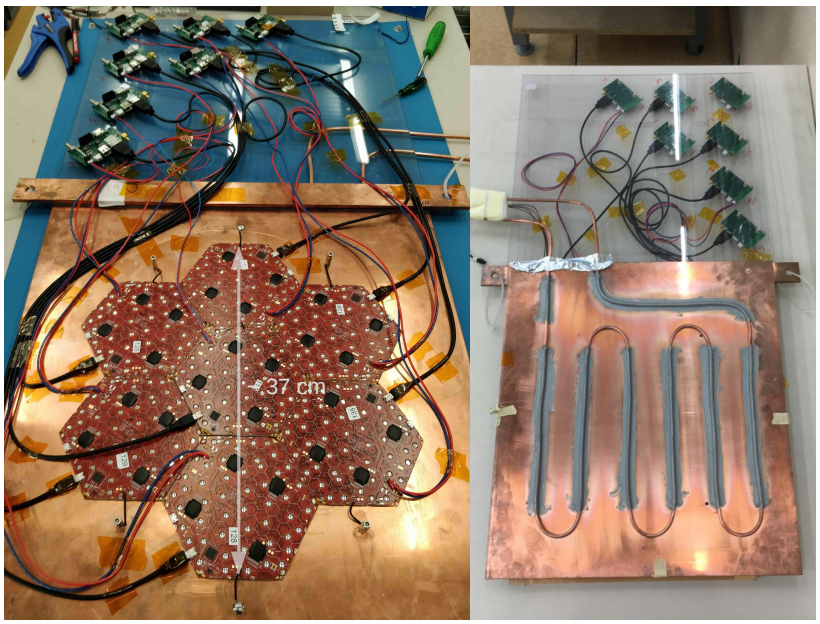

FIGURE 6.9: An image of CE-H layer with seven modules arranged in *daisy-like* structure as shown in the left image. Right image shows the back-side of the layer with cooling pipes visible.

To match the endcap design depth, AHCAL prototype is placed downstream of CE-H. The AHCAL is a scintillator tile and silicon photo-multiplier (SiPM) technology based calorimeter that is proposed to be used at International Linear Collider (ILC) detectors [113]. The technical design opted for AHCAL is very similar to scintillator part of HGCAL detector, therefore it also serves the purpose of acting as scinitillator part of hadronic section of the HGCAL detector.

Figure 6.10 (left) shows an image of AHCAL engineering prototype used in the beam test experiment. There are a total of 39 sampling layers interleaved between 17.7 mm of nonmagnetic steel absorbers. Each active layer containes 3 mm thick scintillator tiles of rectangular shape with lateral dimension of 3 cm $\times$ 3 cm, mounted on SiPM as shown in Figure 6.10 (right). The SiPMs and scintillator tiles are mounted on HCAL Base Unit (HBU) of size 36 cm $\times$ 36 cm, with each layer housing 4 HBUs. Each HBU contains 144

tiles + SiPM. Thus each layer in AHCAL contains 576 channels with a total active area of 72×72 cm^2. Second-last layer of AHCAL, however, is a special layer containing tiles of transverse size 6 cm × 6 cm. It is placed to test the effect of different tile size on the energy reconstruction for AHCAL detector. There are ∼22000 channels in AHCAL that are read out independently. The total depth of AHCAL prototype is approximately 4.5 λ_{int}. The scintillator part of proposed CE-H will have 10-16 sampling layers (depending on $|\eta|$) therefore the AHCAL with its 39 sampling layers provides a unique opportunity to probe hadronic shower development more finely in longitudinal direction.

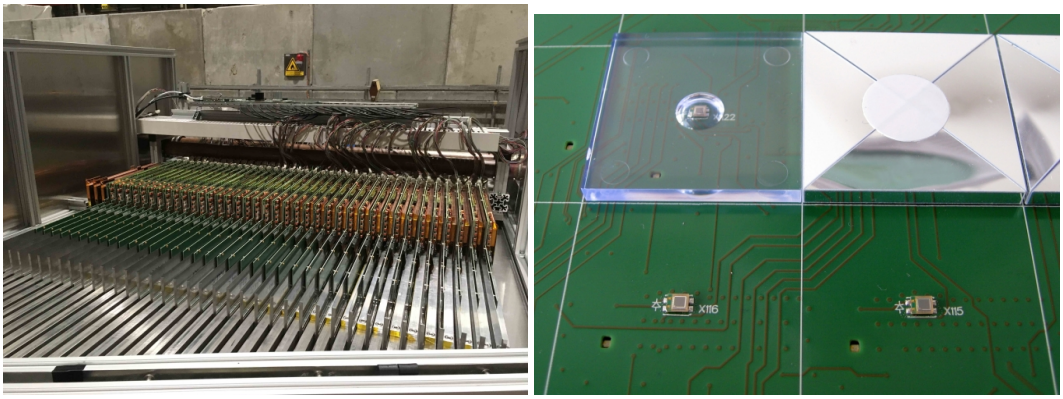

FIGURE 6.10: A close-up view of CALICE AHCAL engineering prototype is shown in the left image. On the right, the tiles mounted on SiPM is shown with tiles wrapped and unwrapped along with SiPM in the middle..

The HGCAL + AHCAL combined prototype results in a total depth of ∼ 9.6 λ_{int} with varying degree of transverse coverage based on the sub-detector prototype. Figure 6.11 shows the absorber and cumulative depth in λ_{int} as a function of layers. The absorber depth in CE-E is observed to fluctuate by 50% in alternating layers. This situation arises because of different $\frac{X_0}{\lambda_{int}}$ ratio for Pb and CuW [36], the absorber material in CE-E. For Pb $\frac{X_0}{\lambda_{int}}$ is about 0.03 and for CuW $\frac{X_0}{\lambda_{int}}$ is about 0.05. It implies that for similar X_0, λ_{int} will be different by ∼ 40%. These variations, however, are not super critical for the energy measurement of hadronic showers as CE-E is only 1.4 λ_{int} deep.

The jump observed at CE-H layer 7 in Figure 6.11 is because of extra material of the walls of iron boxes (9 mm + 9 mm = 18 mm Fe) which adds extra ∼ 0.1 λ_{int} as seen by CE-H layer 7. Similarly, the first layer of AHCAL sees the wall of iron box as well as material from Cu cooling plate + baseplate of last module in CE-H, thus adding extra material (9 mm Fe + 7.2 mm Cu) of 0.1 λ_{int}. The last active layer of AHCAL is placed three slots downstream of the penultimate layer, therefore it sees three times more absorber as compared to other layers of AHCAL and results in a jump of absorber thickness (3×0.108 λ_{int}) at AHCAL layer-39 in Figure 6.11.

6.3.3 Auxiliary detectors in the beam-line

The beam test experiment made use of several other detectors along with HGCAL and AHCAL prototype to assist in triggering, vetoing, identifying, and tracking the beam particles. Figure 6.12 shows the schematics of different detectors employed in the beam test experimental hall. Two Cerenkov detectors are placed (not shown in the image) upstream of HGCAL prototype for particle identification. Depending on the gas pressure inside the Cerenkov detectors, particle of different species can be tagged and saved for offline analysis. However, the efficiency of both the Cerenkov detectors were found to be extremely

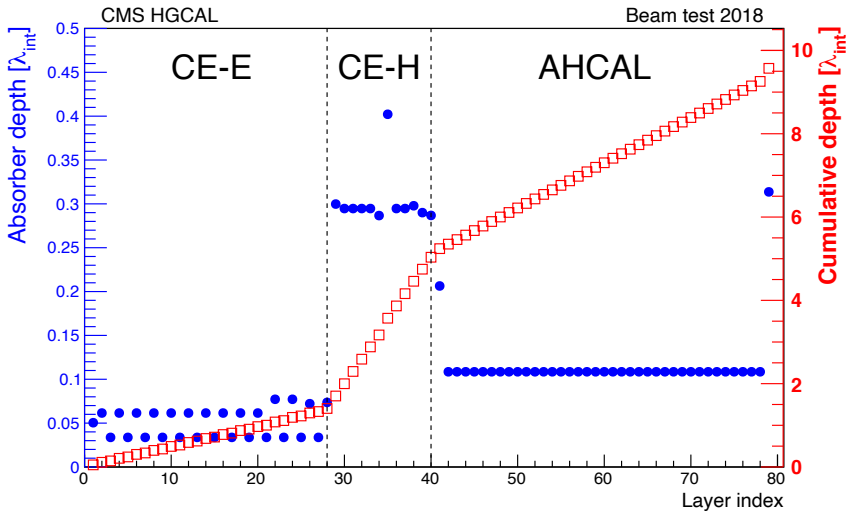

FIGURE 6.11: The absorber depth (blue solid circles) and cumulative depth (red open squares) in terms of λ_{int} for CE-E, CE-H and AHCAL. The values are obtained from simulated beam test geometry.

low (< 1 %) therefore it is not used in the data analysis, and other method are employed for particle identification, discussed in chapter 7.

Four delay wire chambers (DWCs) are placed upstream HGCAL prototype to help reconstruct the track of single particle beam. Details of the DWC working can be found in [114]. In summary, a delay wire chamber produces four signals (two for each coordinate) on its output. The time separation between the signals in a pair depends on the impact position (*hit*) in the wire chamber. The time separation are saved with the help of Time-to-Digital (TDC) converters and are calibrated to provide corresponding (x, y) coordinate with a resolution of 200μm. An independent study on these DWCs showed that for the beam of single particles from SPS CERN, the average occupancy of DWCs is one, i.e. each wire chamber records only one hit on an average [115]. In the beam test experiment, TDCs corresponding to only leading hit were saved from each wire chamber for each event. Based on the independent hits in four available DWCs, a straight line is fitted and extrapolated to each active layer of HGCAL as well as AHCAL prototype, on event-by-event basis. The track is not reconstructed if no hits are available. The DWC tracking information provides an important input for data analysis such as selection of one silicon cell per layer through which MIP-like particles pass.

Two micro-channel plate (MCP) detectors are placed in front of CE-E to provide timing reference. The information from MCPs are used in the timing-calibration and timing-analysis (not discussed in this thesis). Two scintillators of sizes 4 cm $\times$ 4 cm (or 2 cm $\times$ 2 cm) and 10 cm $\times$ 10 cm are placed in front of CE-E prototype. The coincidence signal from these two scintillators provided triggers to HGCAL and AHCAL DAQ for data taking. Another scintillator of larger size was placed behind CE-H to generate veto-trigger and used only for e^+ runs.

6.4 H2 beamline description

A dedicated beam of single particles is required to perform beam test experiments in order to assess the detector performance. For HGCAL and AHCAL prototype, beams of e^+,

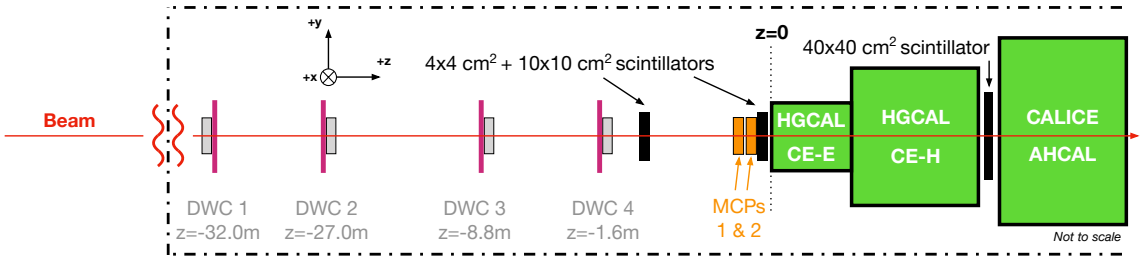

FIGURE 6.12: The image shows an schematic view of all the detectors deployed in the beamline upstream of HGCAL + AHCAL prototype experimental setup.

π^- and μ^- of known momenta range was used to calibrate and measure the detector performance to electromagnetic and hadronic showers. These beams are extracted from the accelerator complexes which can be the primary beam from accelerator, e.g. at DESY and SLAC, or can be obtained from the interaction of primary beam to a fixed target, e.g. at CERN and FNAL.

At CERN, high energy beam of protons circulating in the SPS ring makes it ideal to obtain secondary beams spanning a large momentum range. The North Area Facility at CERN, Figure 3.2, is used for this purpose where protons of 450 GeV/c momentum from SPS ring are used to generate secondary beams. The proton beam from SPS is split and made to interact with multiple fixed target stations, as shown in Figure 6.13, by adjusting their bending angles. The beam of secondary particles generated in this interaction, are transported to the experimental halls via parallel beam-lines. At North Area facility, there are four beam-lines; namely H2, H4, H6, and H8 as shown in Figure 6.13. This is done in order to serve multiple experiments in parallel and to maximize SPS beam time utilization. The beam test experiment of HGCAL + AHCAL was held at the H2 beam-line experimental hall [116, 117].

For H2 beam-line, Beryllium of 500 mm thickness is used as the fixed target material due to its roughly equal characteristic lengths i.e. $X_0 \sim 35.28$ cm and $\lambda_{int} \sim 42.1$ cm, as compared to other material. Having equal X_0 and λ_{int} implies equal effective thickness seen by hadronic and electromagnetic showers, thus it would not selectively suppress or attenuate any particular particle species. A fraction of highly energetic protons in a given bunch from SPS interact with the Beryllium target and produces secondary particles of different types. The remaining intense beam of protons that do not undergo any interact with the Beryllium is dumped on the *Target Attenuators* (TAXs).

The secondary beam is filtered according to the desired momentum. This is done by placing multiple bending magnets in the beam-line. The focal length of the bending magnets are adjusted such that the resolution of beam momentum is $\frac{\delta p}{p} \sim 2 \times 10^{-4}$. The magnets that are deployed in the beam-line is also used to steer, bend and collimate the beam. The primary proton beam from SPS is extracted every 30 seconds towards the target, and the extraction lasts for about 5 seconds. It is called a *spill*. The particle count in each spill ranges from 10^3 to 10^7. At H2 beam-line, typically, particles of momentum 10 GeV/c to 400 GeV/c $e^{\pm}$, charged hadron and muons can be transported with varying degree of composition purity, depending on the particle type.

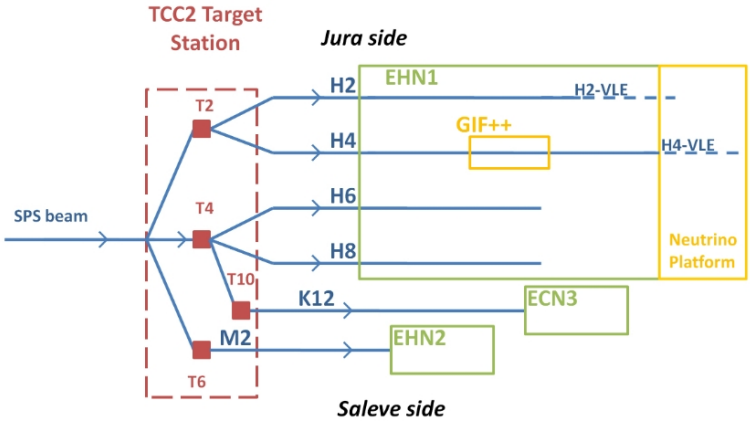

FIGURE 6.13: The schematic for the North Area with all four beam-lines indicated. The secondary beams produced from primary proton beams are transported to the experimental hall via these four beam-lines [116].

The secondary beam further passes through a cascade of multiple collimating and bending magnets that act as spectrometer as shown in Figure 6.14 . The particle type selection of beam is done by a series of techniques, viz. *selected attenuation of particle*, *particle tagging* and *separation*. In the first step, i.e. *"selected attenuation of particle"*, the beam is exposed to a secondary target. The material of secondary target is chosen such that it can selectively absorb particles of unwanted type. For example, 40 cm thick copper ($\sim 30\ X_0$, $\sim 3\ \lambda_{int}$) is used as an absorber to produce hadronic beam since 30 X_0 of copper absorbs and attenuates almost all of electromagnetic shower induced by $e^{\pm}$ or γ. To obtain $e^{\pm}$ beam, 4 mm thick lead ($\sim 1\ X_0$, $<< 1\ \lambda_{int}$) is used as 1 X_0 of lead does not absorb $e^{\pm}$ fully. This, however, does not get rid of hadronic particles. Therefore, further cleaning of the beam is performed. This tertiary beam is passed through Cerenkov detectors (CEDARs), where *particle tagging* is done based on the gaseous pressure set in the CEDARs. In the next step, called *separation*, the beam is further cleaned by using electrostatic and RF electromagnetic separators which are placed downstream of the CEDARs and it segregates the beam to desired and undesired particle beam type. These techniques are discussed in detail in [117].

After momentum and particle type selection, the beam is focused in order to reduce the spot size and to minimize beam divergence. The beam is finally delivered to the experimental setup. The total distance between primary target and the experimental setup is about 600 meters.

Table 6.1 summarizes the particle types and their corresponding momenta used in configuration-1 of data taking during the October 2018 beam test experiment.

	20	30	50	80	100	120	150	200	250	300
μ^-	–	–	–	–	–	–	–	✓	–	–
e^+	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
π^-	✓	–	✓	✓	✓	✓	–	✓	✓	✓

TABLE 6.1: Table of beam of single particle with their corresponding momenta used in the configuration -1 of data taking.

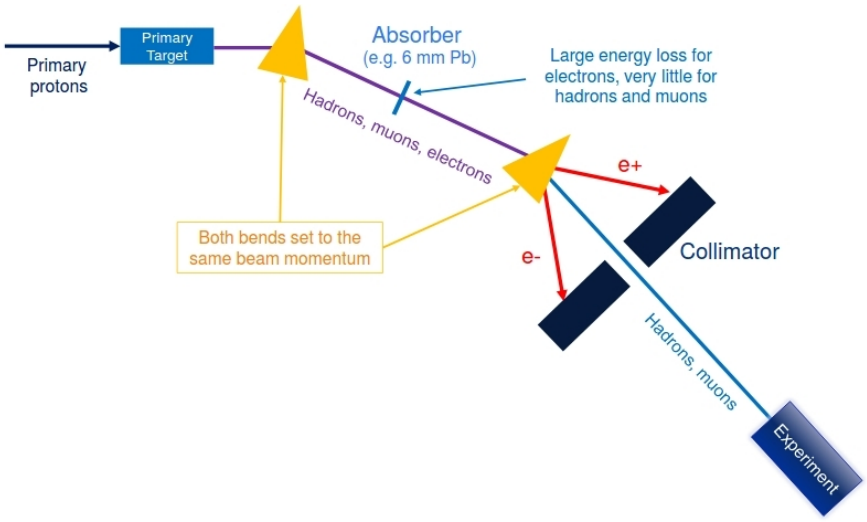

FIGURE 6.14: A simplified schematic of secondary beam production and transportation for H2 beam line. Various beam optics components depicted here are used to separate and select particle of desired type and momentum.

6.5 Event reconstruction in HGCAL

On receiving the trigger, both HGCAL and AHCAL recorded the digitized signal collected from the silicon or scintillator+SiPMs sensors, and saved on the disk in compressed *raw* format for offline analysis. The raw data is essentially a compressed version of data-stream received from each readout chip for each event. The data saved on the disk can not be directly used for any analysis and it needs to be converted into meaningful observables such as position and the energy recorded by the sensor. This process is called *event reconstruction*. The raw event data has to be synchronized and converted into reconstructed hits, called *rechits*, to perform further studies. In the following section, the event reconstruction in HGCAL and AHCAL prototype is described for the beam test data as well as simulation.

6.5.1 Reconstruction of data events in HGCAL

As mentioned in previous sections, SKIROC2-CMS ASIC provides three measurements, i.e. HG, LG and ToT. The HG is sensitive to very small induced charge in the silicon ($\sim$ 2 fC to 150 fC), LG is sensitive to a wide range of moderately induced charge ($\sim$ 100 fC to 900 fC), and ToT is sensitive to a very large induced charge ($\sim$ 600 fC to 10 pC). A MIP induces about $\sim$ 3 fC of charge in silicon sensor of 300 μm active thickness. Whereas, the intense core of high energy electron shower can have hundreds of MIP-like secondary particles. Thus, three measurements ensure large dynamic range for energy measurements. When the ASIC receives the trigger, the charge stored in 13 SCAs from two gain modes are passed to the ADC which digitizes the pulse in 13 corresponding time-samples. The digitized ADC data, ToT and ToA values are stored in the raw format on the disk for offline analysis. During the data-taking run, the information from all the silicon cells are stored.

The raw file is first unpacked and is created into an intermediate file in the CMS standard *Event Data Module* (EDM) format using the CMSSW framework. During the unpacking, each cell is assigned (x, y, z) coordinate based on electronic map. The electronic map contains the location of each cell center with the IDs of corresponding chip and channel. This intermediate file is called *DIGI* file and contains HG and LG ADC data (of all 13 time samples) as well as ToT and ToA values along with their physical location in the detector for each event, called *rawhits*. In the next step, pedestal and common-mode noise is estimated

and subtracted for each of the cell.

As mentioned in section 5.3, in the absence of any traversing particle, silicon cells show a DC shift in the ADC data distribution, called *pedestal*, and have fluctuations about this DC shift, called *pedestal noise*. The pedestal and noise is the result of multiple sources which includes thermal fluctuations in the silicon sensor, picked up stray signals, and accumulated electronics noise in the DAQ chain. Due to its random nature, it manifests as the Gaussian ADC distribution with mean at pedestal value and the width corresponding to noise. The noise itself is the sum of following two components:

- **Intrinsic noise:** Thermal fluctuations in the silicon sensor. The intrinsic noise is the inherent characteristic of a silicon cell and it is uncorrelated between any two sensors.
- **Common-mode noise:** It can originate from the power source or nearby electromagnetic interference. The common-mode noise, although random in nature event-by-event, is correlated across different sensor cells in a given module that can be estimated and subtracted from the ADC data.

In order to bring baseline of ADC distributions to zero, it is important to estimate and subtract the pedestal level. The method for pedestal and common-mode noise estimation is described as follows:

Pedestal estimation: In the ADC data from 13 time-samples, the 1^{st} time-sample (TS0) usually remains signal-free due to the time lag between trigger generation and reception, while the signal peaks in 3^{rd} or 4^{th} time-sample [111]. Since the noise and signal are uncorrelated quantities, it allows us to use ADC data from TS0 to estimate the average pedestal value of a given cell. The pedestal is estimated by taking the median of TS0 population for each run. The pedestal is estimated for HG and LG separately which is calculated as follows:

$$Ped_{i,j}^{\mathrm{HG/LG}} = \mathrm{Median}\left\{\mathrm{ADC}_{TS0}^{\mathrm{HG/LG}}:\ \mathrm{ADC}_{TS0}^{\mathrm{HG/LG}} \in \text{Run } i \text{ and cell } j\right\} \tag{6.1}$$

where $Ped_{i,j}^{\mathrm{HG/LG}}$ is the pedestal value for i^{th} run and j^{th} cell for HG or LG channel, and $\mathrm{ADC}_{TS0}^{\mathrm{HG/LG}}$ is an element of TS0 ADC data population of HG or LG channel for corresponding run and cell. During reconstruction, these median values are subtracted from all the time-samples in each channel for each event in order to bring the DC level to zero.

Common-mode noise estimation: A dedicated study using beam test data from October 2018 beam test campaign shows that pedestal subtracted ADC data from different cells, of a given gain channel and time-sample, are 30% to 90% correlated in an event [111]. However, no correlation is observed between different time-samples of a channel in an event. Similarly, no correlation is observed among same channels of different modules. This indicates that the common-mode noise is "common" for all cells within same module and same times-sample. Therefore, common-mode noise is estimated from channels in a single module for each time-sample separately, on an event-by-event basis. The estimation of common-mode (CM) noise is described in the expression 6.2 below:

$$\begin{aligned} CM_{i,j,k}^{\mathrm{HG/LG}} = \mathrm{Median}\{\mathrm{ADC}^{\mathrm{HG/LG}} : (\mathrm{ADC}^{\mathrm{HG/LG}} \in \text{TS } i, \text{module } j, \text{event } k, \forall \text{ Si cells}) \\ \text{and } (\mathrm{ADC}^{\mathrm{HG/LG}} < 2 \text{ MIPs})\} \end{aligned} \tag{6.2}$$

where $CM_{i,j,k}^{\mathrm{HG/LG}}$ is common-mode noise for HG or LG ADC for i^{th} time sample of j^{th} module for k^{th} event. The $CM_{i,j,k}^{\mathrm{HG/LG}}$ is then subtracted from all cells. The cells that have ADC counts corresponding to more than 2 MIPs are not considered for common-mode noise estimation in order to ensure that there is no *"signal"* contamination during estimation. The ADC data is converted into physics scale of MIP equivalent of energy deposit in a silicon cell (discussed in section 6.8). Typically, 40 HG ADC counts and 5 LG ADC counts correspond to 1 MIP for HG and LG channel, respectively.

After pedestal and common-mode noise subtraction from HG and LG channel, the ADC data from 13 time samples are fitted with function to extract the pulse. Since all the CE-E and CE-H channels were readout in the beam test experiment for offline analysis, a pre-selection based on prior knowledge of typical pulse shape corresponding to an ionizing particle traversing silicon cells, is applied to reject noise-like hits. This helps in reducing the computing resources as most of the channels do not have real particles passing through it (called *zero-suppression*). The expected signal pulse-shape from HG channel is an asymmetric bipolar pulse which is positive for first few time-samples and negative for next few time samples as shown in Figure 6.15. After reaching the peak, the pulse starts to diminish and undershoots $y = 0$ axis and finally settles to zero. The undershoot is a consequence of using CR-RC pulse shaping network which makes the pulse to go below zero. The exact pre-selection cut applied on HG ADC data is defined as follows:

$$\begin{aligned}\text{preSelection} = [(TS_3 > 500 \times \text{maxADCscale}] \ \texttt{OR} \ [(TS_1 < TS_3) \ \texttt{AND} \ (TS_3 > TS_6) \ \texttt{AND} \ (TS_4 > TS_6) \\ \texttt{AND} \ (TS_3 > 20 \times \text{maxADCscale})]\end{aligned} \tag{6.3}$$

where TS_i stands for HG ADC counts in i^{th} time sample and and maxADCscale = 1 for 300 μm silicon cells and 2/3 (~ 0.67) for 200 μm silicon cells. If the ADC count in TS_3 is too high or ADC counts in neighbouring time samples follow the typical pulse trend, the pre-selection flag is set to `True` for the channel otherwise it is rejected from the further analysis.

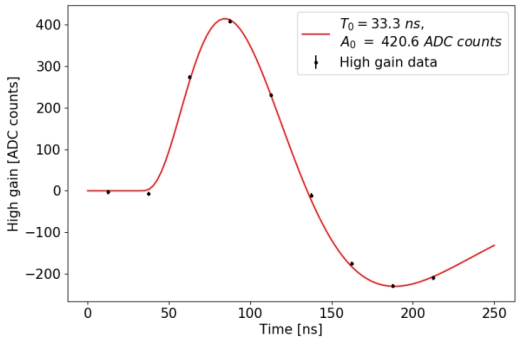

FIGURE 6.15: Example high gain pulse shape for a channel corresponding to 300 GeV positron shower [111].

For each cell passing the pre-selection cut, ADC counts from all 13 time samples are used to reconstruct the pulse for each event, shown in Figure 6.16. The pulse is fitted with the following function:

$$S(t) = \begin{cases} A_0 \left[\left(\frac{t-t_0}{\tau} \right)^n - \frac{1}{n+1} \left(\frac{t-t_0}{\tau} \right)^{n+1} \right] e^{-\alpha(t-t_0)/\tau} & \text{if } t > t_0 \\ 0 & \text{otherwise} \end{cases} \tag{6.4}$$

where A_0 is the amplitude of the pulse. The values τ, n and α are calculated in laboratory based tests [111]. Other parameters are let to float during the fitting procedure. The positive maximum of the fitted function is extracted as the *signal amplitude*, separately for the HG and the LG stages of each cell.

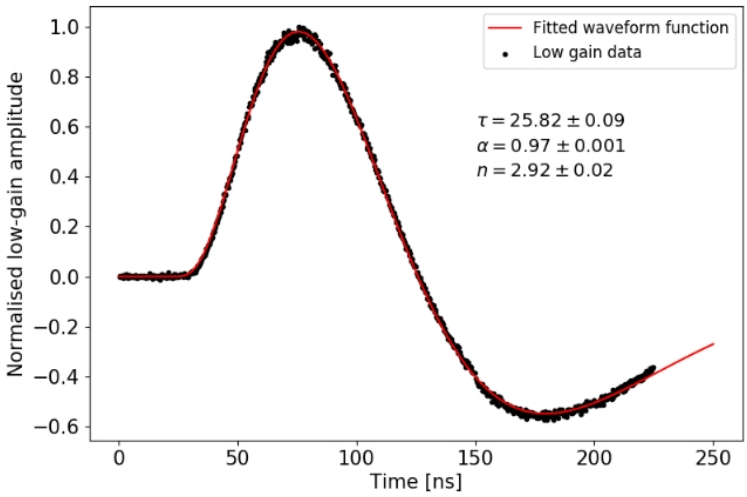

FIGURE 6.16: Pulse shape fitted with function [111].

The HG, LG and ToT stages are intended to be operated in different signal regime to cover a large dynamic range. An intergain calibration procedure is used to combine the HG, LG and ToT signal amplitudes for the energy measurement across the intended large dynamic range. For the beam test experiment, two methods were adopted to achieve this: data-driven method and charge injection method which are described below.

Charge-injection method: The charge injection method is based on injecting a charge of known value to each channel of SKIROC2-CMS, and use the digitized HG, LG and ToT signal thus obtained to find the intergain calibration constants. A dedicated test stand was fabricated to perform system tests in the laboratory which was equipped with an inbuilt 12-bit digital-to-analog converter (DAC). For each channel, 1000 values of increasing charge values are injected in which the DAC is programmed to send voltage pulse with a linear increment for each event, to an on-board capacitor (on the silicon sensor module) that converts the voltage to charge. The HG and LG signal amplitude along with ToT is plotted against event on (x-axis) as shown in Figure 6.17 (left). The figure shows that HG, LG and ToT increases with increasing event (or injected charge). The x-axis is converted to equivalent MIP value using the slope of HG with event on x-axis and the HG to MIP conversion factor (C_{MIP}) for corresponding channel. The x-axis is transformed from event ID to injected charge (in MIPs) as shown in Figure 6.17 (right).

To obtain inter-gain calibration factors, a straight is fitted to HG and LG ADC data in the linear region. The ToT data is fitted with following function that also accounts for non-linearity for low injected charge and turn-on threshold.

$$ToT = \begin{cases} 0 & x \leq ToT_{thres} \\ p_0 + x \cdot p_1 + \frac{c}{x^{\alpha} - ToT_{thres}} & \text{Otherwise} \end{cases} \tag{6.5}$$

where p_0 is the ToT offset, p_1 is the slope of linear region, and the last term represents non-linear part of ToT curve. All the parameters are let to float while fitting the function. The saturation point (HG_{sat}, LG_{sat}) for a given gain stage is defined as the value at which

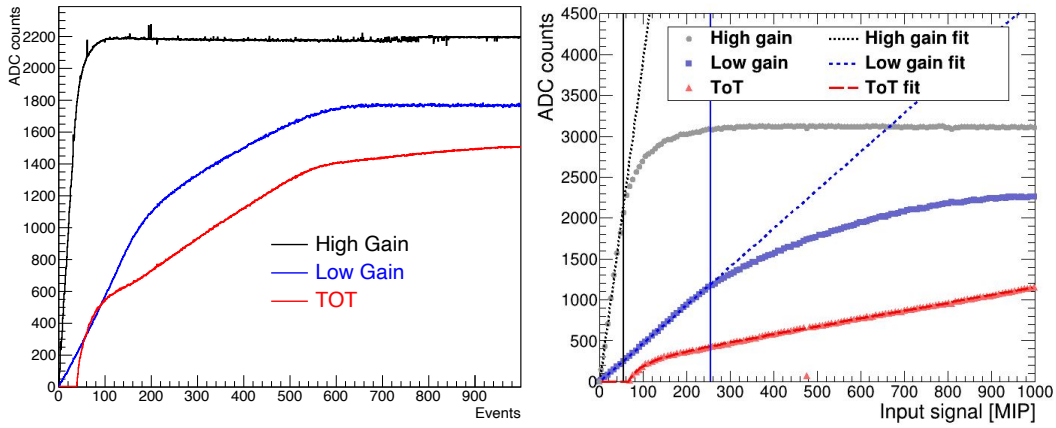

FIGURE 6.17: High-gain, low-gain and ToT ADC counts (in different coloured markers) as a function of events taken during the run (left) and injected charge in terms of MIPs (right). In the rigt plot, a straight line is fitted to high-gain and low-gain ADC data (dotted coloured lines) in order to extract inter-gain calibration factors. The vertical solid lines show saturation point for HG (black line) and LG (blue line).

ADC counts deviates from the fitted straight line by 3%. The slopes of the fitted line (c_{HG}, c_{LG}, c_{ToT}) is extracted and the conversion factor between gains is obtained by taking the ratios as follows: LG-to-HG conversion factor c_{HL} as c_{HG}/c_{LG}; ToT-to-LG conversion as c_{LT} as c_{LG}/c_{ToT}. This procedure is done for all the silicon cells to obtain HG_{sat}, LG_{sat}, c_{HL} and c_{LT} is used to reconstruct the energy according to equation 6.6.

Data-driven method: In data-driven method, in-situ calibration is obtained using 300 GeV e^+ and π^- beam data for CE-E and CE-H respectively. High energy incident particle beam ensures that the shower particles are energetic enough to be able to trigger ToT threshold (i.e. few hundreds of fC charge collected), that can be used for calibration purposes. The HG as a function of LG and LG as a function of ToT is plotted to obtain a correlation among different gain stages. Figure 6.18 shows an example of such correlation plots between HG-LG and LG-ToT for a silicon cell. A straight line is fitted in linear region, which is determined using a spline fit, to obtain inter-gain calibration values. A deviation of HG or LG values from linearity larger that 3% is taken as an indication of saturation which defines the ranges of validity of HG, LG and ToT measurements. The slope of the fitted straight-line is used as the conversion factors between these gain settings such that both LG and ToT counts are expressed in terms of HG amplitude. Finally, HG ADC counts are converted into MIP equivalent of energy deposit (discussed in section 6.8) as shown below:

$$E = \begin{cases} E_{HG} = C_{MIP} \cdot A_{0,HG} & \text{, if } A_{0,HG} < HG_{sat} \\ E_{LG} = C_{MIP} \cdot C_{HL} \cdot A_{0,LG} & \text{, if } A_{0,HG} > HG_{sat} \text{ and } A_{0,LG} < LG_{sat} \\ E_{ToT} = C_{MIP} \cdot C_{HL} \cdot C_{LT} \cdot \left(ToT - ToT_{offset}\right) & \text{, otherwise} \end{cases} \tag{6.6}$$

where $A_{0,HG}$ and $A_{0,LG}$, ToT are high and LG signal amplitude and ToT value respectively. C_{MIP}, C_{HL}, C_{LT}, ToT_{offset} are high-gain to MIP, low-gain to high-gain, ToT to low-gain conversion factor, and ToT offset, respectively. The HG_{sat} (LG_{sat}) is high (low) gain saturation point.

Using data-driven calibration method, a total of $\sim$ 17% of HGCAL prototype silicon cells are fully calibrated, of which $\sim$ 44% are from CE-E and $\sim$ 5.5% cells are from CE-H

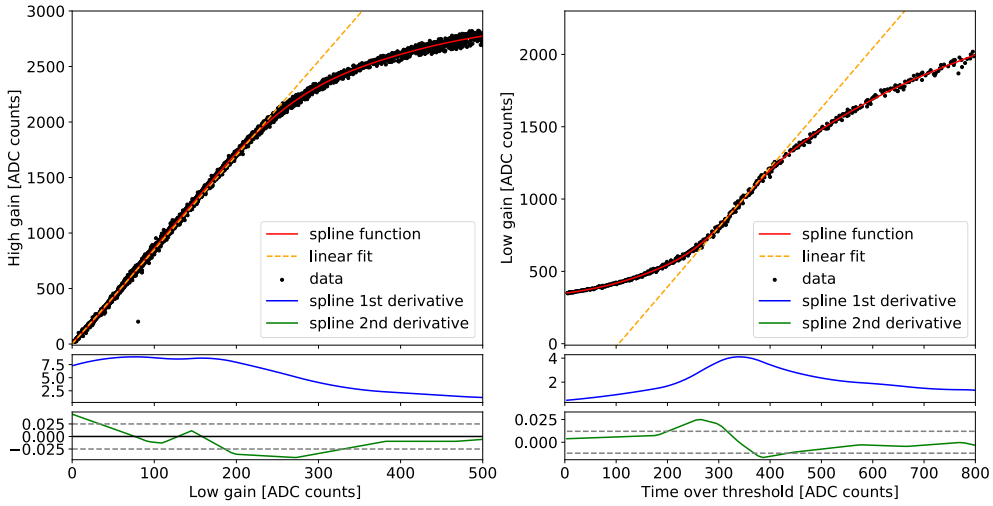

FIGURE 6.18: Correlation between high and LG is shown in the left plot and LG and ToT is shown in the right plot. A straight line is fitted in the linear region of overlap to extract inter-gain calibration constants [111].

prototype depending on which cells dominantly get exposed to shower particles. An average calibration value per ASIC is used for the cells located at the periphery or off-centre modules which were not exposed to particles, or a global average is used if none of the cells fired for that ASIC. Figure 6.19 shows a visualization of the location of cells which are fully calibrated using data-driven method. The visualization indicates that the population of the cells that are fully calibrated are in central modules in both. We do not expect ToT to fire in pion showers except in the core of high energy showers. The pedestal noise across the module have been found to be stable [111], hence using average in CE-H is good enough for our purpose.

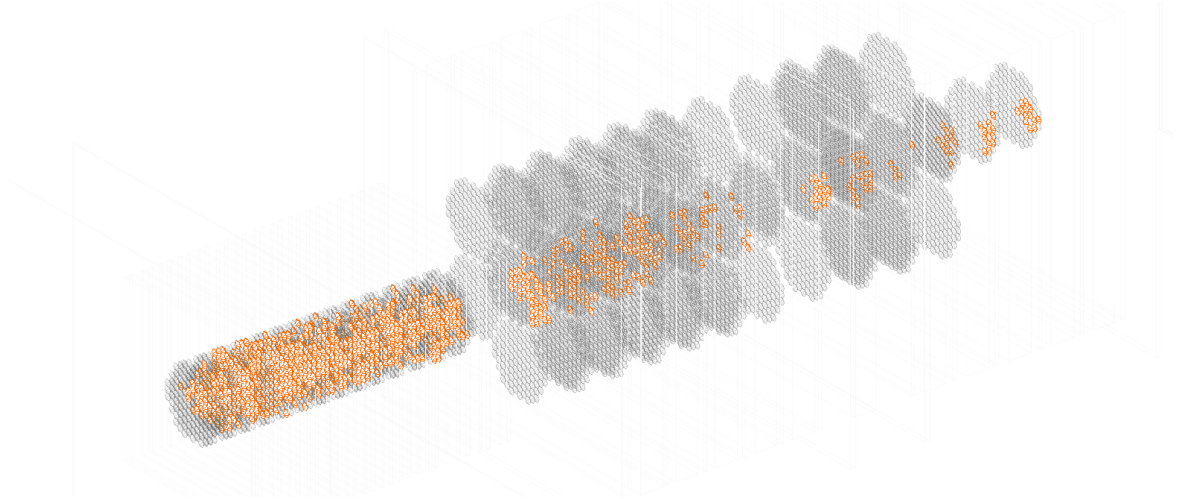

FIGURE 6.19: The yellow-coloured cells show the location of silicon cells that were fully calibrated using data-driven method. The cells lie almost near the center of the layer in both CE-E and CE-H prototype where the core of the shower deposits its energy.

As shown in Figure 6.19, the inter-gain calibration can not be achieved for all the cells using data-driven method. Similarly, for the foreseen HGCAL detector, data-driven calibration will not be possible for all the cells and modules. Therefore, the inter-gain calibration factors shall be obtained using charge injection method.

The inter-gain calibration obtained using data-driven method show better energy linearly for e^+ [118], therefore the results presented in this thesis uses data-driven inter-gain calibration, the option agreed upon by the collaboration. However, more studies to better understand the charge injection calibration are continuing.

6.5.2 Reconstruction of simulated events in HGCAL

A reliable simulation framework for the detector as well as particle interaction with the detector, is an important ingredient for the experiments to perform physics data analysis. The beam test experiments provide a good opportunity to validate the simulation framework and make improvements in the physics modelling of various interactions in the detector. For October 2018 beam test experiment, the events were simulated using GEANT4 (described in section 2.7) toolkit integrated in CMSSW software.

The detector geometry is simulated using GEANT4 version 10.4.3 with corresponding absorber and active material in CE-E, CE-H and AHCAL sections. Figure 6.20 shows simulated HGCAL + AHCAL experimental setup for configuration-1. The materials, absorber and active layer thicknesses are taken from the beam test experiment, in order to accurately simulate detector geometry. The silicon (for CE-E and CE-H) and scintillator tiles (AHCAL) are made to be *"sensitive"* elements, i.e. the energy deposited by the particles in these materials are recorded and saved for energy reconstruction.

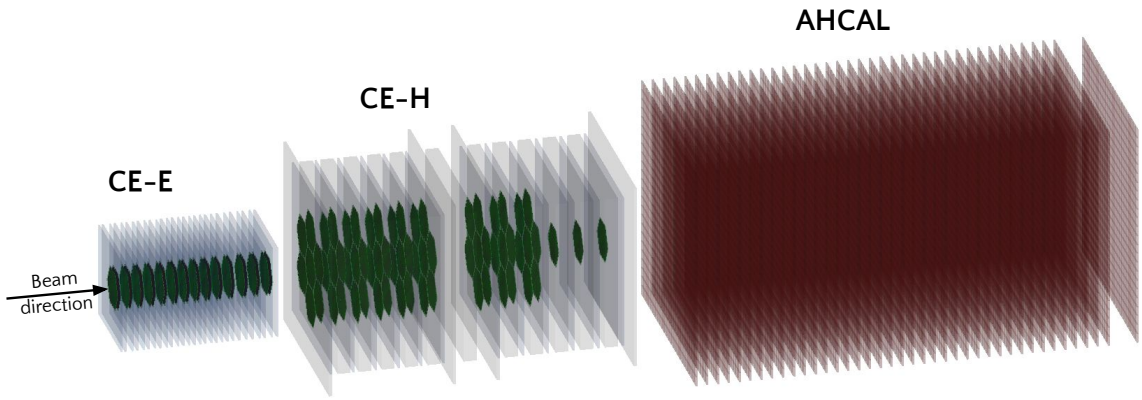

FIGURE 6.20: The simulated geometry for CE-E, CE-H and AHCAL detector prototype.

The material of various detectors and beam optics, instrumented in H2 beamline that are installed upstream of HGCAL + AHCAL experimental setup, affects the beam of single particle reaching the calorimeter setup. The particles interact with the beamline elements such as dipole and quadruple magnets, beam pipe, scintillators, DWCs etc., and result in different processes such as particle scattering, synchrotron radiation and shower initiation. It especially affects beam of lighter particles such as e^+ beam. To simulate such effects from beam-line elements, GEANT4 provides a dedicated package, called *G4beamline*, which

physicists can incorporate in the simulation framework. For October 2018 beam test experiment, the beam-line elements were simulated using G4beamline version 10.3. Figure 6.21 shows different beamline elements from the secondary particle production till the experimental setup for beam test. The beam of single particles with momentum ranging from 20 to 300 GeV/c were shot from $\sim$ 600 m upstream of the HGCAL prototype front and is propagated through the beamline elements. The interaction of particle with the beamline elements were simulated with physics list FTFP_BERT_EMZ. The original particle may interact with the beamline elements and may produce secondary particles. All the particles reaching CE-E front is saved along with their particle IDs, three-momenta and position vector in a ROOT ntuple format.

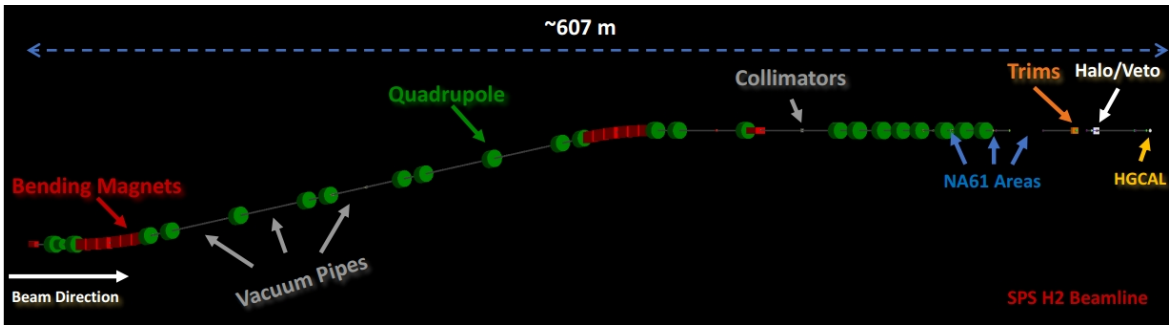

FIGURE 6.21: The image shows various beamline elements upstream the experimental setup simulated using *G4beamline* package of GEANT4 toolkit [119].

The information from full beamline simulation is used as an input to the second step of the simulation process. In the second step, the particle gun is placed at z = 0 cm , i.e. at the front of CE-E. For each event, all the particles arriving at the HGCAL front are shot into the detector setup with their corresponding species, momentum, and (x, y) coordinates. The position of particle gun in the transverse plane, i.e. (x, y) coordinates, is shifted by an offset from the origin and is restricted to 2x2 cm^2 window. Figure 6.22 and 6.23 shows the position of particle-gun in x and y coordinate for charged pion beam simulation. The offset is obtained from data by computing mean value of beam profile in x and y direction (see appendix D), and is introduced in order to reproduce beam profile as close to data as possible. The window of 2x2 cm^2 is used to mimic the acceptance window of the trigger scintillator used during the beam test experiment. It is worth mentioning that in simulation all the active layers are aligned with respect to each other, i.e., the origins of their local coordinates lie along the z axis as opposed to real beam test experiment. In the real experiment, the individual layers are not perfectly aligned and needs to be corrected as described in section 6.7.

For beams of e^+ and μ^- FTFP_BERT_EMN physics list is used, and for the beam of π^- FTFP_BERT_EMN as well as QGSP_FTFP_BERT_EMN physics lists are used. The details of these lists are summarized in section 2.7.

The energy deposited by the shower particles in the silicon are saved in terms of GeV, called *"simHits"*, along with their corresponding *"detectorID"*. The *"detectorID"* determines the location of a particular silicon cell within the sub-detector compartment such as active layer number and (x, y) coordinates of cell center. The detectorID is assigned in an encoded format for each cell that is based on its location in the detector. The detectorID is a unique

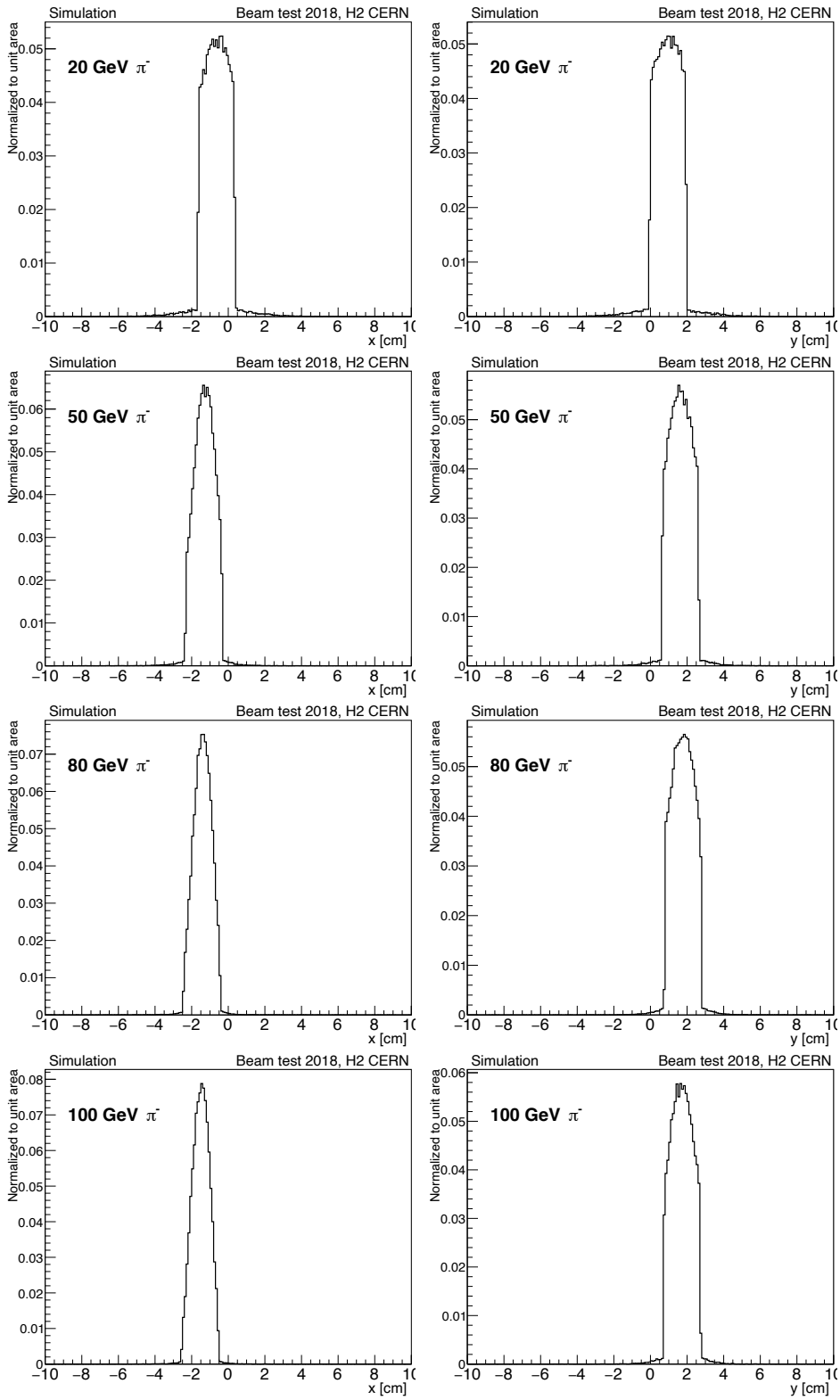

FIGURE 6.22: Position of particle gun in transverse plane (left column: x position, right column: y position) for **20, 50, 80, and 100 GeV** pion simulation.

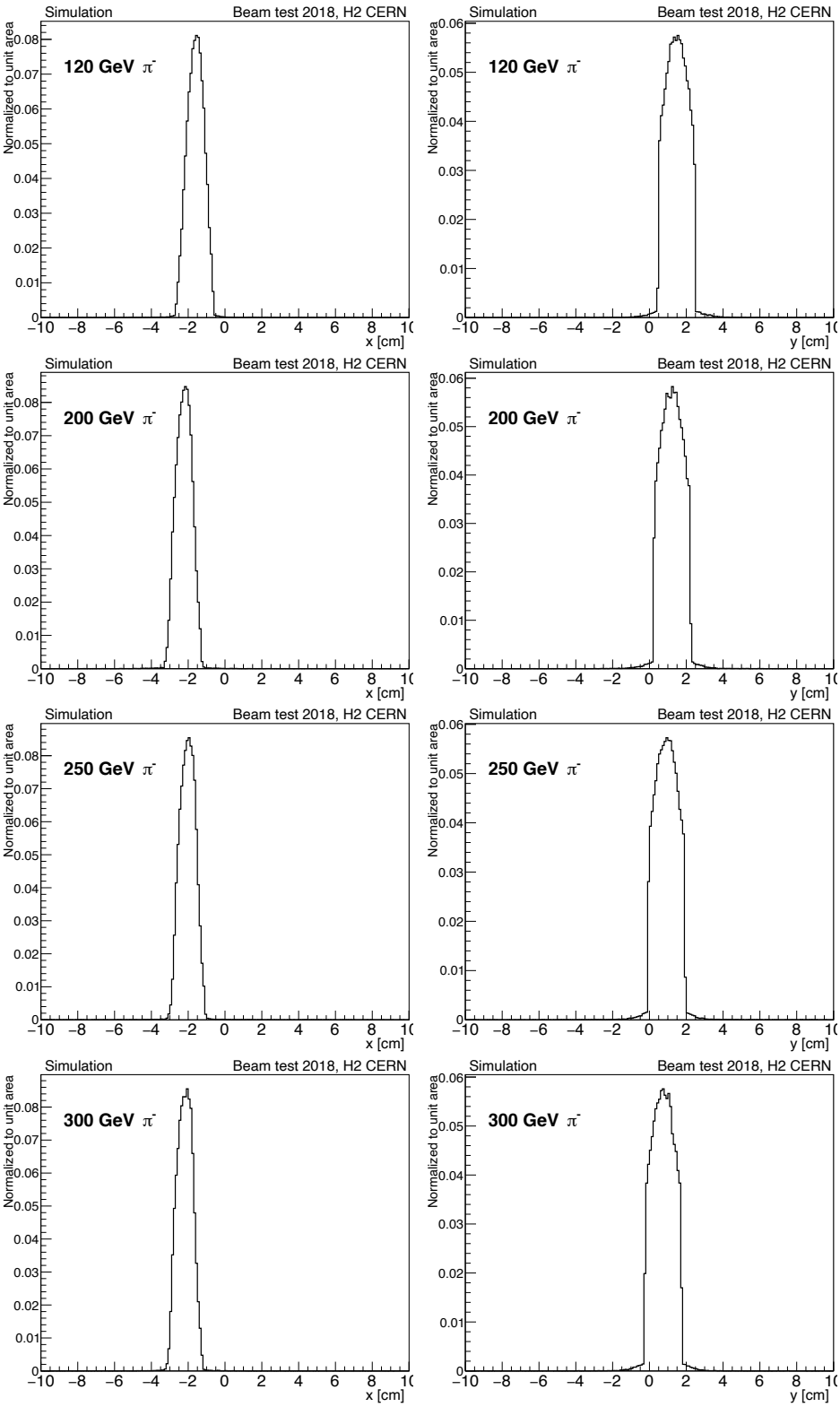

FIGURE 6.23: Position of particle gun in transverse plane (left column: x position, right column: y position) for **120, 200, 200, and 300 GeV** pion simulation.

number for a given cell.

The simhits generated from GEANT4 needs to be reconstructed into a proper format or observables in order to make proper comparison with data. This is done in the following steps: the detectorID corresponding to each simHit in an event is unpacked to obtain the cell location in the different sub-detectors with the help of electronic map. In the simulation setup, *"digitization"* is not implemented i.e. the effects of electronics such as ADC counts corresponding to collected charge, pulse shaping, time-sampling and electronics noise, is not reproduced. Therefore, to convert energy of simulated hits from GeV to MIP, a hard-coded factor is used for silicon. Figure 6.24 shows energy deposited (simHit energy) by 200 GeV muons in silicon cells of different thicknesses in terms of keV. The energy distribution shows Landau shape as expected for energy-deposit distribution in thin detectors, discussed in section 5.4.

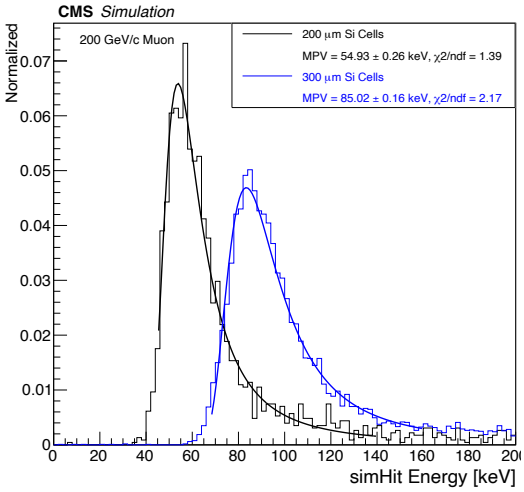

FIGURE 6.24: Simulated Energy distribution for 200 GeV/c muon passing through 300um (blue distribution) and 200um (black distribution) silicon cell with a landau fitted around the core of the distribution

To take electronics noise into account the converted energy is smeared with a Gaussian distribution with $\mu = 0$ and $\sigma = 1/6$ MIP. The smearing value i.e. $\sigma = 1/6$ MIP is taken from the fact that signal-to-noise ratio (discussed in later section) of silicon cells were found to be around 6 - 7 and is a reasonable approximation for the intrinsic noise of silicon cells. The cells through which no shower particles pass, no corresponding simhits are created and hence no smearing is done for those silicon cells. Therefore, simulation sample for beam test experiment is free of any spurious noise.

Smearing Landau distribution with a Gaussian distribution has a non-trivial effect on the final distribution. Even if the μ of the smearing Gaussian distribution is fixed at zero (as is done in our case), the convolution with Landau results in not only widening the Landau distribution, the desired effect, but also shifting the MPV of Landau distribution, an undesired effect. Figure 6.25 shows cell energy in terms of MIPs, using 85 (55) keV per MIP conversion factor for 300 (200) μm as shown in Figure 6.24, before and after applying Gaussian smearing. The MPV of resulting distribution shifts by 4% after applying Gaussian smearing for 300 and 200 μm thick silicon. The shift in MPV is larger if the MPV of parent Landau distribution is larger and/or the σ of Gaussian smearing is larger. To take this effect into account, the GeV-to-MIP conversion factor is corrected by 4% for 300 μm silicon cells i.e. $\simeq$ 89 keV. The GeV-to-MIP conversion factor for 200 μm thick silicon is not corrected as

the MPV lies approximately at one after smearing. Correction for this effect led to improved agreement of energy scales in data and simulation to better than 3% (from initial $\sim$ 7.5%) for electromagnetic as well as hadronic showers.

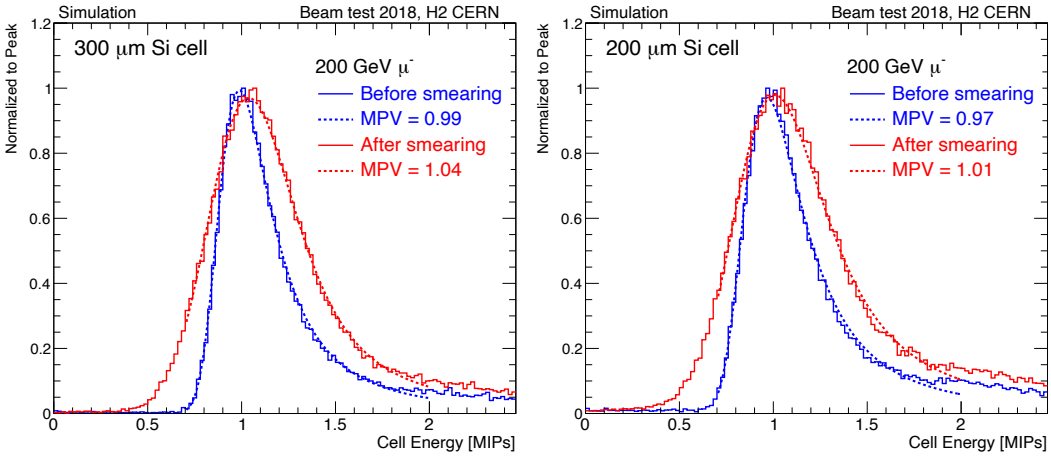

FIGURE 6.25: Simulated Energy distribution in terms of MIPs for 200 GeV muon passing through 300 μm (left) and 200 μm (right). The blue and red curve corresponds to energy distribution before and after applying Gaussian smearing, respectively. The dotted lines correspond to Landau and Laundau convoluted with Gaussian fitted on before and after applying smearing, respectively.

Thus, for each particle beam about 100k events are generated in simulation and the energy is reconstructed as follows:

$$E^i_{MIP} = \left(c_{\text{GeV\_to\_MIP}} \cdot E^i_{GeV}\right) * Gaus(\mu = 0, \sigma = 1/6) \tag{6.7}$$

where E^i_{MIP} is the reconstructed energy in terms of MIPs, $c_{\text{GeV\_to\_MIP}}$ is the GeV-to-MIP conversion factor, E^i_{GeV} is the simhit energy in GeV and converted energy is convoluted with *Gaussian* distribution with mean (μ) fixed at zero and ($\sigma = 1/6$).

6.6 Event reconstruction in AHCAL

The data from AHCAL detector were read out using a copy of coincidence trigger from scintillators as used by the HGCAL DAQ. The beam test data was stored offline for event reconstruction. In the following, the event reconstruction for AHCAL data and simulation is described.

6.6.1 Reconstruction of data events in AHCAL

The very front-end electronics used by AHCAL is SPIROC2E read out chip developed by OMEGA group [120]. It is a low power, low noise, large dynamic range ASIC. Each SPIROC2E ASIC can read up-to 36 SiPM channels. Each channel offers a preamplification stage, followed by pulse shapers of two gains (HG and LG) that provide energy measurement corresponding to 160 fC to 320 pC induced charge at SiPM. The HG shaper is also connected to a fast shaper with a discriminator and acts as a trigger-line for channel-wise self-triggering. The self triggering can be achieved for a signal as small as 50 fC. All channels of the ASIC are read out if a single channel is triggered. The pulse is stored in 16 deep analog memory array and is digitized by a 12-bit Wilkinson ADC which can digitize only

two measurements at a time (among HG, LG and time). During the beam test campaign, the SPIROC2E was operated in auto-gain mode where the ADC data from one of the gain and the timing information was saved. Special calibration runs were taken where both the gain values were stored to determine inter-calibration (*IC*) factor between the two.

The ADC data stored for each event is converted to MIP equivalent of energy deposit, similar to HGCAL event reconstruction, according following prescription:

$$E_i[\text{MIP}] = \frac{(A_i[\text{ADC}] - Ped_i[\text{ADC}]) \cdot IC_i}{MIP_i} \times f^{-1}_{sat,i}(A_i[\text{pixel}]) \tag{6.8}$$

$$A_i[\text{pixel}] = \frac{(A_i[\text{ADC}] - Ped_i[\text{ADC}]) \cdot IC_i}{Gain_i} \tag{6.9}$$

where $A_i[\text{ADC}]$ is the ADC data of measured gain of i^{th} cell, $Ped_i[\text{ADC}]$ is the pedestal value of i^{th} cell, IC_i is the inter-calibration value, MIP_i is ADC to MIP conversion factor for i^{th} cell. The $f^{-1}_{sat,i}$ is the inversion function to account for saturation of number of pixels fired in a SiPM. The f is parametrized by an exponential function that asymptotically reaches the value of maximum number of pixel in the SiPM, i.e. 2668 pixels. For very small values of fired pixel, the f is practically one. The $A_i[\text{pixel}]$ is the number of fired pixel calculated using factor $Gain_i$ which is essentially the ADC counts corresponding to a single fired pixel. This value depends on the SiPM specification.

The MIP calibration factors are determined using dedicated muon beams of 40 and 120 GeV extracted from SPS at CERN. The cells through which muons are not traversing, are free of "*signal*", therefore the pedestals are also estimated from muon runs. By analysing the beam test data, the MIP conversion is found to be ~ 240 ADC counts, the inter-calibration is ~ 20, the *Gain* is ~ 16 ADC per pixel. For a single muon traversing through the tile, about 15 pixels are found to be fired. The cells with energy less than 0.5 MIP are not saved for offline analysis, hence, not considered in the event reconstruction. More details can be found [112].

6.6.2 Reconstruction of simulated events in AHCAL

The AHCAL and HGCAL events are simulated using common simulation setup as described for HGCAL simulation. The simHit energy and cell ID for AHCAL is stored in conjunction with HGCAL cells. The event reconstruction in simulation, however, differs from HGCAL simulation reconstruction. The cell ID is *unfolded* using dedicated AHCAL electronic map. In the reconstruction framework, the digitization step is implemented to take care of various detector effects, and is explained briefly in the following.

The simHit energy is converted from GeV to MIP using hard-coded GeV-to-MIP conversion factor. This factor is obtained by shooting 40 GeV muons on 3 mm scintillator tile. The corresponding energy is found to be ~ 0.48 MeV per MIP and is used as GeV-to-MIP conversion factor for AHCAL, referred to as `simHitMIP`. The converted `simHitMIP` is used for the first estimation of fired pixel using the fact that ~ 15 pixels are fired corresponding to a MIP, as mentioned in previous section. This, however, does not take the finite number of SiPM pixels (m) into account, therefore the saturation function (f_{sat}) is applied to calculate mean number ($< n >$) of fired pixel. The pixel fired in SiPM corresponding to incident photon is subject to statistical fluctuation. This stochastic effect is modelled by calculating the expected number of fired pixels (n) using Binomial distribution $B(n_B, p_B)$, where $n_B = m$ and $p_B = < n > /m$. A further Gaussian smearing with mean = 0 and rms = 0.55, is applied to account for electronics noise. The resulting final number of fired pixels is converted to

corresponding ADC counts, referred to as `simHitADC`, using *Gain* value as mentioned in previous section. In simulation, the inter-calibration is assumed to be one, i.e. only HG is used and no inter-calibration factor is required. However, a protection against very large energy deposit is applied by limiting the maximum value that `simHitADC` can have. The rest of energy reconstruction scheme is same as data by using `simHitADC` as input to equation 6.8 with $IC = 1$.

To check the validity of simulation energy reconstruction, the energy distribution of MIP is compared between data and simulation. Figure 6.26 shows example of MIP energy distribution comparison between data and simulation in CE-E, CE-H and AHCAL layers. The figure shows that the MIP peak lies at one and the overall energy is well reproduced by simulation. However, the distribution width is slightly wider for CE-E and CE-H in simulation as compared to data. This indicates the Gaussian smearing applied in simulation a bit aggressive for the HGCAL cells. With a more realistic digitization implementation in HGCAL simulation, the agreement is expected to improve.

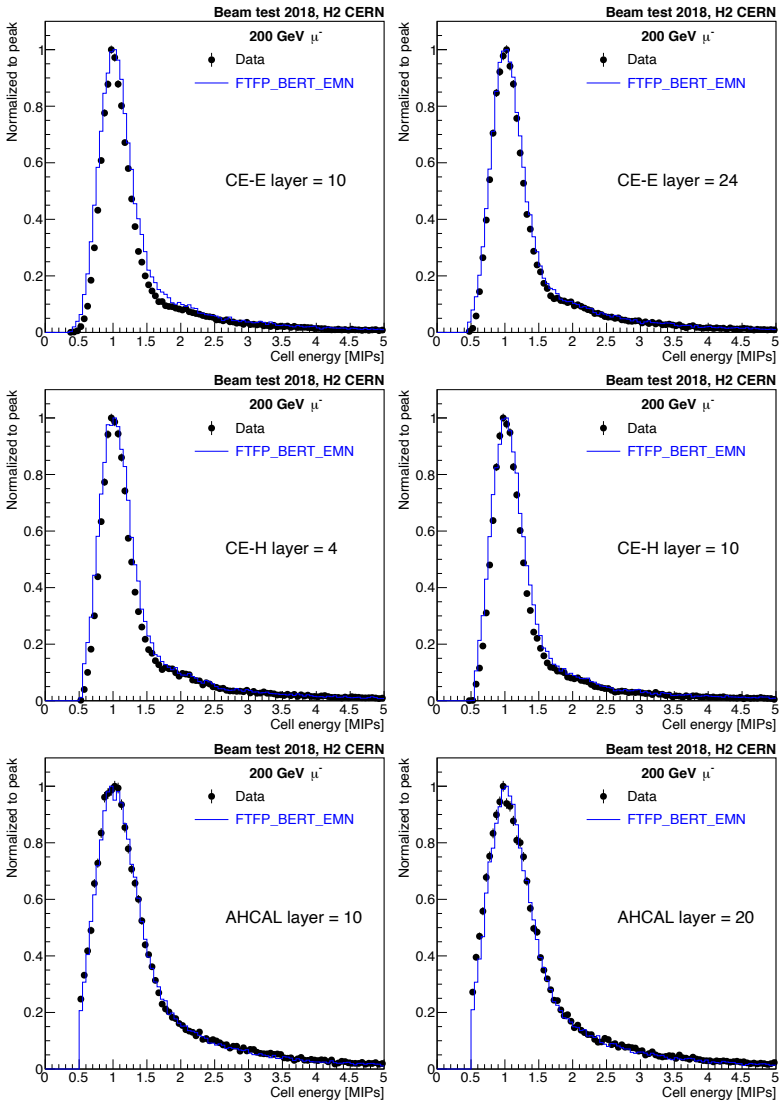

FIGURE 6.26: MIP energy distribution comparison between data (black points) and simulation (blue line) in different layers of CE-E (top row), CE-H (middle row) and AHCAL (bottom row).

6.7 Detector alignment using muon beam

All the detectors that were employed during the beam test experiment such as delay wire chambers, HGCAL and AHCAL prototype have their own local coordinate system. It is important to make sure that these coordinate systems are aligned with each other otherwise it may result in incorrect measurement of particle track and hit location in the detector as well as the shower axis.

The coordinate system of DWCs are defined with respect to the center of the DWC as origin. Similarly, the coordinate system of each layer of HGCAL and AHCAL prototype are defined using the center of the central cell as the origin. During the experimental setup installation at H2 beamline, efforts were made to properly align different detectors as precisely as possible using commercially available hand-held laser system. This, however, does not guarantee a proper alignment of all the detector systems. Hence, the alignment is done at the analysis level by correcting the hit position (by appropriate coordinate transformation) according to the level of difference between the detector coordinates, if any.

The alignment correction factors are obtained by measuring the distance between the hit position in the detector and the track, and use this distance as an offset for correction. For this purpose, muon beam data from configuration-1 is used. The reason for using muon beam is because the probability of developing a shower by 200 GeV μ^- is extremely small and it acts as a MIP leaving just one hit in each layer. To begin with, the correlation between hit position in HGCAL and AHCAL layers and the extrapolated impact position of reconstructed track by DWCs, is investigated. Figure 6.27 shows the correlation between track impact position (`track_X`, `track_Y`) and hit position (`hit_X`, `hit_Y`) at the first layer of CE-E, CE-H and AHCAL. The position correlation plot indicates that the DWC and HGCAL (CE-E and CE-H) coordinates are anti-correlated. It was later found to be the result of wiring swap of DWC detector. The coordinate system for AHCAL is opposite to that of CE-E and CE-H layers. By inverting track coordinates for CE-E and CE-H, a correct correlation between track impact and hit positions is obtained as shown in Figure 6.28. The AHCAL layers do not require this correction.

In order to estimate the degree of misalignment, the distance parameter dR between track impact point and hit position for each layer in HGCAL and AHCAL prototype is calculated as defined in equation 6.10 below:

$$\begin{aligned} dX &= x_{track} - x_{hit} \\ dY &= y_{track} - y_{hit} \\ dR &= \sqrt{(dX)^2 + (dY)^2} \end{aligned} \tag{6.10}$$

Figure 6.29 shows two-dimensional distribution of distance between hit position and track impact position at two layers of CE-E, CE-H and AHCAL prototype. The bright spot in Figure 6.29 shows a shift from the origin i.e. $(0, 0)$. This indicates that the level of misalignment between DWC and HGCAL prototype is of the order of a couple of cm. The correction factors is obtained by fitting a Gaussian function on corresponding one-dimensional distribution of dX and dY as shown in Figure 6.30, for each layer of detector.

The mean (μ) of the Gaussian fit is taken as the offset for each layer in x and y coordinates. Using these offsets, the HGCAL and AHCAL coordinates are transformed with

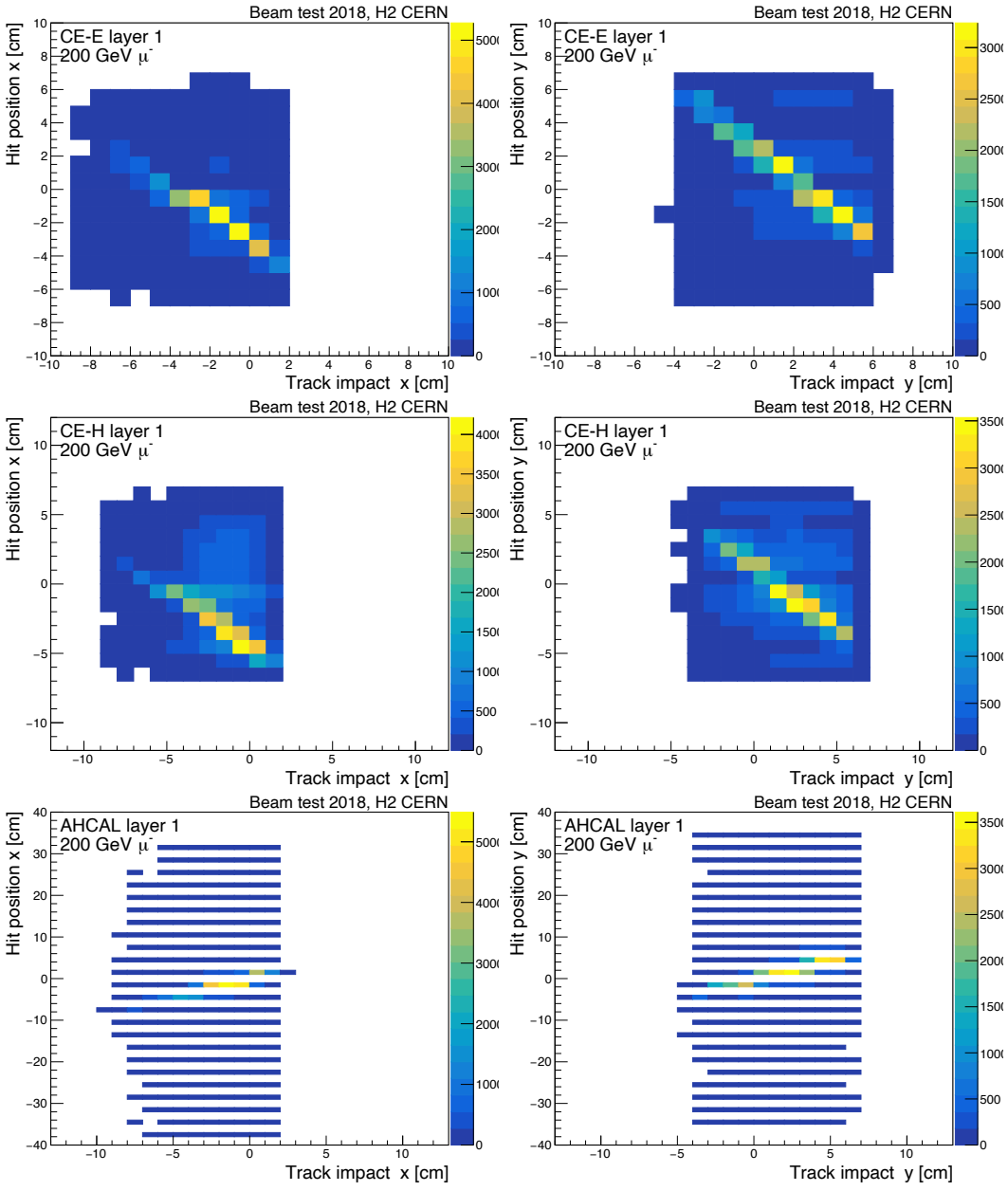

FIGURE 6.27: Correlation plot between track impact position and hit position in x-coordinate (left column) and y-coordinate (right column) at layer 1 of CE-E (top row), CE-H (middle row) and AHCAL (bottom row). The plot shows anti-correlation between DWC and CE-E/CE-H prototype coordinate whereas AHCAL coordinate is already opposite to HGCAL.

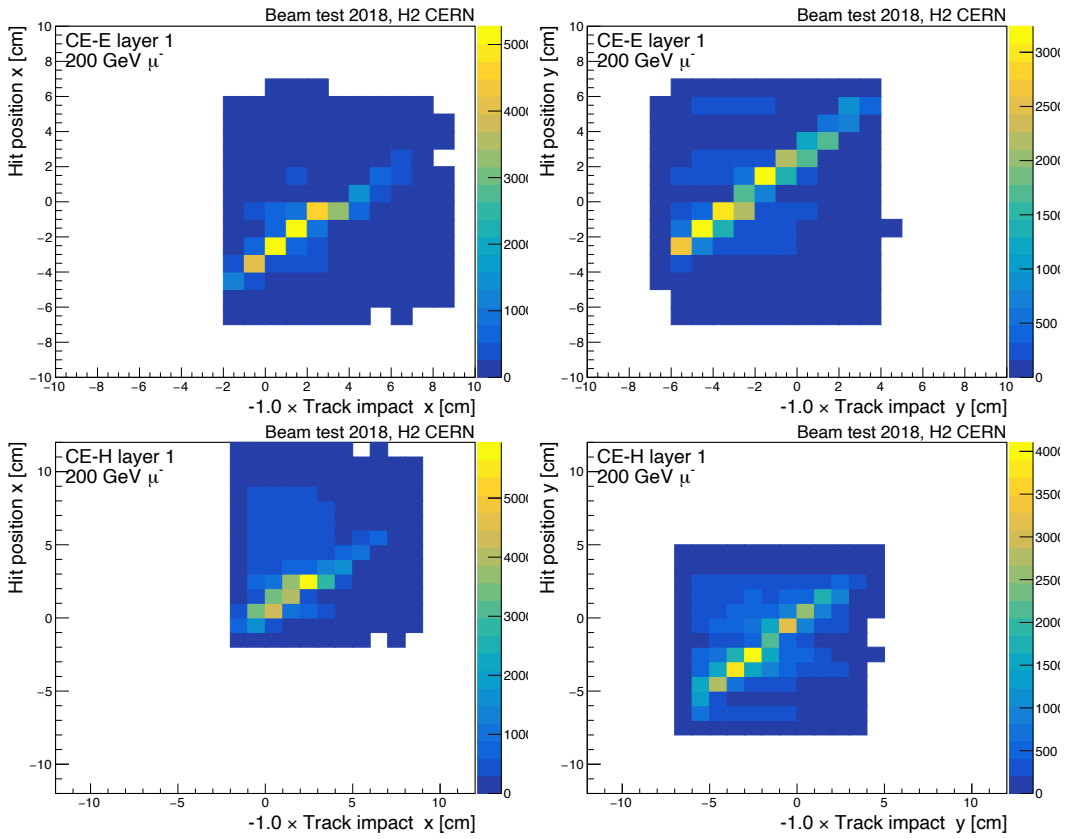

FIGURE 6.28: Correlation between track impact position and hit position in x-coordinate (left column) and y-coordinate (right column) at layer 1 of CE-E (top row), CE-H (bottom row) after inverting track impact position. The anti-correlation is fixed by inverting the track impact positions extrapolated to each layer of CE-E and CE-H layers.

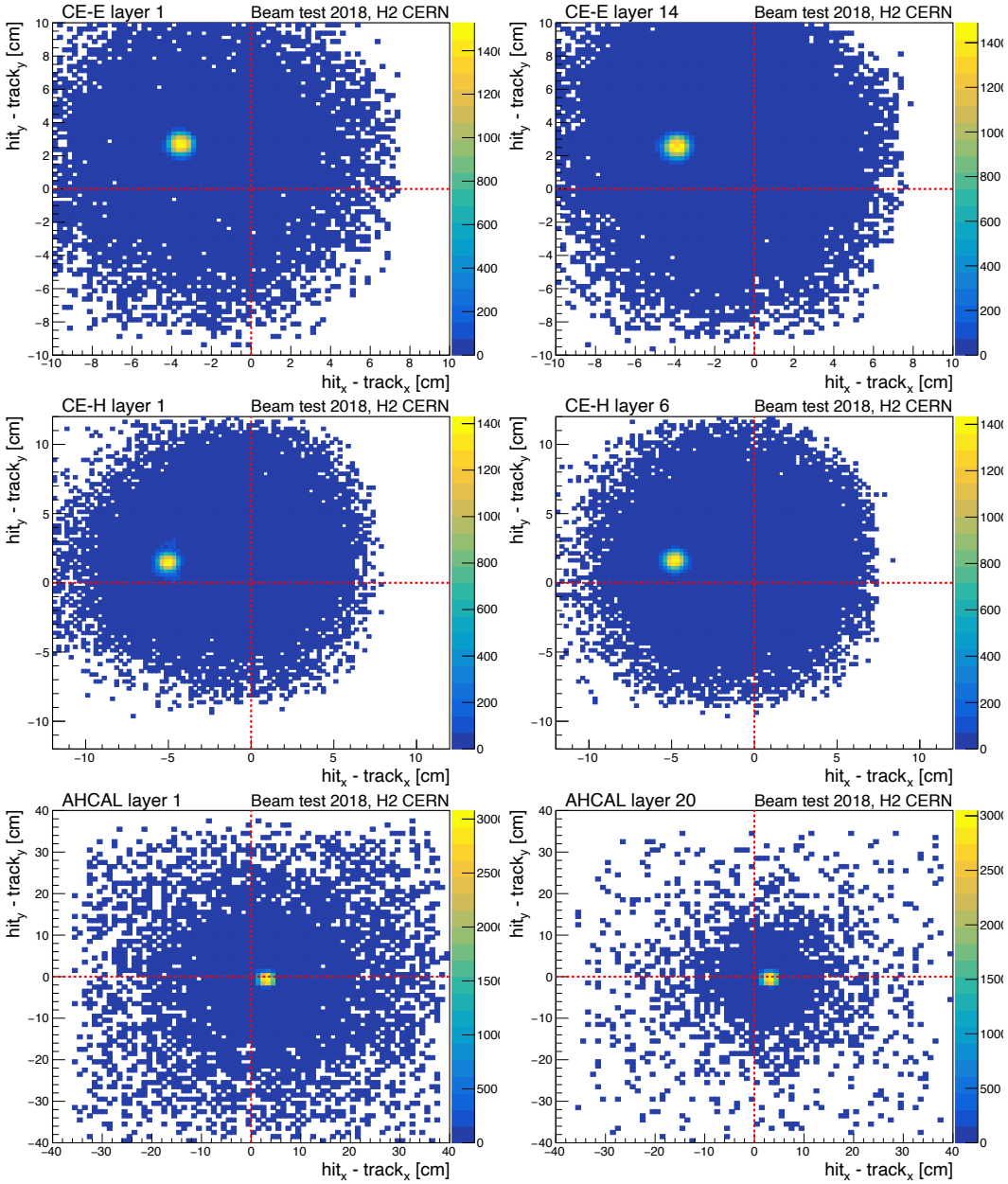

FIGURE 6.29: The distance between hit position and track impact position at example layers of CE-E (top row), CE-H (mid row) and AHCAL (bottom row). The red dotted line shows $x = 0$ and $y = 0$ axis i.e. origin point of corresponding layer.

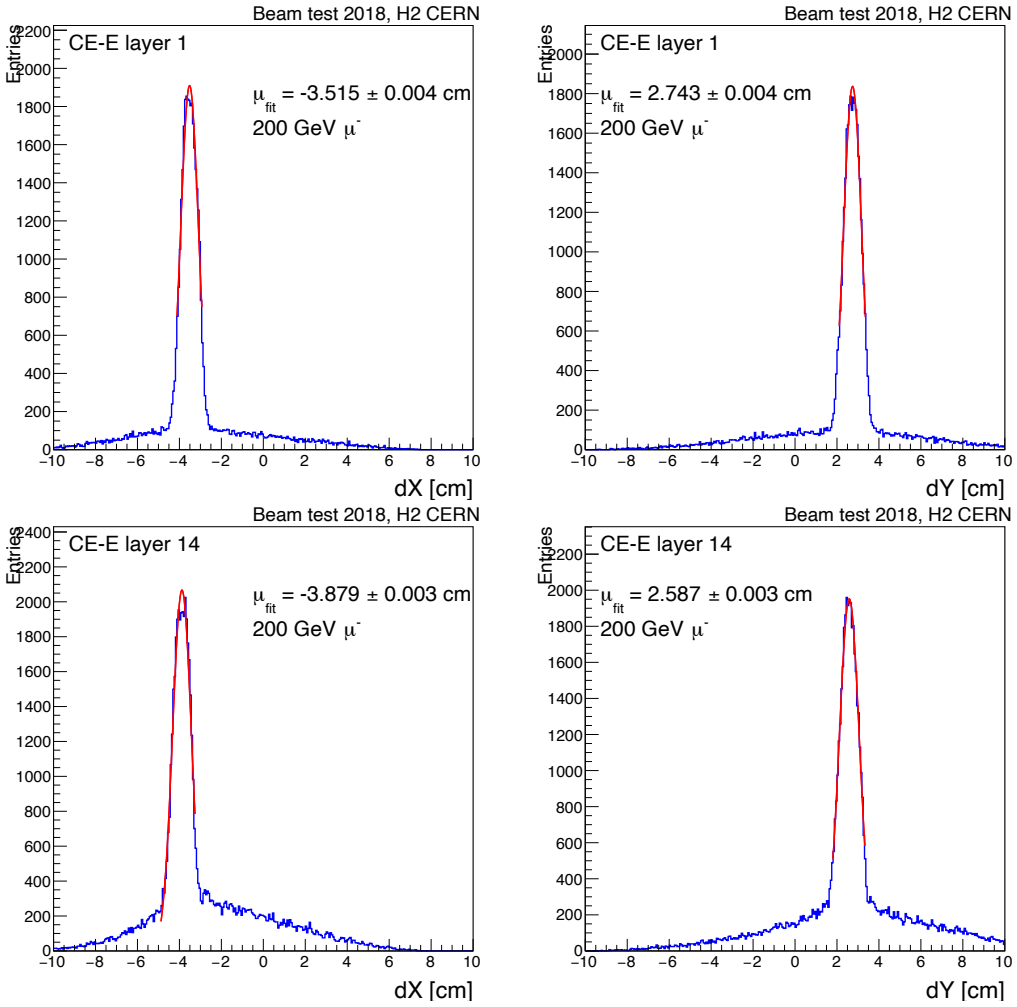

FIGURE 6.30: One-dimensional projection of dX (left column) and dY (right column) (distance between hits and track impact position) for two example layers of CE-E (bottom plots). A Gaussian function is fitted around the core of the distribution to obtain residuals to be used for alignment correction.

respect to DWCs as follows:

$$\begin{aligned} hit_X &\rightarrow hit_X - \mu_X \\ hit_Y &\rightarrow hit_Y - \mu_Y \end{aligned} \tag{6.11}$$

After applying the corrections according to equation 6.11, the distance parameter between track impact and hits restores to zero as shown in Figure 6.31. The alignment correction factors for all the layers are given in appendix C.

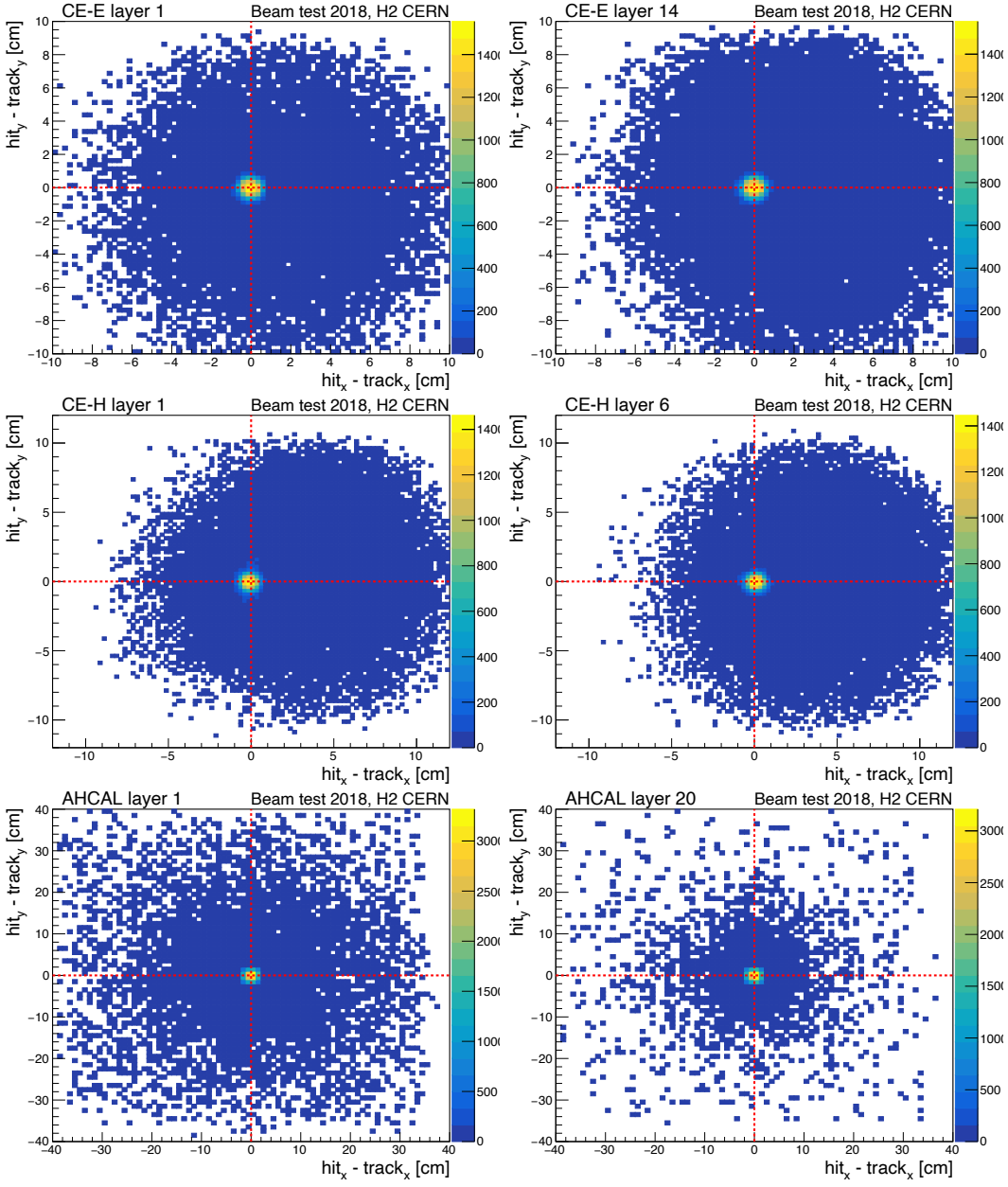

FIGURE 6.31: Two-dimensional distance plot between hit position and track impact position after applying alignment correction for two example layers of CE-E (top row), CE-H (mid row) and AHCAL (bottom row).

6.8 Channel-to-channel response equalization of silicon sensor cells using muon beam

As explained in section 5.4, different sensor cells might respond differently to the same ionizing particle traversing through it. This situation may arise due to different active thicknesses of silicon cells (e.g. 300 μm or 200 μm) which results in different charge collection, or due to different gain settings of the readout chip which results in different ADC counts for same charge collected. The choice of detector technology also contributes to unequal response for different sensors, such as silicon versus scintillator+SiPM. Therefore, it is important to equalize the response using a well understood standard candle which is also easily accessible. For this purpose, minimum-ionizing-particles (MIPs) are used for calibration as they deposits almost similar energy per unit length traversed in the detector (irrespective of active layer location in calorimeter depth) via ionization.

For beam test experimental setup, the dedicated muon beam of momentum 200 GeV/c is used. The required energy for muons to act truly like a MIP is $\sim$ 300 MeV [36]. However, the level of difference in terms of energy deposited by 200 GeV (available muon beam energy for October 2018 TB) and 300 MeV muons are expected to be small. To check this expectation, I made use of a stand-alone GEANT4 framework to simulate energy deposited by different energy muons in individual silicon cell. In this private setup, the only sensing material is a 300 μm silicon cell which is exposed to 200 GeV and 300 MeV muon beam. The comparison of energy deposited in 300 μm by the two muon beams is shown in Figure 6.32. Both the energy distributions are fitted with Landau functions, and MPV of 85.5 keV and 80.1 keV is obtained for 200 GeV and 300 MeV muons, respectively. It is found that MPV for 200 GeV muon is not far-off ($\sim$ 5% difference) as compared to 300 MeV muons. We refer to 200 GeV muon as MIPs in the following.

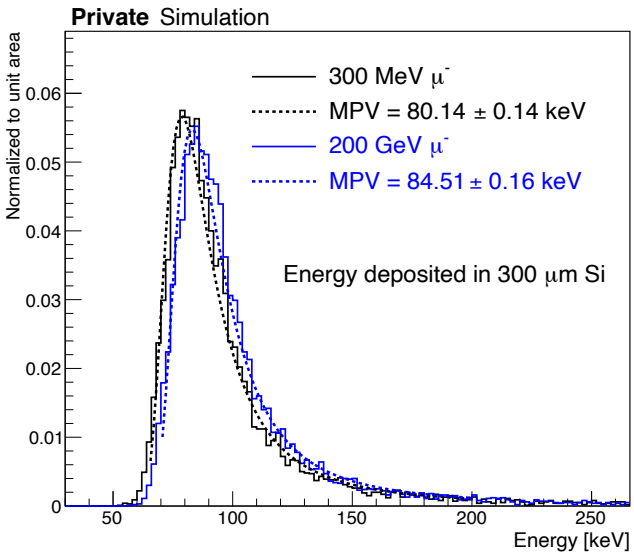

FIGURE 6.32: Comparison of energy deposited by 300 MeV (black line) and 200 GeV (blue line) in 300 μm silicon. The distributions are fitted with a Landau function (dotted line) in order to extract the MPV. The level of difference in energy deposited by two energy muons is $\sim$ 5%.

The strategy for equalizing channel-to-channel response, also referred to *MIP calibration* procedure, is as follows. The muon energy spectrum in terms of HG ADC counts is obtained for each channel. Due to differences in sensors, the MPV of HG ADC counts will be different for different channels. We extract this MPV and use it as MIP calibration factor

by normalizing the energy spectrum such that MPV lies at 1 MIP. In this way, the variation from channel-to-channel is taken away by the MIP calibration constants.

The beam profile of the muon beam, shown in Figure 6.33, is broad enough to cover almost all cells of the layers in CE-E and central modules of CE-H layers. In order to obtain MIP calibration factors, the energy spectrum of MIP is plotted for each cell in terms of HG ADC counts as shown in Figure 6.34. This figure shows MIP energy spectrum for two example HG channels located in CE-E layer 2.

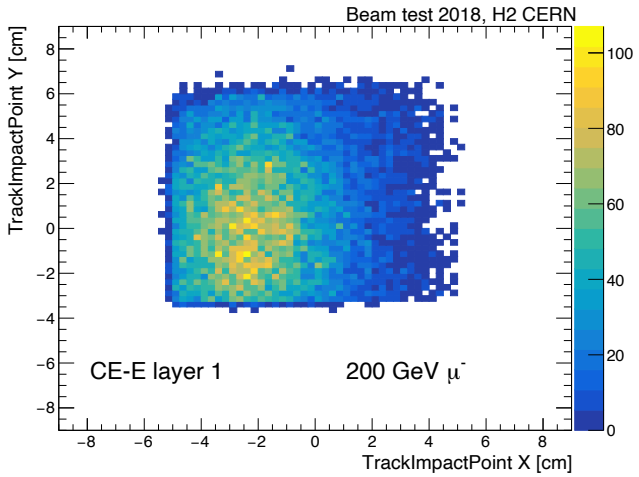

FIGURE 6.33: Beam profile for 200 GeV μ^- beam is shown at CE-E layer 1. The beam profile is wide enough to cover most of cells in the CE-E prototype.

It is observed that HG ADC spectrum shows a peak around 20 HG ADC counts and a second peak around 40 HG ADC counts. Although, there is a pre-selection cut (equation 6.3) applied on the pulse shape with a threshold of 20 HG ADC counts to reject noise-like hits but sometimes upward fluctuations in the intrinsic noise pass the pre-selection cut and will show up in the HG ADC distribution, taken out of the box, as shown in Figure 6.34. Therefore, events corresponding to the first peak are essentially the Gaussian tail of the intrinsic noise. The sharp decline of distribution below 20 HG ADC counts are the artefact of 20 HG ADC counts threshold in the pre-selection cut. The small population below 20 is due to the fact that the pulse is fitted after applying the pre-selection cut, therefore the fitted pulse amplitude can be slightly lower than 20 HG AD counts in some cases.

The MIP energy distributions can be cleaned up by getting rid of events where MIP does not pass through the given cell. In other words, retain the cell only when it has signal-like hit, i.e. signal generated by the muon traversing through the cell. This is done with the help of extrapolated tracks which are reconstructed from four DWCs installed upstream the detector. The cleaning procedure is described as follows:

1. Align track impact position extrapolated from DWC coordinates and rechits from HGCAL prototypes using alignment correction described in previous section.

2. For each layer, calculate the distances dR between the track impact position and the position of all the rechits.

3. Select the rechit with minimum dR, and fill HG ADC count for corresponding cell into a histogram.

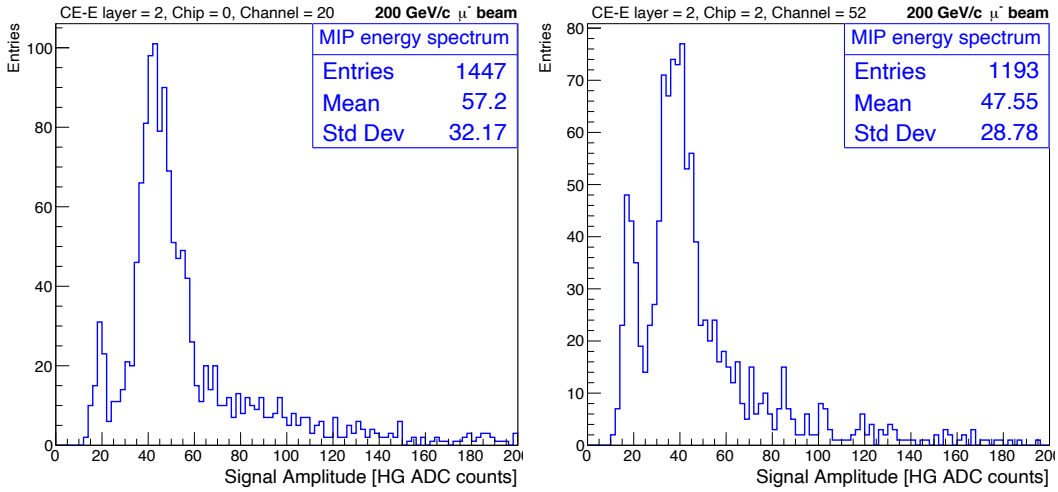

FIGURE 6.34: Figure shows two examples of MIP energy distribution in terms of HG ADC counts.

The dR is defined as follows:

$$dR = \sqrt{(x_{track} - x_{rechit})^2 + (y_{track} - y_{rechit})^2} \tag{6.12}$$

This procedure allows us to select the cell (or hit) through which the muon passes through, thereby rejecting all the other noise-like hits which could have passed the pre-selection cut due to upward fluctuation in the noise. The effect of this cleaning cut is shown in Figure 6.35 on the two example cells located in CE-E layer 2. The figure shows HG ADC counts distributions with and without applying track selection cut. It is observed that the distribution is free of almost all of noise-like hits without affecting MIP distribution.

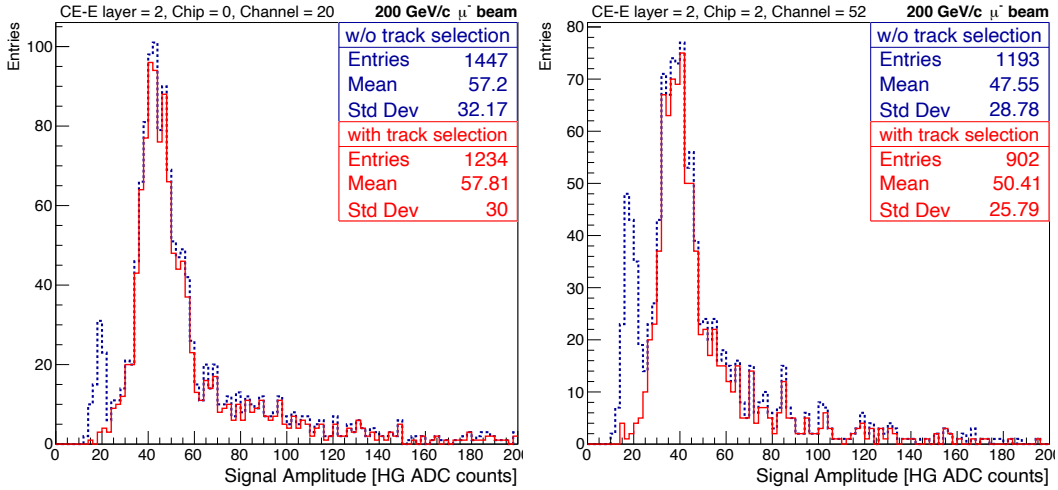

FIGURE 6.35: Figure shows HG ADC distribution before (blue dotted line) and after (red solid line) applying track selection cuts.

After applying the track selection cut, a relatively pure MIP energy spectrum is obtained for each cell. To extract HG ADC to MIP conversion factor, the distribution is fitted with a function involving a Gaussian component (intrinsic noise contribution) and a Landau component (energy deposited by MIP), given as follows:

$$
\begin{aligned}
f(\mathrm{MPV}, \sigma_{noise}, \sigma_{MIP}, a_1, a_2, a_3, a_4) = a_1 \times \mathrm{Gaussian}(0, \sigma_{noise}) \\
+ a_2 \times \mathrm{Landau}_1(\mathrm{MPV}, \sigma_{MIP}) * \mathrm{Gaussian}(0, \sigma_{noise}) \\
+ a_3 \times \mathrm{Landau}_2(2\mathrm{MPV}, \sigma_{MIP}) * \mathrm{Gaussian}(0, \sigma_{noise}) \\
+ a_4 \times \mathrm{constant}
\end{aligned}
\tag{6.13}
$$

The first term $\mathrm{Gaussian}(0, \sigma_{noise})$ of the fit-model corresponds to Gaussian distribution due to intrinsic noise with mean fixed at zero and width as σ_{noise}, second term $\mathrm{Landau}_1(\mathrm{MPV}, \sigma_{MIP}) * \mathrm{Gaussian}(0, \sigma_{noise})$ corresponds to a convolution of Landau with Gaussian distribution to fit first MIP peak of the distribution, third term $\mathrm{Landau}_2(2\mathrm{MPV}, \sigma_{MIP}) * \mathrm{Gaussian}(0, \sigma_{noise})$ corresponds to second MIP peak if two muons pass through a cell. A fourth term, a constant, is also added to the fit model in order to improve χ^2/ndf of the fit. The coefficients a_1, a_2, a_3 and a_4 are the weights assigned to each term.

The fit is performed using *Minuit* minimizer of RooFIT package [121] , while allowing MPV, σ_{noise}, σ_{MIP}, a_1, a_2, a_3 and a_4 to float during the fit. Figure 6.36 shows HG ADC distribution fitted with the fit-model for two representative channels.

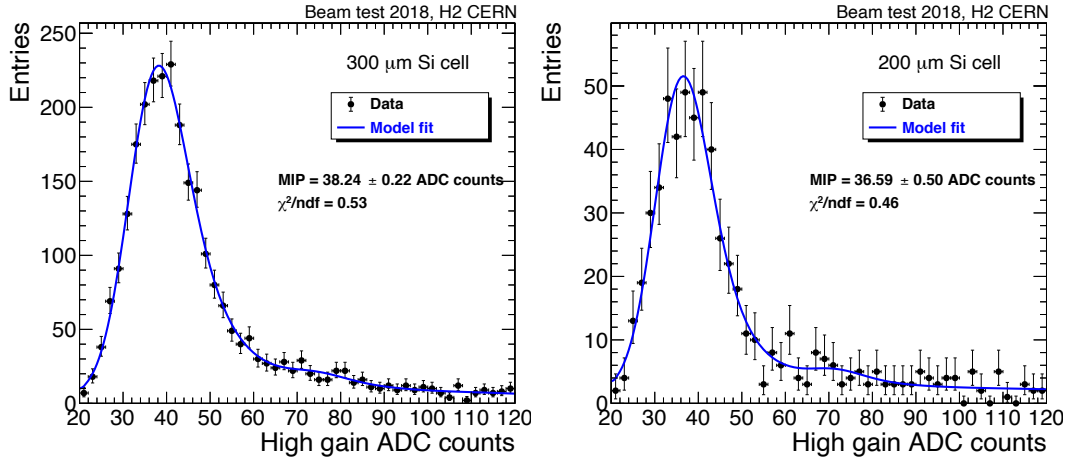

FIGURE 6.36: Figure shows HG ADC distributions fitted with landau convoluted with Gaussian distribution for 300 μm (left) and 200 μm (right) thick silicon cell.

Because of Landau's non-trivial implementation in ROOT, the MPV of fitted Landau does not correspond to true MPV, therefore the maximum of fitted function is used as the HG-to-MIP conversion factor. This procedure is done for all the cells which have enough statistics to perform the fit procedure i.e. histograms with entries greater than 100.

With this procedure, we were able to calibrate most of the channels in the CE-E and the channels in central module of CE-H layers. Figure 6.37 shows a summary plot for HG ADC counts per MIP for a few layers of CE-E prototype. The summary plots show that the HG-to-MIP conversion factor is fairly stable within a SKIROC2-CMS ASIC. Therefore, ASIC average value is used as the calibration constants for the channels which could not be calibrated due to lack of statistics.

The LG channels, although have a lower gain factor than HG channels, also show a good sensitivity to MIP signal. Adopting the similar approach as used for HG channels, MIP energy distributions are obtained in terms of LG ADC counts for silicon cells. Figure 6.38 shows example of MIP energy spectrum in LG ADC counts with the MIP fit-model (equation 6.13) fitted on the data for two silicon cells in CE-E module.

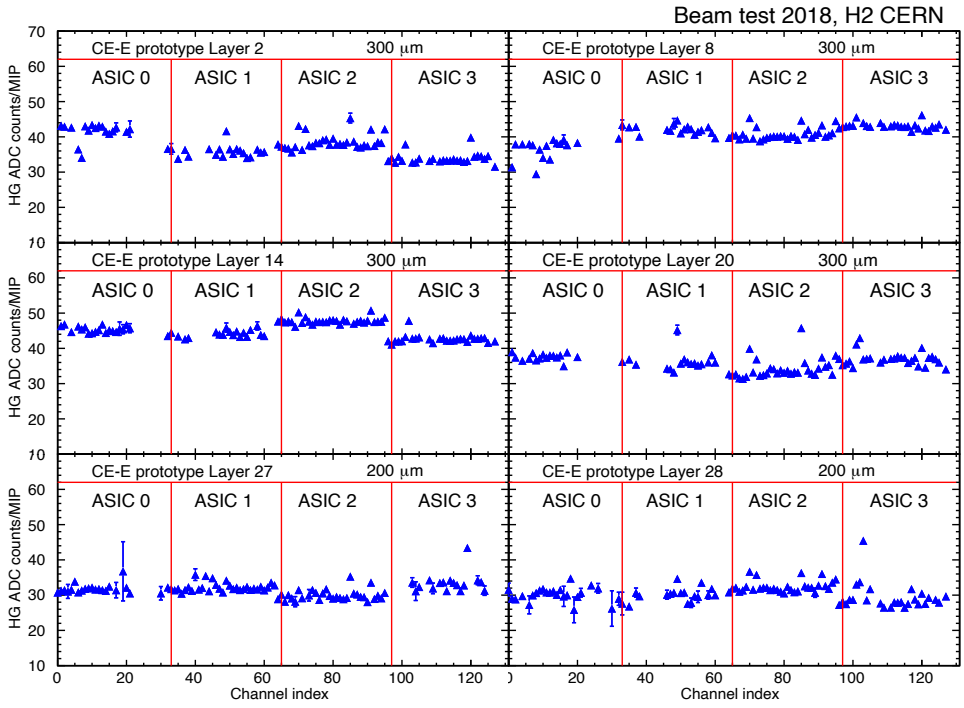

FIGURE 6.37: Summary plot for HG ADC counts per MIP for HG channels for few CE-E layers. Bottom two plots corresponds to 200 μm thick silicon cells.

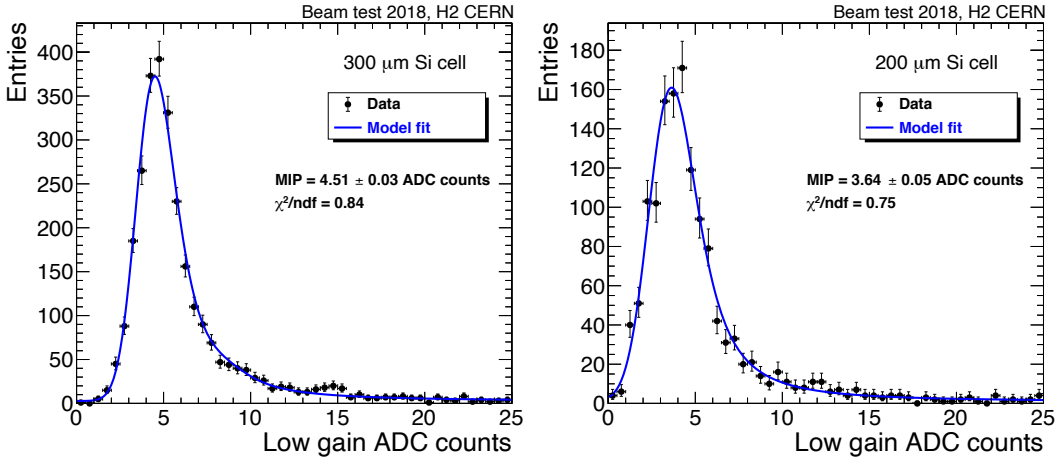

FIGURE 6.38: Figure shows LG ADC distributions fitted with landau convoluted with Gaussian distribution for 300 μm (left) and 200 μm (right) thick silicon cell.

Similar to HG-to-MIP conversion factor, one can extract LG-to-MIP conversion factors from the fitted functions. Figure 6.39 shows the LG ADC counts per MIP values for a few layers in CE-E. As expected, it is observed that LG-to-MIP is lower as compared to HG-to-MIP conversion factor. On an average, 5 LG ADC counts corresponds to one MIP for 300 μm silicon cells.

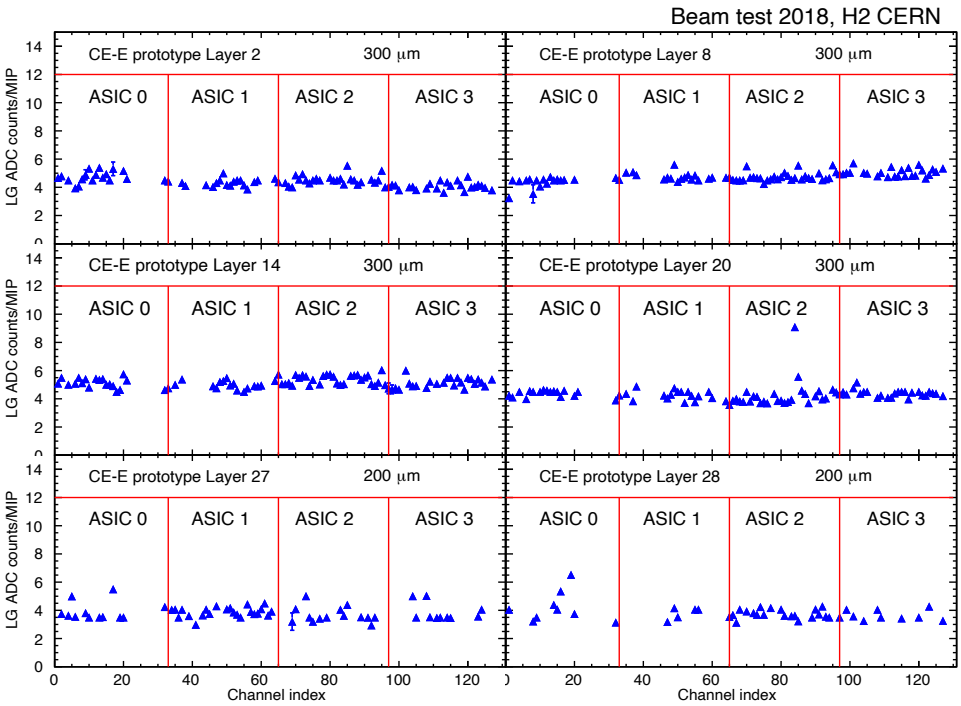

FIGURE 6.39: Summary plot for HG ADC counts per MIP for LG channels for a few CE-E layers. Bottom two plots corresponds to 200 μm thick silicon cells.

The summary plots for HG (Figure 6.37) and LG (Figure 6.39) channels show that MIP conversion factors are smaller for 200 μm thick cells as compared to 300 μm. This is expected because in a silicon cell which has a smaller depletion (active) width, electron-hole pair produced by traversing ionizing particle will be smaller. Thus, the charge collected at the end is smaller, hence a MIP will correspond to lower ADC counts. As mentioned earlier, last two layers of CE-E, were equipped with sensor modules with 200 μm silicon cells. Figure 6.40 shows the distribution of MIP calibration factors for all the calibrated cells in CE-E for HG channels (left) and LG channels (right). The distribution clearly shows that on an average MIP conversion factors for 200 μm are smaller by a factor $\sim 2/3$ as compared to 300 μm silicon cells.

The channels in CE-H which are located on the off-center modules could not be calibrated with the dedicated beam of muons as the beam profile is not wide to enough to cover these modules. These channels are calibrated with the remnant particle beam. The HGCAL prototype was left operating for another week after the end of three weeks of data-taking. During this operation, the HGCAL prototype was exposed to particles of unknown energy and species. A larger scintillator of size 40×40 cm^2 was placed in between CE-H and AHCAL to form trigger signal. This operation of data taking is called *parasitic run*. Making use of fine longitudinal and lateral granularity of HGCAL prototype, a dedicated tracking algorithm was developed that makes use of hits in the consecutive layers and provides space points for a straight line fit to reconstruct a track. The cells with traversing muons

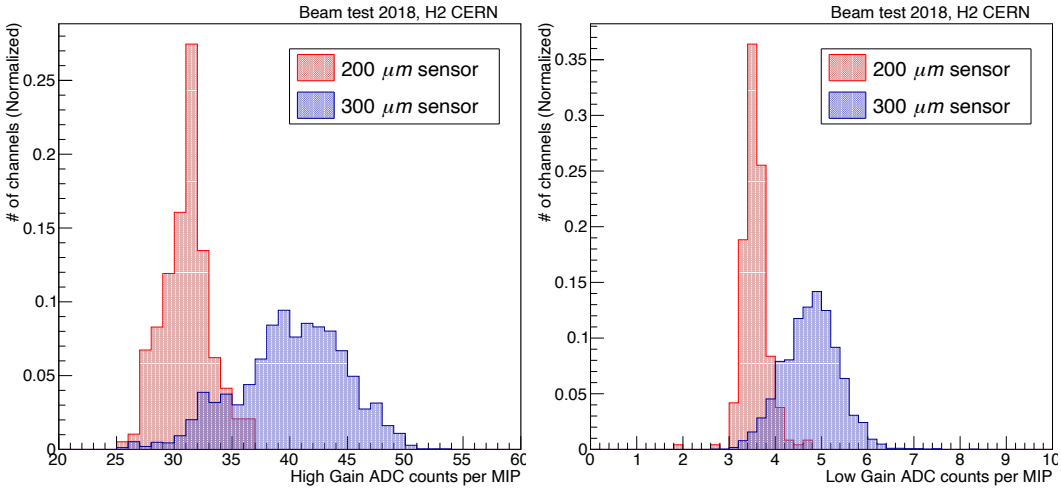

FIGURE 6.40: The distribution of MIP conversion factors for HG (left) and low (right) channels. The distribution for 300 μm thick silicon (blue filled histogram) shows higher a higher mean as compared to 200 μm (red filled histogram) for both high and LG channels.

are selected and the MIP calibration is performed using HGCAL tracking. This procedure is discussed in detail in [115].

Thus, the channel-to-channel response equalization is achieved for $\sim$ 85% silicon cells in CE-E and CE-H prototype using the dedicated beam of muons and parasitic runs.

6.9 Signal to Noise ratio estimation

Signal to noise ratio (S/N) quantifies the level of separation between signal and noise of the detection system used, and is an important aspect of bench-marking its performance. Any detector system employed in modern particle physics experiment needs to be capable of effectively resolve between the noise and the a minimum signal induced by a MIP passing through it, in order to use sensor data for event reconstruction. Quantification of S/N ratio also dictates the threshold to be applied for effective noise rejection in real data-taking scenario. For example, for a detector with a very high S/N ratio one can use a very tight or high noise-rejection criteria without biasing the overall energy measurement, i.e. rejecting hits containing true particle signal. Conversely, a detector with poor S/N ratio, the data will have a significant noise contribution. The effect of such noise-rejection threshold is discussed in the context of hadronic shower performance in HGCAL beam test experiment in chapter 7 and appendix E.

The signal produced by the MIP is determined using muon beam in terms of HG or LG ADC counts per MIP, as discussed in previous section. For noise level measurement, one of the muon runs (Run number 719 : $\sim$ 10k events) is re-reconstructed without applying the pre-selection cut. However, pedestal and common-mode noise subtraction is applied during re-reconstruction. Figure 6.41 shows number of reconstructed hits with and without applying pre-selection cut. After removing the pre-selection cut, the plot (right) shows that all the channels in CE-E and CE-H ($\sim$12000) are reconstructed, majority of which are noise-like hits.

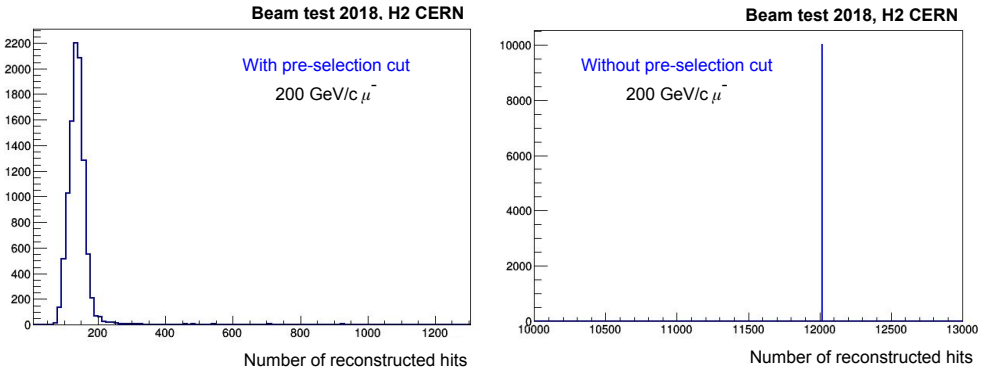

FIGURE 6.41: Number of reconstructed hits with (left) and without (right) pre-selection cut applied on muon data.

The pulse shape studies show that the signal is expected to peak in time-sample 2 (TS2) or time-sample 3 (TS3), e.g. see Figure 6.15, therefore the noise is estimated using same time-sample to be consistent. The extraction of ADC data is done before pulse fit in order to allow negative fluctuations to also populate HG ADC distribution. The distribution of HG ADC counts shows a Gaussian-like distribution peaking around zero shown in the Figure 6.42.

Few channels show an unusual behaviour in pedestal distributions, for example as shown in Figure 6.43. These channels are found to be in the neighbourhood of μHDMI connection at the sensor module. Due to digital coupling between these channels and the digital bus of μHDMI connection, the ADC data gets contaminated with signal interference. For all further data analyses, these three cells are masked in all the modules.

After masking faulty channels, the consistency of pedestal distribution is examined in:

1. **Different time-samples:** To check the consistency of pedestal noise across different time samples.
2. **Different data-taking runs:** To check the consistency of pedestal noise across different data taking runs.

These checks are important to make sure that no bias is introduced by studying the noise behaviour from only one time-sample and one run. To check the noise consistency across different time-samples, HG ADC distribution is compared between TS3 and TS0 for each channel for a muon run 719. The TS0 is the time-sample which remains free of signal even in the case of signal contaminated pulse because of delay in trigger received by SKIROC2CMS chips. Figure 6.44 shows the HG ADC count distribution for the two samples for two example cells. The HG ADC distribution comparisons show good agreement between TS3 and TS0.

To check the consistency across runs, a different muon run (Run number 718 : $\sim$ 10k events) is re-reconstructed and noise distribution is checked in TS3 for corresponding channels. Figure 6.45 shows HG ADC count distribution in TS3 for these two runs. The distribution comparisons show good agreement between run number 719 (red dotted line) and run number 718 (blue dotted line).

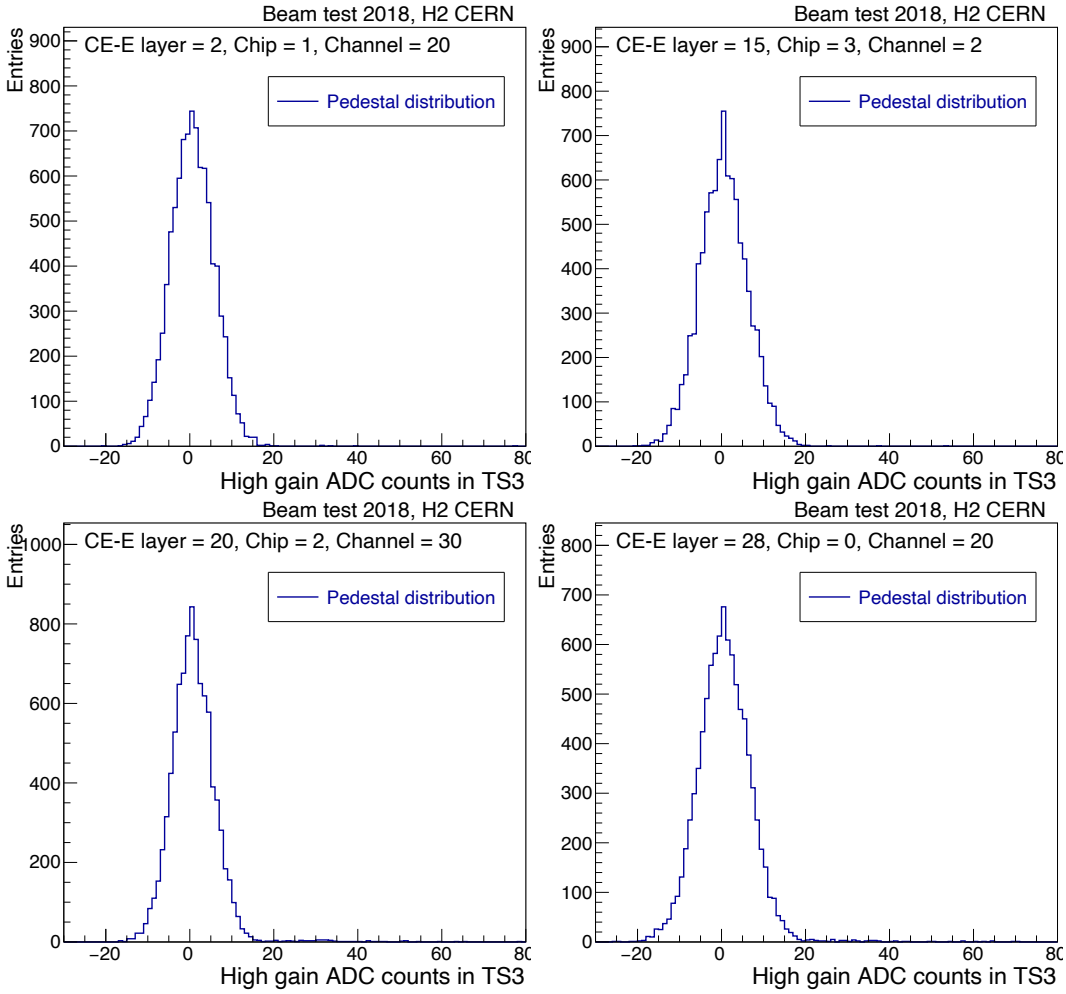

FIGURE 6.42: HG ADC distribution in time sample 3 (TS3) for a few channels in different CE-E prototype layers. The distribution shows a Gaussian-like behaviour with mean at zero, which is expected for random noise.

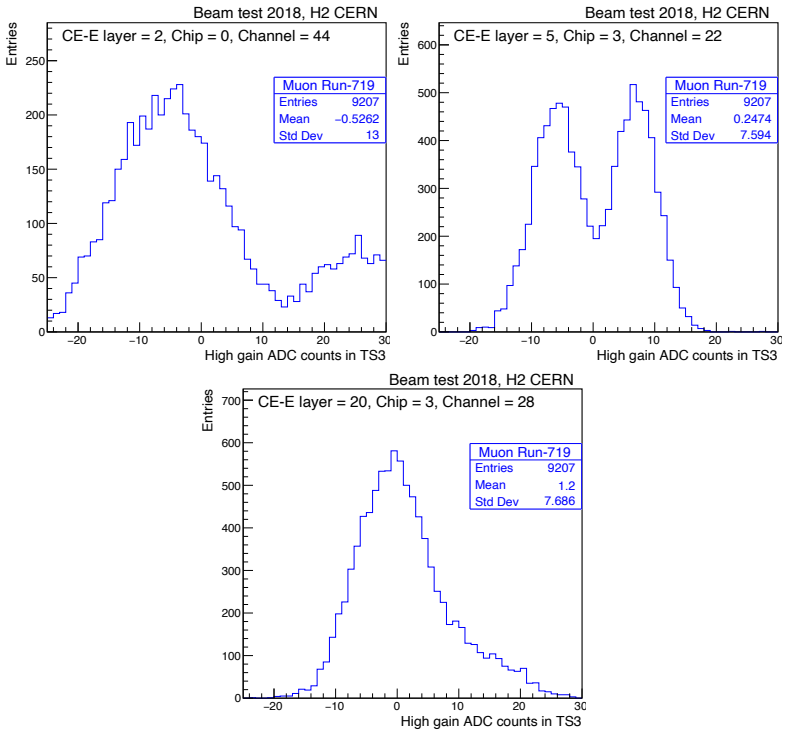

FIGURE 6.43: HG ADC distribution for two channels in the neighbourhood of μHDMI connection. These channels showed abnormal behaviour noise behaviour therefore were masked (not used) for further data analysis.

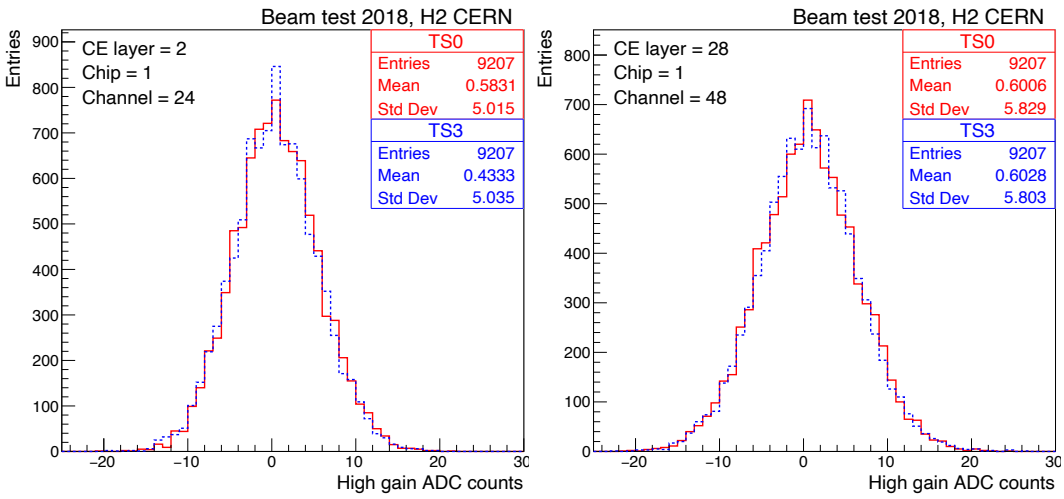

FIGURE 6.44: Pedestal distribution comparison for two different samples (TS0 in red, TS3 in blue) for HG channels and both distributions show consistent behaviour.

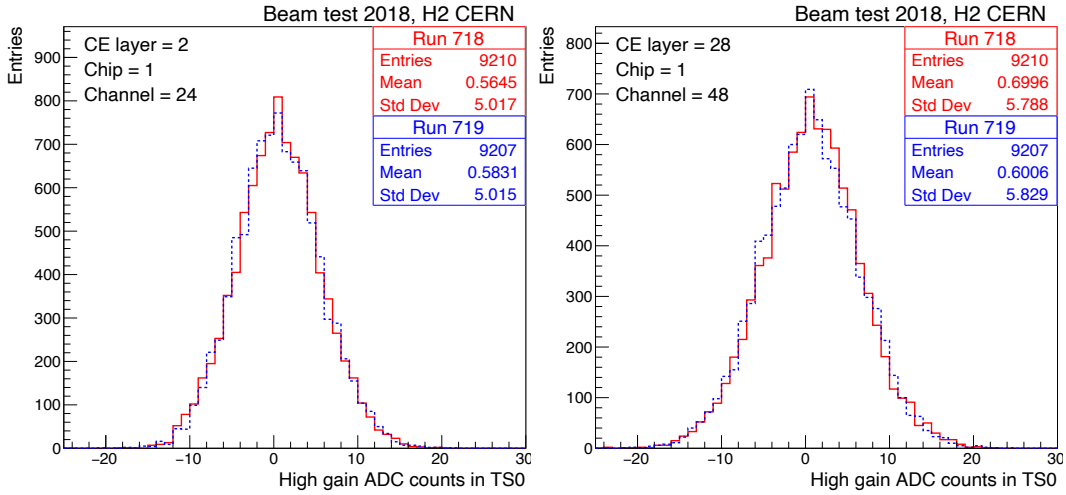

FIGURE 6.45: Pedestal distribution comparison for two different muon runs (run 718 in red, run 719 in blue) for HG channels and both distributions show consistent behaviour.

For the HGCAL prototype, the stability of noise was studied in detail and found to be consistent across the whole data taking campaign [111]. After ensuring that the noise distribution is consistent for different runs and different time-samples, TS3 HG ADC distribution for each cell is fitted with a Gaussian function as shown in Figure 6.46.

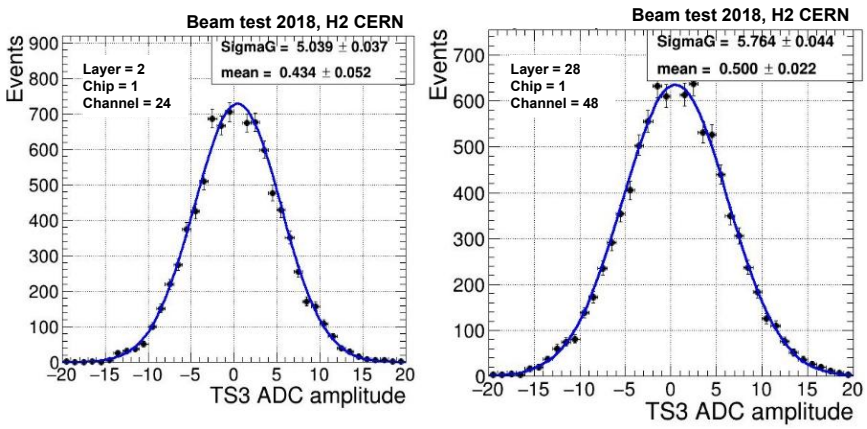

FIGURE 6.46: Figure shows pedestal distribution for two HG channels with a Gaussian function.

The width σ_{noise} of the fitted Gaussian is extracted for each cell in order to define noise level in terms of HG ADC counts. With the same procedure, the noise level for LG channels are also estimated. Figure 6.47 shows the LG ADC distribution of TS3 with a Gaussian fit. Similar to HG noise estimation, σ_{noise} from the Gaussian fit is defined as the noise level.

The study of noise level shows a few interesting, albeit expected, behaviour. The noise level of the silicon cells depends on the cell geometry, such as depletion width and surface area. Figure 6.48 shows the distribution of noise level in terms of HG (left) and LG (right) ADC counts for 200 μm and 300 μm active thickness. Silicon cells with 200 μm active thickness exhibit higher noise level as compared to 300 μm cells. Similarly, if one compares the overall behaviour of noise for cells of different types (i.e. different surface areas), one observes that the cells with larger surface area exhibit larger noise as compared to cells

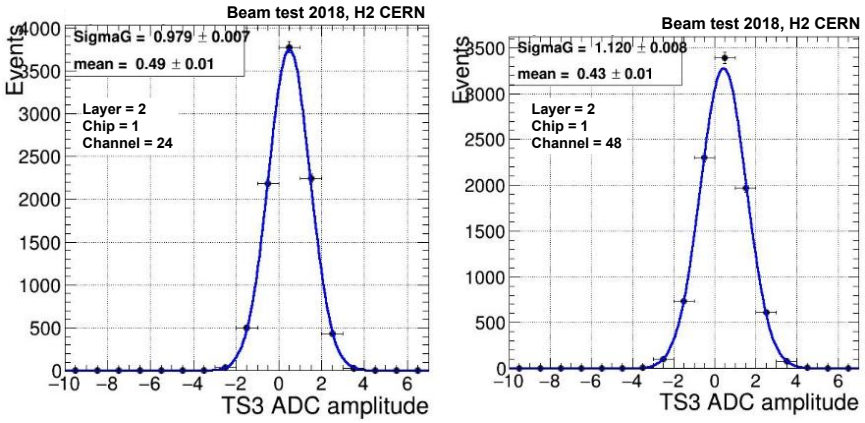

FIGURE 6.47: Figure shows pedestal distribution for two LG channels fitted with a Gaussian function.

with smaller surface area, as shown in Figure 6.49. The different cell types as depicted on the x-axis of Figure 6.49 correspond to cells of different area as was shown in Figure 5.2 (right). The box in the Figure 6.49 represents 50% inter-quantile range of the distribution, the horizontal line within the box corresponds to the median, the whiskers on the bottom and top corresponds to 5^{th} and 95^{th} percentile, respectively. The points above and below the whiskers are the outliers that make up 0.3% of the overall distribution.

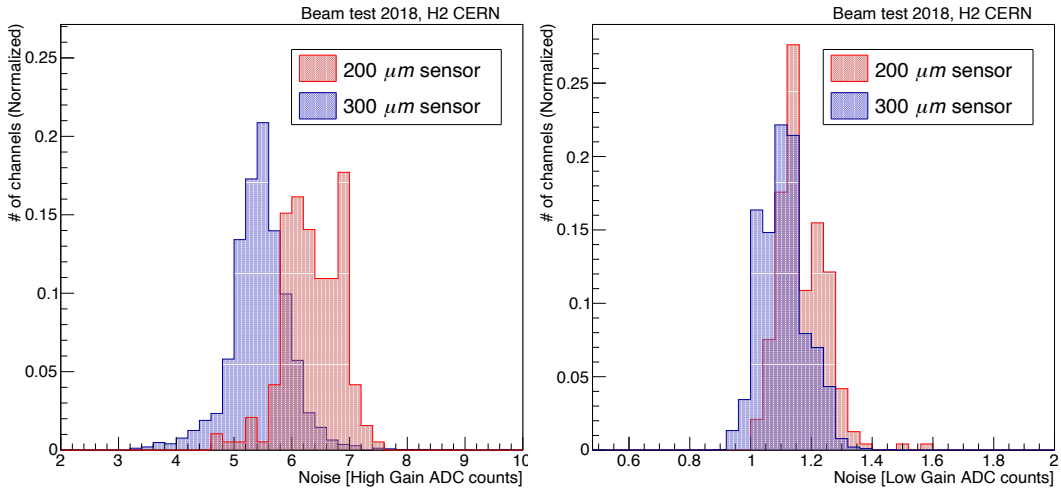

FIGURE 6.48: Noise distribution for HG (left) and LG (right) channels. The distribution shows noise for silicon cells with two different thicknesses i.e. 200 μm (red) and 300 μm (blue).

Such behaviour of geometry dependence of noise can be understood in terms of capacitance and its effect on the noise. Silicon sensor cells are a typical p-n junction diodes. A diode forms a capacitance across the depletion width which can be expressed in the following equation 6.14:

$$C = \frac{\epsilon A}{W} \tag{6.14}$$

where ϵ is the permitivity of the medium, A is the area and W is the depletion width of the diode. The capacitance (C) is directly proportional to the area (A) of the silicon cell and inversely proportional to depletion width (W). Thus, higher the surface area or lower

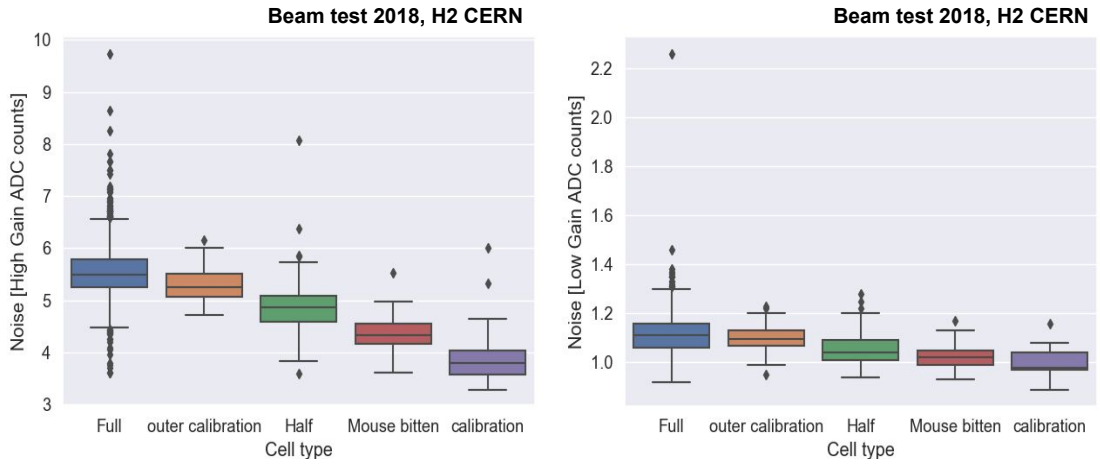

FIGURE 6.49: Noise as a function of cell type for HG channels (left) and LG channels (right). Cells with larger area show high noise as compared to cells with smaller smaller area.

the depletion width, higher will be the capacitance. This has an effect on the impedance of the system. The impedance (Z_C) across the capacitor can be expressed as shown in equation 6.15 below:

$$Z_C = \frac{1}{2\pi f C} \tag{6.15}$$

From equation 6.14 and 6.15, one can conclude that for a high frequency (f) pulse, a capacitor with higher capacitance will offer lower impedance as compared to a capacitor with smaller capacitance. The electronic noise, which is part of intrinsic noise for silicon sensor cells, generally has a high frequency component and will encounter higher impedance across a capacitor with relatively smaller surface area. Therefore, silicon cells with smaller surface area such as half hexagonal cells, have lower noise level as compared to silicon cells with larger surface area such as full hexagonal cell as observed in Figure 6.49. Applying similar reasoning with the distance across capacitors, higher noise level is expected in silicon cells with smaller depletion width as shown in Figure 6.48.

Now that, signal level is estimated as ADC counts per MIP and noise level as one σ_{noise} in terms of ADC counts, one can estimate S/N ratio of silicon cells used in beam test experiment, by using following expression:

$$S/N = \frac{\text{ADC counts per MIP}}{\sigma_{noise}} \tag{6.16}$$

The S/N is estimated for each cell that is well calibrated using muon beam. Figure 6.50 shows S/N ratio distribution for HG channels (left) and LG channels (right) with 300 μm and 200 μm thick silicon cells employed in CE-E prototype. The S/N ratio is on an average ~ 7 for 300 μm in CE-E prototype. As expected, 200 μm thick silicon cells show lower S/N as compared to 300 μm thick silicon cells. Similarly, S/N ratio for smaller cell sizes, e.g. half hexagon cells, shows shows higher average value as compared to larger cell sizes, e.g. full hexagon cells, as shown in box-plot format in Figure 6.51. The reason for S/N value dependence on cell geometry is due to the noise dependence on cells, as discussed already.

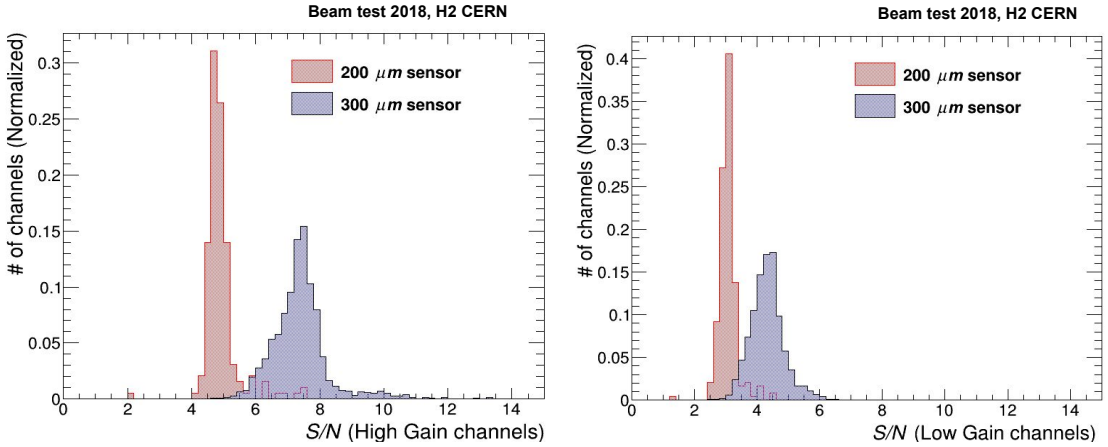

FIGURE 6.50: Signal to noise ratio for HG (left) and LG (right) channels. The distribution shows noise for silicon cells with two different thicknesses i.e. 200 μm (red) and 300 μm (blue) employed in CE-E prototype.

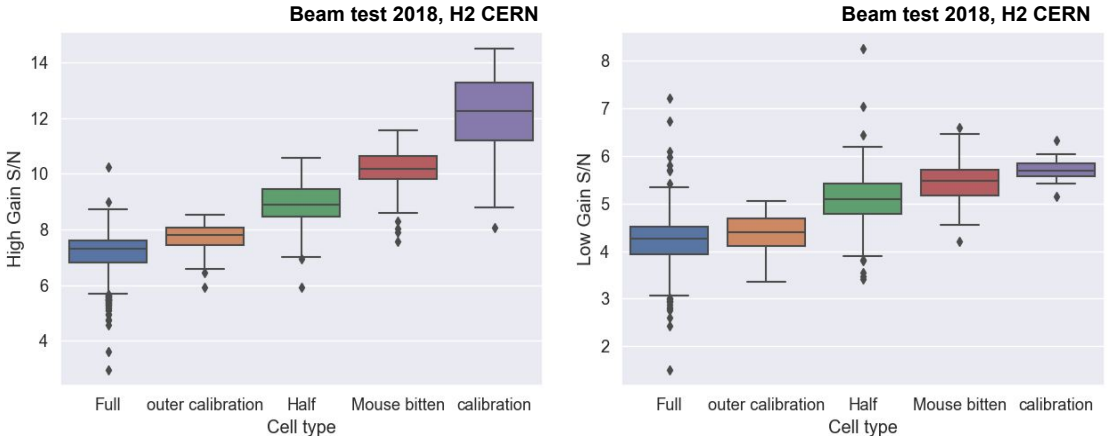

FIGURE 6.51: S/N as a function of cell type for HG channels (left) and LG channels (right). Cells with larger area show low S/N as compared to cells with smaller area.

For the silicon cells of modules employed in CE-H prototype, the intrinsic noise is estimated separately (not discussed in this thesis) and is found to be slightly higher as compared to silicon cells in CE-E prototype. The S/N for these cells is estimated using MIP calibration factors (as obtained from parasitic runs) and a separate noise study. The S/N for CE-H is found to be approximately ~ 5 for 300 μm full hexagon cells. This information is indispensable for pion analysis in order to effectively reject noise-like hits, different noise-rejection threshold is applied for modules in the CE-E and the CE-H sections (appendix E).

Chapter 7

Physics performance of combined HGCAL and AHCAL prototype detectors to hadronic showers

Quarks and gluons, produced via strong interactions, are the dominant outcomes of the p-p collisions at the LHC. As explained in chapter 2, these particles manifest themselves as jets of charged and neutral hadrons, most of which are $\pi^{\pm}$ and π^{0}, and deposit their energies as hadronic and electromagnetic showers while traversing the detector material. In the foreseen HL-LHC operation, the detectors will see copious amount of these hadrons due to increased instantaneous luminosity, especially in the high pseudorapidity regions. Hence, a detailed understanding and efficient reconstruction of hadrons leads to better jet performance which is the key probe to well founded analyses for new physics searches as well as for many SM precision measurements.

Keeping this requirement in mind, one of the major goals of October 2018 beam test experiment was to measure the performance of a realistic HGCAL setup close to final configuration of the detector with real particles. In this chapter, performance of combined HGCAL and AHCAL detector prototype to the hadronic showers initiated by charged pions of energies 20-300 GeV is presented in data and simulation. The analysis includes event selection and cleaning, identifying the location of shower start in the detector, shower energy reconstruction, measurements of energy response and resolution, longitudinal and transverse shower development. These measurements are also compared against GEANT4 based simulation in order to validate the hadronic shower models as well as simulation framework developed for the HGCAL beam test experimental setup. It is to be noted that, because of incomplete lateral shower containment in the detector, here the goal is not to determine the best achievable performance of the detector prototype to pion showers in terms of resolution. Rather, the work shown in this chapter presents a first performance study of HGCAL prototype to hadronic showers. It provides foundation to build upon more sophisticated machine-learning based methods (not discussed in this thesis) to make use of event-by-event detailed information and further improve the performance.

7.1 Event selection and cleaning

A set of cleaning cuts are applied in order to ensure a high purity sample of charged pions in data. The selection cuts remove undesired events such as contamination from other particles, out-of-acceptance particle incidence, energy contribution from spurious noise, and faulty channels. The cleaning cuts are listed as follows:

1. **Channel masking:** A few of the channels that had issues at the hardware level are masked. The channel masking at different layers are described as follows:

- In all silicon sensor module prototypes, the three channels which were in the neighbourhood of HDMI connection showed irregular behaviour in terms of pedestal and noise as discussed in the last chapter (see Figure 6.43).
- One of the off-center modules in CE-H layer-8 had encountered a hardware issue and was not operational in the entirety of the experiment.
- All 32 channels connected to chip 0 of CE-E layer 1 showed high activity even in the absence of particle shower due to technical fault, e.g. as shown in Figure 7.1. The figure shows an event display of pion acting as a MIP in CE-E and starting shower in the CE-H that extends to AHCAL. Unusual high activity is visible in chip-0 of first layer of CE-E and is not used in the analysis.
- Channel number 16 in chip 1 of module number 39 (situated at off-center position in CE-H layer 8) was masked because of unnaturally high activity.
- Layer 38 of AHCAL had a relatively larger scintillator tile area as compared to other layers. This fact was not been fully realized in the simulation, therefore in order to have a consistent data-simulation comparison, this layer is masked.

In total about ∼3.6% channels in HGCAL and ∼0.65% channels in AHCAL prototype are masked.

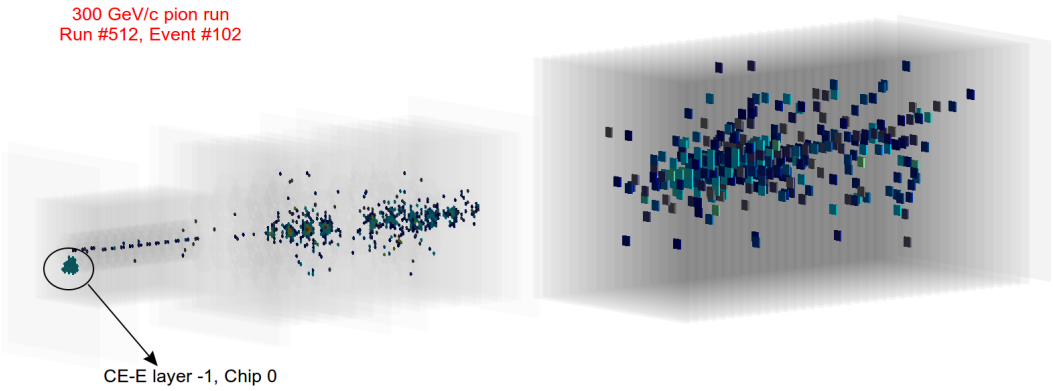

FIGURE 7.1: Event display of a 300 GeV/c pion event. Pion initiates a shower in CE-H prototype compartment which continues to develop in AHCAL section. Chip 0 of the first layer of CE-E shows full occupancy. This behaviour was observed in other runs as well therefore this whole chip is masked at the analysis level.

2. **Noise rejection:** In addition to masking the faulty channels, the noise-like hits are also rejected. For this purpose, the pedestal noise study performed during signal-to-noise ratio estimation, is used. One standard deviation ($1\sigma_{noise}$) noise values are obtained in terms of HG ADC counts for each silicon cell and is averaged over a chip. The hits that are below $3\sigma_{noise}$ ($4\sigma_{noise}$) HG ADC counts corresponding to their respective chip average, are rejected in CE-E (CE-H) prototype. The optimization of noise rejection threshold in CE-E and CE-H are discussed in detail in appendix E. For AHCAL prototype, the hits that are below 0.5 MIPs were already rejected during event reconstruction, therefore no additional noise rejection cut is applied for AHCAL hits.

These noise rejection cuts were applied on the rechit level, thus does not affect the total event counts . However, during the analysis it was observed that two modules at CE-H layer 9 showed very high occupancy as compared to its neighbouring layers. One of

the modules is located at the central position, therefore masking it would have costed major loss in energy measured along the shower axis. Further investigation showed that such high occupancy is manifested in only a few events (less than 1%). Therefore, the events are rejected where the occupancy in either of these two modules (module #59 and #55) were greater than 80. The hit threshold 80 is chosen after the making the occupancy comparison with simulation which is free of any spurious noise.

3. **Track quality cut:** To ensure a well reconstructed track, a quality criteria cut is applied on tracks reconstructed using DWC hits (section 6.3.3) . The track quality cut is as follows: at least 3 out of 4 wire-chambers should record a hit, and $\chi^2/ndof$ of the reconstructed track should be less than 10. These two simultaneous conditions also ensure that there is a reconstructed track in the event. The events which do not satisfy these conditions, are rejected.

4. **Muon Veto:** The beam of pions suffered from particle contamination. For example, Figure 7.2 shows the energy and event fraction of the leading particle arriving at the front of HGCAL surface for different pion beam energies as predicted by the full beamline simulation (section 6.5.2). It is observed that the beam purity for π^- is more than 80% for low energy beam and more than 90% for high energy pion beam. The contribution from electromagnetic particles (e.g. γ) is either $\sim$ 1% or less. The maximum contamination is due to muons which is higher for low beam energy and reduces for high beam energy. Therefore, a muon veto is applied to clean the data of muon contamination.

   Muon contamination could be introduced either during π^- beam generation or in-flight decay of π^- to μ^-. Since there was no dedicated muon-veto scintillator installed downstream of the experimental setup to veto such events, a muon-veto cut is devised at the analysis level. The muon-veto cut is based on the energy deposited in HGCAL prototype and transverse spread in AHCAL prototype. Since muons with few hundreds of GeV energy act as MIP, therefore they deposit a small energy and have smaller lateral spread as compared to showering particles. Thus, these variables can be used to tag *muon-like* events.

   Figure 7.3 shows energy comparison in CE-E (left column) and CE-H (middle column) between 200 GeV muons and pions of two energies (20 and 100 GeV) from beam test data . The energy distribution of muons show Landau-distribution with a MIP peak, whereas energy distribution of pions (with muon contamination) has two components:

   - MIP peak: from non-showering pions as well as from muon contamination.
   - Higher energy tail: from showering pions.

   It is observed that at $E_{CE\text{-}E} \sim 100$ MIPs and $E_{CE\text{-}H} \sim 60$ MIPs, the two distributions start to deviate from each other, thus helping us decide the threshold for MIP-like signature in CE-E and CE-H.

   In conjunction with such energy thresholds, the lateral shower spread in AHCAL is used to arrive at the final muon-veto criteria in the combined detector setup. The energy sum comparison is not used in AHCAL because AHCAL receives mostly tails of the pion shower and hence shows an exponentially falling spectrum of energy. Therefore, applying a dedicated energy cut will reject a lot of genuine pion showering

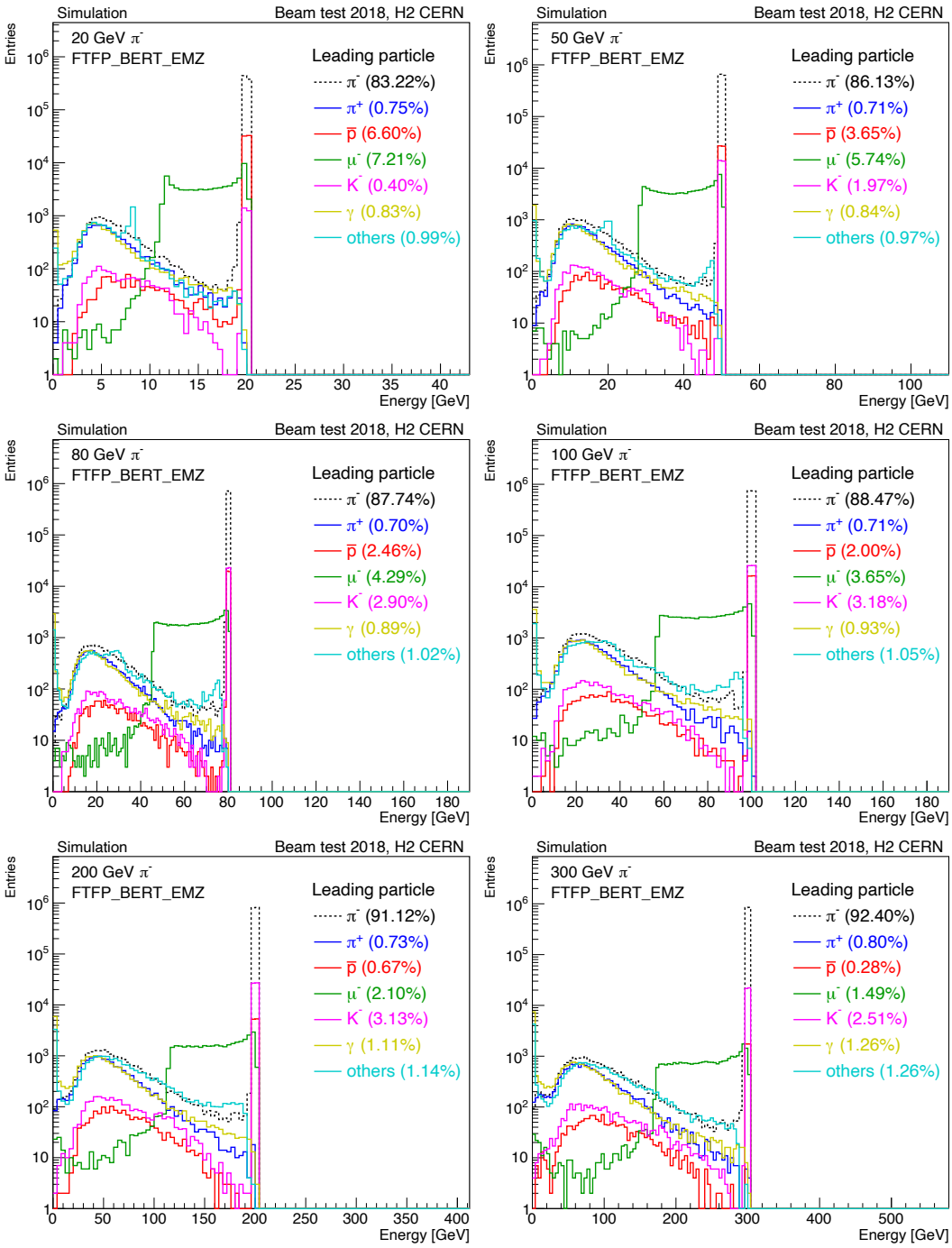

FIGURE 7.2: Energy distribution of leading particle and their event fractions (as indicated in the legends) reaching HGCAL surface for different pion beam energy. The fraction and energy distributions are obtained from full beamline simulation. The beam purity is more than 80% or 90% depending on beam energy.

events.

For lateral shower spread, the lateral energy ratio variable S1/S25 is calculated and used, which is defined as follows:

$$\frac{S1}{S25} = \frac{\sum_{\text{AHCAL-layers}} E1}{\sum_{\text{AHCAL-layers}} E25} \tag{7.1}$$

where E1 is defined as the cell with the maximum energy deposited in a given layer and E25 is the energy sum of 5x5 neighbouring cells around the maximum energy cell (E1). In the case of MIPs, this variable peaks at one because only central cell records the traversing particle whereas for the showering particles it deviates significantly from one. The comparison of S1/S25 distribution between pions and muons show a distinct population with a very little overlap as shown in Figure 7.3 (right column). The peak around one for pions is mostly due to muon contamination. The two distributions deviate from each-other at S1/S25 $\sim$ 0.8 and is chosen as the threshold for this variable. Thus, the muon veto criteria is defined as follows: if $E_{\text{CE-E}} < 100$ MIPs AND $E_{\text{CE-H}} < 60$ MIPs AND S1/S25 > 0.8, then the event is rejected.

No effort is made to reject other type of hadrons since the contamination level is relatively smaller and their energy response and fluctuation are not expected to deviate very significantly from the charged pions in the energy range studied here.

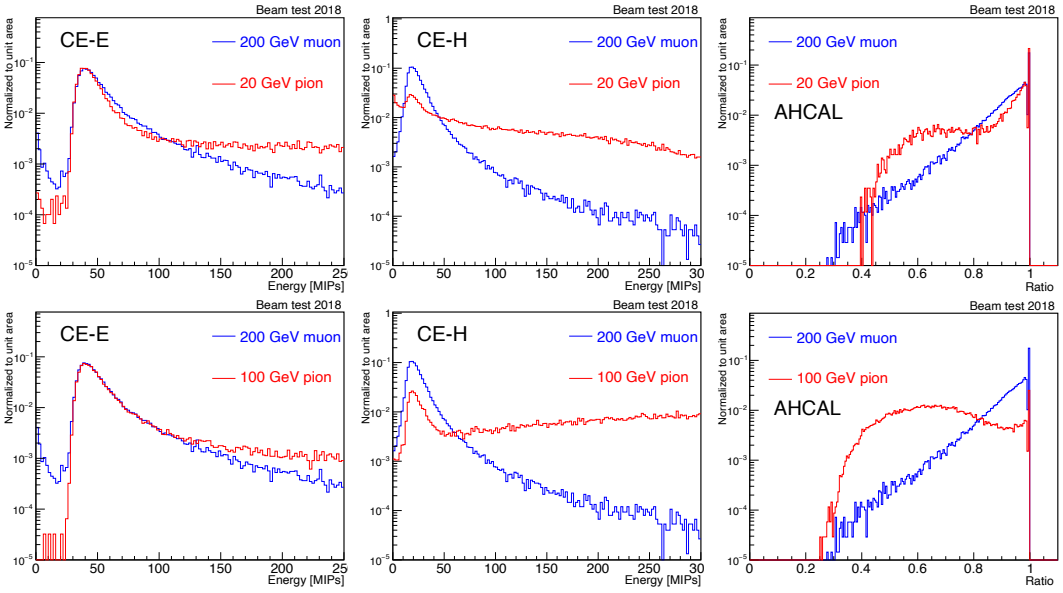

FIGURE 7.3: Energy comparison in CE-E (left column) and CE-H (middle column) and lateral energy ratio in AHCAL, *i.e.* S1/S25 (right column). The top (bottom) row correspond to 20 (100) GeV pion (red line) compared with 200 GeV muon (blue line). The point of deviation is taken as the threshold to define muon veto criteria.

5. **Track window cut:** To reject the events where impinging pion is too far away from the center of the layer, the reconstructed track from DWCs is used. This cut ensures that the shower is mostly contained in the detector, thus minimizing the lateral leakage and also makes the particle impact position consistent in data and simulation.

The beam profile of pions corresponds to the track impact point on the face of HGCAL prototype, i.e. the first layer of CE-E. Since the extrapolated track at CE-E layer 1 is used, therefore it is corrected for misalignments between DWC and HGCAL coordinate system (as explained in chapter 6). The beam profile is compared between data and simulation as shown in Figure 7.4 (corresponding one-dimensional plots are shown in appendix D).

The comparison shows that for lower energies the spread in data is more than the simulation. The reason for this difference is that in simulation the trigger scintillators (acceptance area) was set to 2x2 cm^2 for all beam energies (section 6.5.2); whereas during the data taking, the size of trigger scintillator was 4x4 cm^2 for the runs $E_{beam} <$ 100 GeV. However, the difference is of the order $\leq$ 2 cm which corresponds to one silicon cell distance. Therefore, a cut of 4x4 cm^2 window is applied for lower beam energies and 2x2 cm^2 for higher beam energies to reject out-of-acceptance pions. The values are listed in table 7.1.

TABLE 7.1: Table shows the value of track window threshold to reject the particles impinging outside this window

E_{beam} [GeV]	x_{min} [cm]	x_{max} [cm]	y_{min} [cm]	y_{max} [cm]
20	-2.8	1.65	-1.20	3.2
50	-2.8	1.65	-1.25	3.2
80	-2.8	1.60	-1.15	3.2
100	-2.8	1.50	-1.15	3.2
120	-2.8	1.50	-1.15	3.2
200	-3.0	-1.0	0.2	2.2
250	-2.8	-0.8	-0.2	2.0
300	-3.0	-1.0	-0.3	1.8

6. **Pre-showering pions rejection:** The beam-line elements contribute to an effective material between secondary beam production and the experimental setup. Though the effective material is very small in terms of λ_{int}, however there is a non-zero probability of pions starting their shower before reaching the calorimeter. For such events, shower will not be fully contained within the calorimeter and result in the worsening of pion energy resolution measurement. Therefore, such events are not included in the data analysis. With the help of shower-start-finder algorithm (described in section 7.3 later in this chapter), the events are rejected where shower-starting location is in first two layers of CE-E. The cut is imposed on not the first layer but till second layer of CE-E because a whole chip is masked in the first layer resulting in information loss from layer 1.

Figure 7.5 shows the effect of each cleaning cut (applied in succession) on the beam test data in the form of reconstructed energy in the calorimeter. For example, *muon veto* cut removes the peak at the left of the distribution and *pre-shower reject* and *track-window selection* cuts reduce the width of the distribution. Table 7.2 shows the number of recorded events for charged pion beam during beam test experiment in configuration-1 and table 7.3 shows the cut flow, i.e., fraction of surviving events after each cut, for each beam energy in data. These selection cuts are also applied to the simulated events. Table 7.4 shows the cut flow for FTFP_BERT_EMN physics list with 100k generated events. The selection efficiency is almost identical for QGSP_FTFP_BERT_EMN physics list.

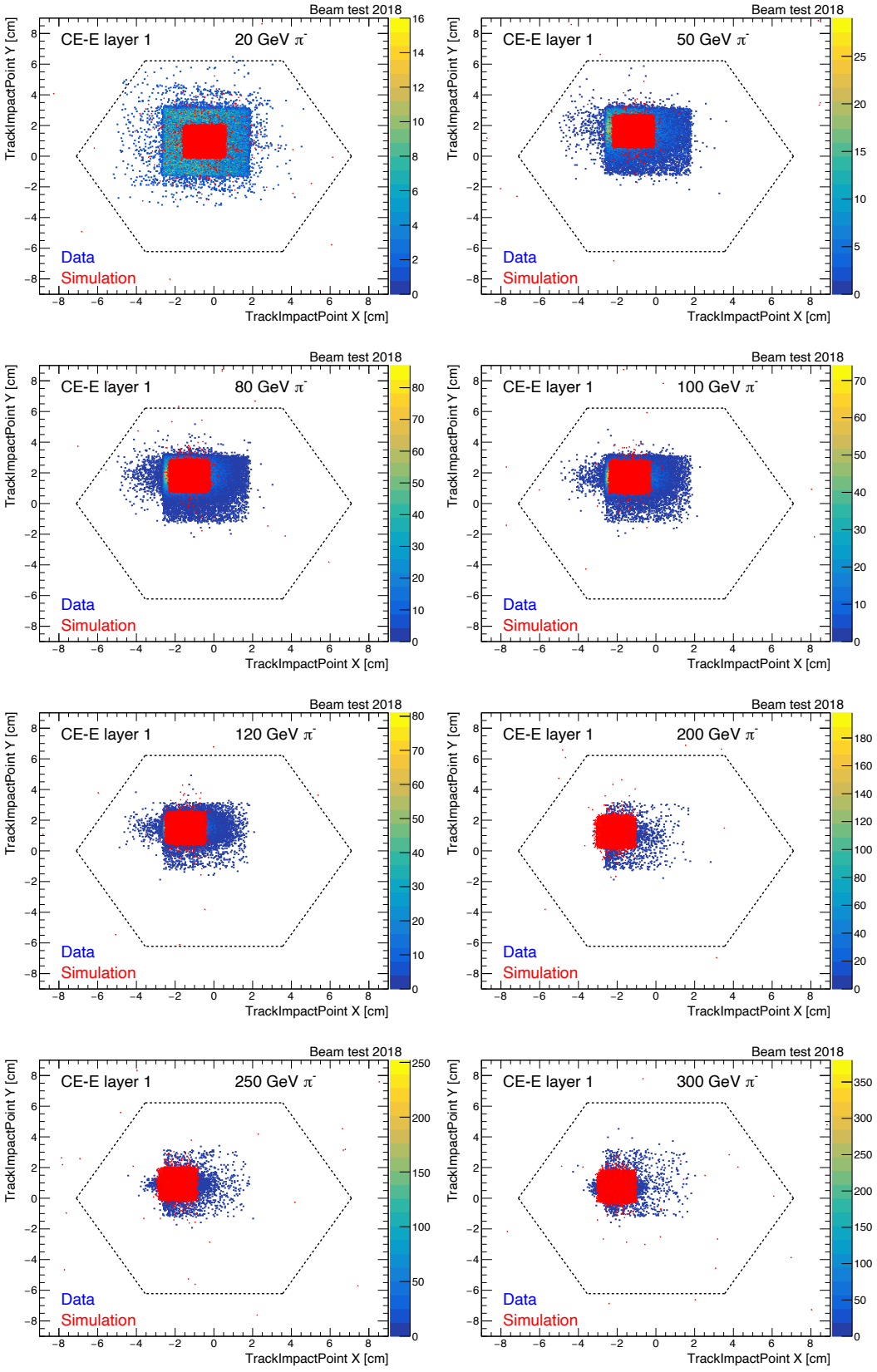

FIGURE 7.4: The track impact position at CE-E layer 1 for different pion beam energies. The blue and red points correspond to data and simulation, respectively. The dotted hexagonal line corresponds to the module-border at CE-E layer 1.

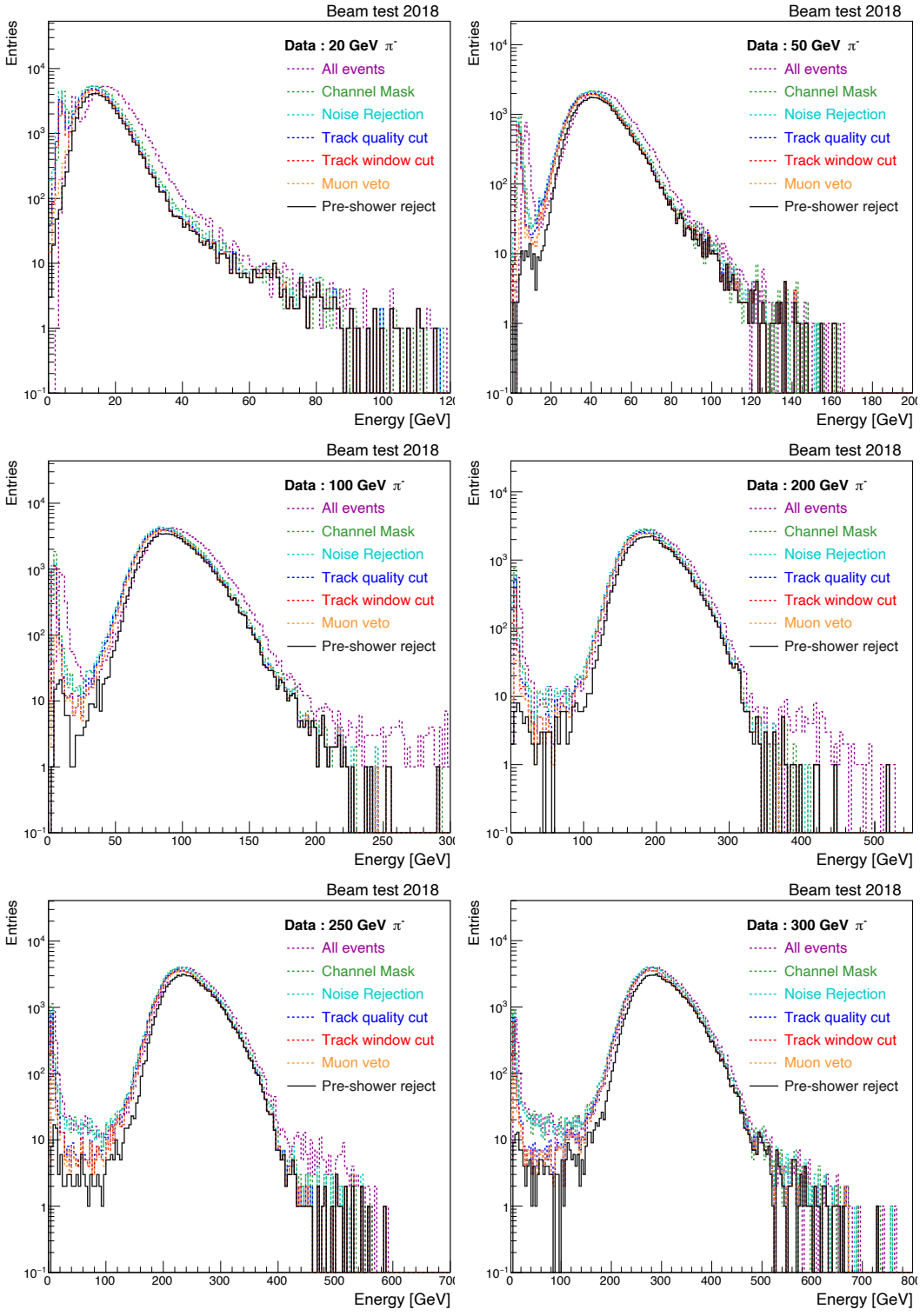

FIGURE 7.5: Energy in the calorimeter after applying each cleaning cut in succession for different energy of pions in beam test data.

TABLE 7.2: Table for recorded events for each pion beam energy corresponding to configuration-1 of beam test experiment.

Energy [GeV]	Recorded events
20	80432
50	61114
80	143729
100	102600
120	74945
200	56473
250	95814
300	108667

TABLE 7.3: Table shows the cut flows for beam test data after applying each selection criteria sequentially.

Cuts	20 GeV	50 GeV	80 GeV	100 GeV	120 GeV	200 GeV	250 GeV	300 GeV
Total events	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
Track quality cut	88.6%	92.9%	91.9%	92.2%	91.1%	91.0%	90.4%	89.5%
Track Window	82.4%	89.7%	89.5%	90.0%	88.7%	88.3%	88.2%	88.0%
Muon veto	75.5%	87.2%	87.2%	88.0%	86.8%	87.5%	87.4%	87.3%
Pre-shower reject	65.8%	75.7%	75.3%	76.0%	74.7%	75.8%	74.4%	74.0%

TABLE 7.4: Cut flow table for simulation FTFP_BERT_EMN physics list after applying each selection criteria sequentially.

Cuts	20 GeV	50 GeV	80 GeV	100 GeV	120 GeV	200 GeV	250 GeV	300 GeV
Total events	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
Track quality cut	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
Track Window	97.0%	98.5%	98.6%	98.8%	98.9%	98.1%	97.9%	98.0%
Muon veto	95.0%	97.1%	97.3%	97.5%	97.5%	96.9%	96.9%	97.1%
Pre-shower reject	82.9%	84.9%	85.2%	85.2%	85.2%	85.0%	84.5%	84.5%

After applying the cleaning cuts, the out-of-the-box energy sum in different compartments are compared between data and simulation. Figure 7.6 and 7.7 shows the reconstructed energy sum comparison in terms of MIPs in CE-E, CE-H and AHCAL for all pion energies. It is observed that both the physics lists show almost similar performance at low energies. However, at higher energies QGSP_FTFP_BERT_EMN physics list shows slightly better agreement with data in CE-E compartment.

7.2 Categorization of events based on energy deposit in sub-detector prototype

The application of cleaning cuts removes the unwanted events from the pion runs. We then proceed to analyse the performance of detector prototype to hadronic showers. The first goal towards this step is to reconstruct the energy deposited by pions showering in the calorimeter. As discussed in chapter 4, hadronic showers are wider and penetrate deeper into the calorimeter as compared to electromagnetic showers, therefore the energy of the pion shower is expected to be shared between electromagnetic (CE-E) and hadronic (CE-H + AHCAL) section. Figure 7.8 shows such energy sharing between CE-E (on x-axis) and CE-H+AHCAL prototype (on y-axis), in terms of MIPs for all π^- beam energies.

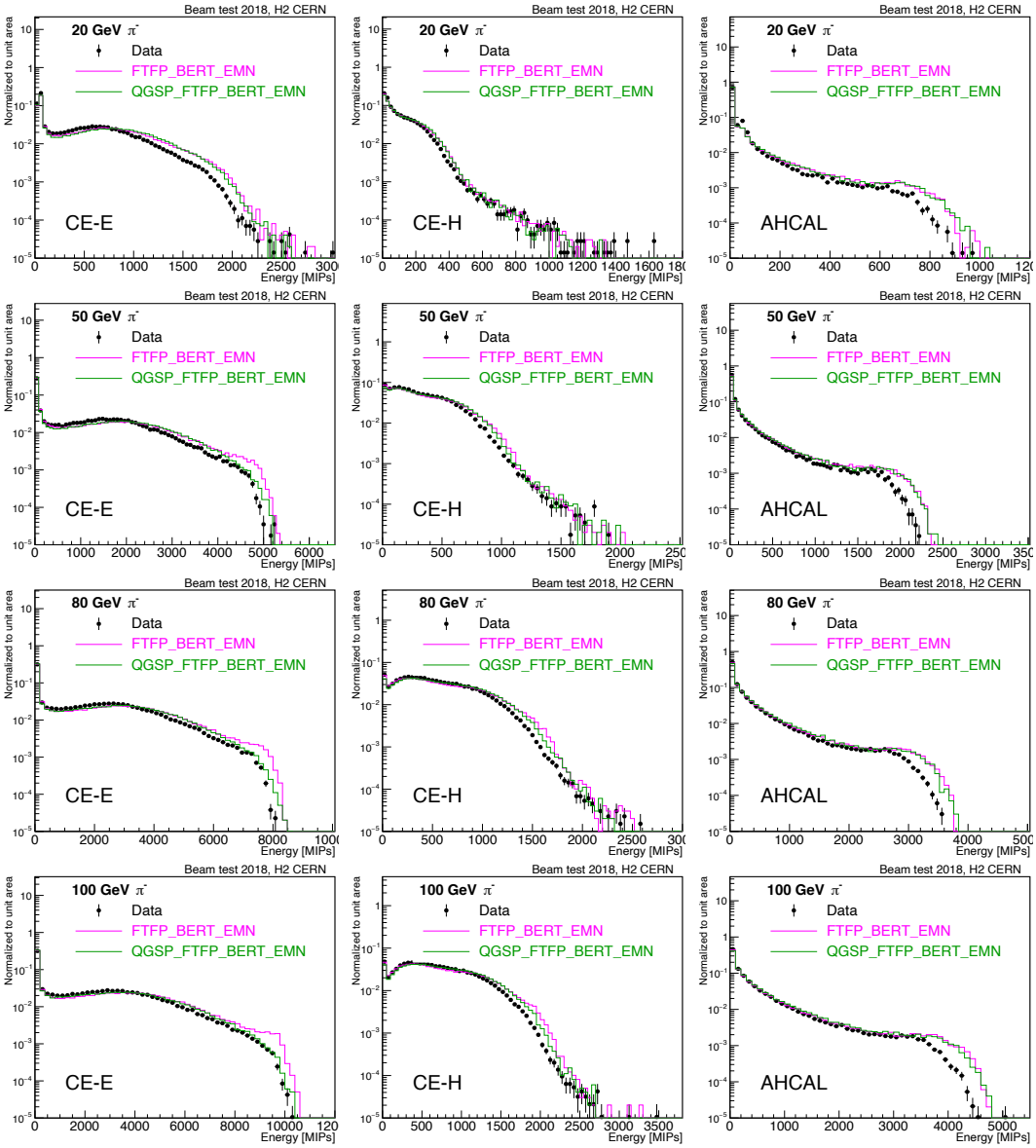

FIGURE 7.6: Energy comparison between data and simulation in CE-E (left column), CE-H (middle column) and AHCAL (right column) for **20, 50, 80 and 100 GeV pions**. The black dots correspond to data points and magenta and green line corresponds to simulation with FTFP_BERT_EMN and QGSP_FTFP_BERT_EMN physics list, respectively.

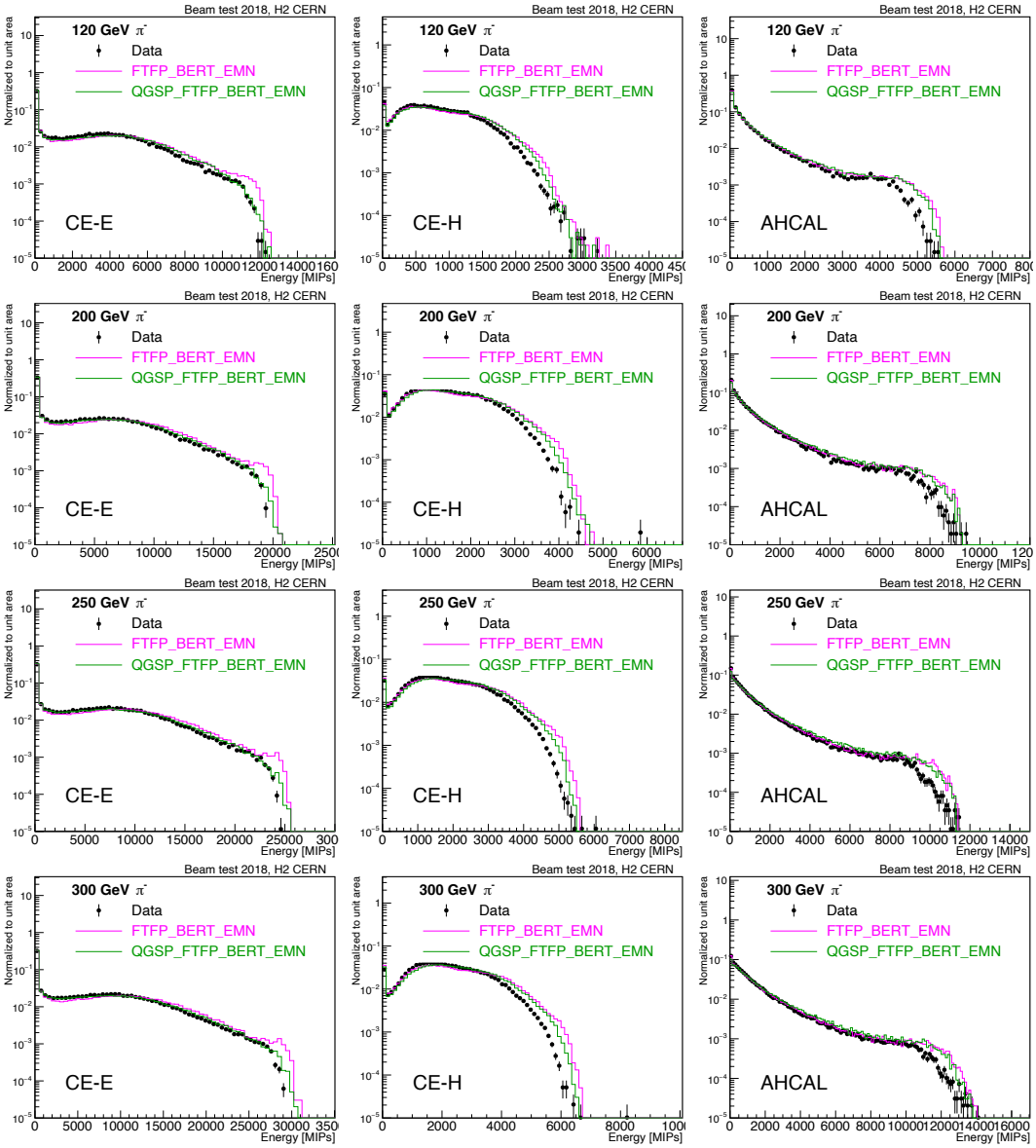

FIGURE 7.7: Energy comparison between data and simulation in CE-E (left column), CE-H (middle column) and AHCAL (right column) for **120, 200, 250 and 300 GeV pions**. The black dots correspond to data points and magenta and green line corresponds to simulation with FTFP_BERT_EMN and QGSP_FTFP_BERT_EMN physics list, respectively.

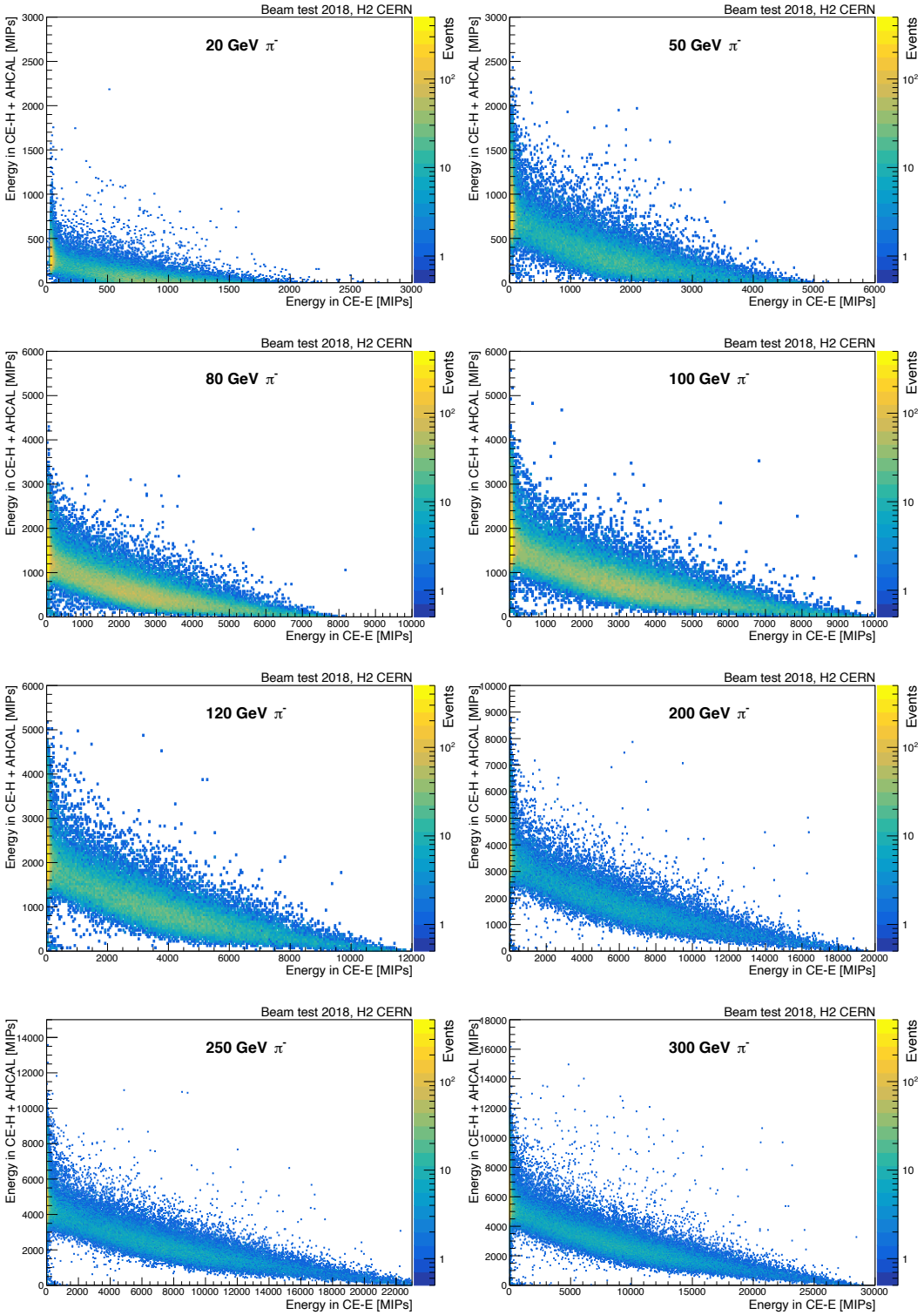

FIGURE 7.8: Energy sharing between electromagnetic (CE-E) and hadronic (CE-H + AHCAL) compartments for all pion beam energies.

From Figure 7.8, it is observed that the energy sum in CE-E has a significantly larger range than energy sum in CE-H + AHCAL. This feature can be understood by considering the fact that the electromagnetic and hadronic sections of the beam test experimental setup have different geometry and characteristics. The electromagnetic section i.e. CE-E, consists of 28 sampling layers in a depth of $\lambda_{int} \sim 1.4$, whereas CE-H consists of 12 sampling layers in $\lambda_{int} \sim 3.4$ and AHCAL consists of 39 sampling layers in $\lambda_{int} \sim 4.4$. Since the production of energetic π^0 is more probable at the first hadronic interaction which develops an electromagnetic shower cascade (the em component of hadronic shower) [40]. Therefore, for pions that start showering earlier in CE-E, the electromagnetic component will be fully contained in CE-E which is $\sim 27\ X_0$ deep. As discussed in section 2.3.2, the electromagnetic shower has high particle multiplicity as compared to pure hadronic shower. Hence, with its 28 sampling layers, the CE-E records more energy in terms of MIPs. However, that does not necessarily mean that more energy is deposited in CE-E in terms of GeV as it is only 1.4 λ_{int} deep.

Also, CE-E uses **Pb/Cu/CuW** as the absorber whereas CE-H and AHCAL uses **steel (Fe)** as the absorbers, which results in different *e/h* ratio for electromagnetic and hadronic sections. Therefore, the hadronic shower of pions that start showering in CE-E will encounter different absorber in electromagnetic and hadronic section. Hence, there is an additional fluctuation due to the difference in π^0 production [37] as well as their different response to electromagnetic versus hadronic component of the shower because of different *e/h* ratio (section 2.6.2). Therefore, in order to properly combine the energies of different compartments, it is necessary to segregate the events into two categories i.e. pions that start showering in CE-E (referred to as ***CE-E pions***) and pions that are MIPs in CE-E (referred to as ***CE-H pions***). The simplest way to categorize the events into these two categories is by making use of pion and muon energy comparisons to define CE-E pions and CE-H pions, as is done for muon-veto criteria (Figure 7.3). Thus, energy based event-categorization is defined as follows:

- If $E_{CE\text{-}E} > 100$ MIPs and $E_{CE\text{-}H} > 60$ MIPs, then the pion is referred to as **CE-E pions**
- If $E_{CE\text{-}E} < 100$ MIPs and $E_{CE\text{-}H} > 60$ MIPs, then the pion is referred to as **CE-H pions**
- If $E_{CE\text{-}E} < 100$ MIPs and $E_{CE\text{-}H} < 60$ MIPs, then the pion is referred to as **MIP-like pions**

Events which fall into the third category but pass the muon-veto cut, are rejected. Such type of events are the ones where pion starts to shower in AHCAL prototype. Such events are expected to be very low ($\leq$ 1.4%), hence it does not have significant impact on the shower energy measurement.

Figure 7.9 shows energy sum distribution in terms of MIPs for CE-E and CE-H pions as well as inclusive in both (called all pions in the figure). The energy distribution corresponding to all pions shows a highly non-Gaussian distribution. The energy distribution for CE-E pions exhibit a long tail and has a higher mean energy in terms of MIPs, whereas CE-H pions show a narrower and lower mean energy. The reason for such feature is exactly as discussed above for Figure 7.8, i.e. higher sampling frequency in CE-E and different e/h ratios of absorbers.

A more precise categorization of the events can be done by exploiting high longitudinal and lateral granularity of HGCAL prototype detector. An algorithm is devised that

identifies the location of shower start within the calorimeter. Based on the shower start location, the events can be categorized as either CE-E pions or CE-H pions. The algorithm is explained in detail in the next section.

7.3 The algorithm to identify the shower start location

The initiation of hadronic shower is the result of strong interaction of incident hadron with the nucleus of the detector material, as described in chapter 4. The incident hadron can initiate a shower anywhere in the detector with an associated probability as given below:

$$N = N_0 \times exp\left(\frac{-z}{\lambda_{int}}\right) \tag{7.2}$$

where N is the number of hadrons surviving without undergoing any hadronic interaction at a distance z into the detector, N_0 is initial number of incident hadron, and λ_{int} is the nuclear interaction length of the detector.

It follows from equation 7.2 that the probability of hadronic interaction is strongly dependent on λ_{int} of the detector. Precise identification of the shower starting location of hadrons within the detector opens up plethora of avenues to study and optimize the detector performance, such as optimized energy reconstruction, hadronic shower development, validation of hadronic physics modelling of the simulation etc.

7.3.1 A preliminary shower start finder algorithm

The simplest and most obvious choice to develop such an algorithm is to take advantage of the fact that a hadron acts like a MIP and deposits very low energy in the detector before it starts showering. A simple trial algorithm based on the trend of energy deposited in the consecutive layers is employed to identify the shower start location. It uses the energy deposited at the i^{th} as well as previous two layers of the detector. The algorithm is described below:

1. Compute energy sum (E) in a circle with radius R $<$ 2 cm at layer i with track impact position as the center.
2. If at first CE-E layer, the condition ($E_{Layer\ 1}$ $>$20 MIPs) is satisfied, then shower started at layer 1 and exit from the algorithm, otherwise go to next layer.
3. If at second CE-E layer, the conditions ($E_{Layer\ 2}$ $>$ 20 MIPs) AND ($E_{Layer\ 2}$ $>$ 2 $\times$ $E_{Layer\ 1}$) are satisfied, then shower started at layer 2 and exit from the algorithm, otherwise go to next layer.
4. If at next i^{th} layer, the conditions ($E_{Layer\ i}$ $>$20 MIPs) AND ($E_{Layer\ i}$ $>$ 2 $\times$ $E_{Layer\ i\text{-}1}$) AND ($E_{Layer\ i}$ $>$ 2$\times$ $E_{Layer\ i\text{-}2}$) are satisfied, then shower started at layer i and exit from the algorithm, otherwise go to next layer.
5. If none of the above condition is satisfied in any of the HGCAL layers, then shower start is NOT FOUND and exit from the algorithm.

The algorithm can only predict the layer downstream of true shower start point. For example, if the true shower start point is between layer 7 and 8, then the very best estimation of algorithm will point to layer 8. Therefore, the events get integrated (ΔN) for the shower start between two consecutive layers with absorber thickness Δz as shown below:

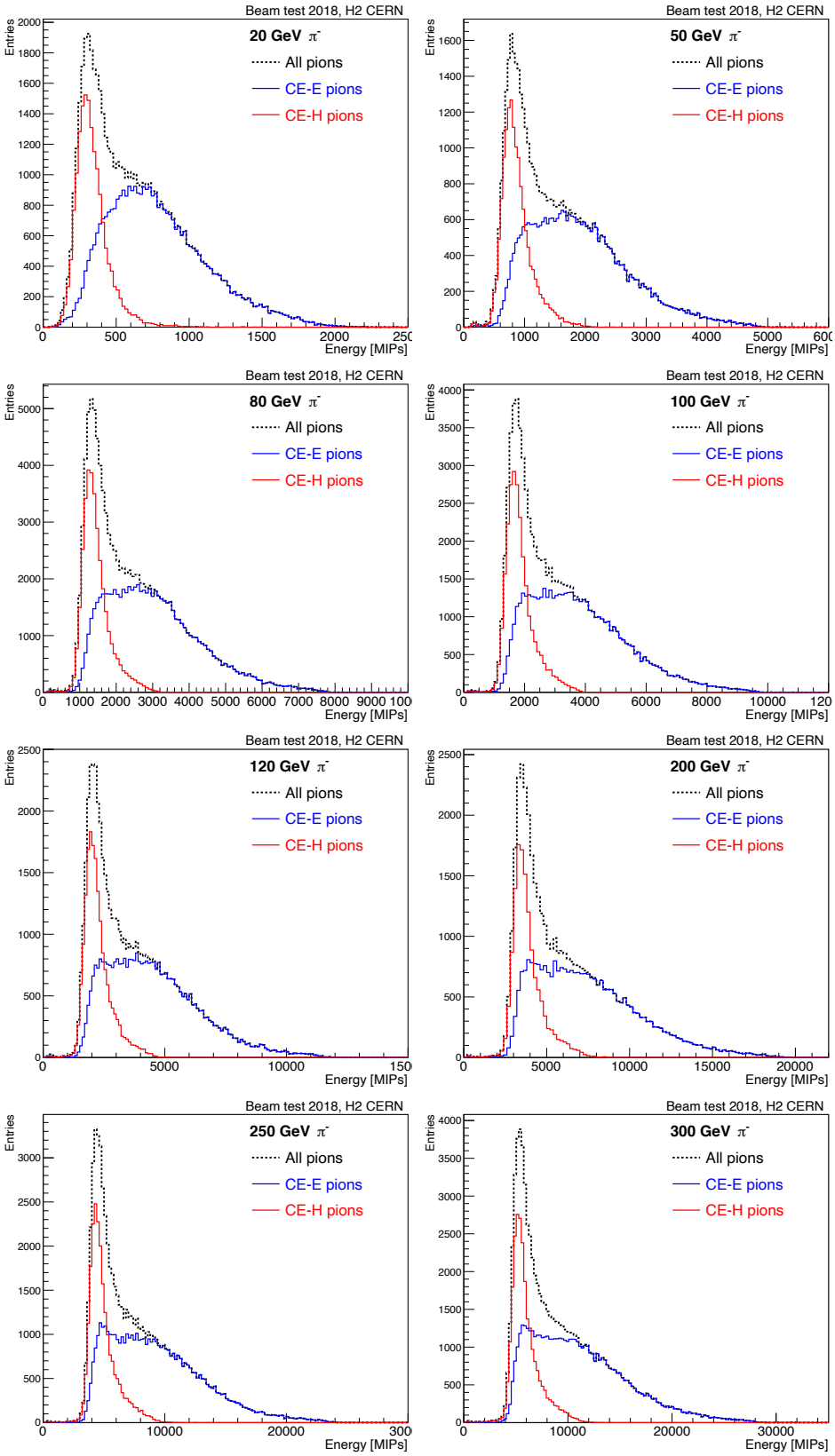

FIGURE 7.9: Figure shows energy distribution for all pion beam energies. Blue and red solid line corresponds to energy distribution for events when pion starts to shower in CE-E (CE-E pions) and when it is MIP-like in CE-E (CE-H pions), respectively. Black dotted line corresponds to both types of pions combined.

$$\begin{aligned}\Delta N &= \int_{z}^{z+\Delta z} N_0 e^{-z} dz \\ &= N_0 \cdot e^{-z} \cdot \left(1 - e^{-\Delta z}\right)\end{aligned}$$

Taylor expanding $e^{-\Delta z}$ gives us,

$$\begin{aligned}\Delta N &= N_0 \cdot e^{-z} \cdot \{1 - (1 + (-\Delta z) + \mathcal{O}(\Delta z))\} \\ \Delta N &= N_0 \cdot e^{-z} \cdot \Delta z \\ \Delta N &= N \cdot \Delta z \\ \implies \Delta N &\propto \Delta z\end{aligned}$$

Therefore, if the distance between two active layers (i and i-1) are larger in terms of λ_{int}, then the predicted number of pions start showering at layer i will be larger. This behaviour is evident if the shower starting events is plotted as a function of layer (in λ_{int}) as shown in Figure 7.10 (left). For example, a jump in the number of shower starting events is observed at the interface of CE-E and CE-H which is an artefact of higher absorber thickness in CE-H as compared to CE-E (see Figure 6.11).

To account for this fact, the predicted number of shower starting at i^{th} layer is normalized by $\Delta\lambda^{i}_{int}$. The entries corresponding to the first layer of CE-E is not re-weighted because of the fact that there is no active layer before it and $\Delta\lambda_{int}$ from particle production till first layer is very small. Entries in back-to-back layers in CE-E are also summed up to make the distribution more compact. Thus, it results in expected exponentially falling behaviour of number of showering pions as a function of calorimeter depth, shown in Figure 7.10 (right). The algorithm predicts that for $\sim$ 70% events, the pion start showering in CE-E, which is consistent with the fact that CE-E is 1.4 λ_{int} deep.

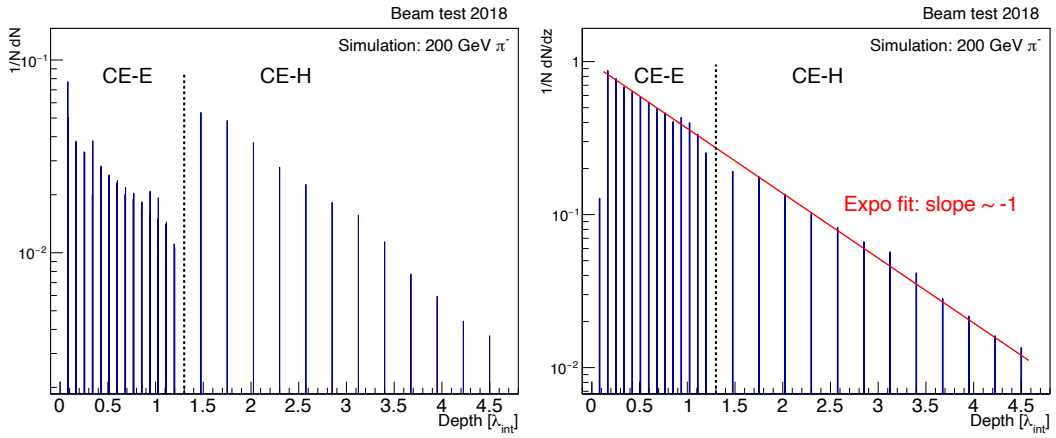

FIGURE 7.10: The number (normalized by total number of events) of surviving pions without initiating shower as a function of calorimeter depth (in λ_{int}) as predicted by preliminary version of shower start finder algorithm. The plot on the left shows the numbers before weighting with $\Delta\lambda_{int}$ at each layer and right plot shows after weighting it along with a exponential fit.

Though the preliminary version of algorithm performs reasonably well for high energy pions, it shows sub-optimal performance for low energy pions. This shortcoming becomes evident when the algorithm is used for the event categorization purpose. Table 7.5 shows fraction of events in different categories as predicted by preliminary algorithm.

TABLE 7.5: Fraction of events segregated into different categories after employing preliminary version of shower start finder algorithm in beam test data.

Beam energy	CE-E pions	CE-H pions	Shower start not found
20 GeV	56.0%	13.7%	30.2%
50 GeV	67.2%	24.1%	8.7%
80 GeV	68.4%	25.9%	5.6%
100 GeV	69.33%	25.9%	4.8%
120 GeV	69.8%	25.8%	4.3%
200 GeV	71.1%	26.0%	2.8%
250 GeV	72.3%	25.0%	2.6%
300 GeV	73.1%	24.6%	2.2%

It is observed that for more than 30% events in 20 GeV pion data, this algorithm fails to find shower start location whereas it falls to below 3% for 200 GeV pion data. Such events do not fall into either CE-E or CE-H pion categories, hence are rejected. About 1.4% pions are expected to survive without undergoing hadronic interaction by the end of the last layer of CE-H ($\lambda_{pion} \sim 4.26$). Figure 7.11 shows energy comparisons for 20 and 200 GeV pion data in CE-E (left) and CE-H (right) for the events where the algorithm fails to find any shower start layer. Energy distribution for 20 GeV pion clearly shows a longer tail and a bump in CE-E and CE-H, respectively. It points to the fact that there is a significant energy deposit in the compartments and therefore these events correspond to showering pions.

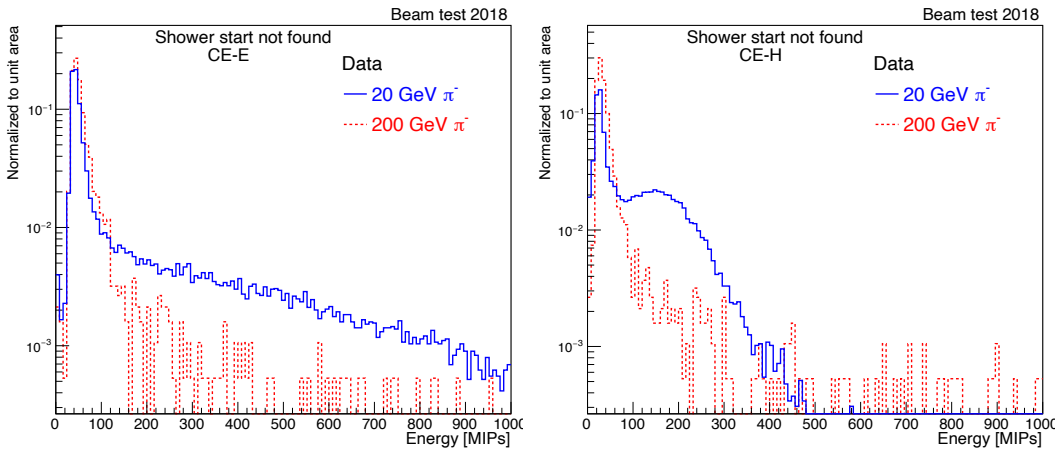

FIGURE 7.11: Energy distribution comparison between 20 GeV (blue solid line) and 200 GeV (red dotted line) pions for which preliminary version of shower start finder algorithm fails to find shower start location in beam test data.

Such sub-optimal performance of the preliminary version of algorithm warrants an improvement, especially for low energy pions. In order to assess the performance and optimization of the algorithm, the predicted shower start layer is compared against the true first hadronic interaction (*FHI*) layer extracted from simulation.

7.3.2 Extraction of true first hadronic interaction in simulation

The extraction of the location of FHI ($z^{\text{shower-start}}$) is possible in the GEANT4 simulation package of CMSSW framework using `simWatcher`, as discussed in section 2.7.3. For the purpose of hadronic shower analysis, if the primary `G4Track` (i.e. the particle shot from

the particle gun, π^- in our case) undergoes hadronic interaction (`HadInelastic`, `HadCapture`, `HadFission`, `HadAtRest` processes in GEANT4 framework), then following information is extracted and saved:

- (x, y, z) coordinates of the G4step
- number of secondaries produced at the interaction point
- particle IDs, charge, (x, y, z) coordinates of each secondary particle.
- kinetic energy of each secondary particle

If the primary track (π^-) does not undergo any of the hadronic interactions within HGCAL+AHCAL detector, then the event is rejected. About 40000 events are generated for each beam energy using FTFP_BERT_EMN physics list. The fraction of the events for which no hadronic interaction occurs within HGCAL+AHCAL is listed in Table 7.6. It is to be noted that it also contains events where π^- decays to μ^- in flight.

TABLE 7.6: Fraction of events in which pions survive without undergoing any hadronic interaction in HGCAL + AHCAL calorimeter.

Beam energy	Event fraction
20 GeV	2.80%
50 GeV	1.18%
80 GeV	0.64%
100 GeV	0.61%
120 GeV	0.56%
200 GeV	0.24%
250 GeV	0.24%
300 GeV	0.18%

The $z_{\text{true}}^{\text{shower-start}}$ coordinate of the interaction point is a continuous distribution, whereas the shower start finder algorithm gives shower start location in terms of layer number. Since the algorithm only predicts the layer just after the shower start location, therefore it is important to propagate $z_{\text{true}}^{\text{shower-start}}$ onto the next active layer in order to make one-to-one comparison. For each $z_{\text{true}}^{\text{shower-start}}$ extracted from simulation, the next nearest active layer ($L_{\text{true}}^{\text{shower-start}}$) is assigned as depicted in the representation in Figure 7.12.

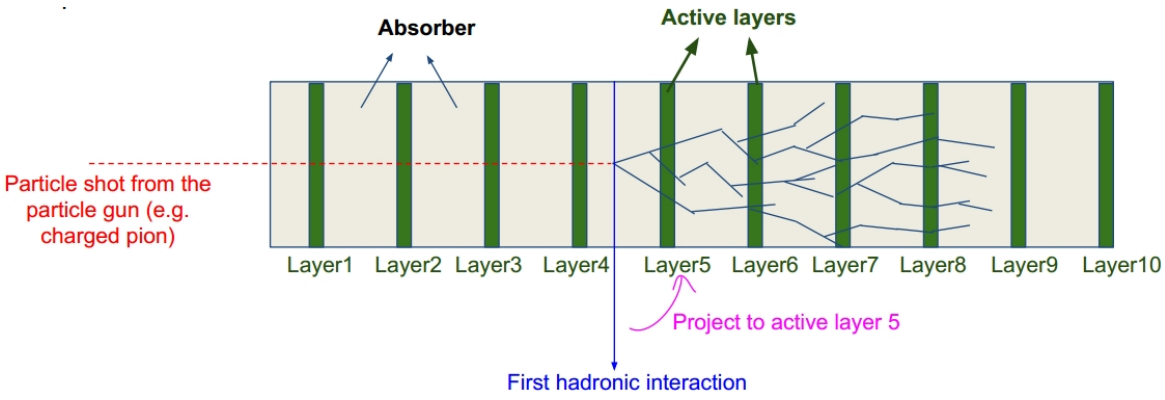

FIGURE 7.12: An image representation of the procedure to project FHI location onto next active layer. In this example, the shower starts between active layer 4 and 5, therefore layer 5 is chosen as FHI location to make one-to-one comparison with the shower start finder algorithm.

Events for which either the shower start is not found or is beyond HGCAL layers, layer = -1 is assigned. In GEANT4, the hadronic interaction includes both soft as well as hard hadronic interaction, therefore the actual shower development might not correspond to $L_{\text{true}}^{\text{shower-start}}$. This scenario is shown in Figure 7.13. The figure shows an event display for a 20 GeV π^- event where truth information indicates $L_{\text{true}}^{\text{shower-start}}$ as CE-E layer 12, whereas it can be observed that the shower does not start until later layers of CE-H.

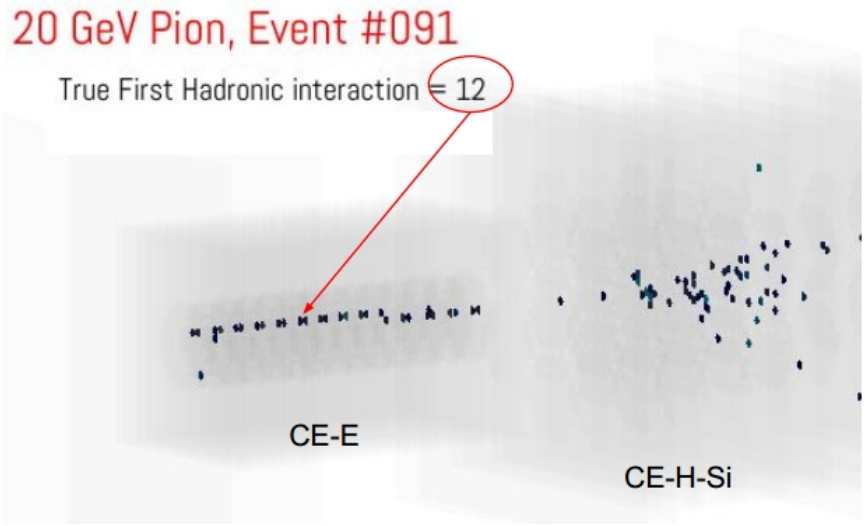

FIGURE 7.13: Event display of an 20 GeV/c simulation event with first hadronic interaction as *soft* hadronic interaction.

The events with such *soft* interactions should be removed not just to assess the performance of the algorithm but also for optimization. In such soft interaction events, the incoming hadron produces a small number of secondary particles at $z_{\text{true}}^{\text{shower-start}}$ and the momentum transfer to secondaries is also expected to be small. Therefore, the truth information is used in order to remove such events, such as the number of secondaries produced at the interaction point and kinetic energies carried by them. However, since GEANT4 treats all the particles as secondary particles after the G4Step undergoes an interaction, therefore one can not distinguish between primary and secondary particles after the interaction. Hence, the kinetic energy sum carried by secondaries is calculated as follows:

$$E_{KE}^{Sec} = \sum E_{KE}^{allSec} - E_{KE}^{leadHadron} \tag{7.3}$$

where $\sum E_{KE}^{allSec}$ is the kinetic energy sum of all the secondary particles produced at the interaction, $E_{KE}^{leadHadron}$ is the kinetic energy of leading hadron of same species and charge as that of incident hadron (π^- in this case) and E_{KE}^{Sec} is the kinetic energy carried by all the secondary particles.

The E_{KE}^{Sec} is further normalized by the beam energy in order to decide a single cut for *soft* hadronic interaction event rejection. Figure 7.14 shows a correlation plot between number of secondary particles produced at $L_{\text{true}}^{\text{shower-start}}$ versus fractional kinetic energy (E_{KE}^{Sec}/E_{beam}) carried by secondary particles for a few beam energy pions. The figure shows that there is a population of events around $(0,0)$. This population corresponds to the events where the number of secondaries as well as the kinetic energy transfer is very small. The type of events as shown in the event display Figure 7.13 belongs to this population.

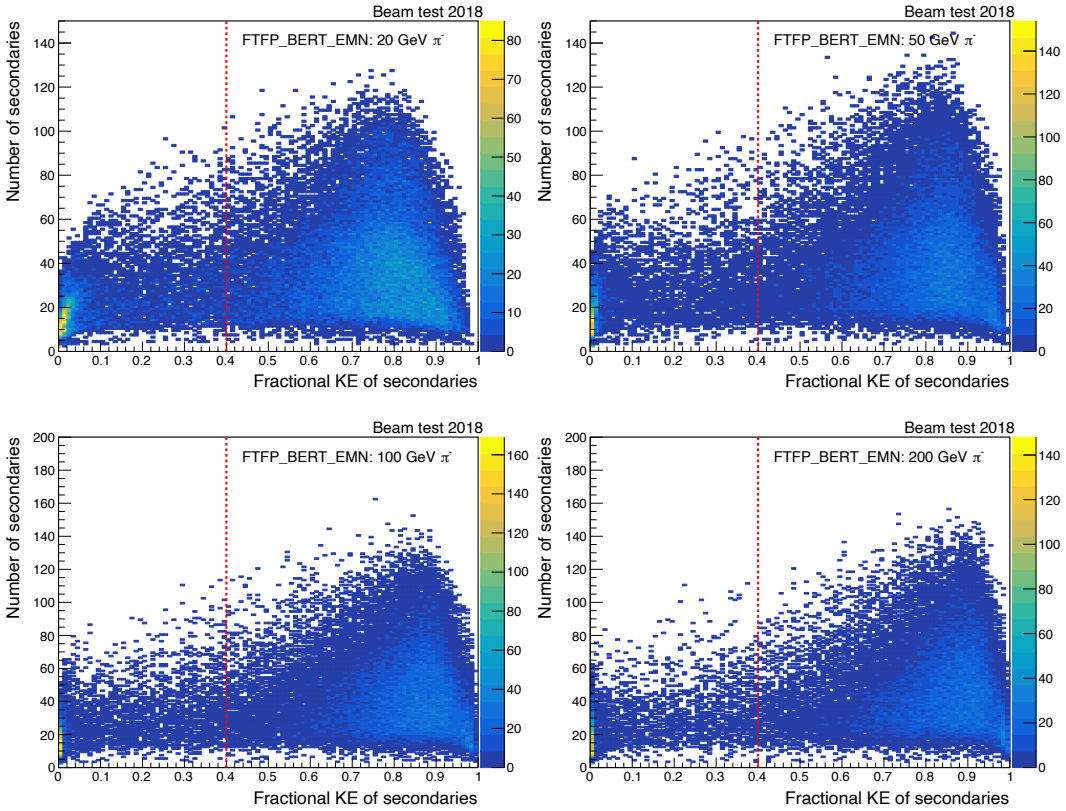

FIGURE 7.14: The correlation between fractional kinetic energy carried by secondary particles versus number of secondary particles produced at the first hadronic interaction for different π^- beam. The population at the left bottom corner of the plot corresponds to interactions where the number of secondary particles produced are small and the fractional energy carried by the incident π^- is large. A cut at 0.4 as indicated by vertical dotted red line is applied to reject such events for optimization purpose.

In order to reject such events, a cut is applied on fractional kinetic energy carried by secondaries. The cut is chosen to be greater than 0.4 (indicated by red dotted line in Figure 7.14) and events below this cut are rejected for all beam energies. This cut ensures that only hard hadronic interaction events are retained for the optimization purpose. Equipped with hard-hadronic interaction events, one can now optimize the algorithm using truth information as a reference. Following reconstructed variables are used for the optimization:

- **Energy threshold** : Since the energy deposited in the detector by hadrons is non-linear, therefore an energy dependent threshold is chosen for consistent performance.
- **Radius of energy sum** : Hadronic shower has a wider lateral spread, therefore a larger radius of the circle is chosen.
- **Number of reconstructed hits** : HGCAL prototype is a highly granular detector, hence the number of reconstructed hits at a given layer is also used.

7.3.3 Algorithm optimization

In the preliminary shower start finder algorithm, the radius to sum up the energy is taken as 2 cm and the center of the circle is defined as the track impact position extrapolated from the delay wire chambers (DWC). Since λ_{int} is $\sim$ 19.9 cm and $\sim$ 20.4 cm in Pb and Fe respectively, therefore the hadronic shower containment in a circle of 2 cm radius is expected be very small. Here, Fe is taken as a proxy for steel absorbers (CE-H absorber). A circle of 10 cm radius is considered as an optimum value which is about half λ_{int} for Pb and Fe and also covers a full silicon module, i.e. a full sampling layer in CE-E and central module in CE-H. The fact is also considered that a slight deflection in the direction of incident hadron, due to scattering, might cause a deviation from the extrapolated track from DWCs. This might make the track impact position to differ from the actual incident hadron impact point later in the layers. Therefore, *center of gravity* (COG) is taken as the center of the circle for summing up the energy and number of reconstructed hits, instead of track impact position. The center of gravity is defined as follows:

$$\begin{aligned} x^i_{\mathrm{COG}} &= \frac{\sum_j x^i_j \times E^i_j}{\sum_j E^i_j} \; ; i \in [1,40] \text{ and } j \in \text{[reconstructed hits at layer i]} \\ y^i_{\mathrm{COG}} &= \frac{\sum_j y^i_j \times E^i_j}{\sum_j E^i_j} \; ; i \in [1,40] \text{ and } j \in \text{[reconstructed hits at layer i]} \end{aligned} \tag{7.4}$$

Three discriminators are used for the shower start finder algorithm optimization, viz. number of reconstructed hits, energy sum within a circle of radius 10 cm around COG and transverse energy spread. The transverse energy spread is calculated in terms of lateral energy ratio, averaged over three consecutive layers, denoted as R_i, which is defined in the equation 7.5 below:

$$\mathrm{R}_i = \frac{\sum_{j=i}^{i+2} En^j_{2cm}}{\sum_{j=i}^{i+2} En^j_{10cm}} \tag{7.5}$$

where the numerator and denominator is the energy summed in a circle of 2 cm and 10 cm, respectively, around the COG in three consecutive layers. For last and second-last layer of CE-H, the summation runs over two and one layer, respectively. The variable R_i encapsulates the shower development information in a moving window of three layers which is

expected to peak at 1 for MIP like particles and lower than 1 for showering particles.

For the threshold optimization of these three discriminators, the truth information of shower start layer is used and the corresponding distributions are compared with μ^- as a reference for MIP-like pions i.e. before pions initiate a shower. The distributions are compared at $L_{true}^{shower\text{-}start}$, referred to as *layer i*. Figure 7.15 and 7.16 shows such comparisons between muons and pions of 20 GeV and 100 GeV, respectively, at *layer i* and *layer i - 2* in CE-E and CE-H. It is observed that two layers before pion starts showering (*layer i - 2*), the distributions of number of reconstructed hits, energy sum and R_i, are similar between muons and pions since both act as MIP-like particle. Whereas, when pion starts showering the distribution show different behaviour as compared to MIP as expected, such as high hit multiplicity, high energy and relatively large transverse shower spread; for both high as well as low energy pions. Based on the distribution comparisons at i^{th} layer, the hit multiplicity and R_i threshold is chosen as greater than 3 and less than 0.96, respectively. The threshold for these two variables is fixed for all beam energies. However, the energy deposition threshold is made energy dependent, described later in this section, in order to achieve consistent performance across all beam energies.

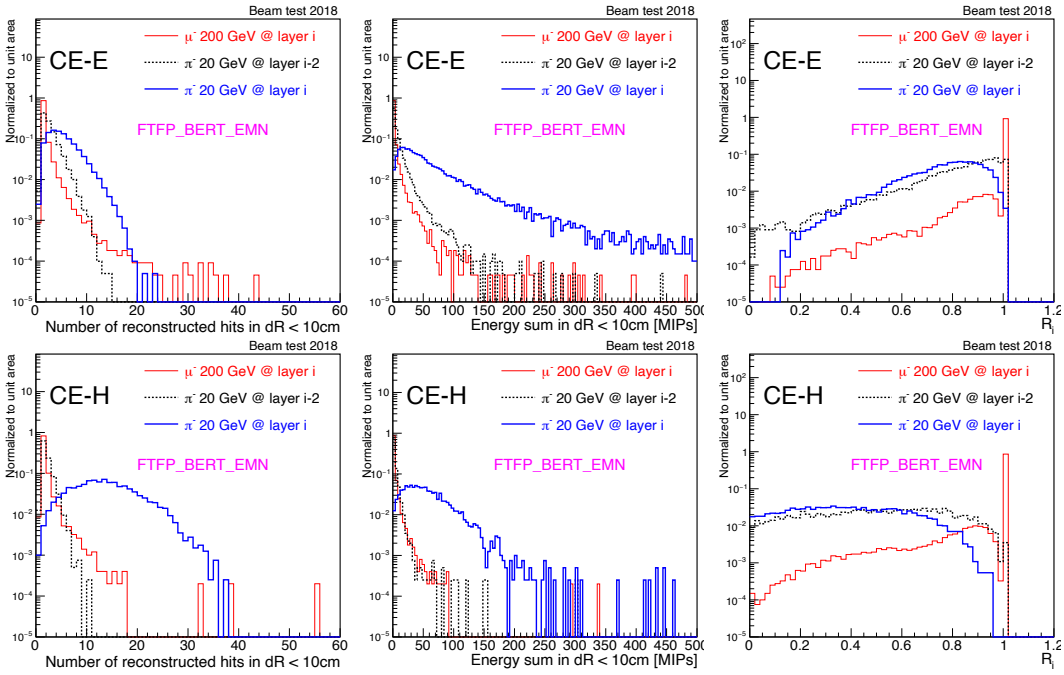

FIGURE 7.15: Comparisons of number of reconstructed hits (left column), energy (middle column) and lateral energy ratio (right column) in CE-E (top row) and CE-H (bottom row) between muons and **20 GeV pions** at shower starting layer and previous-to-previous layer.

Combining all these information, the optimized shower start finder algorithm can be summarised as follows:

1. At the i^{th} layer, calculate number of reconstructed hits (Nrechits^i), energy sum (En^i) in a circle of radius 10 cm around COG, and transverse shower spread (R_i).

2. If the conditions (Nrechits$^i > 3$) AND (En$^i >$ `E_thres`) AND ($R_i < 0.96$) are satisfied, then shower started at layer i and exit from the algorithm, otherwise go to next layer.

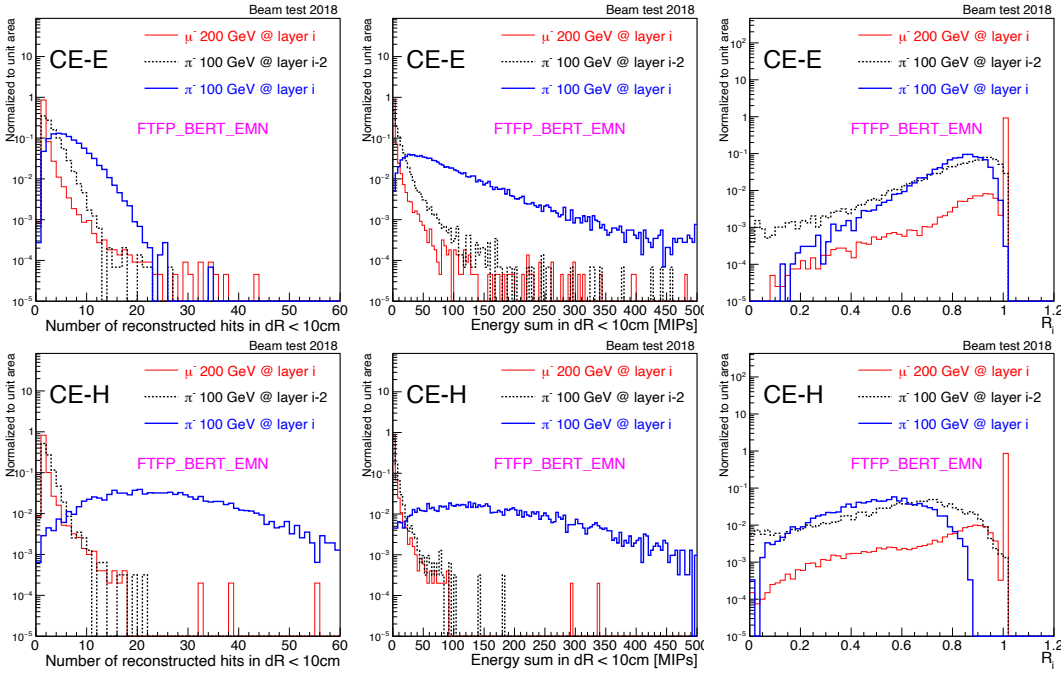

FIGURE 7.16: Comparisons of number of reconstructed hits (left column), energy (middle column) and lateral energy ratio (right column) in CE-E (top row) and CE-H (bottom row) between muons and **100 GeV pions** at shower starting layer and previous-to-previous layer.

3. If above condition is not satisfied in any of the HGCAL layers, then shower start is NOT FOUND and exit from the algorithm.

The value of `E_thres` is determined by scanning different values of `E_thres` with respect to *efficiency*, for each beam energy. The value corresponding to which the *efficiency* is maximum is set as `E_thres` for a given beam energy. The *efficiency* is defined as the fraction of events for which the reconstructed shower start layer ($L_{reco}^{shower\text{-}start}$) falls within $\pm n$ layers with respect to $L_{true}^{shower\text{-}start}$ and is defined as follows:

$$\text{Efficiency} = \frac{\text{Events with } |L_{true}^{shower\text{-}start} - L_{reco}^{shower\text{-}start}| \leq n}{\text{Events with } L_{true}^{shower\text{-}start}} \tag{7.6}$$

The window n for efficiency calculation depends on $\Delta\lambda_{int}$ between active layers of sub-detector. Since $\Delta\lambda_{int}$ in CE-E ($\Delta\lambda_{int} \sim 0.05$) is smaller than CE-H ($\Delta\lambda_{int} \sim 0.28$), therefore for CE-E the window size is chosen as ±1 , ±2 and ±3 layers; whereas for CE-H the window size is chosen as ±1 and ±2 layers.

Figure 7.17 shows `E_thres` versus efficiency in the window of ±2 layers for CE-E and ±1 for CE-H as well as in ±2 layers for both combined, for all beam energies. The value of `E_thres` is chosen for which the efficiency peaks for combined detector. The `E_thres` values are listed in table 7.7 for all beam energies.

While in this case, the beam energy is used to extract `E_thres` value, in the real experiment the information about true energy of incident particle will not be known. In that case, track momentum of the particle can be used as a reference to obtain the value of `E_thres`. Figure 7.18 shows `E_thres` as a function of beam energy fitted with two separate functions as shown in below:

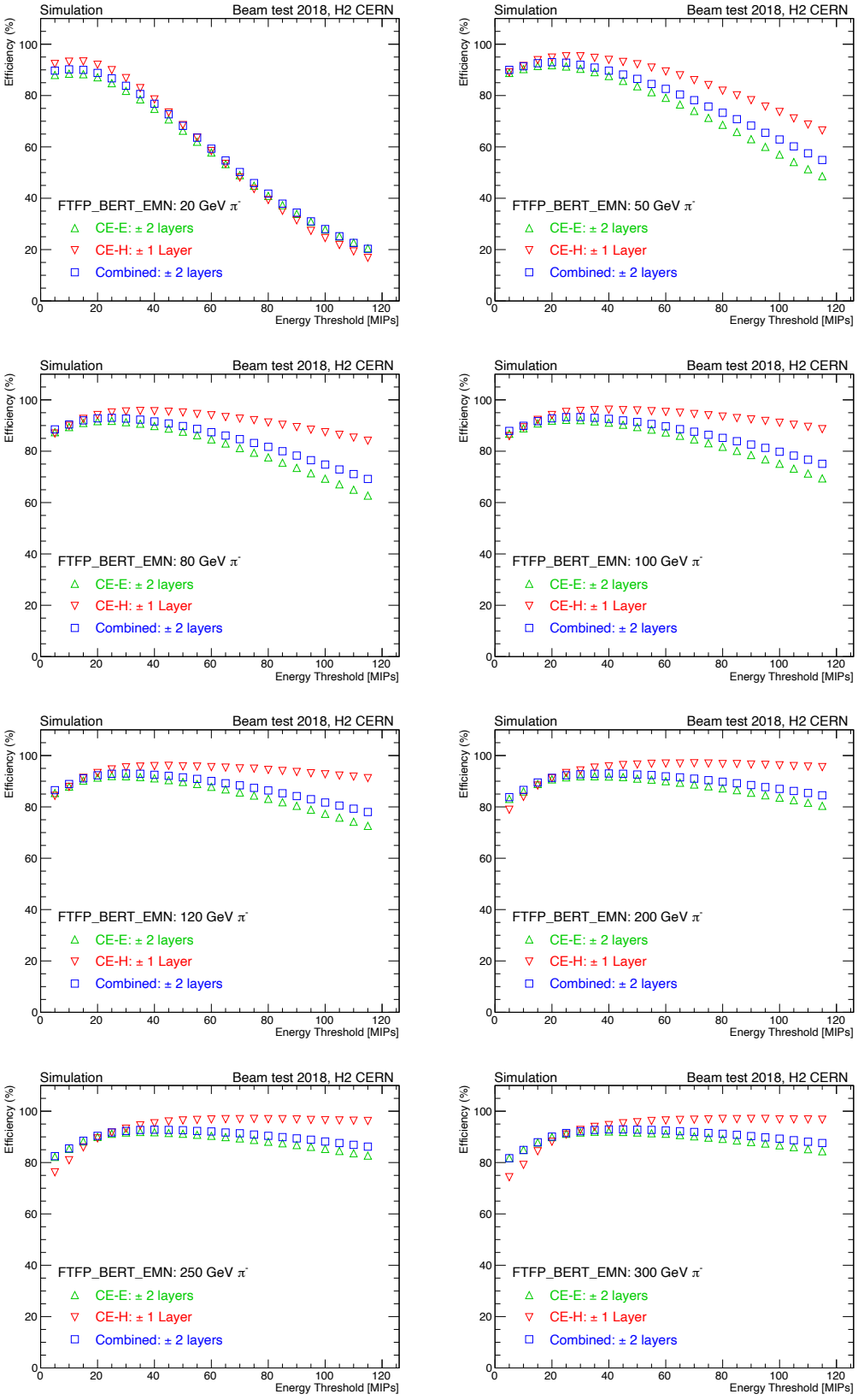

FIGURE 7.17: Efficiency versus `E_thres` values for all beam energies in different window n, The value at which the efficiency peaks is chosen to be `E_thres` value for corresponding beam energy.

TABLE 7.7: The table for `E_thres` value in terms of MIPs for all beam energies.

Beam energy [GeV]	`E_thres` [MIPs]
20	12
50	20
80	25
100	30
120	30
200	40
250	40
300	40

$$E_{thres} = \begin{cases} -6.6 \times 10^{-7} \cdot E^3 - 4.1 \times 10^{-4} \cdot E^2 + 0.275 \cdot E + 6.8 & \text{, if } E < 200 \text{ GeV} \\ 40.0 & \text{, if } E \geq 200 \text{ GeV} \end{cases}$$

The fitted function can be used to extract the threshold values (`E_thres`) based on the reference.

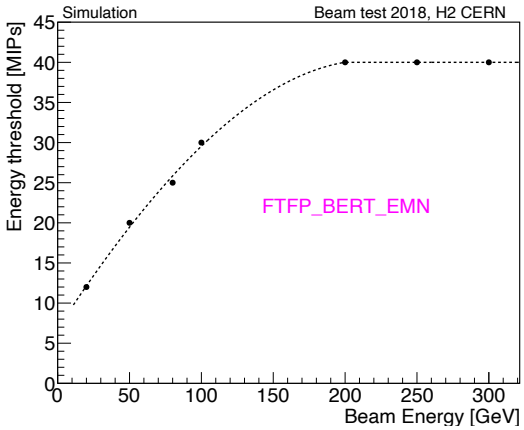

FIGURE 7.18: Optimized chosen threshold for `E_thres` as a function of beam energy fitted with two polynomial functions (dotted line).

7.3.4 Performance of optimized algorithm

The optimized (or new) shower start finder algorithm shows significant improvement over the preliminary version of the algorithm. The comparison between preliminary and new version of the algorithm is shown in tabular format in Table 7.8 as well as in graphical format in Figure 7.19. The preliminary version has a good performance for beam energies greater than 100 GeV, however, it shows very poor performance for lower beam energies where the efficiency is observed to be as low as 52.0% for CE-H within $\pm$ 1 layer. The optimized algorithm shows a consistent performance across all beam energies for both CE-E and CE-H prototype. It should be noted that a slightly lower efficiency of the algorithm is observed in CE-E as compared to CE-H. This is because of the fact that even for a window of $n = \pm$ 2 layers in CE-E, the available depth for the shower to develop is $\sim 0.2\ \lambda_{int}$ (= 0.05×4; i.e. 2 CE-E layers before and 2 CE-E layers after $L_{\text{true}}^{\text{shower-start}}$) whereas for CE-H a window of $n = \pm$ 1 layer is $\sim 0.56\ \lambda_{int}$ (0.28×2). Therefore in CE-H, the thresholds are

easily satisfied. This feature was also observed while tuning the threshold parameters as is shown in Figure 7.15 and 7.16.

TABLE 7.8: Efficiency table in different window n for preliminary version (in black) and new version (in red) of shower start finder algorithm, in CE-E and CE-H prototype. The new version of the algorithm shows higher efficiency for all beam energies.

Energy [GeV]	20	50	80	100	120	200	250	300
CE-E (within ±1 layer)	67.6% (76.5%)	80.1% (81.7%)	82.9% (82.9%)	84.7% (83.5%)	84.7% (83.5%)	85.1% (84.0%)	84.7% (84.1%)	84.8% (84.8%)
CE-E (within ±2 layer)	75.1% (88.6%)	86.6% (91.8%)	88.6% (91.8%)	89.7% (92.1%)	89.4% (91.9%)	89.0% (91.8%)	88.7% (91.8%)	88.5% (92.1%)
CE-E (within ±3 layer)	89.7% (93.3%)	89.2% (95.1%)	90.6% (94.8%)	91.6% (94.9%)	91.2% (94.6%)	90.6% (94.2%)	90.2% (94.1%)	90.3% (94.3%)
CE-H (within ±1 layer)	52.0% (93.4%)	83.4% (94.8%)	88.1% (95.1%)	89.3% (95.7%)	88.4% (95.4%)	88.4% (95.8%)	86.8% (95.3%)	86.7% (94.6%)
CE-H (within ±2 layer)	53.8% (95.7%)	85.7% (96.7%)	89.9% (96.6%)	90.9% (96.8%)	89.8% (96.4%)	89.6% (96.8%)	87.8% (96.2%)	87.7% (95.6%)

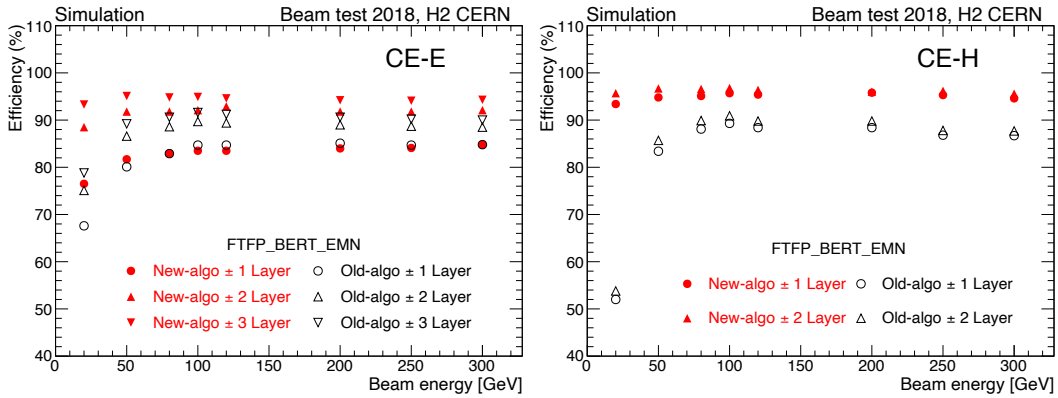

FIGURE 7.19: The efficiency comparison between preliminary and new version of shower start finder algorithm in CE-E (left) and CE-H (right) prototype. Different style of markers depict different window n used for efficiency calculation. The new version of the algorithm shows a consistent performance across all beam energies and shows higher efficiency as compared to preliminary version.

Furthermore, in order to also make sure that the optimized cuts are not biased towards one particular physics modelling, the validity is checked by assessing the performance of the algorithm on the simulation samples produced with QGSP_FTFP_BERT_EMN physics list. The algorithm is employed in the simulation samples without changing any parameter of the algorithm and the efficiency is calculated in different window n of $L_{\text{true}}^{\text{shower-start}}$ in CE-E and CE-H. Figure 7.20 shows the efficiency comparison for different samples, and is found to be in good agreement with each other. The efficiency numbers have been shown in the tabular format in Table 7.9.

Finally, the optimized algorithm is employed in the beam test data and compared with the simulated sample with FTFP_BERT_EMN and QGSP_FTFP_BERT_EMN physics list. Figure 7.21 shows the number of pions, normalized with $\Delta\lambda_{int}$, surviving without undergoing

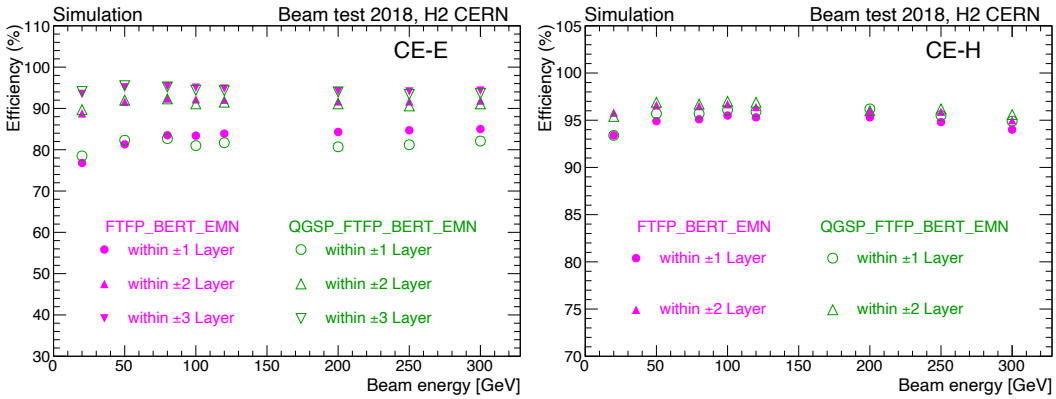

FIGURE 7.20: Figure shows the efficiency of the algorithm for simulation samples produced with two physics lists (shown in different colours) in CE-E (left) and CE-H (right). The algorithm shows similar performance and is in good agreement with each other.

TABLE 7.9: The table lists the efficiency values for three physics lists for different beam, energies in the window n in CE-E and CE-H. Magenta and green colours corresponds to FTFP_BERT_EMN and QGSP_FTFP_BERT_EMN physics list, respectively.

Energy [GeV]	20	50	80	100	120	200	250	300
CE-E (within ±1 layer)	(76.8%) (78.5%)	(81.3%) (82.3%)	(83.5%) (82.7%)	(83.4%) (81.0%)	(83.9%) (91.7%)	(84.3%) (80.7%)	(94.7%) (81.2%)	(85.0%) (82.1%)
CE-E (within ±2 layer)	(88.8%) (89.7%)	(91.6%) (92.0%)	(92.4%) (92.3%)	(92.2%) (91.1%)	(92.1%) (91.5%)	(91.7%) (91.1%)	(91.7%) (90.6%)	(91.9%) (91.1%)
CE-E (within ±3 layer)	(93.5%) (94.1%)	(95.1%) (95.5%)	(95.4%) (95.2%)	(95.0%) (94.4%)	(94.8%) (94.3%)	(94.2%) (94.0%)	(94.1%) (93.4%)	(94.3%) (93.6%)
CE-H (within ±1 layer)	(93.4%) (93.4%)	(94.9%) (95.7%)	(95.1%) (95.7%)	(95.5%) (96.1%)	(95.3%) (95.9%)	(95.3%) (96.2%)	(94.8%) (95.5%)	(94.0%) (94.9%)
CE-H (within ±2 layer)	(95.8%) (95.4%)	(96.6%) (96.9%)	(96.5%) (96.7%)	(96.7%) (97.0%)	(96.4%) (96.9%)	(96.4%) (96.8%)	(95.9%) (96.2%)	(95.0%) (95.6%)

hadronic interaction as a function of calorimeter depth in terms of λ_{int} in beam test data and is compared with two simulated samples for all beam energies. The distribution shows an exponentially falling trend as expected. Both data and simulation (with both the physics lists) show exponentially falling behaviour as expected, and show good agreement especially at higher beam energies. An exponential function is fitted to both data and simulation points to extract the interaction length for pions (λ_{pion}). We also checked that the data and simulation comparisons of λ_{pion} becomes more closer if one fits only the points in CE-H. Furthermore, it is found that λ_{pion} is $\sim 1.2 \times \lambda_{int}$ (see table 7.10). This observation is consistent with the fact that the interaction length λ (equation 2.43) depends on the interaction cross-section of incident hadron and the material nuclei. As discussed in detail in section 2.2.3 that the interaction cross-section for p-p (used in λ_{int} calculation) is higher than π-p interaction cross-section, which effectively translates to π^- seeing less material as compared to proton or neutron. The observation that λ_{pion} is $\sim 1.2 \times \lambda_{int}$ in HGCAL prototype agrees with with cross section measurements on iron [122] [123].

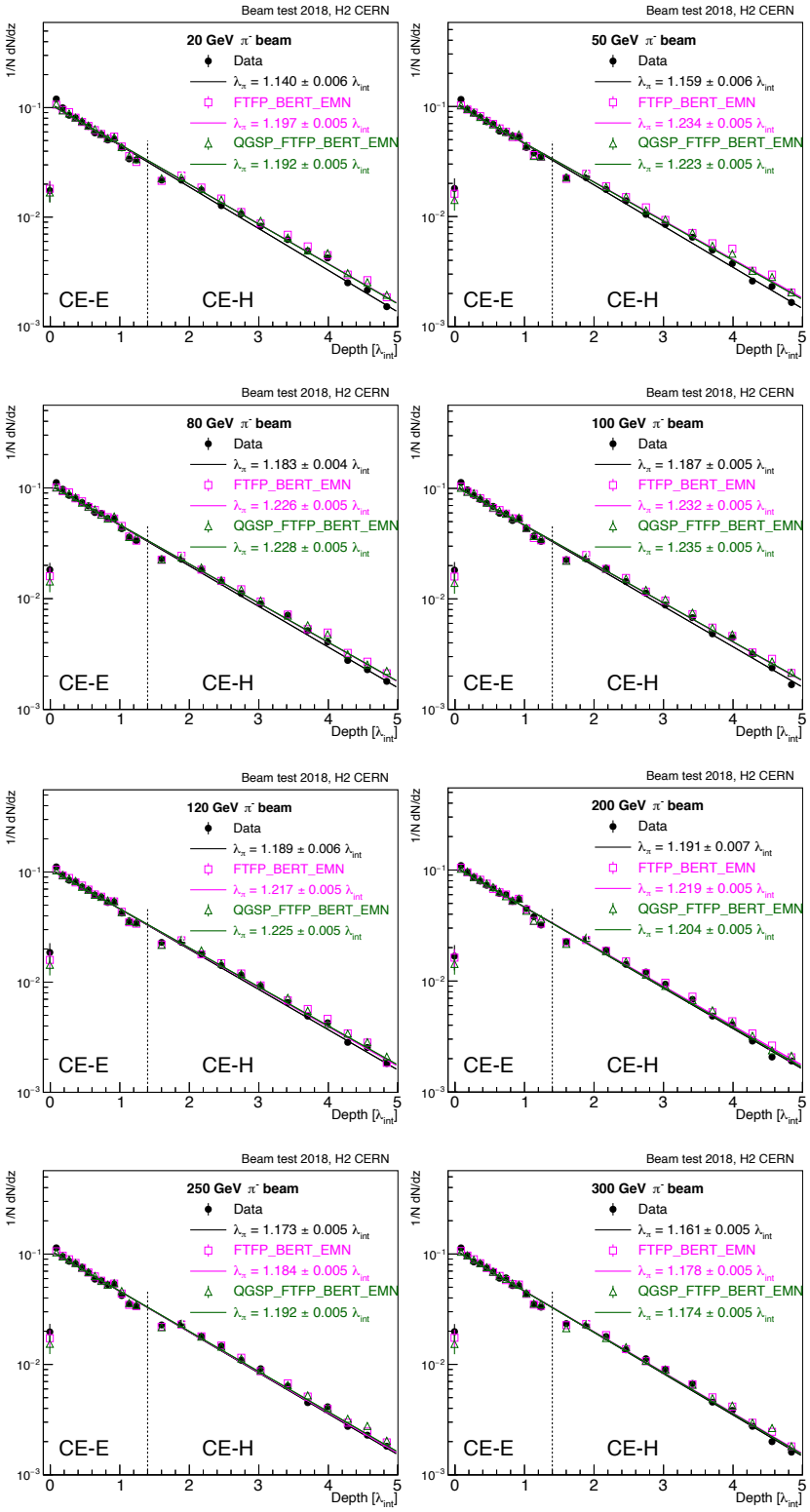

FIGURE 7.21: Figure shows the number of pions, normalized with $\Delta\lambda_{int}$, surviving without hadronic interaction as a function of calorimeter depth in terms of λ_{int} for all π^- beam energies. The distribution is fitted with an exponential function (solid lines), the fit result shows λ_{pion} is about $\sim$1.2 times larger than λ_{int} as expected.

E_{beam} [GeV]	Data [λ_{int}]	FTFP_BERT_EMN [λ_{int}]	QGSP_FTFP_BERT_EMN [λ_{int}]
20	1.140 ± 0.01	1.197 ± 0.01	1.192 ± 0.01
50	1.159 ± 0.01	1.234 ± 0.01	1.223 ± 0.01
80	1.183 ± 0.004	1.226 ± 0.01	1.228 ± 0.01
100	1.187 ± 0.01	1.232 ± 0.01	1.235 ± 0.01
120	1.189 ± 0.01	1.217 ± 0.01	1.225 ± 0.01
200	1.191 ± 0.01	1.219 ± 0.01	1.204 ± 0.01
250	1.173 ± 0.01	1.184 ± 0.01	1.192 ± 0.01
300	1.161 ± 0.01	1.178 ± 0.01	1.174 ± 0.01

TABLE 7.10: Table for λ_{int} values in data and simulation as obtained by fitting an exponential to pions of all beam energies as shown in Figure 7.21.

7.3.5 Event categorization using shower start finder algorithm

Equipped with an optimized algorithm to identify the shower start location in the HGCAL prototype, it can be used for event categorization purpose and segregate the events into CE-E and CE-H pions. Condition for categorizing into these two events are fairly straight forward as described below:

- If $L_{reco}^{shower\text{-}start}$ is within first 28 CE-E layers, then the pion is referred to as **CE-E pions**.
- If $L_{reco}^{shower\text{-}start}$ is in next 12 CE-H layers, then the pion is referred to as **CE-H pions**.
- If shower start is not found, the pion is referred to as MIP-like pions.

Having two methods of event categorization criteria, the correlation between the two methods is checked to make sure the optimized shower start finder algorithm is sufficient for this purpose. To find the correlation, both the methods are employed on each event in beam test data, and the fraction of events in each category is calculated based on given method. The numbers are visualized in a matrix form as shown in Figure 7.22 for all beam energies. The numbers in the columns correspond to shower start finder method and numbers in the rows correspond to energy-based method. The diagonal bins (with blue arrow) correspond to the events which are categorized in the same category by both the methods. Summing the event fractions from these three diagonal bins indicate the total fraction of events for which both the method categorize the event in the same category, hence the total correlation. The matrix shows that the correlation is more than 89% for all beam energies, therefore shower start finder algorithm is a good handle for event categorization and to probe longitudinal shower development for different depth shower start location.

7.4 Shower energy reconstruction using a fixed detector energy scale

As discussed earlier in this chapter that the pion shower energy measured in terms of MIPs yields a non-Gaussian distribution (Figure 7.9) because of the detector design, i.e. different absorbers and different sampling of electromagnetic and hadronic sections. Therefore, to measure the shower energy in terms of GeV, these two facts needs to be taken into account.

One of the methods to achieve this goal is by determining MIP-to-GeV energy scales for electromagnetic and hadronic sections. For electromagnetic section (CE-E), MIP-to-GeV energy scale is determined using e^+ beam and for hadronic section (CE-H + AHCAL), MIP-to-GeV energy scale is determined using π^- that are MIPs in CE-E. The reason to use e^+ to

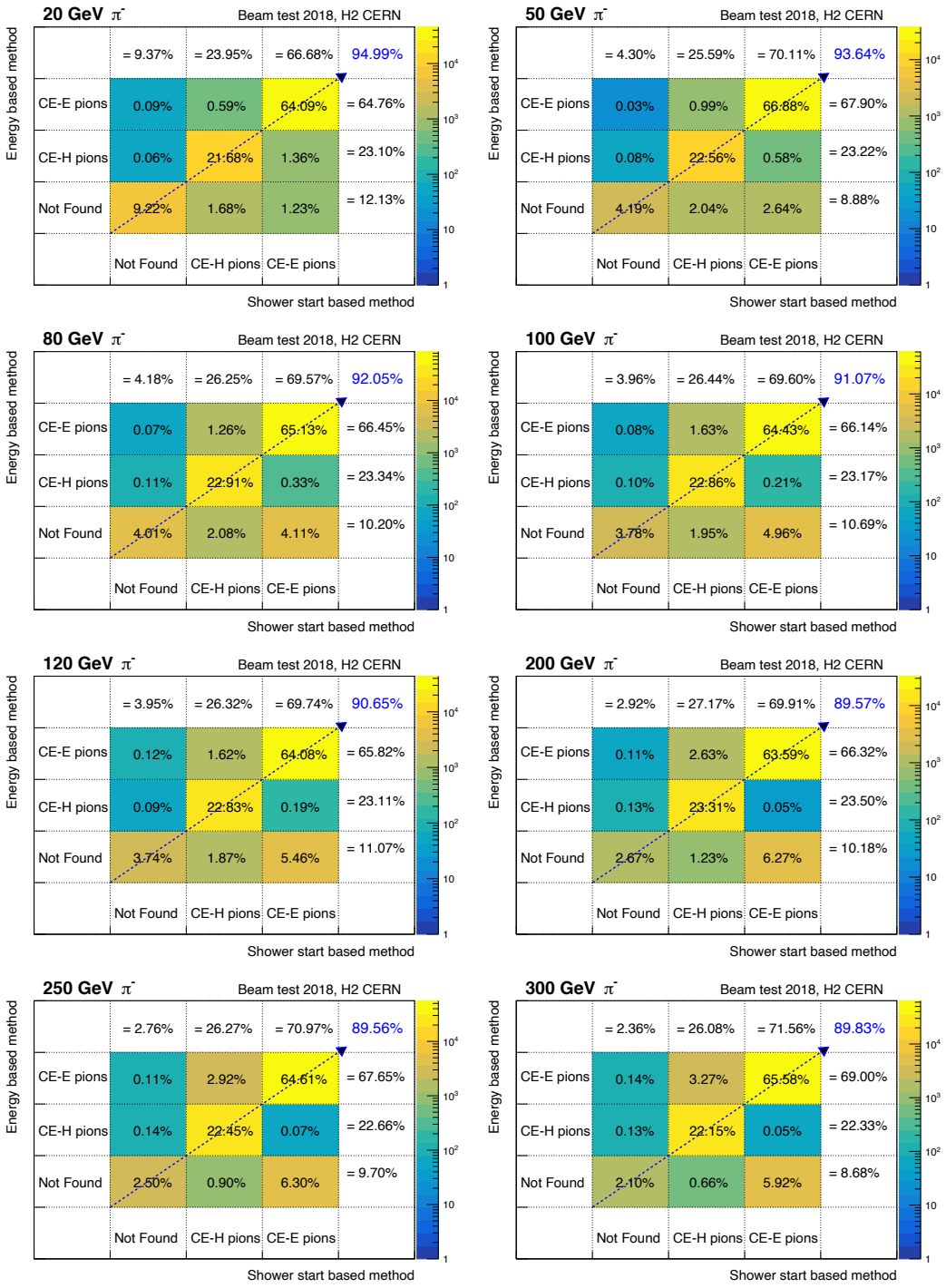

FIGURE 7.22: Figure shows the event fractions (in percent) for which the pion events are categorized in different categories based on shower-start-finder method (in columns) and energy-based method (in rows). The arrow depicts the common events which are categorized in the same categorize by both methods and the number the arrow points to, shows the total sum of such events.

calibrate CE-E is because of the fact that e^+ shower is almost completely contained in CE-E. It facilitates to disentangle the response of the electromagnetic and the hadronic sections. Figure 7.23 shows CE-E energy sum distribution for 50 GeV (left) and 100 GeV (right) e^+ beam test data, with a Gaussian function fitted within $\pm 1.5\ \sigma$ around the mean.

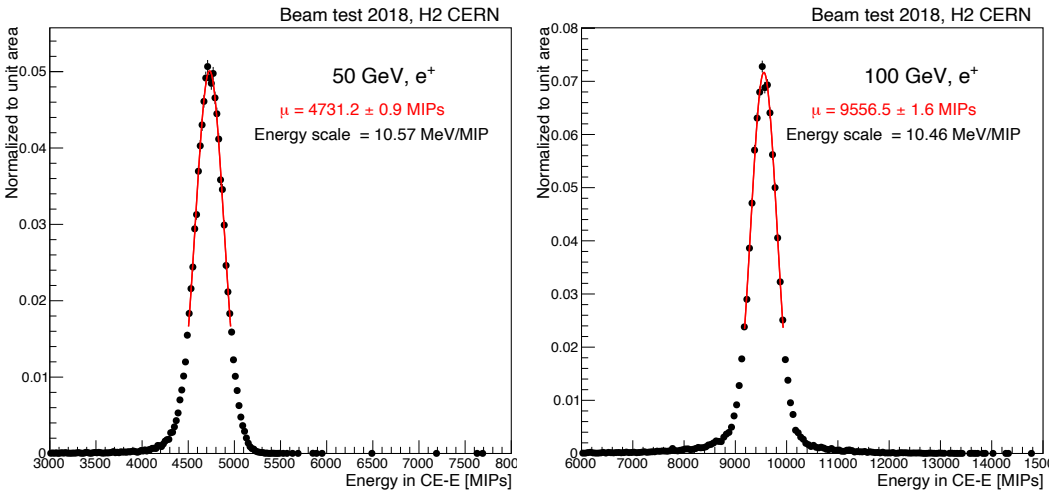

FIGURE 7.23: Energy sum distribution [in MIPs] of 50 GeV (left) and 100 GeV (right) e^+ beam in CE-E prototype, fitted with a Gaussian function.

The MIP-to-GeV energy scale for electromagnetic section (α^{fix}) is obtained by taking the ratio of e^+ beam energy to the mean (μ) of the fitted Gaussian. The energy scale obtained using 50 GeV and 100 GeV e^+ beam is found to be within 1%. This is expected since electromagnetic shower energy scales linearly with the incident energy of positron (section 2.6.1). The MIP-to-GeV energy scale for electromagnetic section α^{fix} is taken as 10.57 MeV/MIP as obtained from 50 GeV e^+ beam test data. This energy scale is consistent with an independent study performed for electromagnetic showers in CE-E prototype [118].

The energy scale of hadronic section, CE-H + AHCAL combined setup, is obtained using 50 GeV CE-H pions. Although the absorber material in both CE-H and AHCAL is steel but there are differences between the two detectors for example active material, sampling frequency, longitudinal and transverse dimensions as mentioned in chapter 6. The differences in the detection technologies, i.e. silicon for CE-H and scintillator on SiPM for AHCAL, is taken care of by converting the ADC data into corresponding MIP equivalent of energy deposit. However, other mentioned differences result in different MIP-to-GeV scales for the two sections. Therefore, it is important to introduce a *relative weight* (δ^{fix}) factor in order to correctly combine the energy shared between CE-H and AHCAL as shown in equation 7.7 below:

$$E^{\text{Had}}_{\text{MIPs}} = E^{\text{CE-H}}_{\text{MIPs}} + \delta^{fix} \times E^{\text{AHCAL}}_{\text{MIPs}} \tag{7.7}$$

The value of δ^{fix} is determined by minimizing the resolution of $E^{\text{Had}}_{\text{MIPs}}$ distribution by varying it from 0.02 to 1.0 in the steps of 0.02. For each value of δ^{fix}, a Gaussian function is fitted within $\pm 1.5\ \sigma$ around the mean of $E^{\text{Had}}_{\text{MIPs}}$ energy distribution. Figure 7.24 shows the example of such distributions of $E^{\text{Had}}_{\text{MIPs}}$ for different values of δ^{fix}, fitted with Gaussian function. Using mean (μ) and width (σ) of the fitted function, the resolution (σ/μ) is plotted as a function of δ^{fix} for all beam energies as shown in Figure 7.25 (left). The minima of resolution versus weight scan for all beam energies are observed at $\delta^{fix} \simeq 0.4$, therefore the δ^{fix} **is set to be 0.4**. After fixing the value of δ^{fix}, the overall MIP-to-GeV factor (β^{fix}) for the

hadronic section is obtained by taking the ratio of 50 GeV π^- beam energy to mean (μ) of the fitted Gaussian and is found to be $\simeq$ 78.9 MeV/MIP, as shown in Figure 7.25 (right).

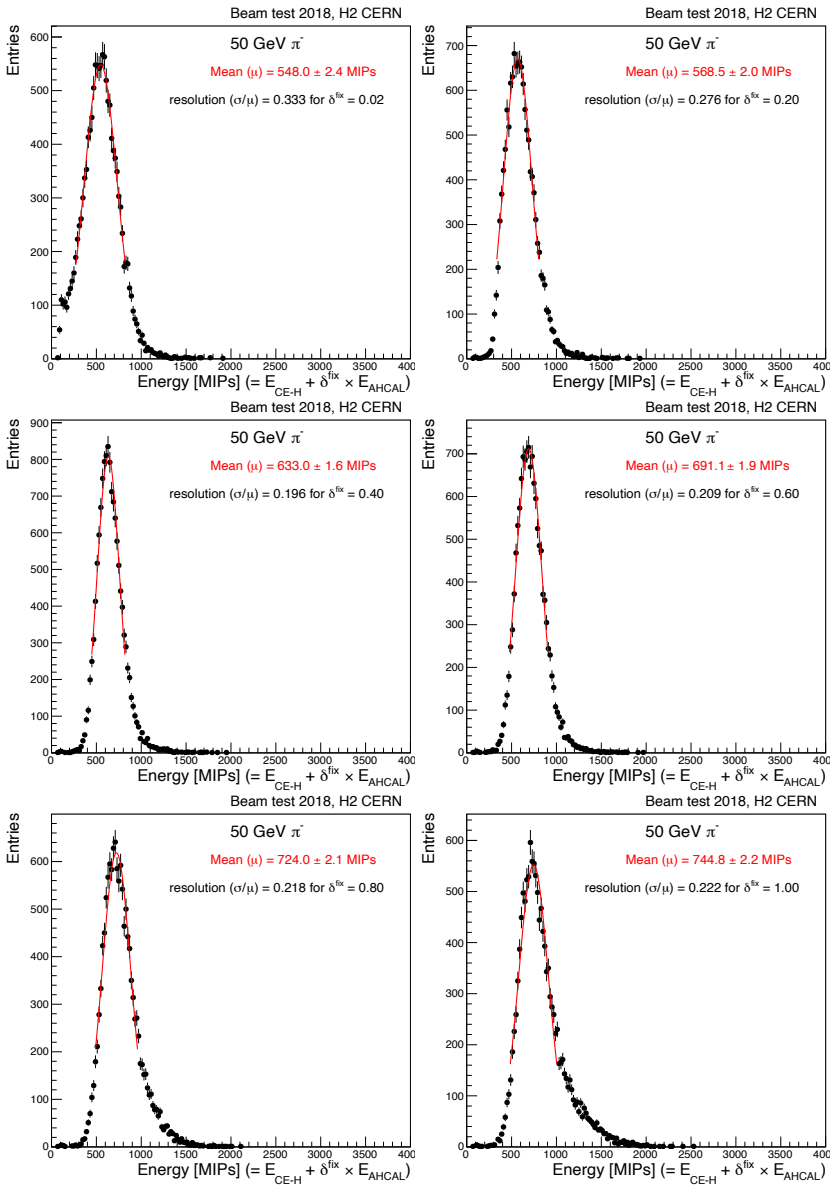

FIGURE 7.24: Figure shows CE-H + AHCAL energy sum distribution with different relative weights (δ^{fix}) and fitted with a Gaussian function to determine relative resolution for 50 GeV π^- beam test data.

Using MIP-to-GeV energy scale for electromagnetic and hadronic section of the calorimeter, the energies are combined in the following manner:

$$E_{\text{measured}}[\text{GeV}] = \alpha^{fix} \cdot E^{\text{CE-E}}_{\text{MIPs}} + \beta^{fix} \cdot (E^{\text{CE-H}}_{\text{MIPs}} + \delta^{fix} \cdot E^{\text{AHCAL}}_{\text{MIPs}}) \tag{7.8}$$

where α^{fix} (= 10.57 MeV/MIP) and β^{fix} (= 78.9 MeV/MIP) are MIP-to-GeV scales for CE-E and CE-H + AHCAL combined system, respectively. And $E^{\text{CE-E}}_{\text{MIPs}}$, $E^{\text{CE-H}}_{\text{MIPs}}$, $E^{\text{AHCAL}}_{\text{MIPs}}$ are the total energy measured in MIPs in different compartments.

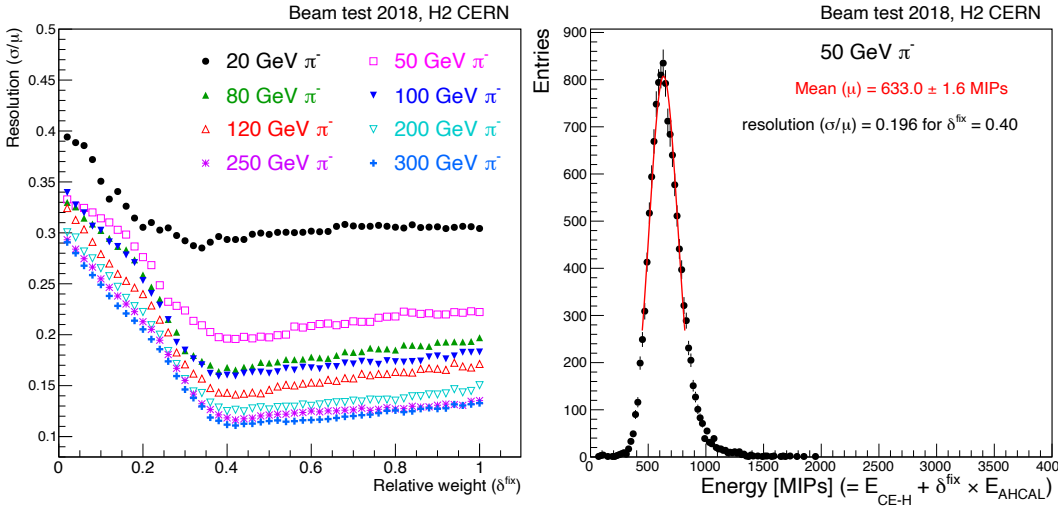

FIGURE 7.25: Figure on left shows CE-H + AHCAL energy resolution as a function of *relative_weight* (δ^{fix}) for all pion beam energies. The minima lies at $\delta^{fix} \simeq 0.4$. The right figure shows CE-H + AHCAL energy sum distribution with $\delta^{fix} = 0.4$, fitted with a Gaussian function for 50 GeV CE-H pions.

To obtain the energy distributions in terms of GeV, energy is combined from CE-E, CE-H and AHCAL prototypes using the prescription described above in equation 7.8. Figure 7.26 shows energy distributions in terms of GeV for CE-E pions (blue solid line) , CE-H pions (red solid line) as well as inclusive of both categories (black dotted line) for all beam energies in data. It is observed that the energy distributions exhibit Gaussian form for both the categories as expected. Energy distributions corresponding to 50 GeV CE-H pions shows the peak at 50 GeV, whereas CE-E pions shows peak lower than 50 GeV. A Gaussian function is fitted iteratively to energy distributions in order to obtain energy response and resolution. In each iteration, the fit range is restricted to $\mu \pm 1.5\ \sigma$ of the previous fit. Typically, three iterations are sufficient to converge the fit. The energy response is defined as the ratio of mean (μ) of the fitted Gaussian function to the beam energy (μ/E_{beam}) and energy resolution is defined as the ratio of σ and μ of the fitted Gaussian function.

Figure 7.27 shows the response and resolution as a function of beam energy obtained using above described method for CE-E pions, CE-H pions and all pions. Corresponding numbers can be found in appendix F. The response plot (left) shows a non-linear trend for all the categories. Such non-linearity in the hadronic response indicates that the detector is non-compensating calorimeter (see section 2.6.2). The energy response is unity for 50 GeV CE-H pions by construction because of the fact that the energy scale for hadronic section , i.e. CE-H + AHCAL, is obtained using 50 GeV CE-H pions. The energy response for CE-E pions are consistently less than one. Since the energy scale for CE-E is obtained using pure electromagnetic shower therefore the response trend points to the under-compensating nature of this calorimeter set up, i.e. e/h ratio is greater than one.

The resolution is fitted with the standard parametrized form (discussed in section 2.6.3) as shown in equation 7.9 below:

$$\left(\frac{\sigma}{E}\right) = \sqrt{\left(\frac{s}{\sqrt{E}}\right)^2 + c^2} \tag{7.9}$$

where, s is the stochastic term and c is the constant term. The noise term is found to be very small therefore not included in the parametrization. Table 7.11 lists the fitted value for

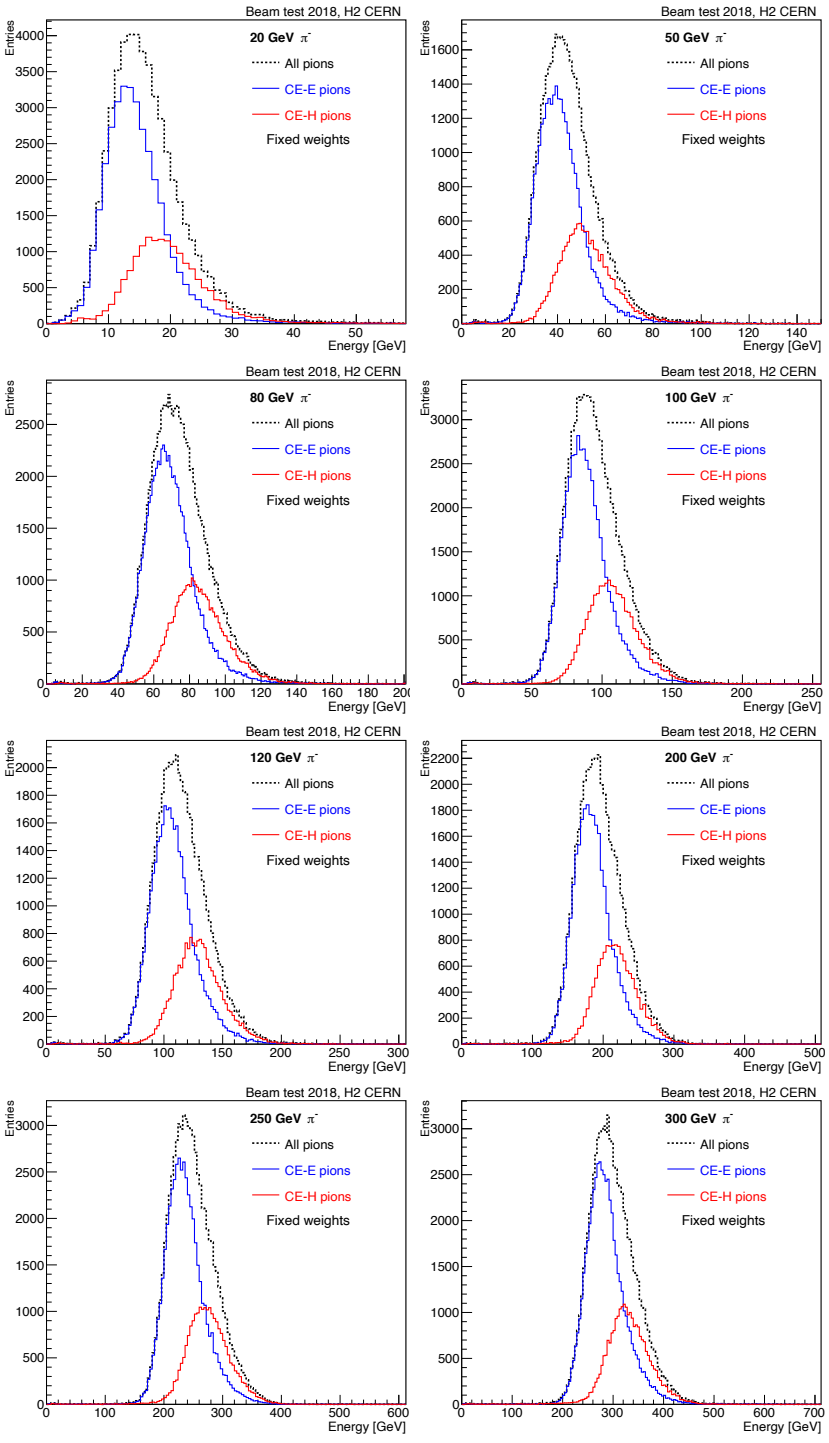

FIGURE 7.26: Reconstructed shower energy distribution in terms of GeV using fixed weights. Solid line in blue color corresponds to CE-E pions, solid line in red corresponds to CE-H pions and dotted black line corresponds to inclusive events.

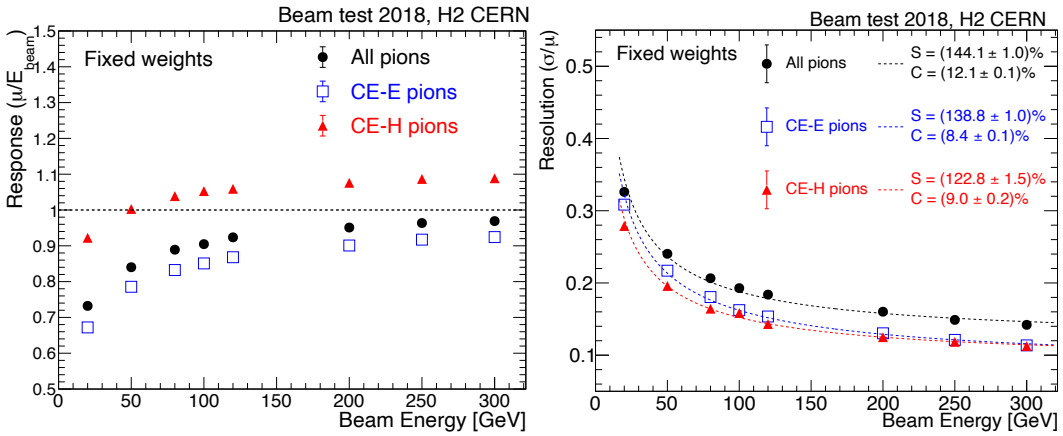

FIGURE 7.27: Energy response (left) and resolution (right) as a function of beam energy for CE-E pions (blue open square), CE-H pions (red full triangles) and inclusive in both categories (black full circles).

all the categories.

Figure 7.27 (right) shows that the energy resolution for CE-H pions is better as compared to CE-E pions for low beam energies. This can be explained by the fact that the pions that start showering in CE-E i.e. CE-E pions, deposit their energies in both electromagnetic as well as in hadronic section which are built with absorbers with different e/h ratios. This has following two implications:

1. Different Z materials have different energy threshold (E_0) for secondary pion (π^0) production, for example E_0 for Fe is 0.7 GeV whereas for Pb, E_0 is 1.3 GeV as documented by Gabriel et al [124], therefore the fraction of electromagnetic component is different for different absorber material.

2. Additionally, different e/h ratio means different response to electromagnetic and hadronic component of the hadronic shower.

Thus, the fluctuation in π^0 production and different response to it in different absorber material results in worsening of energy resolution when combining the energies from electromagnetic and hadronic section. This is reflected in higher stochastic term for CE-E pions as compared to CE-H pions in table 7.11.

It is to be noted that the fluctuation in electromagnetic component due to different e/h ratio does not follow Poisson statistics therefore it does not scale as $1/\sqrt{E}$, instead it follows a power law as explained in [37]. However, as a crude approximation it can be absorbed in the stochastic term and has been used here in the discussion for relative comparison purposes.

The constant term for CE-E pions is found to be slightly better than CE-H pions. This is explained by considering the fact that the shower-containment in longitudinal direction is more for CE-E pions as compared to CE-H pions as CE-E pions see $\sim 1.4\ \lambda_{int}$ (i.e. CE-E) more detector than CE-H pions. Hence, longitudinal shower leakage will be more for late showering CE-H pions, thereby increasing the constant term.

When considering pions of both categories inclusively, it is observed that the response is closer to the CE-E pions which is expected as the fraction of events for CE-E pions is $\sim$ 70% as mentioned earlier. The resolution for all pions is considerably worse than both CE-E

TABLE 7.11: Values of fitted parameters to resolution for events in different categories of pions in beam test data. The errors correspond to statistical uncertainty only.

	CE-E pions	CE-H pions	All pions
Stochastic $[GeV^{-\frac{1}{2}}]$	138.8 ± 1.0%	122.8 ± 1.5%	144.1 ± 1.0 %
Constant	8.4 ± 0.1 %	9.0 ± 0.2 %	12.1 ± 0.1 %

and CE-H pions. This is an artefact of combining energies from two categories that uses different energy scales. For example, for 50 GeV pion in Figure 7.26, the energy distribution for CE-E pions peaks at lower energy while for CE-H pions the distribution peaks at 50 GeV. When combining the two, the overall energy distribution for all pions shows broader width and it is reflected in the energy resolution plot as a function of beam energy in Figure 7.27. One way to get rid of this feature is to artificially set the energy scale for CE-E pions such that the energy response is one at 50 GeV, however, this only shifts the overall scale of CE-E pions and does not fix the non-linearity. A better method to fix the response non-linearity as well as the worse resolution of pions in combined category is by using optimized weights and is discussed in detail later in this chapter.

7.4.1 Comparison with simulation

Using the prescription described in equation 7.8 the energy is reconstructed in terms of GeV in simulation and is compared against beam test data. The weights are not re-derived for simulation, instead fixed weights derived from data are applied on the simulated events to check the agreement.

Figure 7.28 shows energy distribution comparison between data and simulation with two physics lists for different pion beam energies for CE-E pions (left column) and CE-H pions (right column) after applying fixed energy weights. Both the physics lists reproduce the shower energy distribution shape well. However, it is observed that the energy distribution in simulation is shifted to a higher value as compared to data. The mean and width of the distributions is extracted using Gaussian fit around the core of the distribution as described earlier, to obtain response and resolution.

Figure 7.29 shows response and resolution comparison between data and simulation as a function of beam energy for CE-E pions and CE-H pions. Corresponding numbers can be found in appendix F. The response plot shows that both the physics lists reproduce the non-linearity trend very well as observed in data for both the categories of pions. However, the response is consistently higher in simulation as compared to data. This indicates that the simulation is over-predicting shower energy for different beam energies. The over-prediction in the simulated shower energy was foreshadowed by the out-of-the-box energy distributions (Figure 7.6 and 7.7). However, the level of over-prediction of energy is almost constant for all beam energy points, and is about $\sim$ 5% for CE-E pions and $\sim$ 10% for CE-H pions. The reason for such energy scale difference could be due to imperfect description of detector material or differences in the modelling of hadronic showers in simulation. These comparisons with data and results are useful in order to tune the simulation to match as closely to real data as possible.

The resolution comparison shows good agreement between data and simulation at higher energies but shows a slight difference at lower energies for CE-E pions with FTFP_BERT_EMN

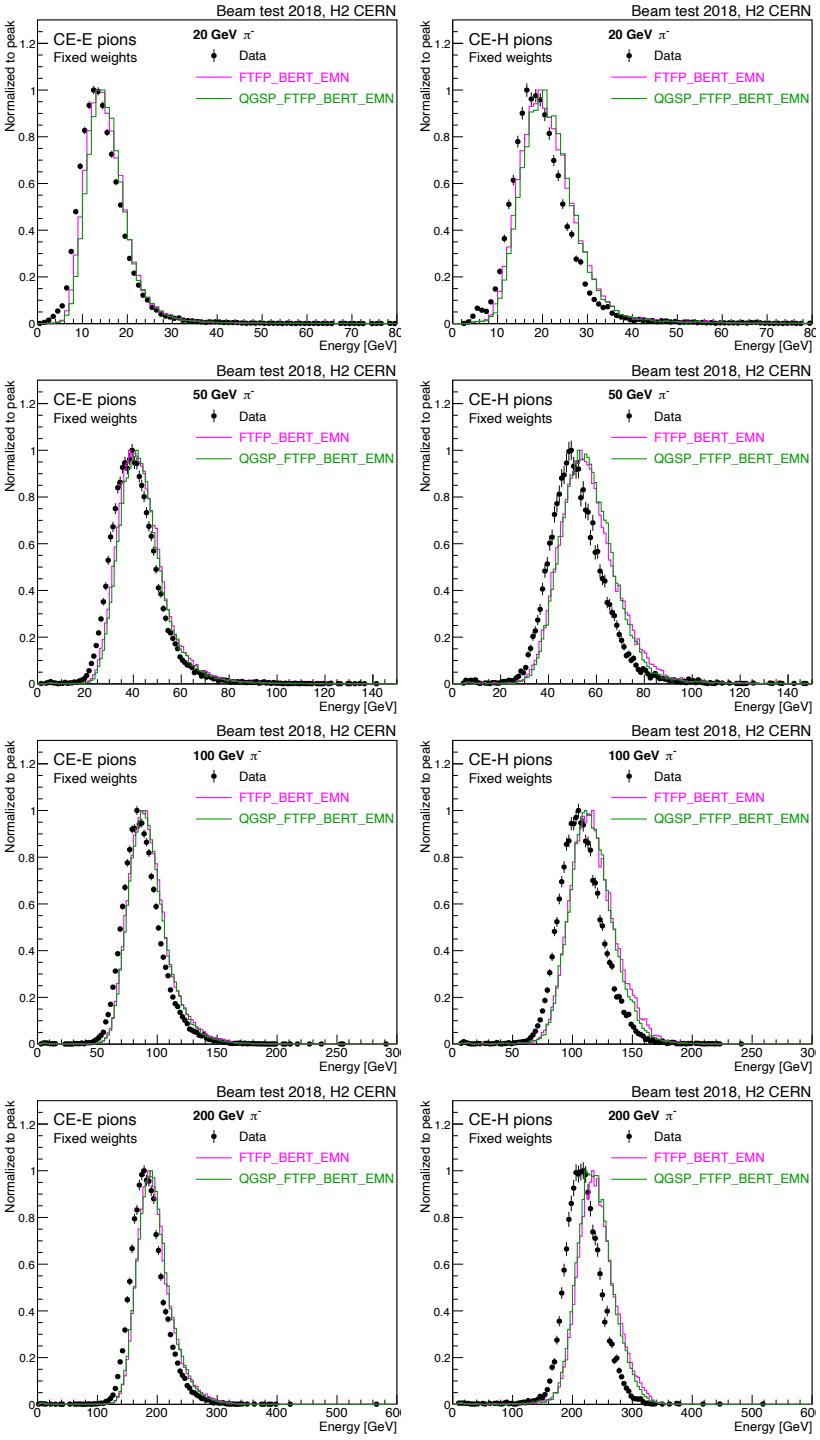

FIGURE 7.28: Energy distribution comparison between data (black dots) and simulation with two physics lists (coloured lines), after applying MIP-to-GeV energy scale for different beam energies CE-E pions (left column) and CE-H pions (right column).

being closer to data as compared to QGSP_FTFP_BERT_EMN. The overall agreement is within $\sim 10\%$ between data and simulation.

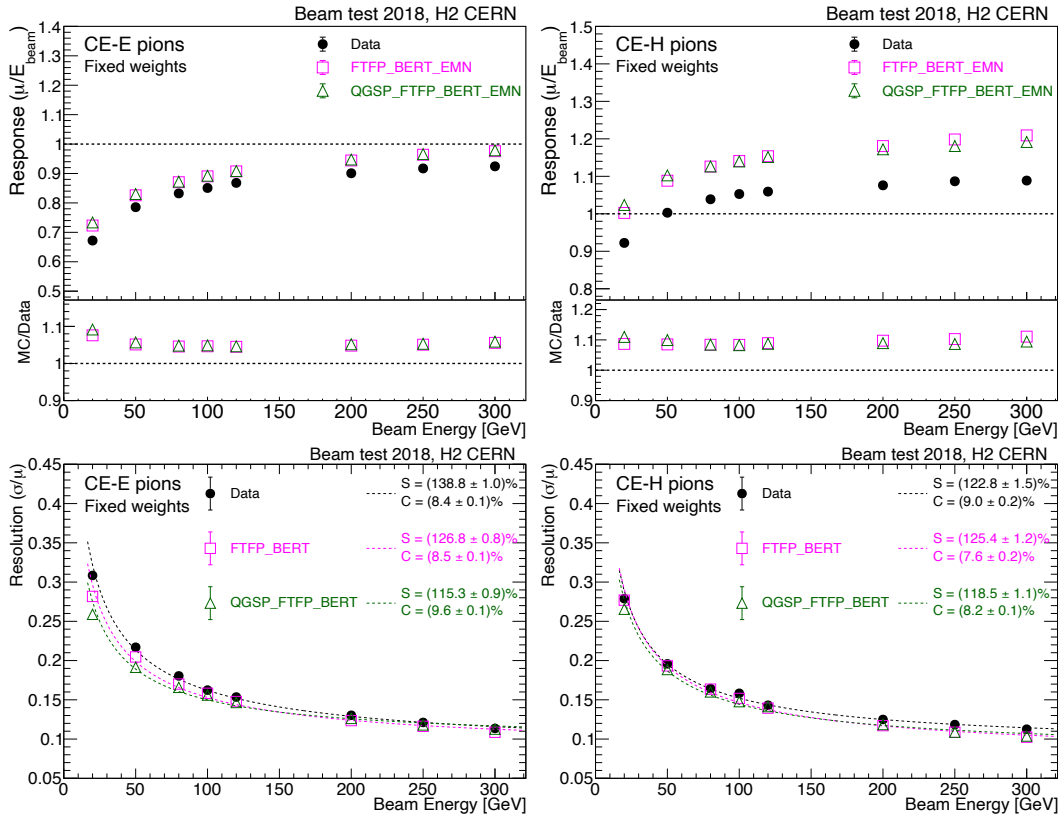

FIGURE 7.29: Energy response (top row) and resolution (top row) comparison between data and simulation (with two physics lists) as a function of beam energy for CE-E pions (left column) and CE-H pions (right column). The bottom panel of response plots shows ratio between simulation and data, and is systematically higher for simulation. The resolution comparison shows good agreement for for higher energy for CE-E pions (bottom left) but shows slight disagreement at lower energy. For CE-H pions (bottom right), the agreement is good across all the energies.

7.4.2 Fixing energy scale for simulation

The constant energy scale level difference between data and simulation allows to derive a global scale factor in order to bring down the energy response in simulation to match with data without affecting the energy resolution. This global scale factor is obtained for electromagnetic and hadronic section separately.

The independent study of the performance of CE-E prototype to electromagnetic shower using e^+ data collected during October 2018 beam test experiment, shows that the simulation over-predicts electromagnetic shower energy in CE-E by 3.5% as compared to data [118]. This energy scale difference is observed to be flat as a function of e^+ beam energy. Therefore for electromagnetic section, the simulation energy is scaled down by 3.5%.

The scaling factor for combined hadronic section is derived by using CE-H pion events and comparing combined energy in CE-H+AHCAL in terms of MIPs between data and simulation. The combined energy is obtained according to equation 7.7 (with $\delta^{fix} = 0.4$), for both data and simulation. Since response comparisons indicate that both physics lists

show almost equal energy response (Figure 7.29), therefore only FTFP_BERT_EMN physics list is used for obtaining the scaling factor. Figure 7.30 shows combined energy distribution comparison in hadronic section between data and simulation for 50 GeV CE-H pions. Both the distributions are fitted with a Gaussian function around the peak of the distribution in a range of ± 1.5 standard deviation. The mean (μ) of both the fitted Gaussian is extracted and ratio $\mu_{\text{sim}}/\mu_{\text{data}}$ is found to be 1.095. Therefore, the simulation energy needs a correction of 9.5% to match with data.

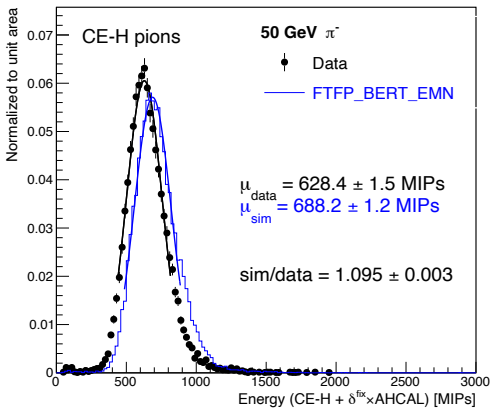

FIGURE 7.30: Combined energy distribution of CE-H and AHCAL for 50 GeV CE-H pions. Black point corresponds to data and blue line corresponds to FTFP_BERT_EMN physics list. A Gaussian is fitted around the core of the energy distributions. The simulation peak is $\simeq$9.5% too high as compared to data.

Thus, energy scale for simulation is fixed by scaling down the energy in all the detector prototypes with following factors:

- **3.5% in CE-E**
- **9.5% in CE-H**
- **9.5% in AHCAL**

The effect of the global energy scale application on simulation energy is shown in energy distribution comparisons between data and simulation in Figure 7.31. The distributions show good agreement with each other. The response and resolution comparison after scaling down simulation energy is shown in Figure 7.32. The energy response is observed to be in agreement between data and simulation with negligible change in the resolution. Corresponding numbers can be found in appendix F.

All results shown from here onwards, the global energy correction factor on simulation has been applied to match the energy scale with data.

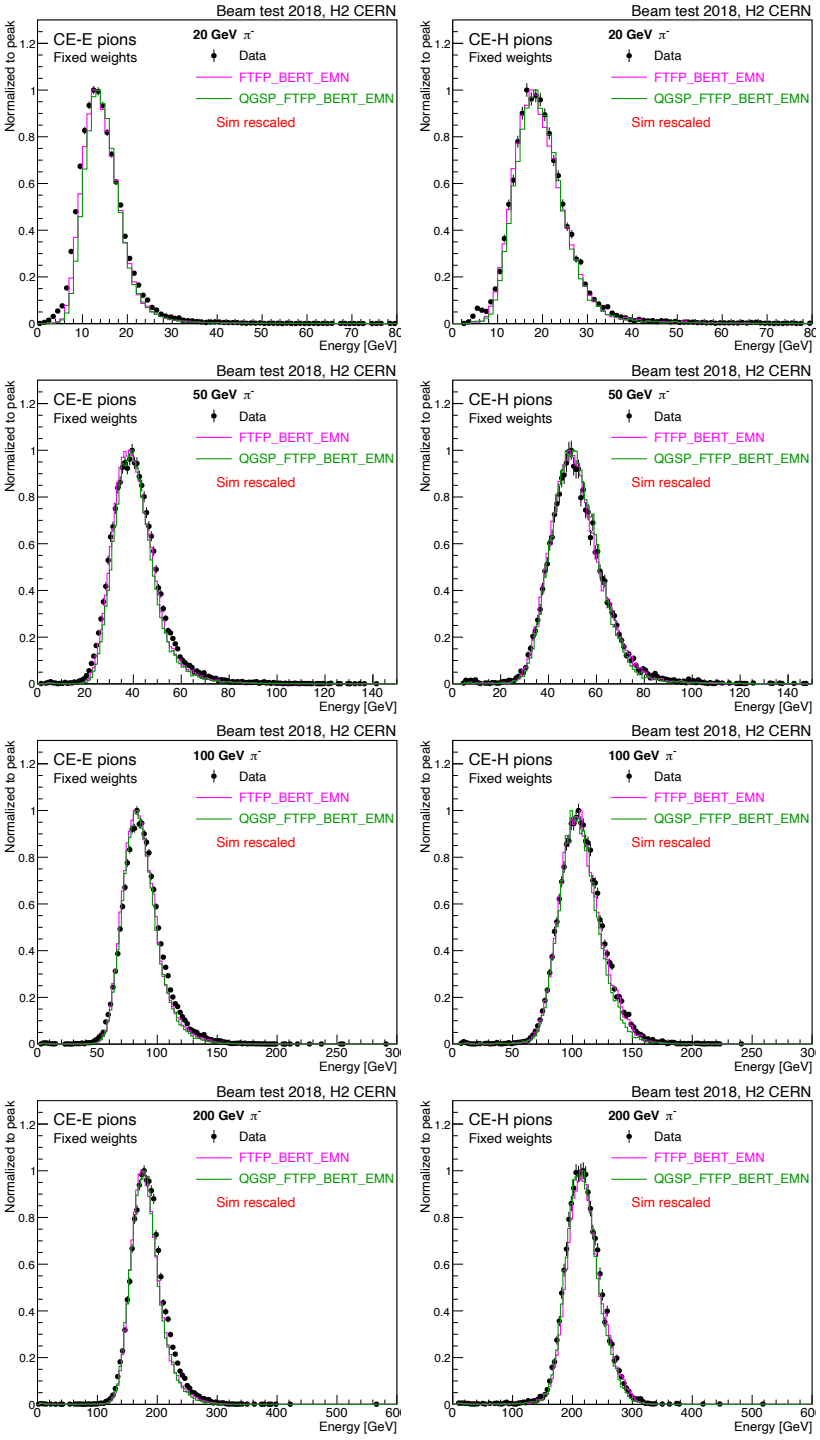

FIGURE 7.31: Energy distribution comparison between data (black dots) and simulation with two physics lists (coloured lines) after applying scale correction on simulation energy. Energy distribution for both physics lists show very good agreement with data.

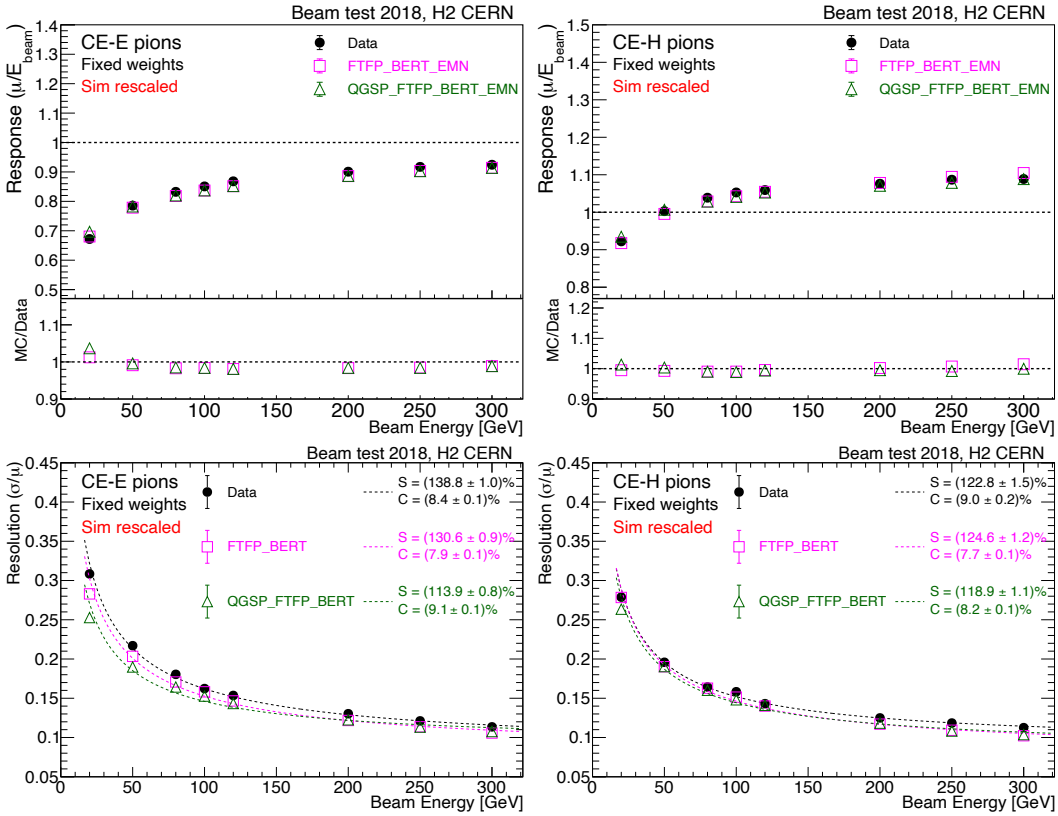

FIGURE 7.32: Energy response (top row) and resolution (bottom row) comparison between data and simulation (with two physics lists) as a function of beam energy for CE-E pions (left column) and CE-H pions (right column) after applying global energy correction factor on simulation. The response matches well data and there is almost no change in resolution.

7.5 Optimization of pion energy reconstruction using energy dependent weights

The non-linearity in the energy response of hadrons is a characteristic of non-compensating calorimeter. There are various ways in which a calorimeter can be made a compensating calorimeter. One way to compensate is at the hardware level where the absorber and active material are chosen such that e/h ratio is one, e.g. HELIOS calorimeter [125]. Another way is to compensate at the software (or offline) level where the response is corrected using software techniques. One of the simplest yet effective method for offline compensation is to use energy dependent weights using χ^2-minimization based technique which is described in the following section.

7.5.1 Energy dependent weights using χ^2-minimization

The response of the hadrons can be linearized by using energy dependent weights. In this method, weights are not fixed but depend upon the energy (or momentum) of the incident hadron such that for low energy pions where hadronic response is low, a higher weight is applied and vice versa. Thus, non-linearity is taken away by the means of non-linear weights. This strategy is inspired by the Particle Flow Cluster calibration method deployed in the CMS reconstruction framework [67] and is discussed in section 3.3.

In order to obtain energy dependent weights, χ^2-minimization based technique is used which is described as follows. The energy measured in different compartments is combined in a linear combination with energy dependent weights in order to obtain the corrected energies, for the two event categories separately. The expression for corrected energies is defined in equation 7.10 and 7.11 below for CE-E pions and CE-H pions, respectively:

$$E_{\text{corr}}[\text{GeV}] = \alpha_1(E_{\text{beam}}) \cdot E_{\text{fix}}^{\text{CE-E}} + \beta_1(E_{\text{beam}}) \cdot E_{\text{fix}}^{\text{CE-H}} + \gamma_1(E_{\text{beam}}) \cdot E_{\text{fix}}^{\text{AHCAL}} \quad \text{...for CE-E pions} \tag{7.10}$$

$$E_{\text{corr}}[\text{GeV}] = E_{\text{fix}}^{\text{CE-E}} + \beta_2(E_{\text{beam}}) \cdot E_{\text{fix}}^{\text{CE-H}} + \gamma_2(E_{\text{beam}}) \cdot E_{\text{fix}}^{\text{AHCAL}} \quad \text{... for CE-H pions} \tag{7.11}$$

where α_1, $\beta_{1,2}$ and $\gamma_{1,2}$ are energy dependent weights, and $E_{\text{fix}}^{\text{CE-E}}$, $E_{\text{fix}}^{\text{CE-H}}$ and $E_{\text{fix}}^{\text{AHCAL}}$ are energies in GeV in different compartments calculated using MIP-to-GeV fixed energy scales as described in previous section 7.4. The first term in equation 7.11 accounts for the energy deposited by the MIP track of CE-H pions. The weights α_1, $\beta_{1,2}$ and $\gamma_{1,2}$ are determined by constructing χ^2 as shown in equation 7.12 and minimizing it with respect to α_1, $\beta_{1,2}$ and $\gamma_{1,2}$.

$$\chi^2 = \sum_i \frac{(E_{\text{beam}} - E_{\text{corr}}^i)^2}{\sigma^2(E_{\text{fix}}^i)} \tag{7.12}$$

The χ^2 is constructed separately for CE-E and CE-H pion events. The summation ($\sum_i$) runs over all the events for a given beam energy. The $\sigma(E_{\text{fix}}^i)$ in the denominator is a preliminary estimate of the uncertainty in the measured energy (E_{fix}^i) using fixed weights. It helps to reduce the contribution from the tails on lower end of energy distributions. The $\sigma(E_{\text{fix}}^i)$ is the parametrized resolution corresponding to Figure 7.27 (right) and is given below:

$$\frac{\sigma(E)}{E} = \frac{138.8\%}{\sqrt{E}} + 8.4\% \text{ .... for CE-E pions} \tag{7.13}$$

$$\frac{\sigma(E)}{E} = \frac{122.8\%}{\sqrt{E}} + 9.0\% \text{ .... for CE-H pions} \tag{7.14}$$

The χ^2 is minimized analytically by requiring the first derivative of equations 7.12 with respect to α_1, $\beta_{1,2}$ and $\gamma_{1,2}$ to be zero simultaneously, for CE-E and CE-H pions separately. The events from the **test beam data** are used for the minimization that fall within ± 2 standard deviation from the peak of energy distribution (distribution corresponding to data as shown in Figure 7.31). It further makes sure that the events from the core of the distribution are picked up and are not affected by the tail effects, such as shower leakage etc., during the optimization. To solve the simultaneous equations, matrix formulation is used which is fairly straight-forward to implement in analysis using C++ code. The matrices obtained by solving the equations are shown below:

$$\begin{pmatrix} \alpha_1 \\ \beta_1 \\ \gamma_1 \end{pmatrix} = \begin{pmatrix} \sum \frac{E_{\text{fix}}^{\text{CE-E}} \cdot E_{\text{fix}}^{\text{CE-E}}}{\sigma^2(E_{\text{fix}})} & \sum \frac{E_{\text{fix}}^{\text{CE-E}} \cdot E_{\text{fix}}^{\text{CE-H}}}{\sigma^2(E_{\text{fix}})} & \sum \frac{E_{\text{fix}}^{\text{CE-E}} \cdot E_{\text{fix}}^{\text{AHCAL}}}{\sigma^2(E_{\text{fix}})} \\ \sum \frac{E_{\text{fix}}^{\text{CE-E}} \cdot E_{\text{fix}}^{\text{CE-H}}}{\sigma^2(E_{\text{fix}})} & \sum \frac{E_{\text{fix}}^{\text{CE-H}} \cdot E_{\text{fix}}^{\text{CE-H}}}{\sigma^2(E_{\text{fix}})} & \sum \frac{E_{\text{fix}}^{\text{AHCAL}} \cdot E_{\text{fix}}^{\text{CE-H}}}{\sigma^2(E_{\text{fix}})} \\ \sum \frac{E_{\text{fix}}^{\text{CE-E}} \cdot E_{\text{fix}}^{\text{AHCAL}}}{\sigma^2(E_{\text{fix}})} & \sum \frac{E_{\text{fix}}^{\text{CE-H}} \cdot E_{\text{fix}}^{\text{AHCAL}}}{\sigma^2(E_{\text{fix}})} & \sum \frac{E_{\text{fix}}^{\text{AHCAL}} \cdot E_{\text{fix}}^{\text{AHCAL}}}{\sigma^2(E_{\text{fix}})} \end{pmatrix}^{-1} \times \begin{pmatrix} \sum \frac{E_{\text{beam}} \cdot E_{\text{fix}}^{\text{CE-E}}}{\sigma^2(E_{\text{fix}})} \\ \sum \frac{E_{\text{beam}} \cdot E_{\text{fix}}^{\text{CE-H}}}{\sigma^2(E_{\text{fix}})} \\ \sum \frac{E_{\text{beam}} \cdot E_{\text{fix}}^{\text{AHCAL}}}{\sigma^2(E_{\text{fix}})} \end{pmatrix} \tag{7.15}$$

$$\begin{pmatrix} \beta_2 \\ \gamma_2 \end{pmatrix} = \begin{pmatrix} \sum \frac{E_{\text{fix}}^{\text{CE-H}} \cdot E_{\text{fix}}^{\text{CE-H}}}{\sigma^2(E_{\text{fix}})} & \sum \frac{E_{\text{fix}}^{\text{AHCAL}} \cdot E_{\text{fix}}^{\text{CE-H}}}{\sigma^2(E_{\text{fix}})} \\ \sum \frac{E_{\text{fix}}^{\text{CE-H}} \cdot E_{\text{fix}}^{\text{AHCAL}}}{\sigma^2(E_{\text{fix}})} & \sum \frac{E_{\text{fix}}^{\text{AHCAL}} \cdot E_{\text{fix}}^{\text{AHCAL}}}{\sigma^2(E_{\text{fix}})} \end{pmatrix}^{-1} \times \begin{pmatrix} \sum \frac{(E_{\text{beam}} - E_{\text{fix}}^{\text{CE-E}}) \cdot E_{\text{fix}}^{\text{CE-H}}}{\sigma^2(E_{\text{fix}})} \\ \sum \frac{(E_{\text{beam}} - E_{\text{fix}}^{\text{CE-E}}) \cdot E_{\text{fix}}^{\text{AHCAL}}}{\sigma^2(E_{\text{fix}})} \end{pmatrix} \tag{7.16}$$

After minimizing χ^2, the weights α_1, $\beta_{1,2}$ and $\gamma_{1,2}$ are obtained for all beam energies as shown in Figure 7.33 for CE-E and CE-H pions. The weights are parametrized by fitting a polynomial function $(p_0 + p_1/\sqrt{E_{\text{beam}}})$ to the weights.

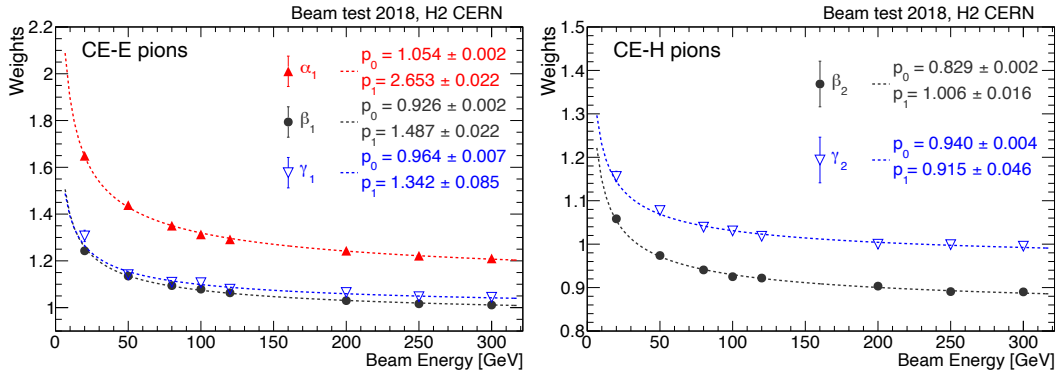

FIGURE 7.33: Weights α_1, $\beta_{1,2}$ and $\gamma_{1,2}$ as a function of beam energy for CE-E pions (left) and CE-H pions (right) with a fitted parametrized function.

The weights are plugged back in equations 7.10 and 7.11 corresponding to respective beam energy to obtain E_{corr} for CE-E and CE-H pions, respectively. Figure 7.34 shows energy distributions for different categories of pions for all the beam energies in beam test data. The energy distribution peaks at the beam energies for all the categories i.e. CE-E pions (blue solid line), CE-H pions (red solid line) as well as for all pions (black dotted line).

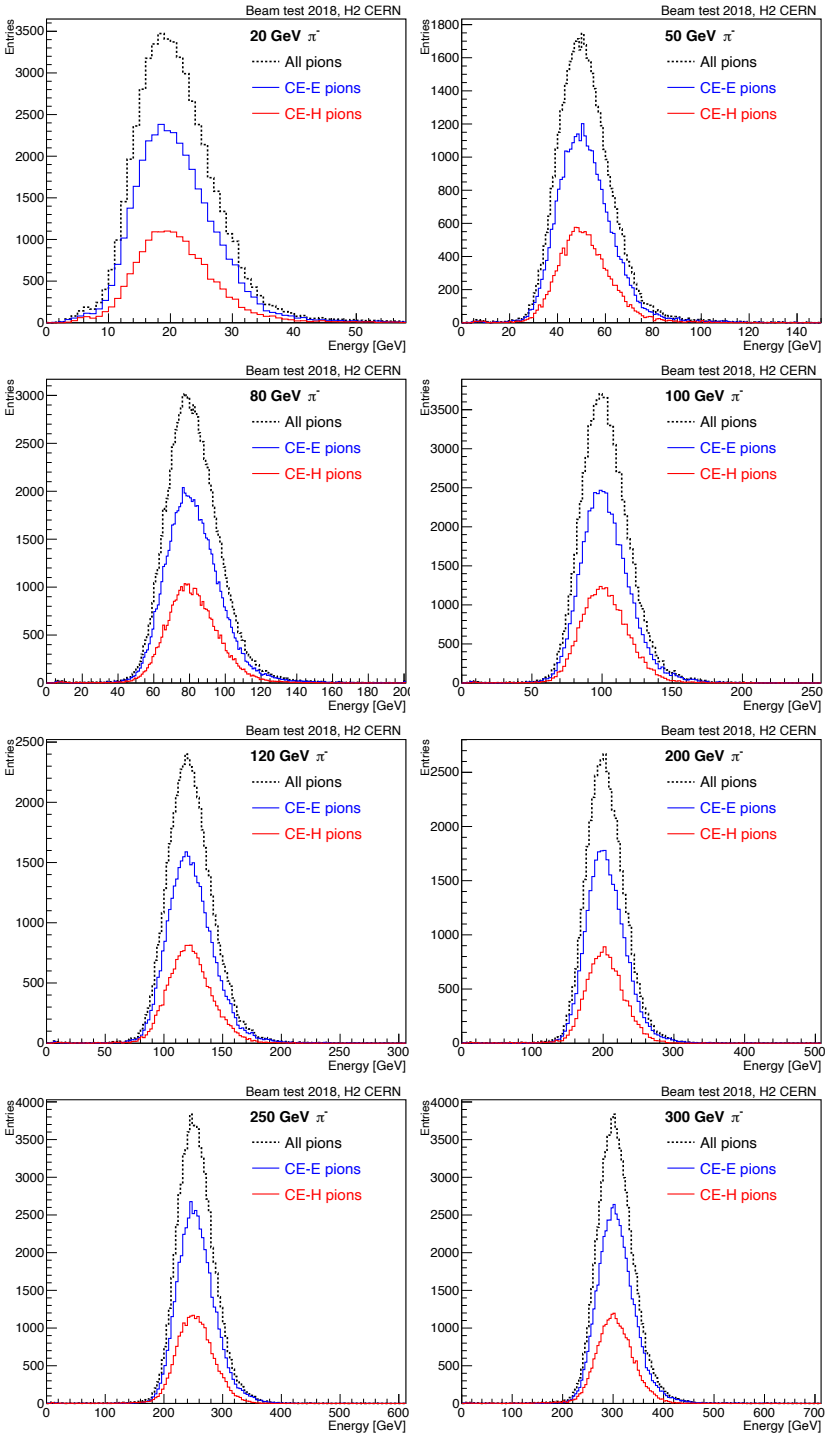

FIGURE 7.34: Reconstructed shower energy distribution using energy dependent weights for different categories of pions in data. Solid line in blue color corresponds to CE-E pions, solid line in red corresponds to CE-H pions and dotted black line corresponds to inclusive events. The peak for energy distributions for all the energies lie at beam energy after applying energy dependent weights.

The effect of energy dependent weights on the response and resolution is shown in Figure 7.35 where a linear response is obtained within a couple of percent. Corresponding numbers can be found in appendix F. A slight improvement in the resolution is also observed for low energy CE-E pions as compared to fixed weights, whereas resolution for CE-H pions remains almost unchanged. Also, different categories of pions show similar response and resolution as a function of beam energy as shown in Figure 7.36 after applying energy dependent weights. It, however, is not the best possible resolution achievable with this set up as we are not making use of detailed information available from CE-E and CE-H on an event-by-event basis (i.e. shower-by-shower details). Extending the work of this thesis, the IISER Pune group (Alpana Sirohi *et al*) have shown that a Graphical-Neural-Network (GNN) based machine learning algorithm shows better performance in terms of energy resolution by using event-by-event information.

It is to be noted that since the weights have been derived using events that lie in the bulk of the distribution (± 2 std. dev. around the mean of data distribution in Figure 7.31) and it is applied on all the events while performing energy reconstruction, therefore, the response does not lie exactly at one that is obtained using Gaussian fits.

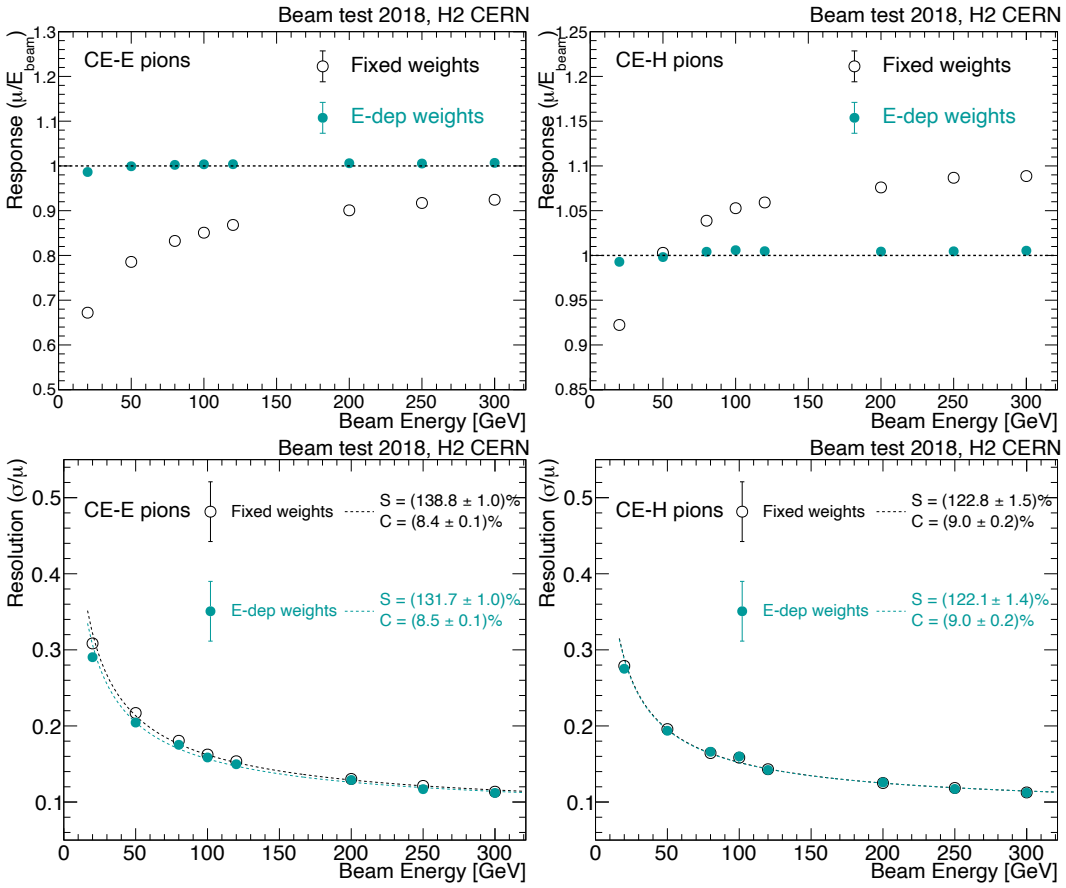

FIGURE 7.35: Response and resolution comparison after applying fixed weights (black open circles) and energy dependent weights (cyan full circles) for CE-E pions (left column) and CE-H pions (right column) as a function of beam energy in data. The response lies at one after with energy dependent weights and a slight improvement is observed in resolution for low energy CE-E pions.

These energy dependent weights, as derived using data (Figure 7.33), are also applied

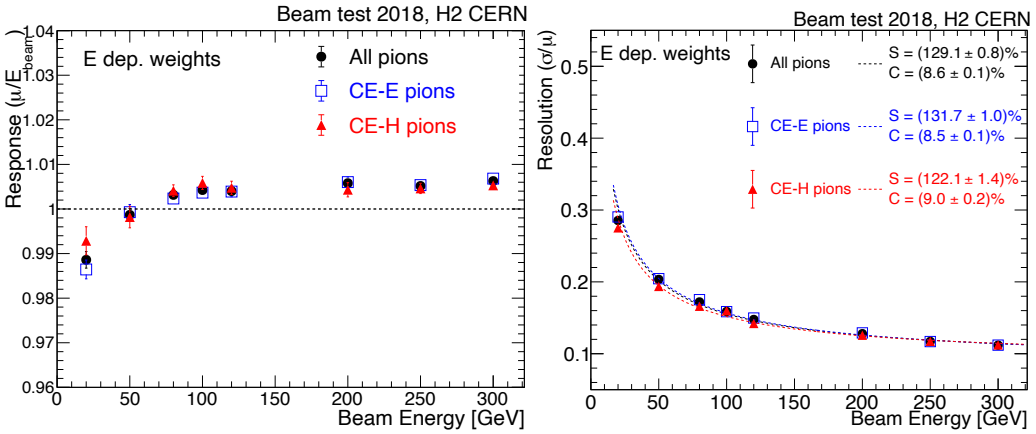

FIGURE 7.36: Response (left) and resolution (right) as a function of beam energies for different categories of pions after applying energy dependent weights.

in simulation to obtain corrected energy distributions. Figure 7.37 shows data-simulation comparison of pion energy distribution for different beam energy of CE-E and CE-H pions. Simulation with both physics lists shows good agreement with data. Figure 7.38 shows response and resolution comparison between data and simulation and it is observed that the linearity for data as well as simulation for both physics lists is within 5%. The resolution comparison shows agreement within 10% between data and simulation for both the categories of pions. The resolution is fitted with the standard parametrized form (equation 7.9). The *stochastic* and *constant* term are obtained for data as $\sim$ 132% (122%) and 8.5% (9.0%) for CE-E (CE-H) pions, respectively. Table 7.12 shows the *stochastic* and *constant* terms for data as well as simulation with both physics lists for CE-E and CE-H pions. The uncertainty quoted in the table 7.37 is statistical only. These numbers also suffer from systematic uncertainties which have not been quoted in these results. The sources of systematic uncertainties include (but not limited to): uncertainty in beam energy momenta, data-simulation beam profile difference, low-level calibrations, shower start finder algorithm efficiency, MIP-to-GeV calibration weights. Of these, systematic uncertainty from low-level calibrations (i.e., MIP-calibration and inter-gain calibration) are expected to influence data-simulation resolution comparison because, in simulation, channel-to-channel variation has not been considered. However, for 85% of HGCAL cells, this variation has been equalized (see section 6.8). For the remaining 15% where an average value (over a chip) is used, the calibration factors are expected to be fairly close to the average value as the calibration trend per chip has been found to be fairly stable, e.g. see Figure 6.37. For inter-gain calibration, such level of precision could not be achieved in CE-H as discussed in section 6.5.1, however, even 300 GeV pion showers have been found to not cross-over to LG or ToT as often as pure electron shower. Therefore, the possible contribution from inter-gain calibration uncertainty is expected to be small.

7.5.2 Energy reconstruction in the absence of tracking or true reference

In the real experiment, the information of true energy of the incident hadron is not available. In such cases the track momentum (which has a momentum resolution of $\sim$ 1% for p_{track} = 100 GeV/c, section 3.2.1) of the hadron is used as the reference, e.g. in PF hadron cluster calibration in the CMS reconstruction framework as mentioned in appendix B. However, it is not always possible to have track momentum measurement, such as for neutral hadrons or charged hadrons which are outside tracker coverage region. In such cases, the measured

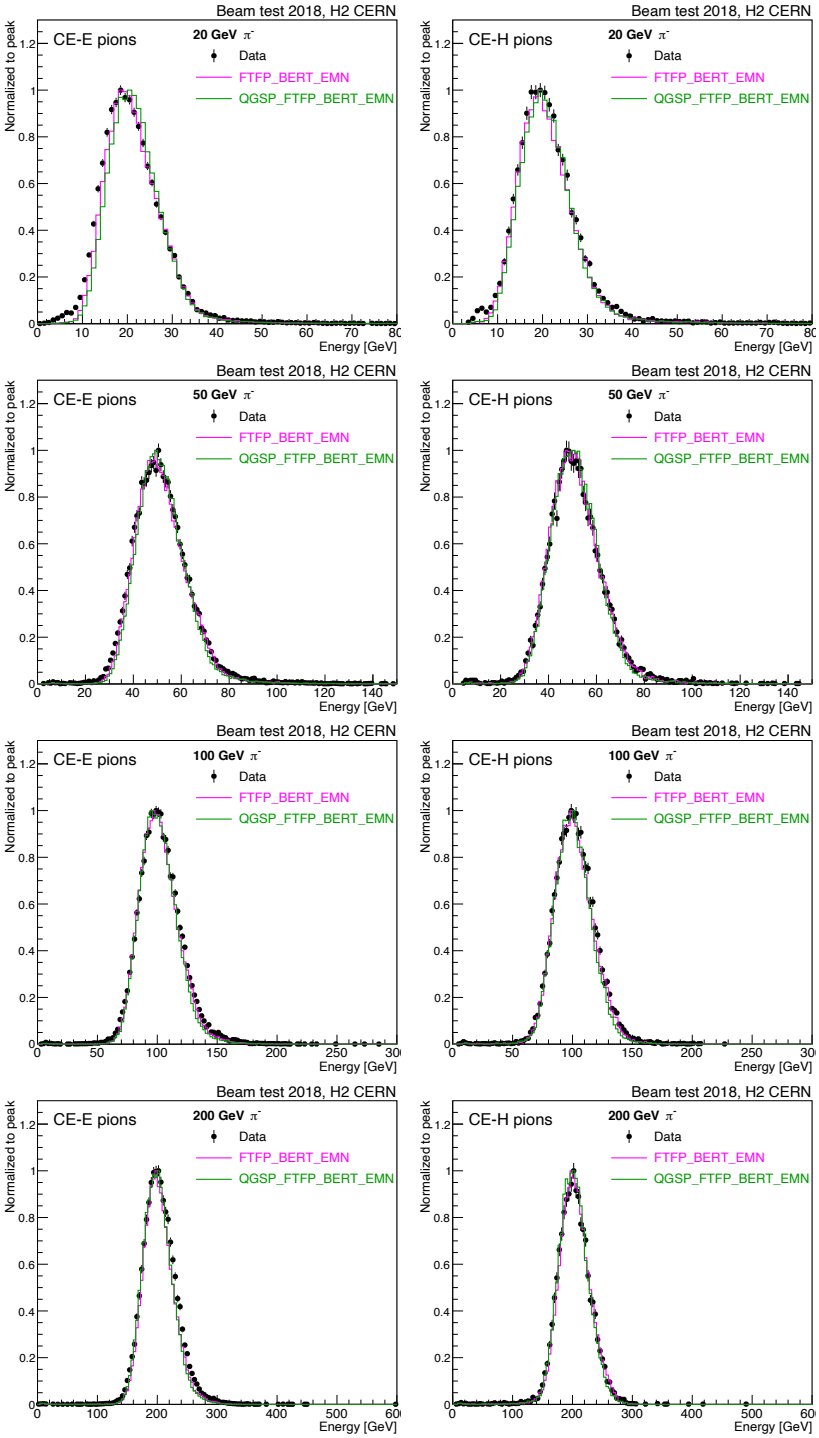

FIGURE 7.37: Energy distribution comparison between data (black points) and simulation (coloured lines) for CE-E pions (left column) and CE-H pions (right column) for different beam energy, after the application of energy dependent weights. The energy distributions of simulation shows very good agreement with data.

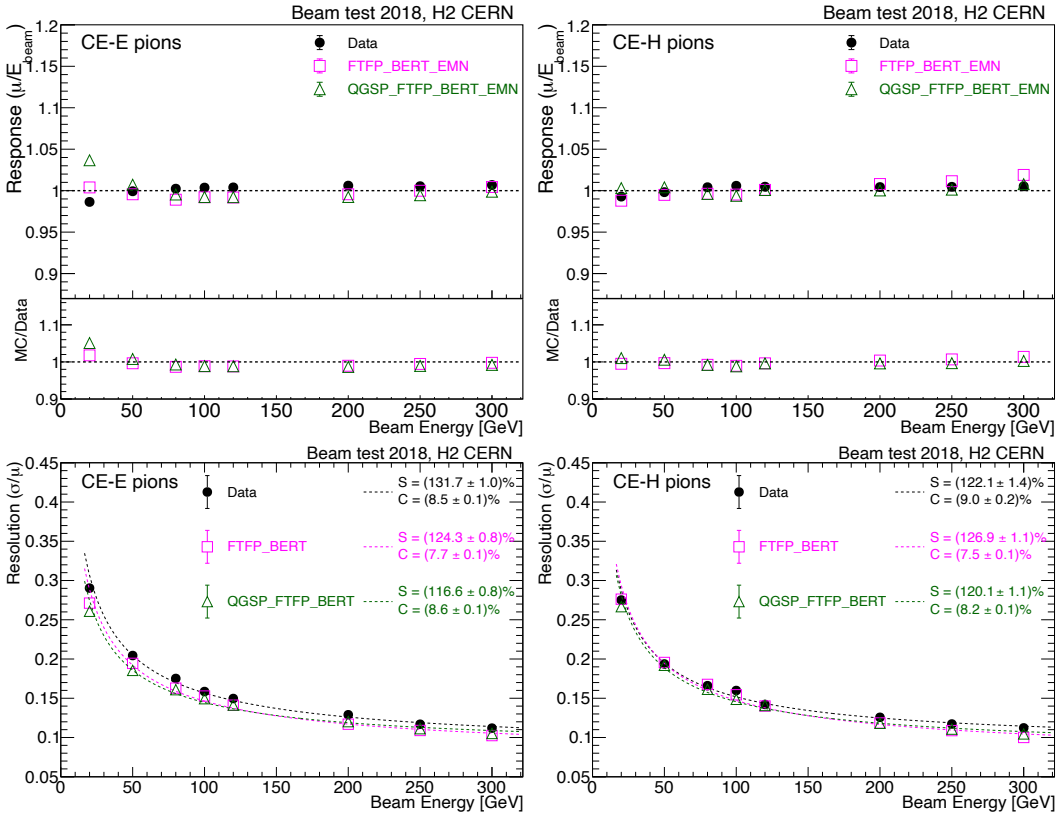

FIGURE 7.38: Energy response (top row) and resolution (bottom row) comparison between data and simulation (with two physics lists) as a function of beam energy for CE-E pions (left column) and CE-H pions (right column) after the application of energy dependent weights. Corresponding numbers can be found in appendix F.

TABLE 7.12: Table for stochastic and constant terms for data and simulation with two physics lists for CE-E and CE-H pions. The errors correspond to statistical uncertainty only.

		Data	FTFP_BERT_EMN	QGSP_FTFP_BERT_EMN
Stochastic $[\sqrt{GeV}]$	CE-E pions	$(131.7 \pm 1.0)\%$	$(124.3 \pm 0.8)\%$	$(116.6 \pm 0.8)\%$
	CE-H pions	$(122.1 \pm 1.4)\%$	$(126.9 \pm 1.1)\%$	$(120.1 \pm 1.1)\%$
Constant	CE-E pions	$(8.5 \pm 0.1)\%$	$(7.7 \pm 0.1)\%$	$(8.6 \pm 0.1)\%$
	CE-H pions	$(9.0 \pm 0.2)\%$	$(7.5 \pm 0.1)\%$	$(8.2 \pm 0.1)\%$

calorimetric energy (section 7.4) , i.e. using energy scale of the calorimeter, is used as the reference to extract the correction factors.

To test the validity and performance of this method for hadronic showers in the October 2018 beam test experiment, the parametrization of optimized weights (α_1, $\beta_{1,2}$ and $\gamma_{1,2}$ aas shown in Figure 7.33) is used. The weights are evaluated by taking reconstructed energy using fixed energy scale (as discussed in section 7.4) as the reference, referred to as E_{reco}. The evaluated weights are plugged back in the equation 7.10 and 7.11 accordingly, in order to reconstruct the energy. The reconstructed energy distributions using E_{reco} as reference is compared against energy distribution using E_{beam} as reference (i.e. the default method). Figure 7.39 shows energy distribution comparison between the two weight extraction methods for CE-E and CE-H pions of different beam energies. The comparison plot shows that for low energy pions (especially CE-E pions), the distribution is slightly narrower and the overall distribution is shifted towards higher side when E_{reco} is used as the reference to evaluate the weights. For higher energy pions, both distributions are almost identical. The picture gets clearer when one looks at the response and resolution obtained from these energy distribution.

Figure 7.40 shows response and resolution comparison between the two references for CE-E pions, CE-H pions, and combined categories. The resolution for the combined categories is observed to get better when E_{reco} is used as the reference. However, it is also observed that the response for low energy CE-E pions start to diverge from 1, and the over-prediction in the energy increases as the pion energy decreases, as observed in energy distributions (Figure 7.39) .

Both of these features can be understood and explained by considering the non-linear trend in the energy response (Figure 7.27) when fixed energy scale is used (i.e. E_{reco}), and the non-linearity in the energy dependent weights (Figure 7.33). In the case when E_{reco} is used as the reference, the weights become dynamic i.e. it depends on measured energy on event-by-event basis. When the measured energy is small, then a larger weight is picked up, and vice-versa for when measured energy is larger. Imagine going left or right on the x-axis of Figure 7.33 and picking up the corresponding weights. Therefore, the resulting energy distribution is comparatively narrower as compared to the default case where a single weight is applied for a given beam energy. This effect is evident in the energy distribution comparison shown in Figure 7.39.

However, its adverse effect is seen in the response at lower energies for CE-E pions. At lower energies, the mean reconstructed energy for CE-E pions is lower than the beam energy, e.g. the energy response is $\sim$ 30% smaller for 20 GeV CE-E pions (Figure 7.27).

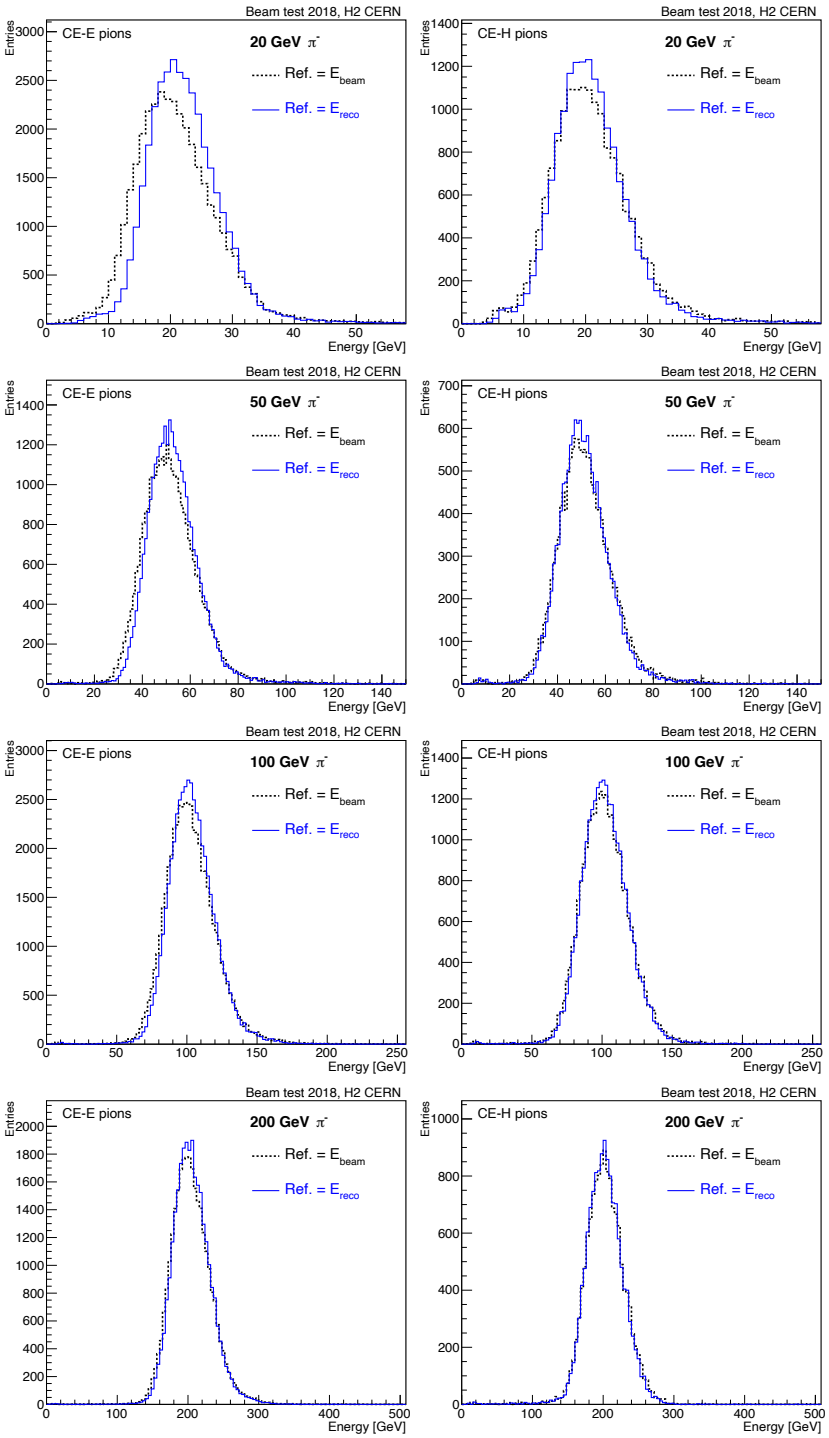

FIGURE 7.39: Energy distribution comparison between two references used to extract the weights (black dashed line and blue solid line corresponds to beam energy and reconstructed calorimeter energy taken as reference, respectively) of energy reconstruction for CE-E pions (left column) and CE-H pions (right) for different beam energies.

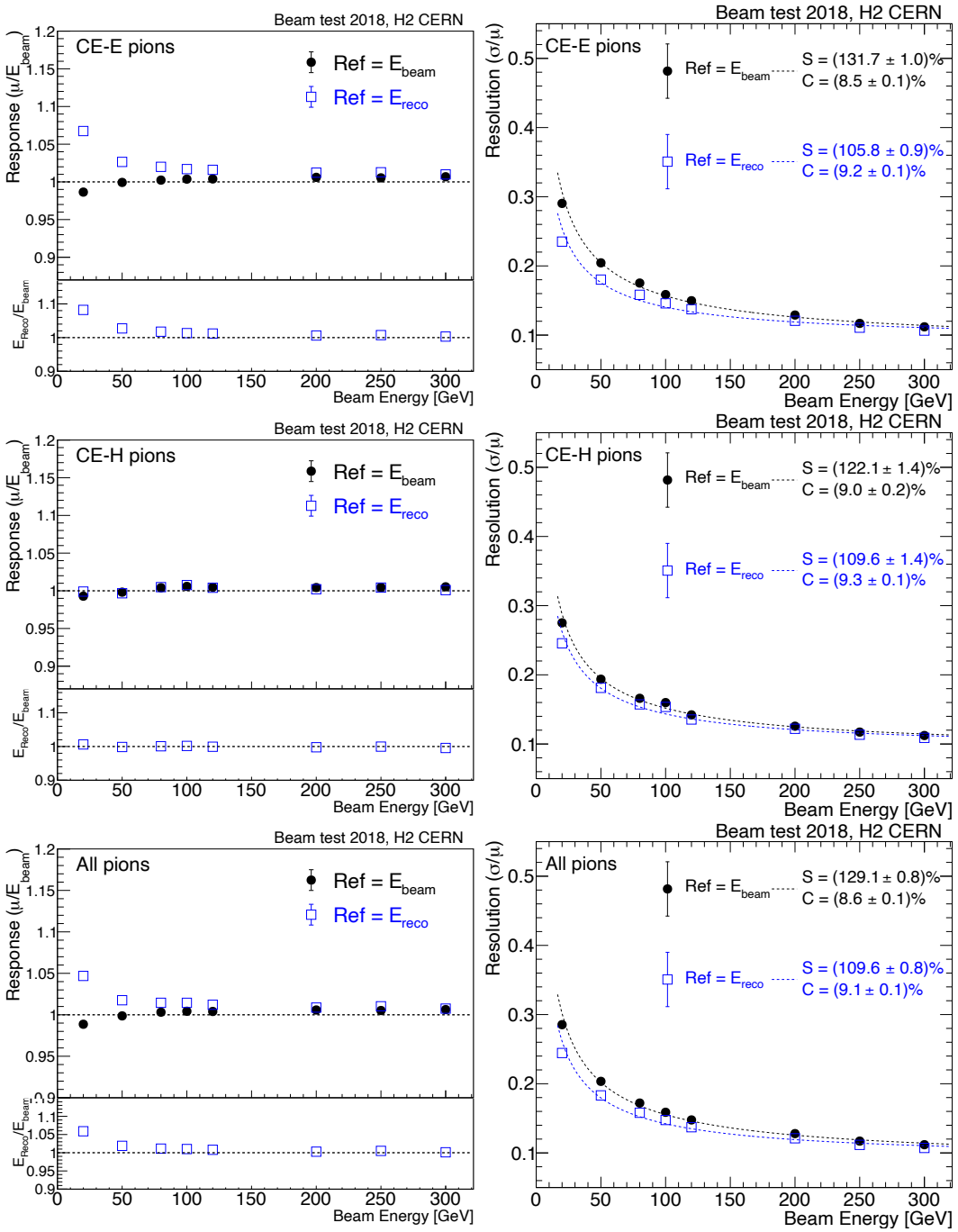

FIGURE 7.40: Energy response (left column) and resolution (right column) comparison between two reference used to extract the weights (black full circle: beam energy, blue open rectangle: reconstructed calorimeter energy) for CE-E pions (top row), CE-H pions (middle row) and all pions (bottom row). The resolution is better for E_{reco} taken as reference, however the response gets non-linear and energy is over-predicted for low energy CE-E pions.

Therefore, the extracted weight (when E_{reco} used as reference) is on average high, resulting in higher reconstructed energy. However, at higher energies, the non-linearity in energy response and weights flattens out, therefore reconstructed energy is almost identical for both the references. In summary, the energy dependent weights show good performance even with calorimetric energy inputs and can reliably be used for neutral and out-of-the-tracker hadrons.

7.5.3 Impact of transverse shower leakage on energy measurement

In the October 2018 beam test experiment, the transverse coverage for CE-E layers and last three layers (see section 6.3) were limited to just one module per layer because of limited availability of silicon sensor modules. Therefore, the lateral coverage for these layers were also limited to the dimensions of a single module which, if approximated, is a circle of radius roughly 7 cm. Such a transverse coverage is reasonable enough (although not ideal) for electromagnetic showers but not for hadronic showers. As discussed in section 2.3 that the longitudinal as well as lateral spread of hadronic showers are substantially larger than electromagnetic showers. Hence, a part of hadronic shower induced by charged pions in the beam test detector setup is guaranteed to be leaked-out in the transverse direction. This transverse leakage results in degradation of pion energy resolution.

To check the effect of the transverse leakage, we make use of simulation samples for pions produced with all layers in CE-E and CE-H containing 7 silicon modules in the daisy-like structure using FTFP_BERT_EMN physics list. The simulation sample is reconstructed similar to default simulation reconstruction procedure as described in section 6.5.2. Since the updated detector geometry has more silicon cells to detect the shower particles, therefore the calibration factors (i.e MIP-to-GeV weights) is updated before making any comparison with the default detector geometry. The same procedure is followed as described in section 7.4. The updated fixed weights are as follows:

$$\begin{aligned} \alpha^{\text{fixed}} &= 9.93 \text{ MeV per MIP ... using 50 GeV } e^+ \\ \beta^{\text{fixed}} &= 77.1 \text{ MeV per MIP ... using 50 GeV } \pi^- \\ \delta^{\text{fixed}} &= 0.36 \text{... relative weight between CE-H and AHCAL} \end{aligned} \tag{7.17}$$

Furthermore, the optimized weights are obtained as a function of beam-energy for the updated detector geometry using χ^2-minimization procedure as described in section 7.5. Equipped with the optimized weights for the updated simulation samples, the reconstructed energy is compared with the default geometry. Figure 7.41 shows comparisons of energy distributions for the two geometries for different beam energies for CE-E and CE-H pions. The energy distribution of CE-E pions for larger TB geometry shows a narrower distribution as compared to CE-E pions in default geometry. It is observed that the left tail gets shorter for larger geometry (blue line), whereas the right tails of the distribution is almost similar. For CE-H pions the differences are minuscule. These observations are consistent with the fact that all the CE-E layers having 7 modules, will contain more transverse shower as compared to default geometry. Also, since only last three layers of CE-H are extended to have 7 modules in updated larger geometry while all other layers already have 7 modules. Therefore, the effect is more pronounced for CE-E pions as compared to CE-H pions.

Also, the improvement is greater for low energy CE-E pions (e.g. 20 GeV) as compared to high energy pions (e.g. 300 GeV). This is because the shower particles for high energy

pions are more boosted along the shower axis, and therefore have higher probability of getting intercepted by initial layers of CE-H compartment as opposed to low energy pions where the shower particles are not as much boosted and leak out of the detector. To check the level of improvement in energy reconstruction with larger geometry, resolution is obtained as a function of beam energy and is compared against the default geometry as shown in Figure 7.42. For larger geometry, both the stochastic as well as constant term improves by about $\sim$ 15% (10%) for CE-E pions (CE-H pions).

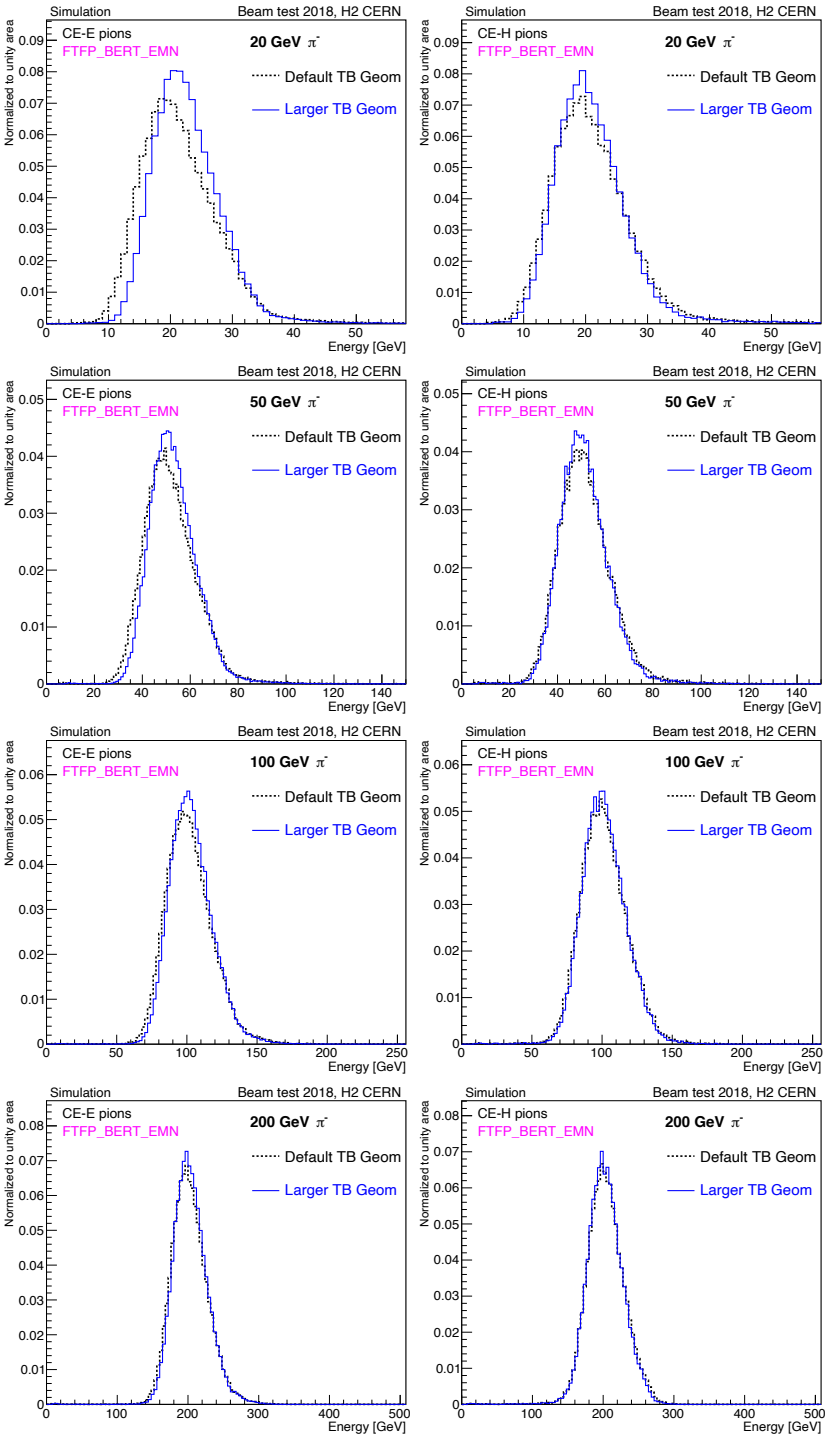

FIGURE 7.41: Energy distribution comparison between default (black dotted line) and larger TB geometry (blue solid line) for CE-E (left column) and CE-H pions (right column) of different energies in simulated samples. The left tail of energy distributions of CE-E pions improve with larger TB geometry indicating more transverse transverse shower containment.

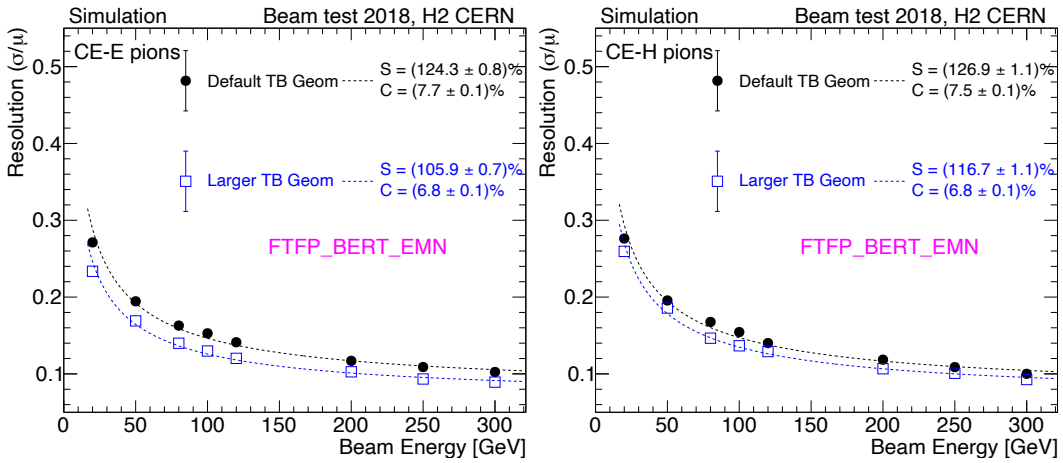

FIGURE 7.42: Comparison of energy resolution with default (black full circle) and larger TB geometry (blue open squares) in simulation. The resolution improves with larger TB geometry.

7.6 Longitudinal shower profiles

Study of shower development of hadronic showers in the detector is one of the important aspect of detector performance studies and simulation modeling. It helps us to understand the detector in a more detailed fashion and plays an important part in particle identification and reconstruction. The fine longitudinal segmentation of the CE-E and CE-H sections could be very useful in mitigating pileup during the HL-LHC operations. Comparing the shower development between data and simulation also helps to establish how well simulation is able to model the hadronic shower physics and make improvements, as needed.

The longitudinal shower development in HGCAL + AHCAL detector prototype is visualized in terms of average energy deposited ($\langle E \rangle$) in each layer. For example, Figure 7.43 shows the distribution of energy distribution in terms of MIPs at at different depths in the calorimeter, i.e. CE-E layer 10, CE-E layer 24, CE-H layer 1, and AHCAL layer 1 for all beam energies in data when pion starts showering at the start of the detector, i.e. at CE-E layer 3. The energy distributions show that for early showering pions, the energy deposition is higher earlier in the detector and is smaller at very later stage of shower development. The mean of these energy distributions at a given layer is plotted as a function of calorimeter depth as shown in Figure 7.44, when shower starts at different depths (shown in different colors) in the calorimeter for different pion beam energy in data. As the shower start location penetrates deeper into the calorimeter, the peak of the shower shape also moves deeper into the detector as expected. The average deposited energy shows a jump at the interface of CE-E and CE-H where a larger $\langle E \rangle$ is observed at the first layer of CE-H as compared to the last layer of CE-E. The jump in $\langle E \rangle$ at the interface is due to the that CE-H (with 7 silicon modules per layer) provides a larger transverse coverage as well as catches the shower particles leaking from CE-E in the lateral direction. It is also observed that significantly lower average energy is deposited in the last three layers of CE-H as compared to its neighbouring layers. This is because the last three layers of CE-H are instrumented with only one silicon module per layer (section 6.3), thus energy recorded is smaller as compared to other CE-H layers.

For a given shower start location, the longitudinal shower development is compared for different energy pions as shown in Figure 7.45. It is observed that with increasing pion beam energy, the peak of the shower profile shifts deeper into the calorimeter as well as

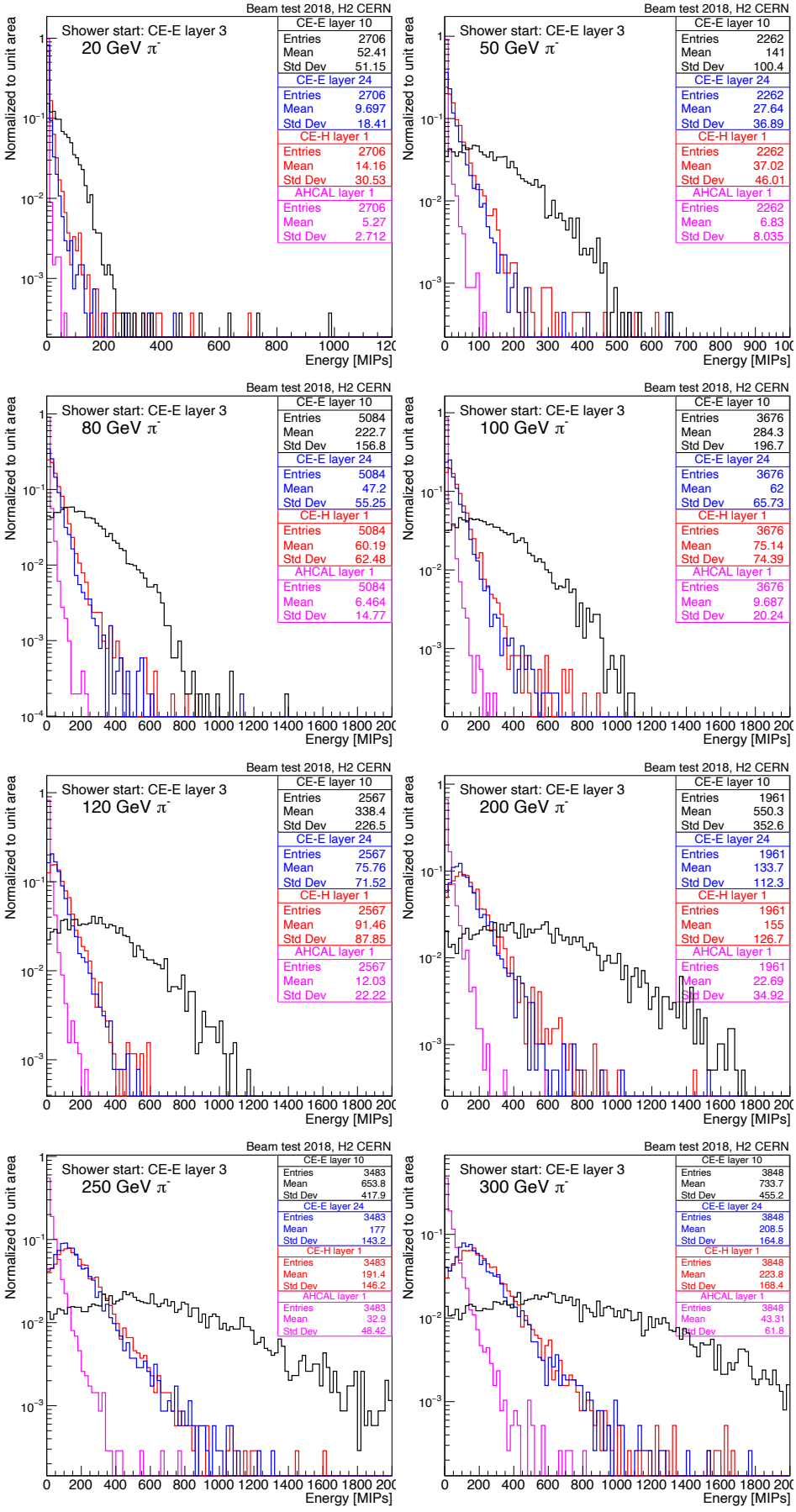

FIGURE 7.43: Distribution of energy deposited at different depths in calorimeter when pion starts showering at CE-E layer 3.

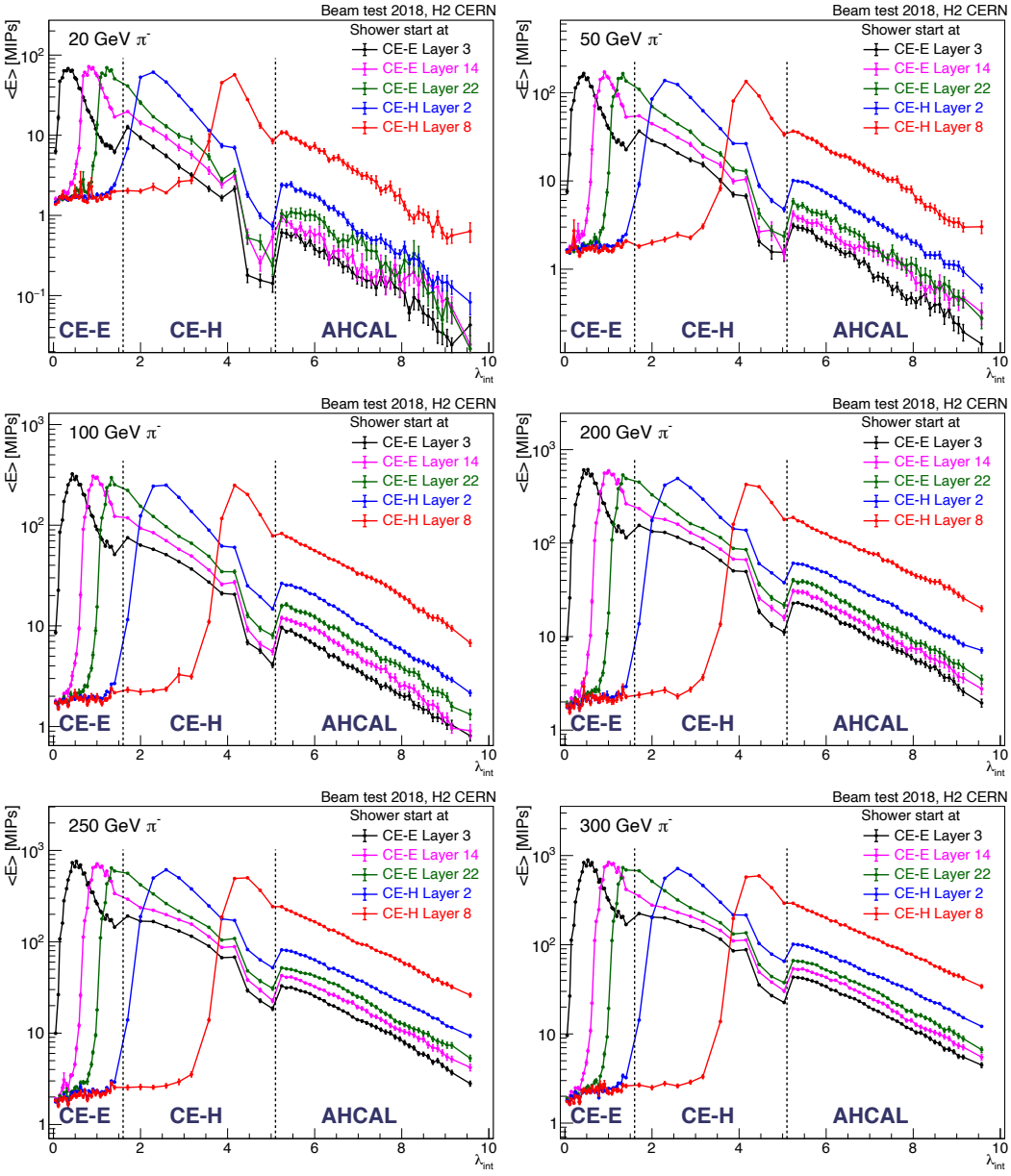

FIGURE 7.44: Average energy deposited (in MIPs) in each layer as a function of calorimeter depth (in λ_{int}) for pions start showering at different layer (shown in different colors) for different beam energies. The shower peak and the tail moves deeper into the detector when pion starts shower in the later layers.

higher average energy is deposited in the layers. This is consistent with the expectation as for higher incident pion energy, the shower particle multiplicity is higher and are more energetic , therefore penetrate deep into the calorimeter.

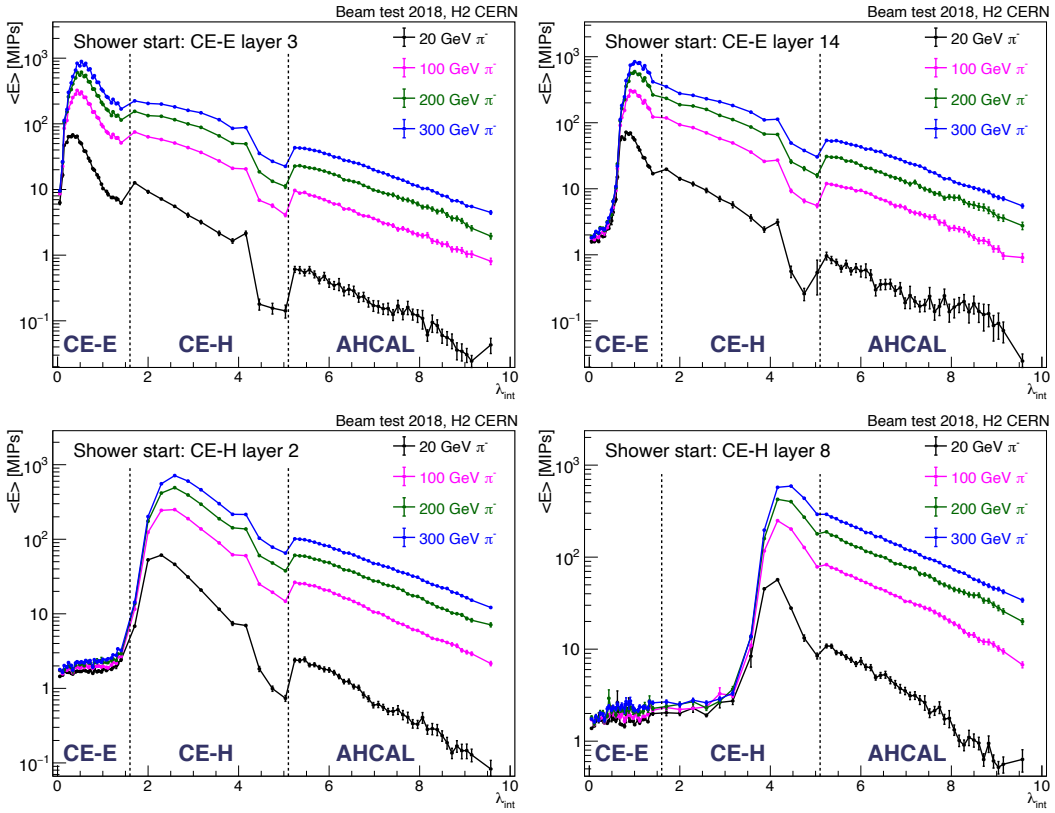

FIGURE 7.45: The comparison of longitudinal shower development among different energy pions (shown in different color marker and line) when start showering at a given layer. The shower peak for higher energy pion lies deeper into the detector with higher energy deposited at each layer, as expected.

7.6.1 Data-simulation comparison (in MIPs)

The modelling of hadronic shower in simulation is compared against that in data. Figure 7.46 and 7.47 shows longitudinal shower profile comparison when pions start showering in CE-E layer 3 (in left column) and CE-H layer 2 (in right column) for 20 to 100 GeV and 120 to 300 GeV pions, respectively. Both the physics lists i.e. FTFP_BERT_EMN as well as QGSP_FTFP_BERT_EMN reproduce all the features of the shower development very well as observed in data. At high energies for early showering pions (Figure 7.47) (left column), it is observed that energy deposited in the AHCAL layers are under-predicted, whereas in CE-E and CE-H layers the energies are over-predicted in simulation by $\sim$ 10-20%. This suggests that while the total energy deposition is consistent with data, the shower length is shorter in simulation which could be improved by tuning the hadronic physics modelling.

7.6.2 Qualitative assessment of neutral pion contribution at first hadronic interaction in CE-E

One of most striking feature that is observed in the longitudinal shower profiles in data (and also well reproduced by simulation), is that the shower peaks rather early ($\lambda_{int} \leq 0.3$) when pions start showering early in CE-E. Also, the peak is quite narrower as compared

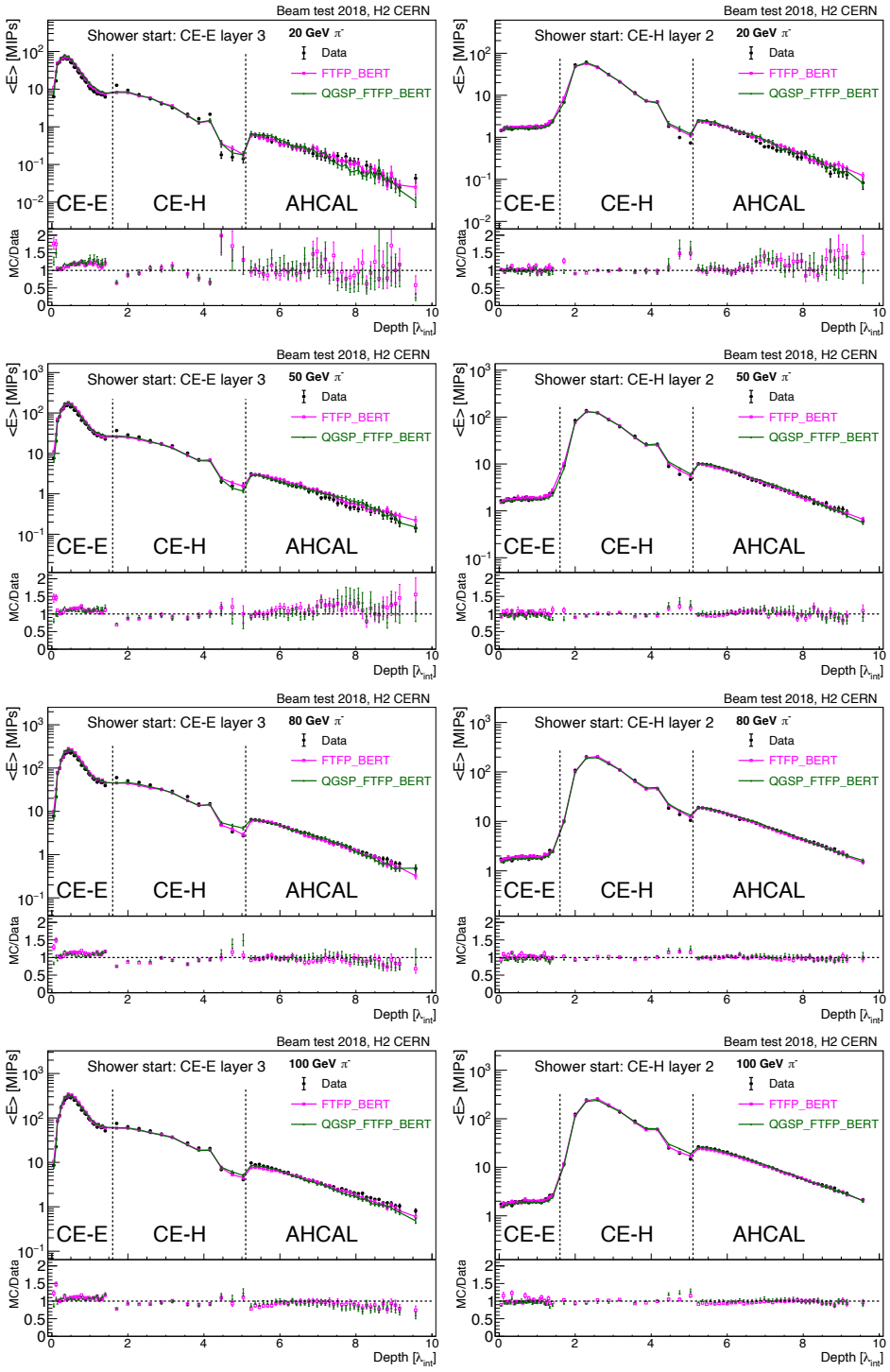

FIGURE 7.46: Data-simulation comparison of average energy deposited in each layer as a function of calorimeter depth (λ_{int}) when pion startrs showering in CE-E layer 3 (left column) and CE-H layer 2 (right column). The data is shown in black marker and the two coloured lines correspond to FTFP_BERT_EMN (magenta line) and QGSP_FTFP_BERT_EMN (green lines) physics lists, for beam energies of **20 to 100 GeV**.

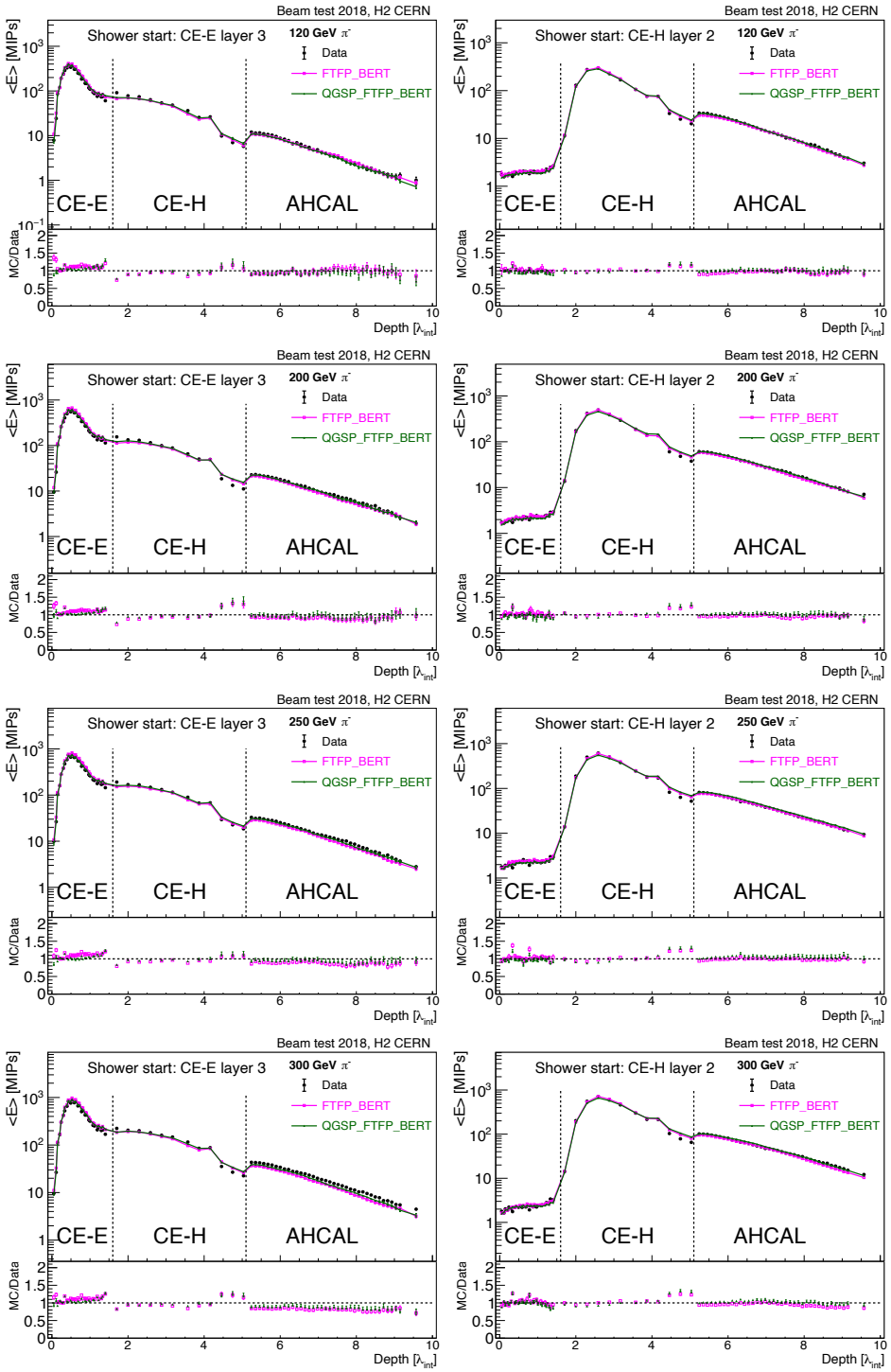

FIGURE 7.47: Data-simulation comparison of average energy deposited in each layer as a function of calorimeter depth (λ_{int}) when pion startrs showering in CE-E layer 3 (left column) and CE-H layer 2 (right column). The data is shown in black marker and the two coloured lines correspond to FTFP_BERT_EMN (magenta line) and QGSP_FTFP_BERT_EMN (green lines) physics lists, for beam energies of **120 to 300 GeV**.

to shower profiles of pions that start showering in CE-H (compare left and right column of Figure 7.46 and 7.47).

Early showering CE-E pions encounter Pb/CuW , i.e. CE-E absorber material whereas, CE-H pions encounter steel, i.e. CE-H absorber material. Both of which have different X_0/λ_{int} ratios (see table 2.1). This has an interesting consequence on the shower development of electromagnetic and hadronic component of the hadronic shower. Steel has $X_0/\lambda_{int} \simeq 0.1$ whereas Pb and CuW has $X_0/\lambda_{int} \simeq 0.03$ which is about 3 times smaller than steel. It means electromagnetic part of hadronic shower sees 3 times more material than the hadronic part. Since CE-E is ~ 27 X_0 deep, therefore all the electromagnetic particles (mainly coming from π^0) produced in the first hadronic interaction, will be well contained in CE-E. Also, since the electromagnetic showers have high particle multiplicity than hadronic showers (section 2.3.2), therefore the peak for early showering CE-E pions in CE-E corresponds to the electromagnetic component of hadronic shower. Whereas for CE-H pions, the shower develops in steel absorber (having higher X_0/λ_{int} ratio), both the electromagnetic as well as hadronic component spreads almost equally in the longitudinal direction as compared to CE-E, thus showing a broader peak later in CE-H.

To verify above reasoning, we make use of the truth information, i.e. secondary particles produced at the first hadronic interaction, from GEANT4 simulation which was used for the optimization of shower start finder algorithm as discussed in section 7.3.2. The estimation of the energy carried by the electromagnetic component at the first hadronic interaction is done as follows:

$$E^{frac}_{\pi^0} = \frac{\sum_i E^i_{\pi^0}}{E_{beam}} \tag{7.18}$$

where $E^i_{\pi^0}$ on the right hand side of the above equation is the energy carried by i^{th} π^0 and is summed over π^0s produced at the first hadronic interaction and E_{beam} is the incident π^- beam. Thus, the fractional energy carried by electromagnetic component as $E^{frac}_{\pi^0}$ is calculated for each event.

The distribution for $E^{frac}_{\pi^0}$ is shown in the Figure 7.48 for all beam energies when the pion starts showering in CE-E layer 3 to 7 inclusively. The inclusive events for shower starting from CE-E layer 3 to 7 are considered in order to increase the statistics (showering in the first two layers of CE-E are rejected during pre-shower cleaning cut).

It is important to give emphasis on the fact the π^0 considered in equation 7.18 and shown in Figure 7.48 are from the **first hadronic interaction only**, and not from all the downstream interactions within the shower development. Therefore, it should not be used for a quantitative inference of total electromagnetic component of hadronic showers.

The fractional energy distribution for π^0 shows that, on an average, neutral pions carry $\sim 40\%$ of the incident pion energy during the first hadronic interaction. Since, the distribution corresponds to only first hadronic interaction, therefore it does not show or convey the message that the electromagnetic component increases with increasing energy of the incident hadron. The probability to produce π^0 later in the shower increases with increased available energy of shower particles for higher incident hadron beam as discussed in section 2.6.2 [39].

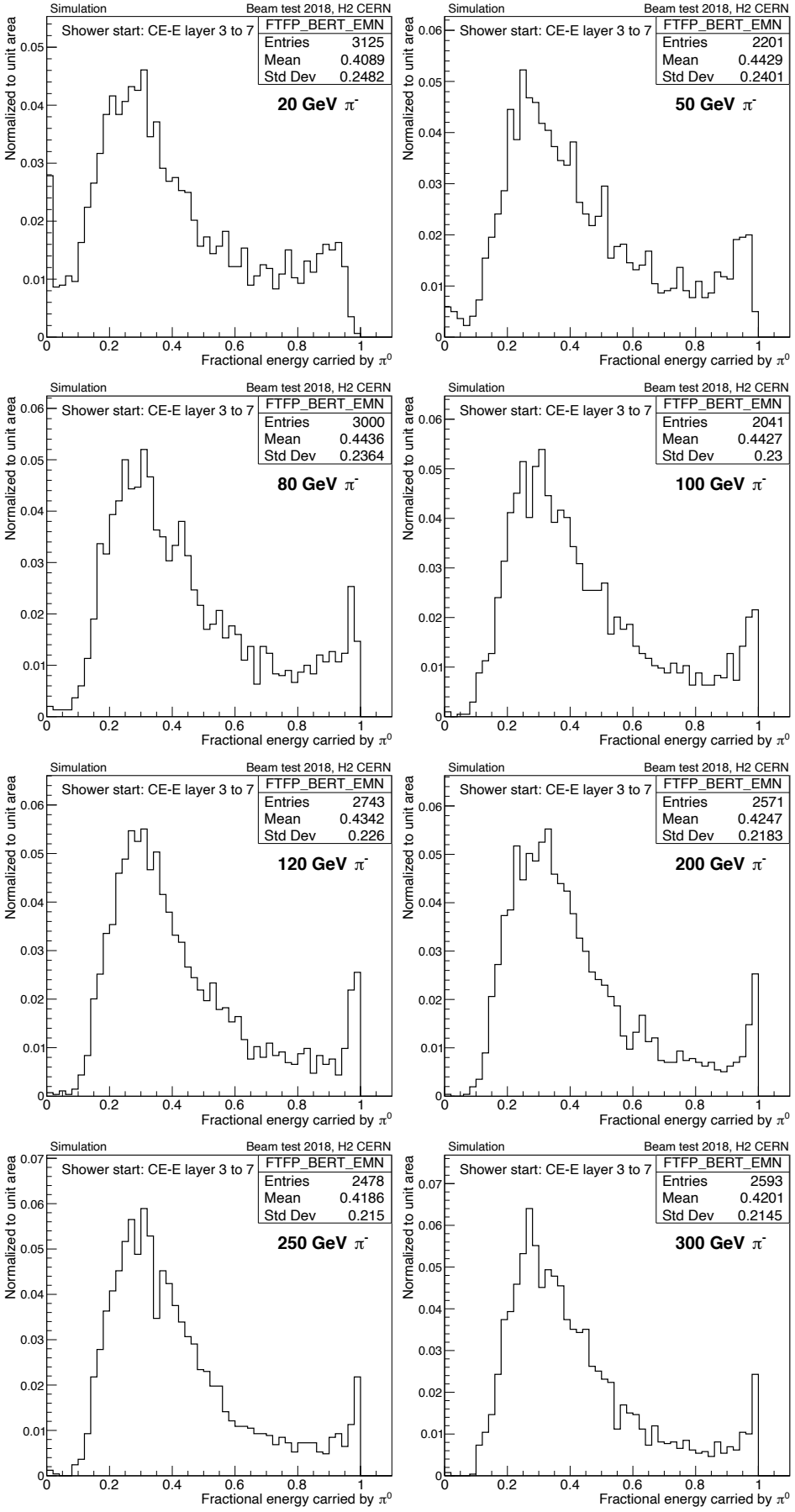

FIGURE 7.48: Fractional energy carried by π^0s produced at first hadronic interaction point in CE-E layer 3 to 7 for all π^- beam energies.

Using the information of $E^{\pi^0}_{frac}$ as shown in Figure 7.48, the events are divided into following three categories based on electromagnetic component.

- Low electromagnetic component : events with $E^{\pi^0}_{frac} < 0.2$
- Mid electromagnetic component : events with $0.2 < E^{\pi^0}_{frac} < 0.4$
- High electromagnetic component : events with $E^{\pi^0}_{frac} > 0.4$

The longitudinal shower profile is plotted for each of above three categories on the same canvas to compare the mean energy deposited as a function of calorimeter depth. Figure 7.49 shows longitudinal shower profile for three categories (shown in different markers and colours) for all beam energies, inclusive in shower starting in CE-E layer 3 to 7.

It is observed that the events where electromagnetic fraction is higher i.e. ($E^{\pi^0}_{frac} > 0.4$), the average energy deposited per layer $\langle E \rangle$ is higher in CE-E compartment but is lower in CE-H compartment as compared to other two categories. Conversely, events where electromagnetic fraction is lower i.e. ($E^{\pi^0}_{frac} < 0.2$), the average energy deposited per layer $\langle E \rangle$ is higher in CE-H compartment but is lower in CE-E compartment as compared to other two categories.

In conclusion, these observations are a clear signature that the peak in CE-E for early showering CE-E pions corresponds to π^0s produced in the shower. The fine longitudinal granularity and detector design of CE-E allows us to probe electromagnetic component of the hadronic shower. This has a far-reaching implications on the overall performance of the HGCAL for hadronic showers, such as event-by-event compensation of electromagnetic component fluctuation.

7.6.3 Longitudinal shower profiles (in GeV)

The longitudinal shower shape is also obtained in terms of energy deposited in each layer in terms of GeV ($\langle E_{GeV} \rangle$). The energy in terms of GeV is obtained by applying energy-dependent weights in corresponding sub-detectors based on shower start location according to equation 7.10 and 7.11.

Figure 7.50 shows data-simulation comparison of $\langle E_{GeV} \rangle$ per layer as a function of calorimeter depth (λ_{int}) for different pion beam energies when pion starts showering in CE-E layer 3 (left column) and in CE-H layer 8 (right column). The narrow peak in CE-E for pions starting shower in early CE-E layer and lower average energy in the last three layers of CE-H is observed for $\langle E_{GeV} \rangle$ as well, consistent with $\langle E_{MIPs} \rangle$ (Figure 7.46 and 7.47).

In Figure 7.50, the jump at the interface of CE-E and CE-H is rather larger as compared to $\langle E_{MIPs} \rangle$ per layer (Figure 7.46 and 7.47). This can be understood by considering the fact that MIP-to-GeV conversion factors or weights, take the absorber thickness into account and reflect the energy deposited into the absorbers. Since $\Delta\lambda_{int}$ for CE-H (~ 0.3) is considerably larger than $\Delta\lambda_{int}$ for CE-E (~ 0.05) (see Figure 6.11), therefore the corresponding energy deposited will also be larger in the absorber layers. The absorber material is also different in CE-E and CE-H, therefore the ratio of MIP energy deposited does not exactly scale as the ratio of $\Delta\lambda_{int}$. This feature is more evident in the plots in right column of Figure 7.50 where pion starts showering in CE-H layer 8 and acts like a MIP till CE-H layer 7. Even if the sampled energy by the active layers are ~ 1 MIP, the energy deposited in the corresponding

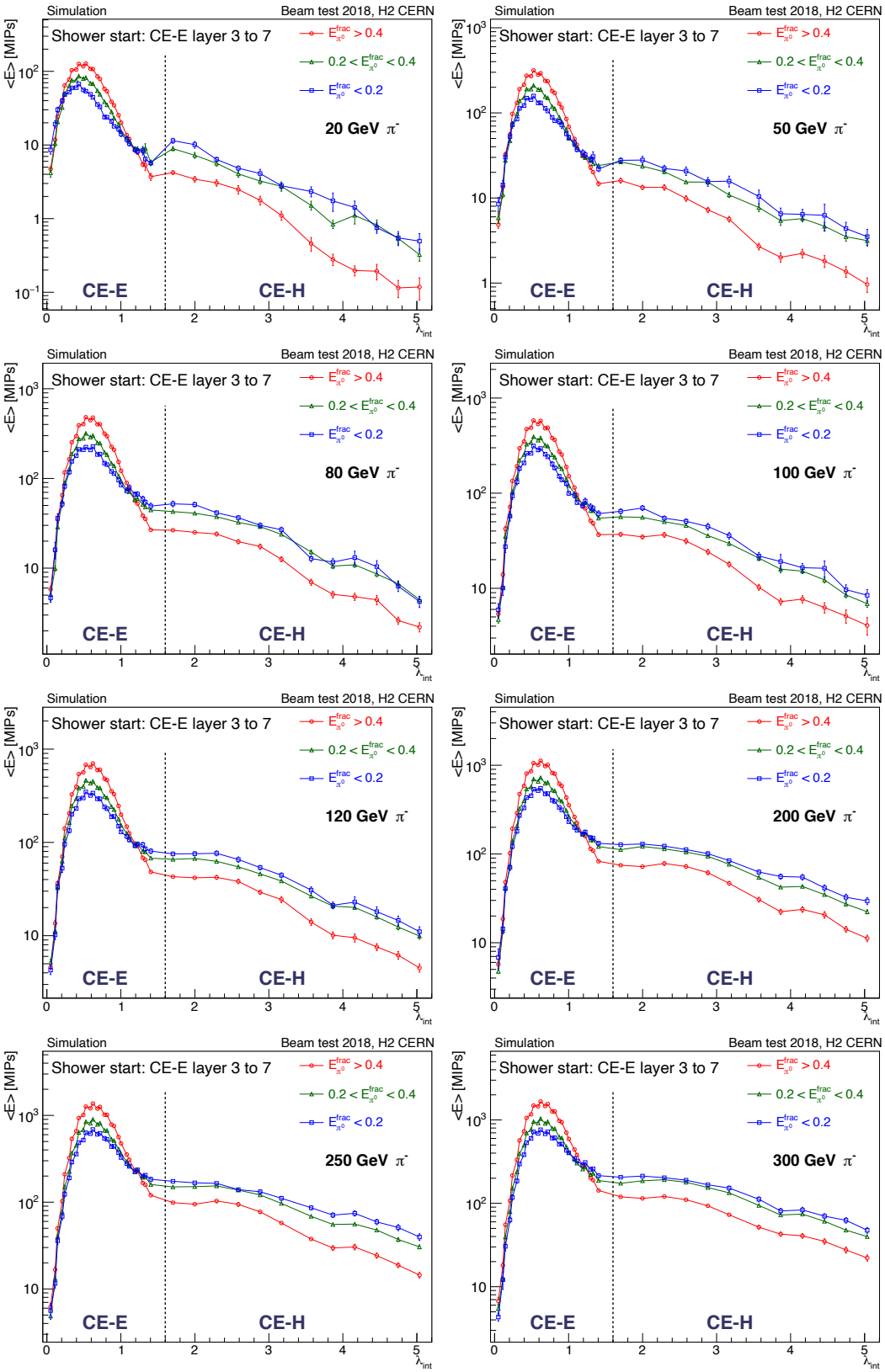

FIGURE 7.49: Mean energy deposited in each layer in terms of MIPs as a function of calorimeter depth (λ_{int}) for high electromagnetic fraction (red open circles connected with red line), medium electromagnetic fraction (green open triangles connected with green line) and high electromagnetic fraction (blue open squares connected with blue line) for pion beam energies.

The events are inclusive for shower starting in CE-E layer 3 to 7.

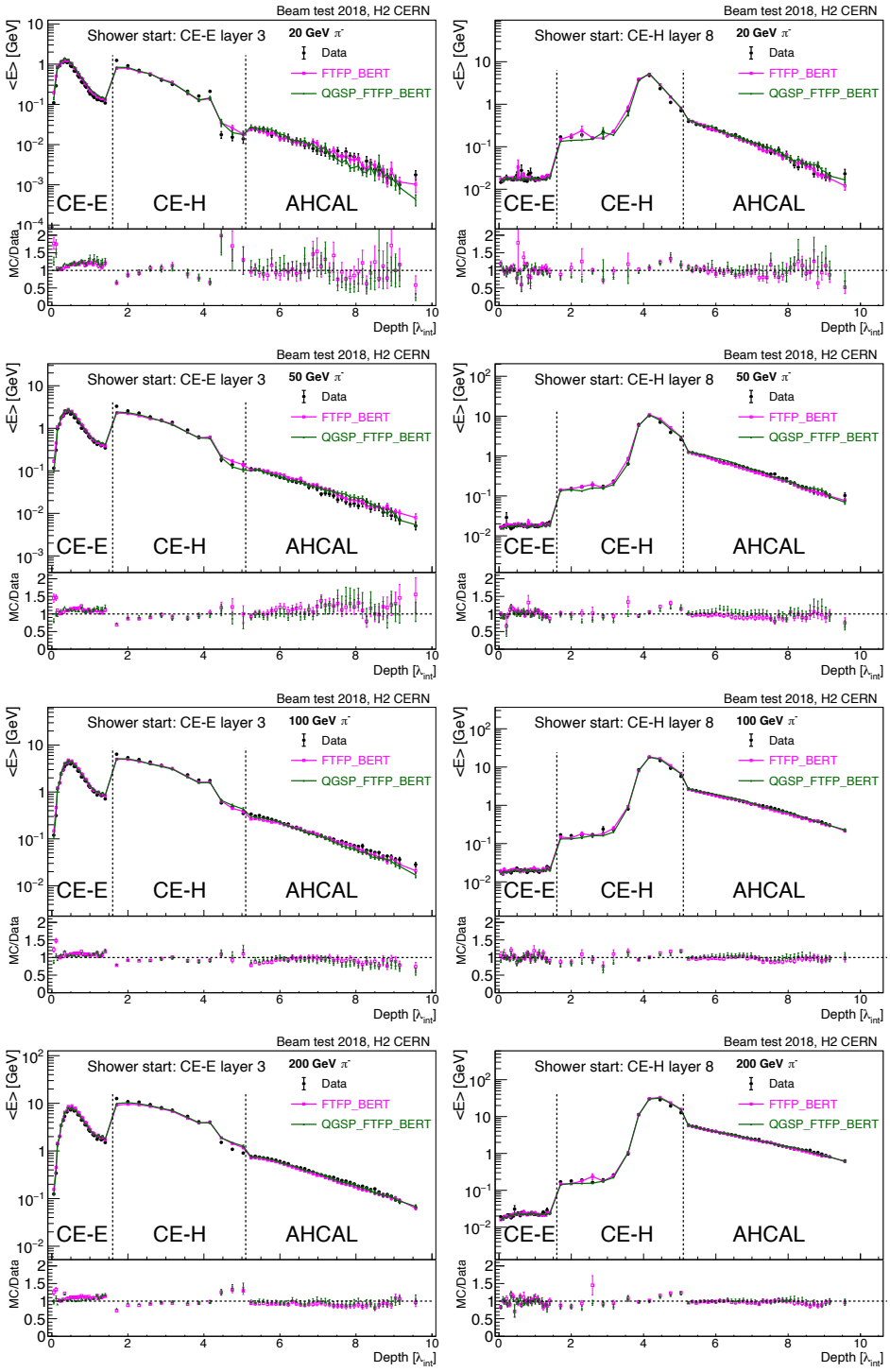

FIGURE 7.50: Figure shows average energy deposited per layer in terms of GeV as a function of calorimeter depth (λ_{int}) for pions start showering in CE-E layer 3 (left column) and CE-H layer 8 (right column) for different pion energies. The black full circle corresponds to data and magenta and green line corresponds to FTFP_BET_EMN and QGSP_FTFP_BET_EMN physics lists, respectively.

absorbers are not the same. Therefore, a jump is observed from $\langle E_{GeV} \rangle \sim 0.02$ GeV per layer in CE-E to $\langle E_{GeV} \rangle \sim 0.16$ GeV in the first few layers of CE-H. All the mentioned features are very well reproduced by the simulation with both the physics lists and a good agreement is observed between data and simulation.

7.7 Transverse shower profile

As the hadronic shower penetrates deeper into the calorimeter, the secondary particles in the shower scatter off of the detector material and move away from the shower axis while depositing energy in the detector. The study of the shower spread in the lateral direction is an important aspect to quantify the performance of the detector as well as to characterize hadronic shower. The development of hadronic shower in lateral direction can be visualized in terms of energy contained in the circles of different radii from the track impact position at various stages (or depth) into the shower. The track are reconstructed using the DWC information as explained in section 6.3.3, which are extrapolated to each layers. Figure 7.51 and 7.52 shows energy contained in the circles of varying radii at layers downstream of shower start location, i.e. CE-E layer 3 to layer 7, for 100 and 200 GeV/c pion beam, respectively.

It is rather trivial to infer from the plots that as the radius is increased, more energy is accounted for within the circle. Also, it is observed that around CE-E layer 9 to 15 (corresponding to the peak in longitudinal shower profile in Figure 7.44), the cells that fall within 2 cm (blue points) from the shower axis receive maximum energy, and beyond that, integrated energy increases relatively slowly. This observation corroborates with the fact that for early showering pions CE-E layer 9 to 15 corresponds to maximum of electromagnetic component of hadronic shower as discussed in previous section. Furthermore, at the later stage of the shower, the integrated energy increases gradually indicating that the lateral spread is larger and is relatively uniform. Plots for other beam energies can be found in appendix H.

The distributions of energy integrated in circles of different radii is compared between data and simulation with two physics lists in order to assess the modelling of hadronic showers in HGCAL TB prototype. Figure 7.53 shows similar energy distribution comparisons between data and simulation for a three example radii at five layers for 100 GeV/c pion. Both physics lists reproduce the shapes well with minor differences in the tails. Plots for other beam energies can be found in appendix I.

Comparing and studying energy distributions in the manner described and shown so far is a cumbersome process as the number of distributions grow linearly with the number of beam energy points. Therefore, instead of using circles to integrate energy at each layer, we use cylinders of different radii to calculate energy contained in CE-E, CE-H, and AHCAL inclusively, thus minimizing number of distributions and encapsulating lateral shower development effectively at the same time. The energy containment in the cylinder of different radii is calculated in following two steps:

- **Step 1:** For all the layers in a given compartment, a histogram is filled with the distance (R) between each reconstructed hit (passing the noise thresholds) and the track impact position, weighted with their respective energies. Weighting the histograms with energy automatically sums-up the energies and makes the procedure easier to implement. In this way, each bin of the histogram holds the total energy summed-up in a ring of dR at a distance R from the track. The radius is increased in the step of dR

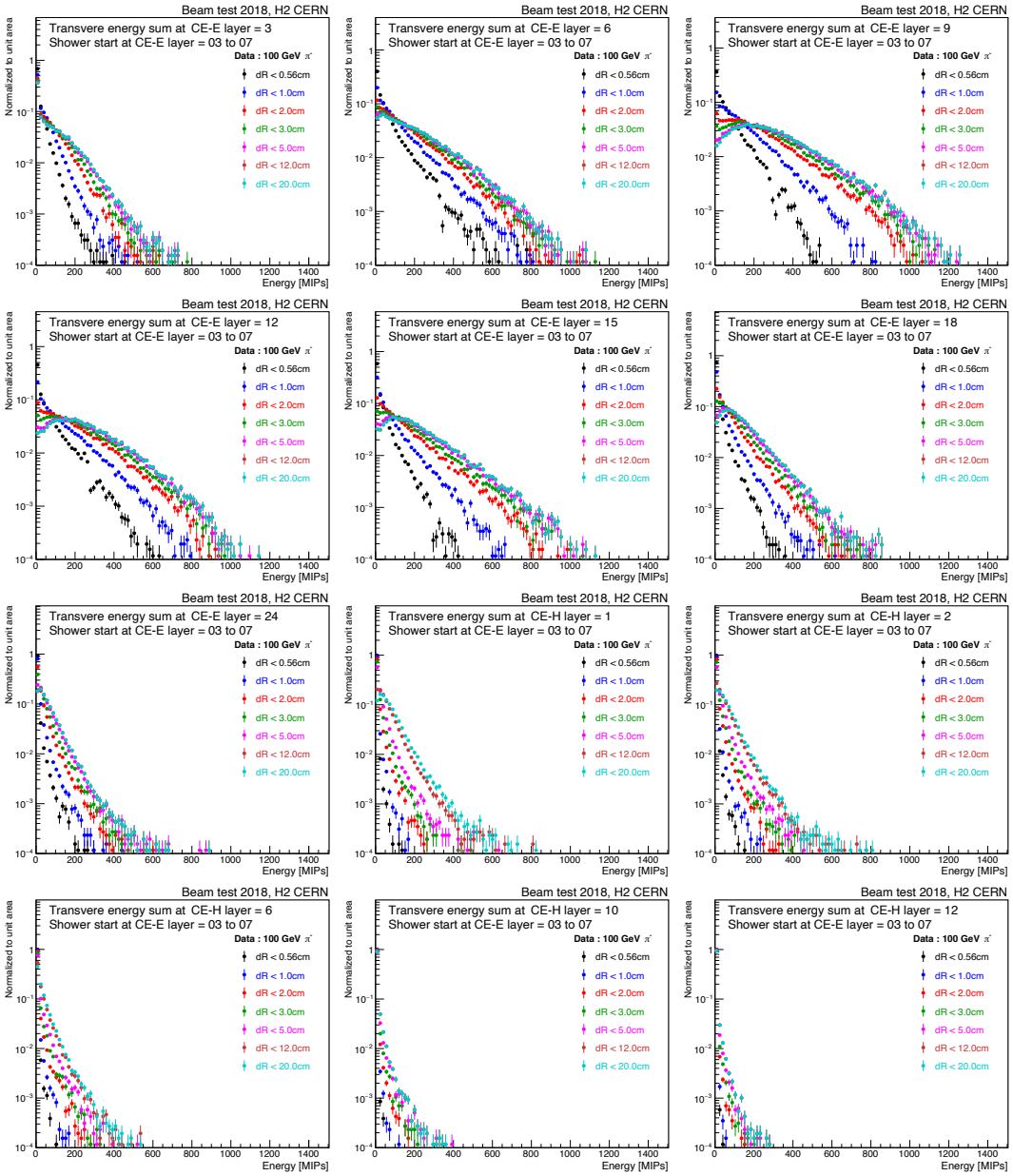

FIGURE 7.51: Energy distributions in the circle of different radii (shown in different colour points) from the track impact position at layers downstream to shower start location from CE-E layer 3 to 7 for **100 GeV/c** pion beam.

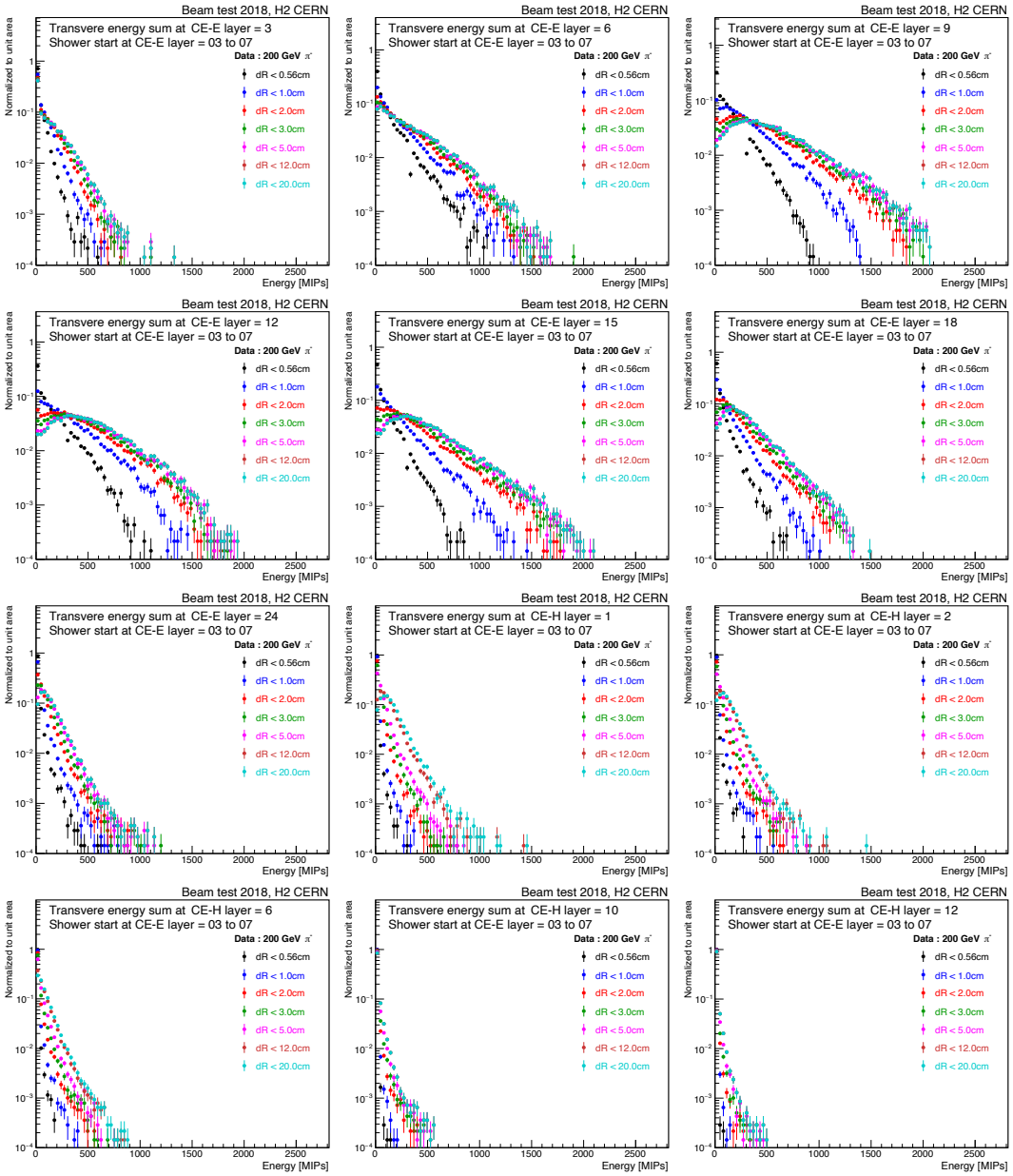

FIGURE 7.52: Energy distributions in the circle of different radii (shown in different colour points) from the track impact position at layers downstream to shower start location from CE-E layer 3 to 7 for **200 GeV/c** pion beam.

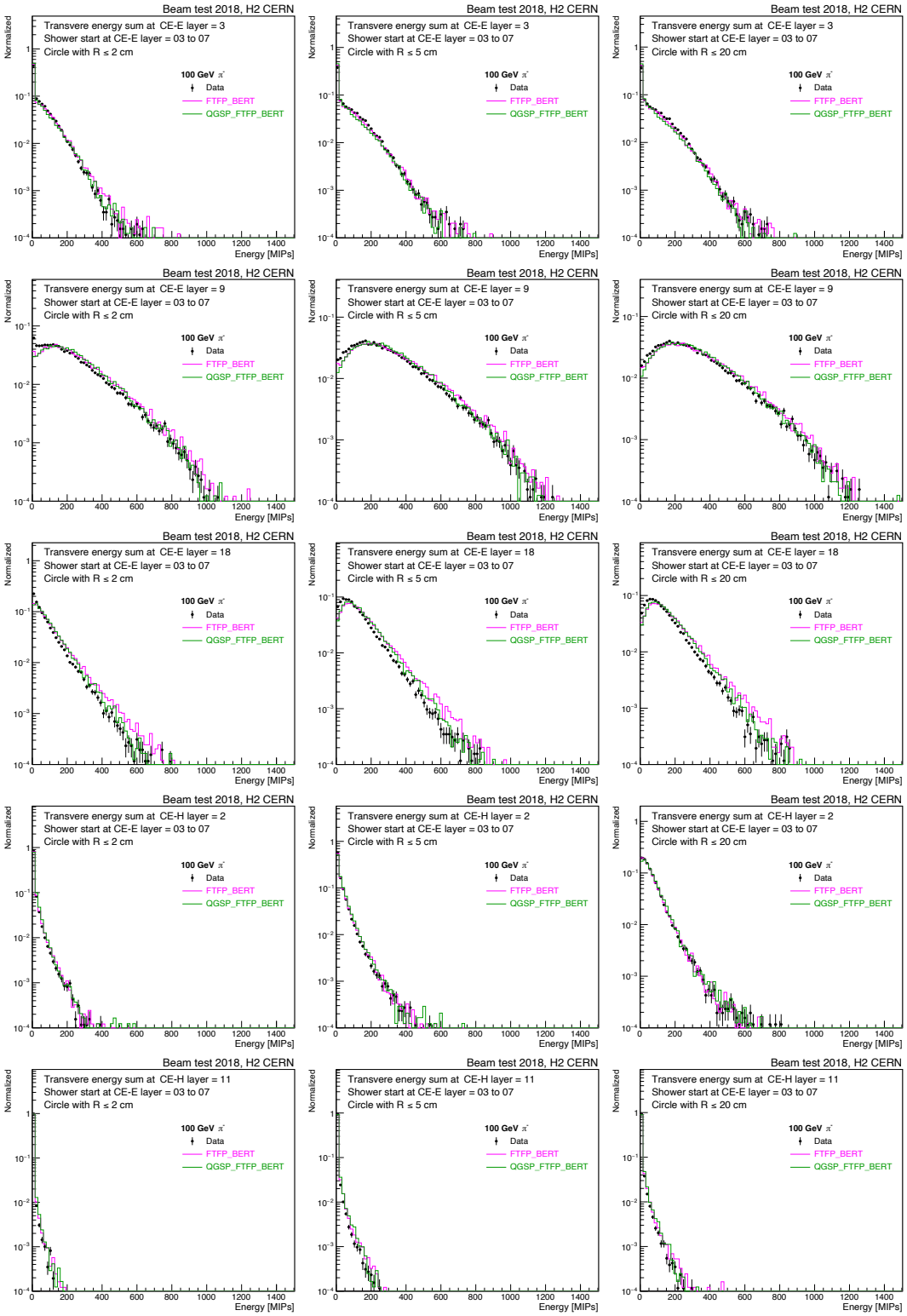

FIGURE 7.53: Data-simulation comparison of energy distributions in the circle of radius 2 cm (left column), 5 cm (middle column) and 20 cm (right column) from the track impact position at layers downstream to shower start location from CE-E layer 3 to 7 for **100 GeV/c** pion beam.

and the histogram is filled accordingly . This is done on event-by-event basis for all the layers of CE-E, CE-H and AHCAL separately.

- **Step 2:** At the end of each event, the histogram is integrated bin-by-bin and is filled into a `TProfile` [126], normalized with the total energy in the corresponding compartment.

Steps 1 and 2 are repeated for all the layers of events; thus, one obtains the "average energy contained in a cylinder of varying radius R". The energy asymptotically reaches one depending on the transverse size of the layer, e.g. $\sim$ 13 cm for CE-E. This procedure has been shown in a cartoon representative image in Figure 7.54. The central cell (shown in red) records the highest energy in this representation. The cells in the first and second ring around it, are second (yellow) and third (green) highest energy cells, respectively. The track impact position has been marked with a blue cross, and it need not coincide with the center of highest energy cell as shown in the image. In fact, it is highly improbable that the center of cell and track impact position will coincide. A ring of thickness dR is thrown around the track impact point at a radius R. The energies of all the cells, whose center falls within the ring, are summed up and filled into the histogram. In this example, three cell centers (yellow cells with the magenta coloured center, marked as numbers 1, 2, and 3) fall within the ring. Therefore, the histogram is filled with R, weighted with their corresponding energies. This is repeated for all the layers of each compartment. The value of R is varied from zero to a large value such that it covers cells in entire layer. At the end of the event, the histogram is integrated and normalized with the total energy of the compartment. Finally, the average energy is obtained by taking the mean of each bin over the events.

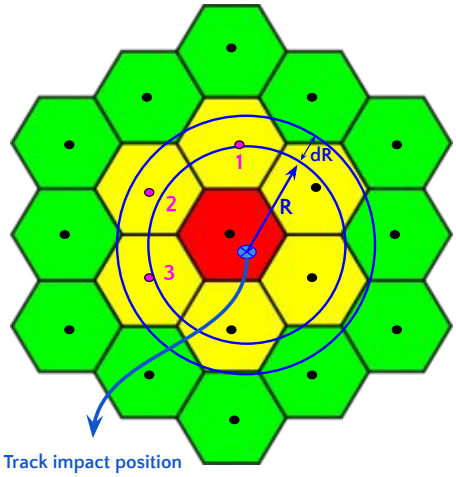

FIGURE 7.54: Pictorial representation of the calculation of energy contained in a cylinder of different radius for a layer in an event. Cells marked as 1, 2 and 3 fall within the ring of dR at a distance R from the track impact point. See corresponding text for explanation.

Figure 7.55 and 7.56 shows the comparison between data and simulation of average fractional energy contained in the cylinders of increasing radii for different beam energies in CE-E (when shower starting in CE-E layers 3 to 7), CE-H and AHCAL (when shower starting in CE-H layer 1 for CE-H and AHCAL) for all beam energies. It is observed that as the radius of the cylinder increases, more energy is contained in the volume and it reaches

to one (i.e. the total energy in the compartment) as we reach the limit of transverse dimension of active layers.

For pions showering in CE-E (CE-H), it is observed that the energy contained in the volume of CE-E (CE-H) increases rapidly as the radius of cylinder is increased. This observation points to the fact that the core of hadronic shower is made up of energetic electromagnetic component for which the volume of containment is smaller as compared to hadronic component. The energy carried by electromagnetic core depends on the incident energy of the hadron and increases with increasing incident energy. That is why the radius (hence volume) required to contain 90% of the total energy in the compartment decreases as the energy of pion beam increases.

The transverse profile in AHCAL for pions showering in CE-H layer 3 shows relatively less steeper curve. This is because AHCAL receives the tail of the showers for these pions which is mostly hadronic component, hence the spread in the transverse direction is larger and requires larger volume to contain the shower energy.

The distance calculation between track and cell has two components: coordinate of track impact point $(x, y)_{track}$ and the center of the cells $(x, y)_{cell}$. Since $(x, y)_{cell}$ are located at a fixed points in the module, hence are discrete, whereas $(x, y)_{track}$ can take any value and hence it is fairly continuous distribution. It makes the resulting distance (R) values continuous as well, which is reflected in the transverse shower profile in Figure 7.55 and 7.56.

The simulation with both the physics lists is observed to reproduce transverse shower development well. However, at the core of the shower axis (R near zero), a disagreement is observed between data and simulation for lower energies. The discrepancy is also observed in AHCAL at higher energies. This could be the effect of differences in beam profiles in between data and simulation (Figure 7.4) and/or in the modelling of the hadronic shower development in simulation which can benefit from further detailed investigations of in the simulation modelling.

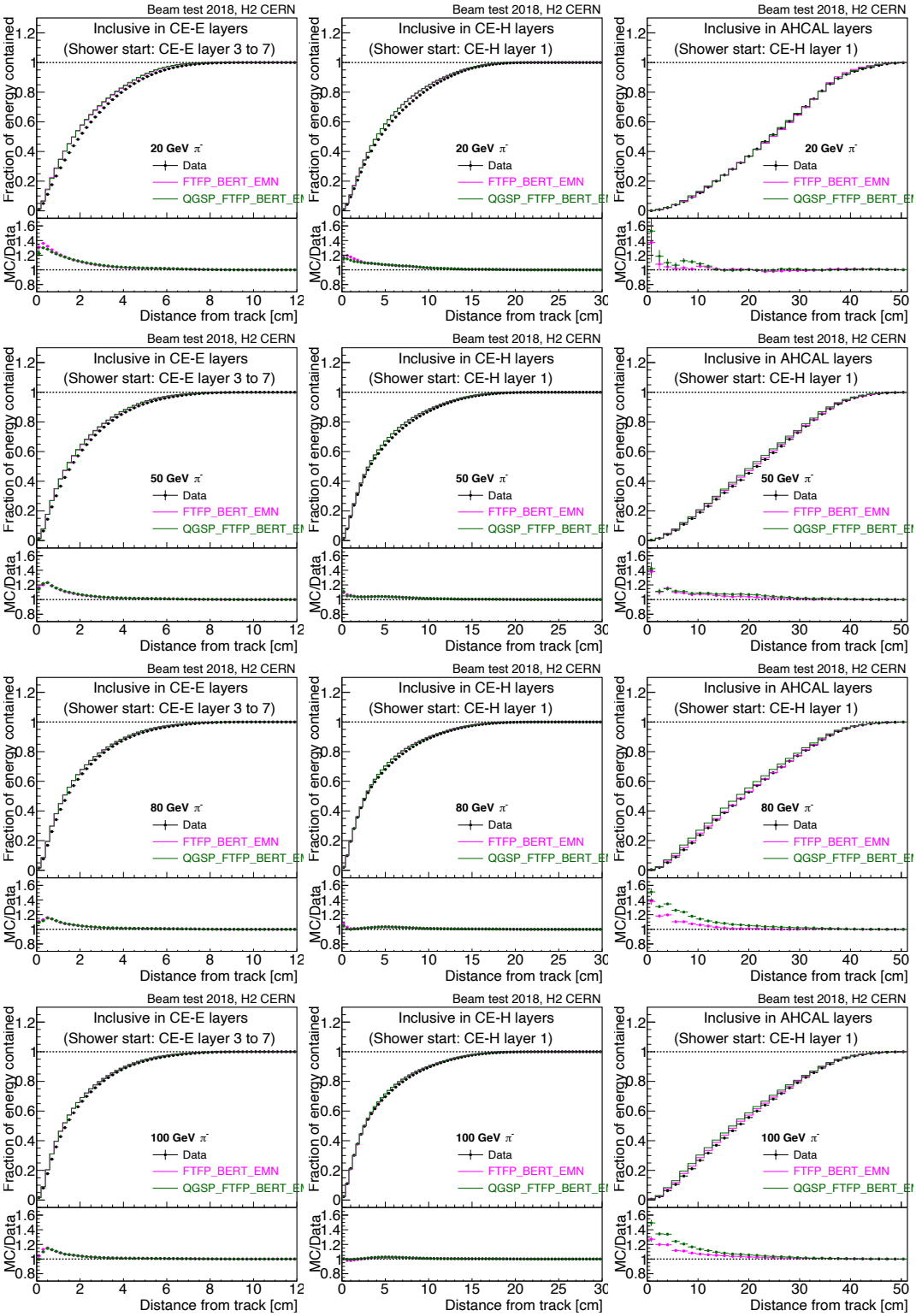

FIGURE 7.55: Transverse shower profile comparison between data (black full circles) and simulation with FTFP_BERT_EMN (magenta line) and QGSP_FTFP_BERT_EMN (green line) for pions of beam energies **20 to 100 GeV**. The left, middle and right column corresponds to CE-E, CE-H and AHCAL, respectively.

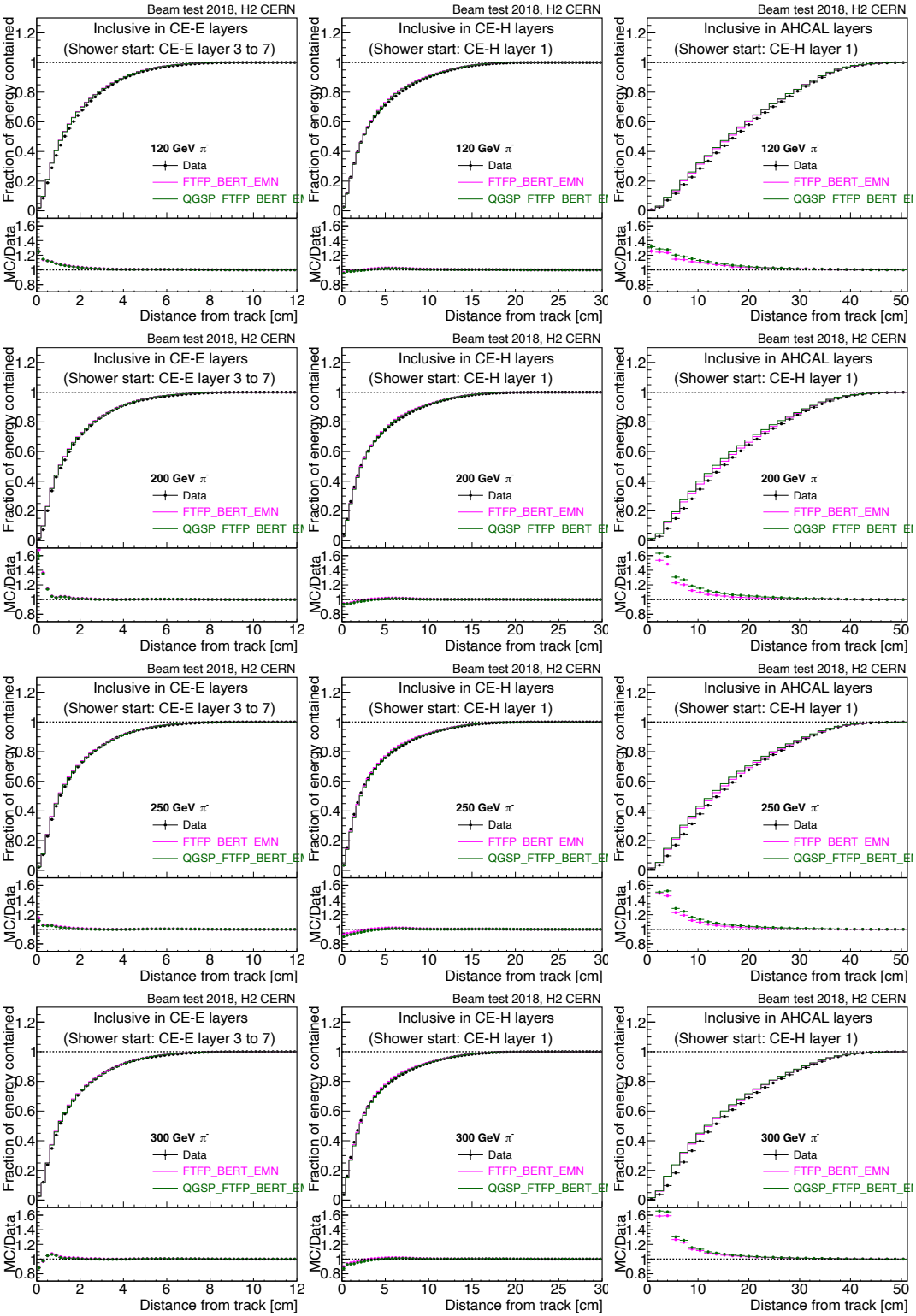

FIGURE 7.56: Transverse shower profile comparison between data (black full circles) and simulation with FTFP_BERT_EMN (magenta line) and QGSP_FTFP_BERT_EMN (green line) for pions of beam energies **120 to 300 GeV**. The left, middle and right column corresponds to CE-E, CE-H and AHCAL, respectively.

Chapter 8

Summary and outlook

The HL-LHC brings opportunities to find new physics in the direct searches as well as through SM precision measurements. However, the detectors need to be ready to perform well in the intense radiation and high pileup conditions to make full use of HL-LHC collision data. The current endcap ECAL and HCAL calorimeters will be replaced by a highly granular and radiation tolerant calorimeter, the HGCAL, as a part of CMS detector upgrade program for Phase-2. The HGCAL is a 5D imaging calorimeter that will provide position, energy, and time measurements, thus helping in efficient *particle-flow*, effective shower separation, particle identification, and pileup mitigation in the harsh environment of HL-LHC while maintaining good physics performance to $e^{\pm}/\gamma$ and *jets*. The detector upgrade studies are being carried out thoroughly in the laboratory-based test benches as well as with the beams of single particles, as discussed in this thesis.

Chapter 5 presents the system test of the most basic unit of the HGCAL detector, the silicon sensor module, and its electronics in the laboratory at CERN. The measurements shown in the chapter for the 6-inch sensor module will be performed for each of the 28000 sensor modules to be used in the final HGCAL detector. The system tests are useful in ensuring the quality of sensors being manufactured. This chapter also gives insight into the working of silicon sensors and the data-acquisition system.

In October 2018, the beam test experiment was performed at H2 beamline at CERN, using silicon sensor-based CMS HGCAL prototype and scintillator+SiPM based CALICE Analogue Hadron Calorimeter (AHCAL) prototype. This was the first large-scale experiment performed by the CMS HGCAL collaboration, where more than 90 modules were employed. The combined prototype was exposed to beams of e^+, π^-, and μ^- of momenta ranging from 20 to 300 GeV. A detailed analysis of the alignment correction, channel-to-channel response equalization, signal-to-noise ratio, and the performance of the combined HGCAL and AHCAL prototype to hadronic shower is presented.

Chapter 6 gives an overview of all the various steps required to prepare the data for final physics measurements. This involves alignment corrections, channel-to-channel gain equalization, and signal-to-noise assessment. The signal-to-noise study shows that the silicon sensor prototypes are in line with the expected performance.

This thesis reports the first performance study of the HGCAL prototype to hadronic showers in the actual particle beam. Chapter 7 presents a detailed performance analysis of the HGCAL and AHCAL combined prototype to hadronic showers. The analysis uses the pion data of 20 to 300 GeV momenta collected during the October 2018 beam test experiment. The data is also compared with GEANT4 based simulation with two hadronic physics lists, namely FTFP_BERT_EMN and QGS_PFTFP_BERT_EMN.

A thorough data cleaning is performed to obtain a relatively pure set of pion data. The longitudinal and transverse segmentation of the HGCAL prototype allows us to develop an algorithm that identifies the layer at which pion stats to shower. The algorithm is optimized using truth information extracted from the GEANT4 simulation. The algorithm's performance in terms of efficiency is found to be $\geq$ 90% (95%) in CE-E (CE-H). The shower start finder algorithm is used to segregate events based on the shower starting location, helping determine the MIP-to-GeV fixed detector energy scale. The MIP-to-GeV energy scale for CE-E is obtained using 50 GeV e^+ data, and for CE-H+AHCAL is obtained using 50 GeV pions that are MIPs in CE-E. The fixed weights are used for pion shower energy reconstruction and allow us to disentangle detector effects. The pion energy response is found to be non-linear, thus pointing to the fact that the HGCAL is a non-compensating calorimeter. The simulation is observed to reproduce the non-linearity well. However, the energy scale is found to be consistently higher than that of data and is fixed by applying a global energy scale factor.

A χ^2-minimization based technique is presented in this thesis to fix the non-linearity in the energy response by applying energy-dependent weights. The energy response of the combined detector prototypes in data is found to be linear within a few percent after applying the energy-dependent weights and shows good agreement with simulation. The energy resolution is obtained in data and compared with simulation. The stochastic term is found to be 131.7 $\pm$ 1.0 % (122.1 $\pm$ 1.4%) for pions that start showering in CE-E (MIPs in CE-E). The constant term is found to be 8.5 $\pm$ 0.1 % (9.0 $\pm$0.1%) for pions that start showering in CE-E (MIPs in CE-E). The energy resolution in simulation with both the physics lists shows agreement within 8-10% with data.

The validity of the energy-dependent weight method is checked considering the fact that the track measurement will not be available for neutral hadrons, or for hadrons outside the tracker coverage region. The calorimetric energy measurement, using fixed weights, is used to estimate energy-dependent weights. The performance of the energy-dependent weight method is found to be reasonable for cases where calorimetric energy measurement is used as the reference for weight extraction and application.

The CE-E and the last three layers of CE-H were instrumented with only one silicon sensor module per layer, resulting in significant hadronic shower leakage in the transverse direction. A simulation-based study is presented with extended geometry of the detector prototype, i.e., all layers instrumented with seven silicon modules arranged in a daisy-like structure. The study shows that with extended geometry, a 10-15% improvement is observed in both stochastic as well as constant terms in energy resolution. This indicates that with better shower containment, the detector's performance is guaranteed to be improved.

The hadronic shower development in the detector in the longitudinal and transverse direction is also presented. The simulation with both the physics lists shows good agreement in the shower development in the HGCAL+AHCAL detector prototype except very close to the shower axis. The discrepancy is observed to be $\sim$ 20%. This indicates that the hadronic shower physics modelling in simulation can benefit from further tuning. One of the most important observations made in this analysis is that fine granularity and detector design of CE-E make it possible to probe the electromagnetic component of early shower pions. This observation is verified with the particle level simulation information. The CE-E information can be exploited with the help of more sophisticated machine-learning-based techniques to improve detector performance to hadronic showers by compensating event-to-event fluctuation in the electromagnetic component. Thus, by extension, the jet performance will be

improved during the HL-LHC operation.

In summary, this thesis presents a detailed description of system tests for the silicon sensor module, and its associated electronics components relevant for the energy reconstruction. It presents the first measurements of the performance of the HGCAL prototype to hadronic showers in the real beam of charged pions. The analyses results presented here provide a strong foundation for developing advanced machine learning tools for hadron energy measurements exploiting the high granularity aspects of the CMS HGCAL as well as for further tuning of GEANT4 based hadronic shower simulation models to be used with the HL-LHC collision data.

Appendix A

Event displays

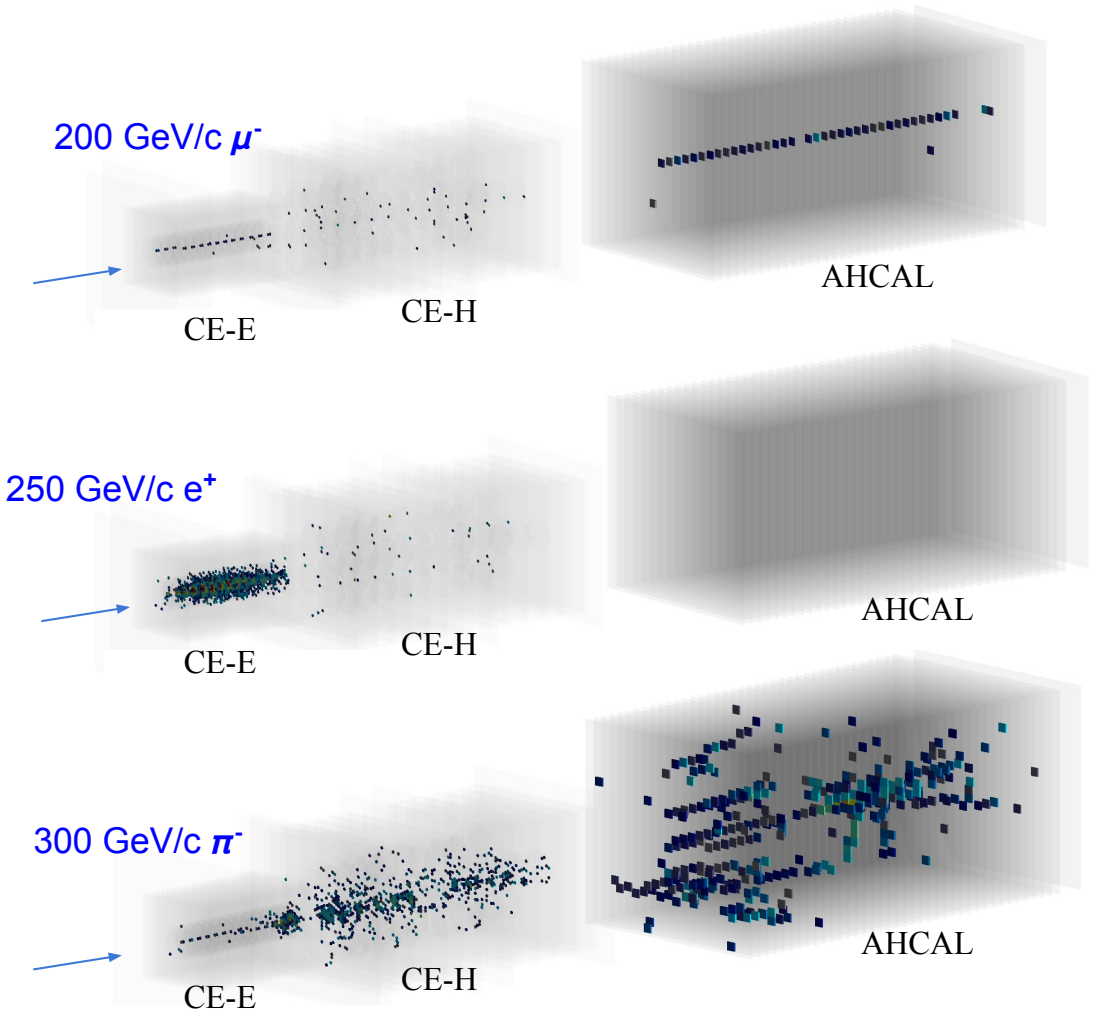

FIGURE A.1: Event display of muon beam (top), positron beam (middle), and charged pion beam (bottom) event in the HGCAL + AHCAL test beam detector setup.

Appendix B

Particle Flow charged hadron calibration

Particle-flow algorithm is used for object and event reconstruction in the CMS detector. The hadrons are reconstructed by combining the information of energy clusters in ECAL and HCAL. Both the calorimeters, being non-compensating, show non-linear energy response to hadronic showers. In order to correctly estimate the energy of the hadrons, energy dependent calibration factors are derived using a simulated flat energy sample (from 2 to 500 GeV) of charged pions. Since the ECAL is $\sim 1.2\ \lambda_{int}$ deep therefore about $\sim 65\%$ hadrons start showering in ECAL have non-zero cluster energy in ECAL. Based on the energy deposited in ECAL, the sample is divided into two categories:

- **EH hadrons:** pions that start showering in ECAL, i.e. pions with non-zero ECAL cluster energy
- **H hadrons:** pions that start showering in HCAL, i.e. pions with zero ECAL cluster energy

Figure B.1 shows the *raw* response distribution, defined as $((E_{ECAL} + E_{HCAL}) - E_{true})/E_{true}$, for EH- and H-hadrons in the true energy bins of 8-12 and 48-56 GeV. The energy response for low energy hadrons show substantially lower response, indicating to the fact that ECAL and HCAL is under-compensating. The energy response and resolution as a function of true hadron energy is obtained by fitting a Gaussian function to the energy response distributions. Figure B.2 shows the *raw* energy response and resolution, as a function of incident energy for EH hadrons and H hadrons. The energy response for both hadrons show non-linear behaviour, a characteristics of hadronic shower in non-compensating calorimeter. By construction, the energy response for 50 GeV H-hadron is at zero because the HCAL is calibrated using 50 GeV pions that do not initiate showering in ECAL.

The response is corrected in two steps: (a) energy-dependent correction, and (b) η dependent correction. The correction (a) corresponds to the inherent non-linearity of hadronic shower. The correction (b) corresponds to the detector effect, i.e. longer path length at larger η hence high dead material traversed by the hadrons. The corrected energy is defined as follows:

$$E^{EH}_{corr} = \alpha(E,\eta)\cdot a(E)\cdot E_{ECAL} + \beta(E,\eta)\cdot b(E)\cdot E_{HCAL} + o_{EH} \quad \text{,for EH hadrons} \tag{B.1}$$

$$E^{H}_{corr} = \gamma(E,\eta)\cdot c(E)\cdot E_{HCAL} + o_{H} \quad \text{, for H hadrons} \tag{B.2}$$

where $a(E)$, $b(E)$ and $c(E)$ are energy dependent calibration and $\alpha(E,\eta)$, $\beta(E,\eta)$ and $\gamma(E,\eta)$ are η dependent calibrations. o_{EH}, o_H are offsets with values = 3.5 GeV and 2.5 GeV, respectively. The offsets are used to take care of various threshold effects applied such as

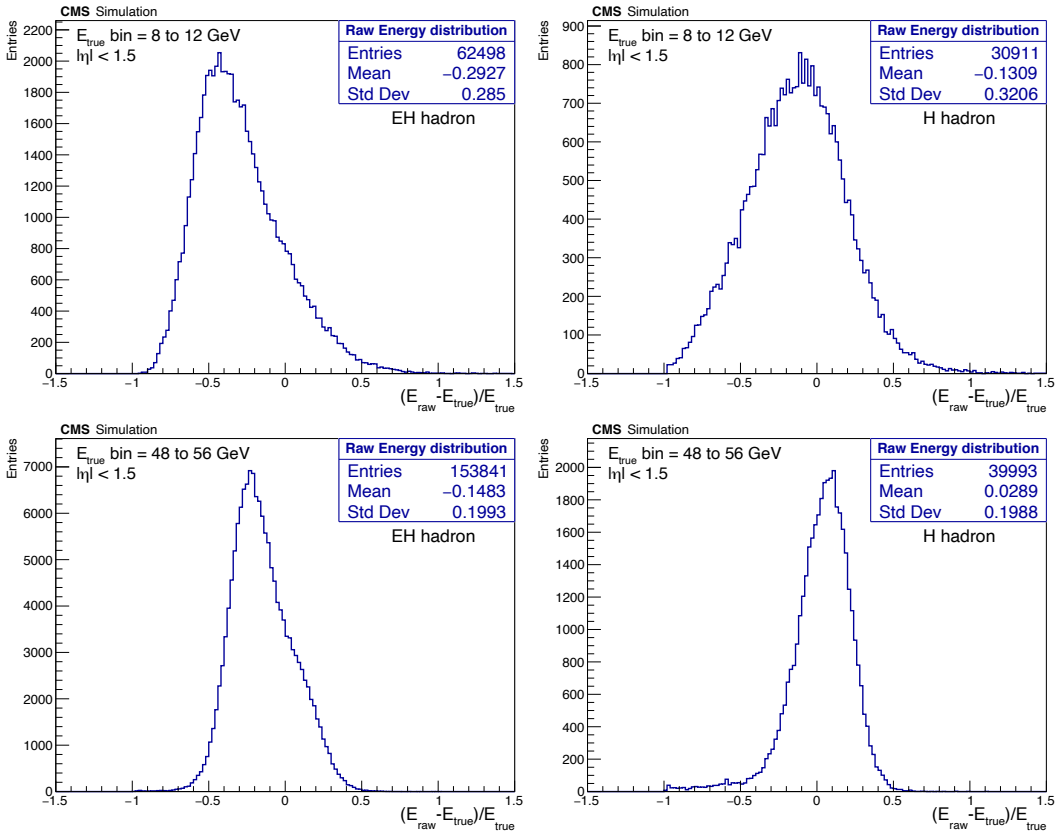

FIGURE B.1: The raw energy response distribution for of EH (left column) and H-hadrons (right column) in the true energy bin of 8-12 GeV (top row) and 48-56 GeV (bottom row) in the barrel region.

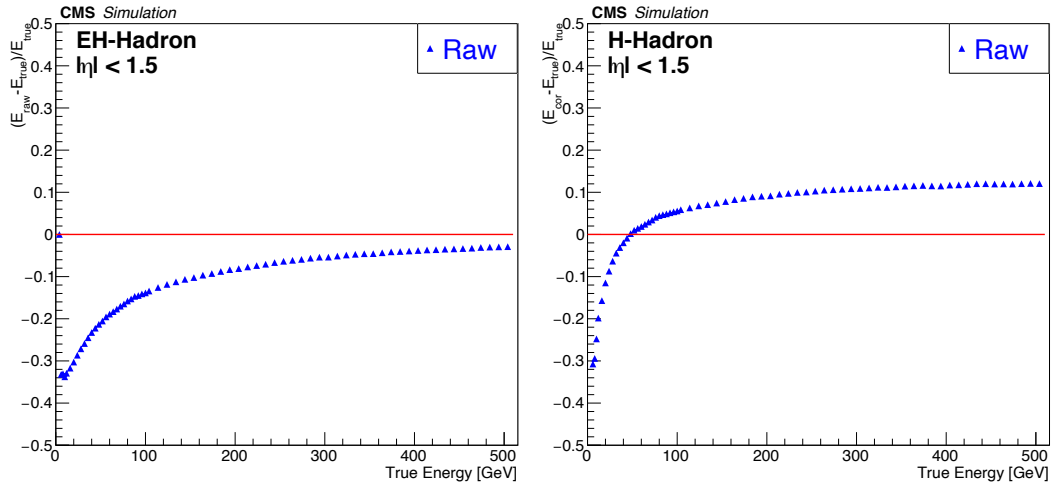

FIGURE B.2: The raw response of EH (left) and H-hadrons (right) in the barrel region.

zero-suppression in ECAL and HCAL and seed thresholds at PF cluster reconstruction. For $\alpha(E,\eta)$, $\beta(E,\eta)$ and $\gamma(E,\eta)$, the *eta* dependence is a polynomial parametrization obtained separately for barrel, endcap region.

The energy dependent calibration parameters are obtained using χ^2-based minimization technique and setting α, β and γ to one,. In this technique a χ^2 is constructed as shown in equation B.3 below, for barrel and endcap region and for EH and H hadrons, separately. The endcap is further divided according to inside and outside tracker coverage region separately.

$$\chi^2 = \sum_{hadrons} \frac{(E_{corr} - E_{true})^2}{\sigma^2(E_{ECAL} + E_{HCAL})} \tag{B.3}$$

where σ is the uncertainty in the measured ECAL and HCAL cluster energies. Note that E_{ECAL} is zero for H hadrons.

The calibration parameters $a(E)$, $b(E)$ and $c(E)$ are obtained by minimizing the χ^2 analytically. These values are plugged back in equation B.1 and B.2, and $\alpha(E,\eta)$, $\beta(E,\eta)$ and $\gamma(E,\eta)$ are obtained using same χ^2-based minimization technique. Figure B.3 shows an example of $a(E)$ and $b(E)$ parameters as a function of true energy of incident pion. A function is fitted on the parameters in order to extract a continuous function.

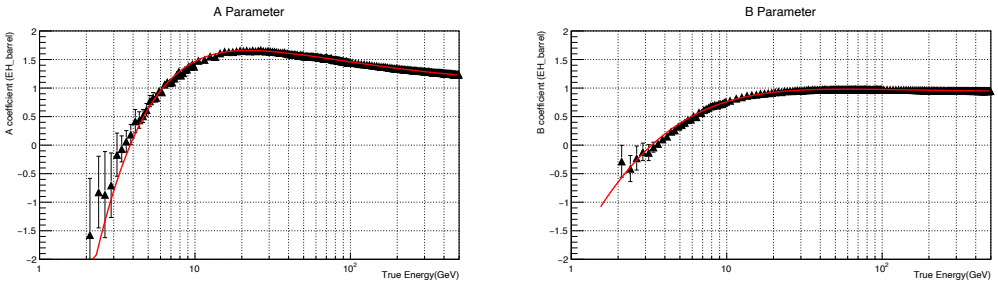

FIGURE B.3: The a (left) and b (right) parameters as function of true energy. The red solid line shows the fitted function.

Using these energy and η dependent calibration parameters, the energy of hadrons are calibrated. Figure B.4 shows the comparison plots for raw and *corrected* response distribution for EH- and H-hadrons in the true energy bins of 8-12 and 48-56 GeV. It is observed that the mean of distribution restores to zero and width gets narrower, especially for EH hadrons. Figure B.5 shows energy response and resolution comparison before and after applying calibration for EH and H hadrons in barrel region. A linear response is obtained for both the categories of hadrons after the application of correction factors. The resolution is also observed to improve, especially for EH hadrons.

In the real experiment track momentum is used as the reference for calibration parameter estimation. Wherever track measurement is not available, such as for neutral hadrons or outside tracker coverage region, raw cluster energy ($E_{ECAL} + E_{HCAL}$) is used as the reference for calibration parameter estimation.

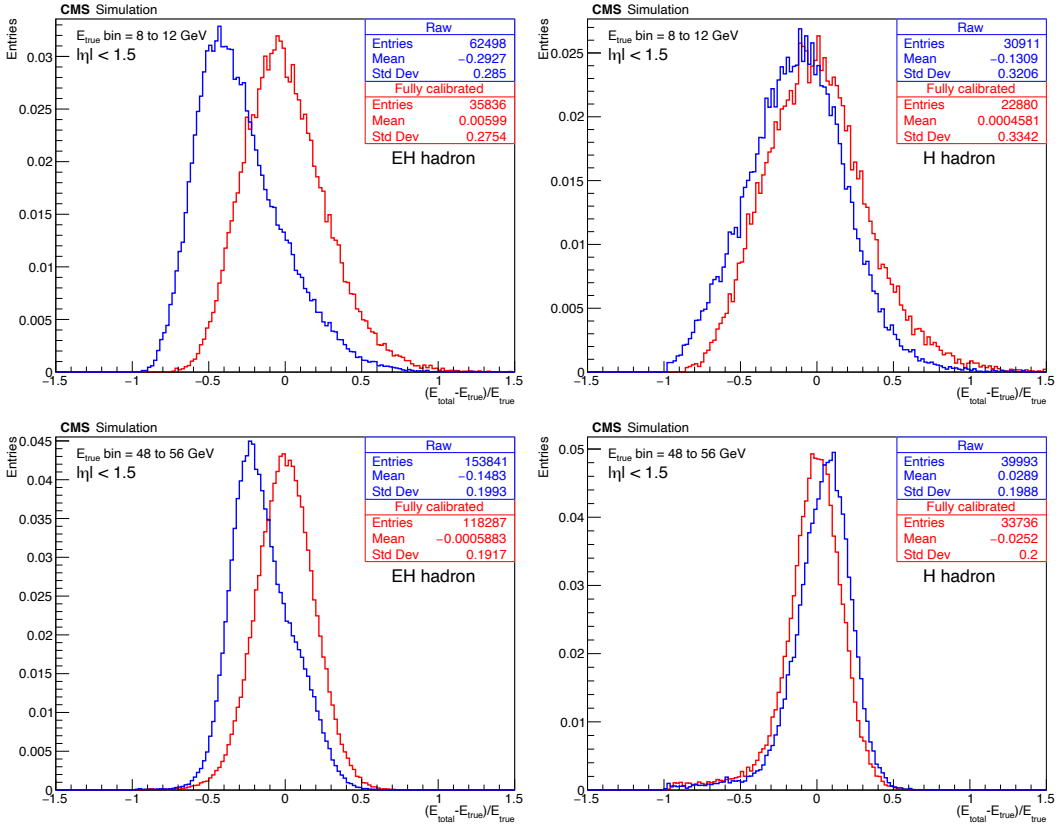

FIGURE B.4: The raw (blue) and corrected (red) energy response distribution comparison for EH (left column) and H-hadrons (right column) in the true energy bin of 8-12 GeV (top row) and 48-56 GeV (bottom row) in the barrel region.

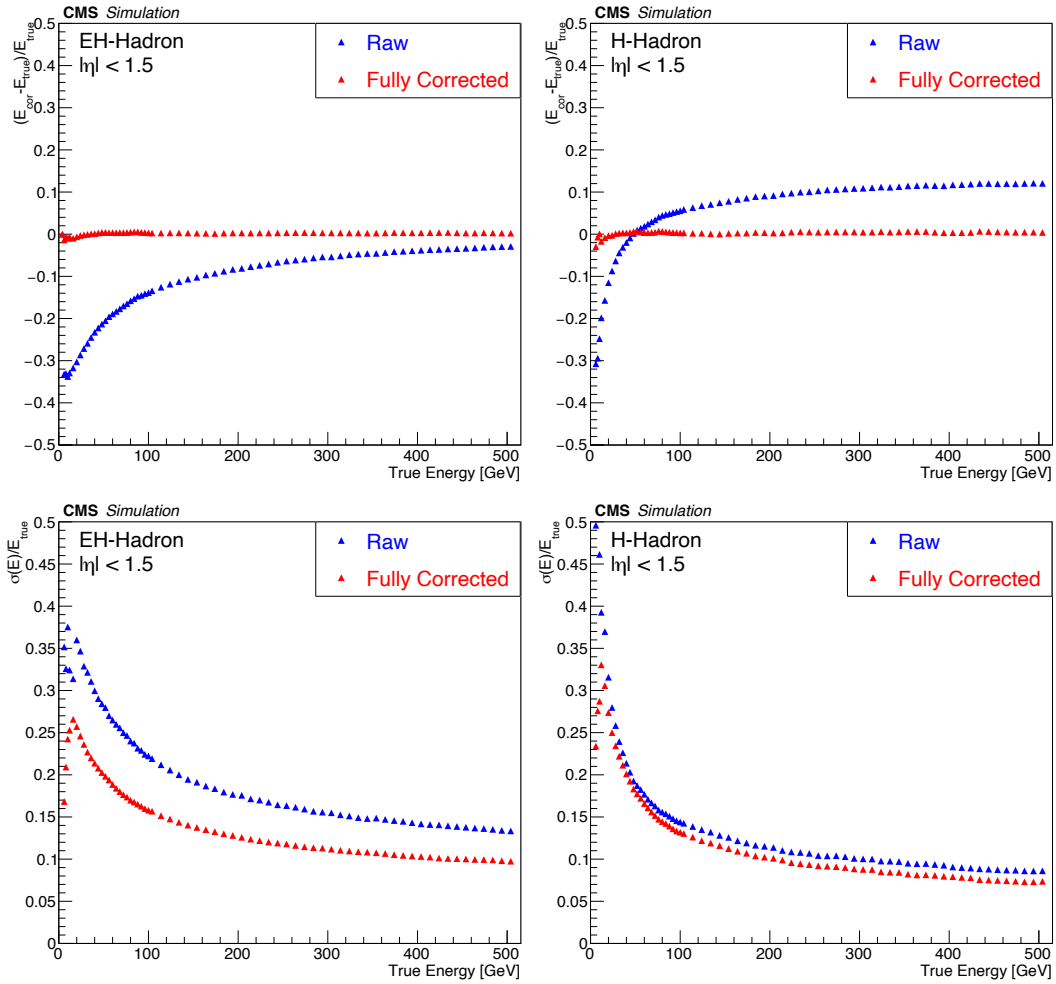

FIGURE B.5: The energy response (top row) and resolution (bottom row) comparison before (blue) and after (red) applying the corrections, as a function of true energy for EH (left column) and H-hadrons (right column) in the barrel region.

Appendix C

Alignment table

Table C.1 shows the alignment correction factors for all layers in CE-E, CE-H and AHCAL as explained in section 6.7. The prescription to correct for misalignment is given in equation 6.11.

Layer	μ_x [cm]	μ_y [cm]
CE-E 1	-3.52	2.73
CE-E 2	-3.62	2.71
CE-E 3	-3.70	2.74
CE-E 4	-3.68	2.70
CE-E 5	-3.53	2.74
CE-E 6	-3.61	2.76
CE-E 7	-3.78	2.71
CE-E 8	-3.78	2.70
CE-E 9	-3.74	2.63
CE-E 10	-3.80	2.64
CE-E 11	-3.82	2.61
CE-E 12	-3.85	2.62
CE-E 13	-3.81	2.65
CE-E 14	-3.89	2.57
CE-E 15	-3.84	2.65
CE-E 16	-3.92	2.62
CE-E 17	-3.85	2.59
CE-E 18	-3.94	2.61
CE-E 19	-3.92	2.53
CE-E 20	-3.96	2.57
CE-E 21	-3.99	2.61
CE-E 22	-4.03	2.63
CE-E 23	-3.97	2.53
CE-E 24	-3.98	2.60
CE-E 25	-3.99	2.59
CE-E 26	-4.02	2.58
CE-E 27	-4.00	2.49
CE-E 28	-3.96	2.50
CE-H 1	-4.94	1.52
CE-H 2	-4.87	1.74
CE-H 3	-4.82	1.75
CE-H 4	-4.95	1.63
CE-H 5	-4.85	1.65
CE-H 6	-4.91	1.65
CE-H 7	-7.17	1.61

CE-H 8	-5.79	1.63
CE-H 9	-6.36	1.53
CE-H 10	-6.18	1.53
CE-H 11	-6.23	1.39
CE-H 12	-6.62	1.29
AHCAL 1	3.11	-0.41
AHCAL 2	2.97	-0.40
AHCAL 3	2.99	-0.41
AHCAL 4	2.99	-0.39
AHCAL 5	3.07	-0.41
AHCAL 6	3.02	-0.40
AHCAL 7	3.08	-0.37
AHCAL 8	3.10	-0.37
AHCAL 9	2.93	-0.37
AHCAL 10	3.06	-0.37
AHCAL 11	3.10	-0.38
AHCAL 12	3.19	-0.40
AHCAL 13	3.04	-0.34
AHCAL 14	2.98	-0.35
AHCAL 15	2.99	-0.33
AHCAL 16	3.02	-0.31
AHCAL 17	3.07	-0.32
AHCAL 18	3.05	-0.36
AHCAL 19	3.08	-0.35
AHCAL 20	3.10	-0.31
AHCAL 21	3.00	-0.31
AHCAL 22	2.92	-0.30
AHCAL 23	3.04	-0.29
AHCAL 24	3.02	-0.24
AHCAL 25	2.98	-0.28
AHCAL 26	3.11	-0.31
AHCAL 27	3.02	-0.30
AHCAL 28	2.98	-0.29
AHCAL 29	2.99	-0.30
AHCAL 30	2.99	-0.30
AHCAL 31	2.99	-0.28
AHCAL 32	2.99	-0.31
AHCAL 33	2.93	-0.33
AHCAL 34	2.99	-0.33
AHCAL 35	3.00	-0.34
AHCAL 36	2.96	-0.36
AHCAL 37	3.01	-0.33
AHCAL 38	3.17	-0.41
AHCAL 39	2.95	-0.36

TABLE C.1

Appendix D

Beam profile distribution comparison between data and simulation

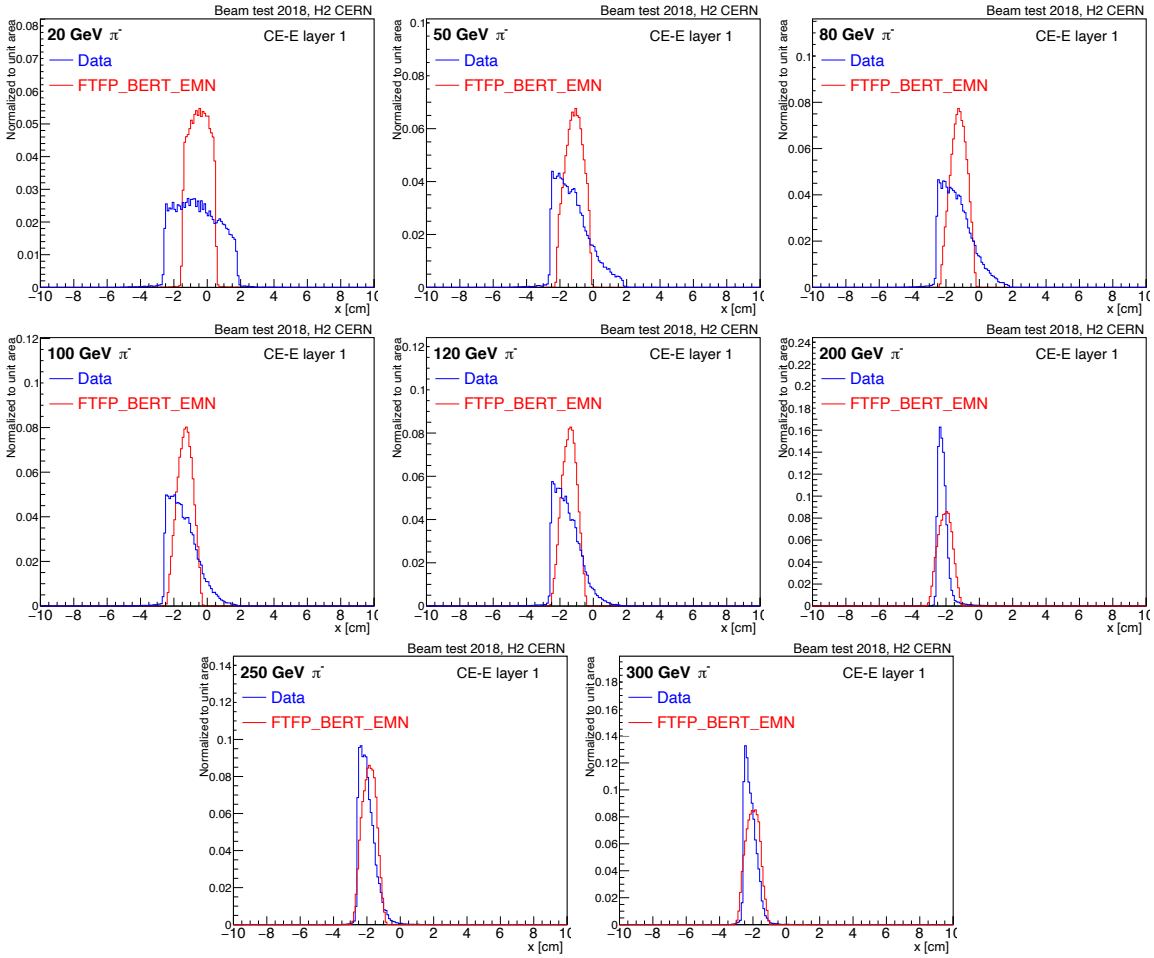

FIGURE D.1: One dimensional track impact position distribution comparisons (corresponding to Figure 7.4) between data and simulation for **x-axis** at CE-E layer 1 for all pion beams.

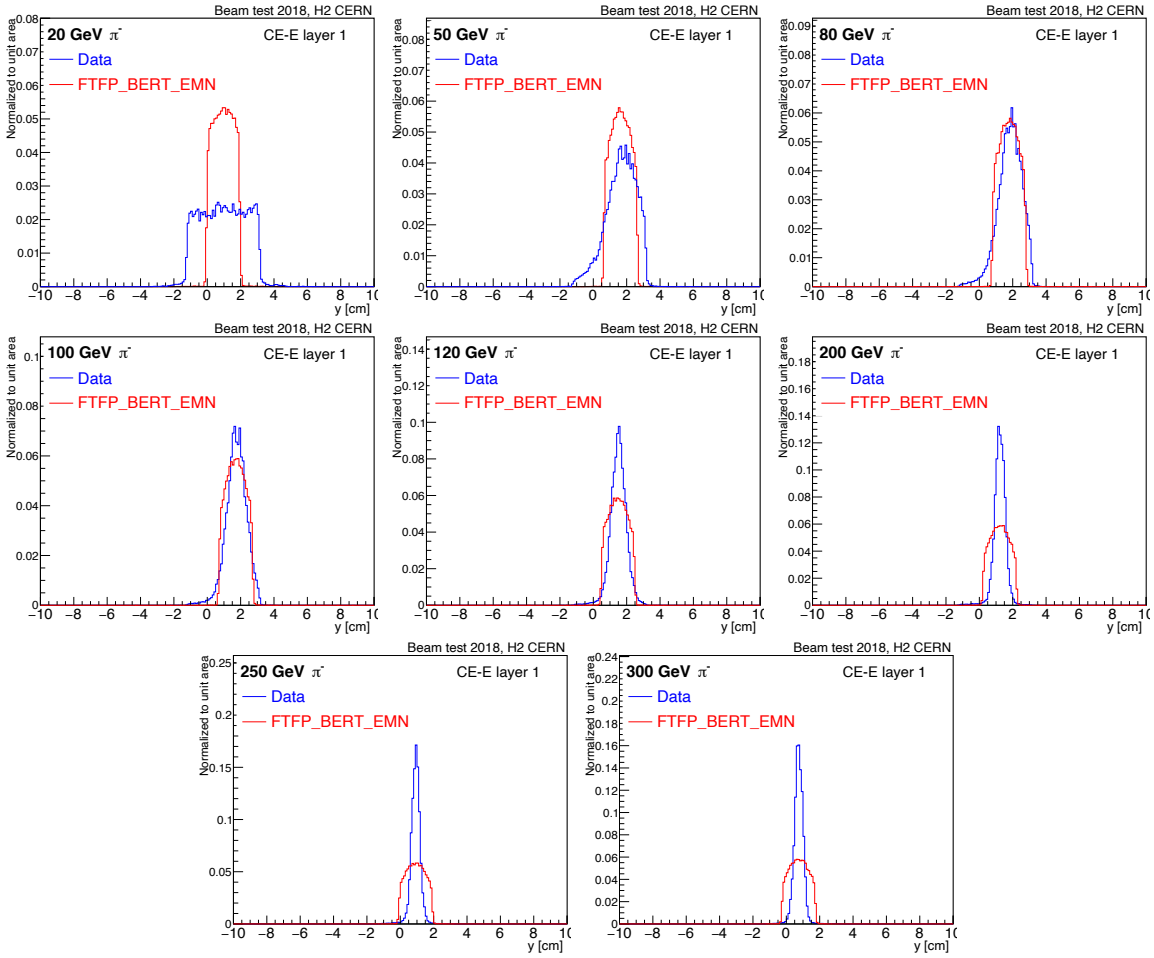

FIGURE D.2: One dimensional track impact position distribution comparisons (corresponding to Figure 7.4) between data and simulation for **y-axis** at CE-E layer 1 for all pion beams.

Appendix E

Noise rejection thresholds

The noise rejection thresholds are optimized by comparing reconstructed energy sum and hit multiplicity, between data and simulation. The digitization step is not implemented in the simulation framework for October 2018 beam test experiment, hence simulation is free of any spurious noise-like hits. Thus simulation is used as the reference for an optimum choice for the threshold to be used in data.

The signal-to-noise ratio studies showed that the SNR is reasonably high for silicon sensor cells in CE-E (SNR $\sim$ 6-7). Therefore, the initial choice for noise-rejection threshold is the standard 3 σ_{noise} and is applied on HG ADC counts. The hit multiplicity and reconstructed energy sum distributions are compared by varying the thresholds, viz. 3, 4 and 5 σ_{noise}. Figure E.1 and E.2 shows such comparisons (with zoomed energy distribution) for 100 GeV pion beam for CE-E and CE-H, respectively. The threshold is applied on both data as well as simulation to make consistent comparisons.

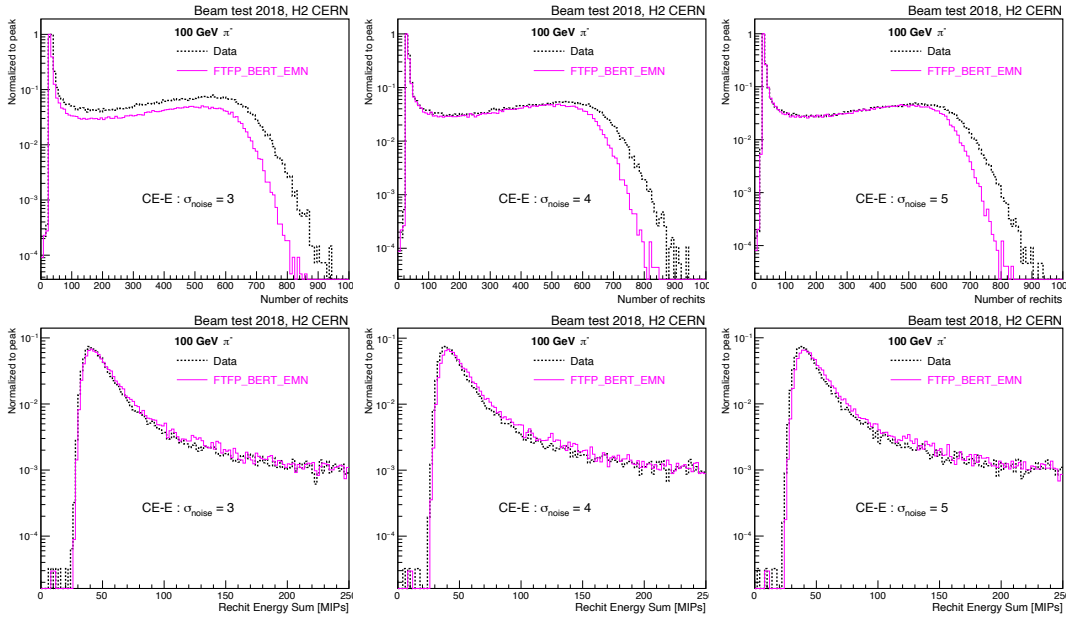

FIGURE E.1: Data-simulation comparison of hit multiplicity (top row) and reconstructed energy (bottom row) distributions for 100 GeV pions in CE-E. The left, middle and right column corresponds to noise-rejection of 3, 4 and 5 σ_{noise} respectively.

It is observed that for CE-H (Figure E.2) the first peak in data which corresponds to MIP-like pions, moves closer to the peak in simulation as more stringent threshold is applied on the HG ADC counts. However, a threshold of 5 σ_{noise} HG ADC counts corresponds to almost 1 MIP energy for few cells in CE-H, therefore *the threshold is chosen as 4* σ_{noise} *HG*

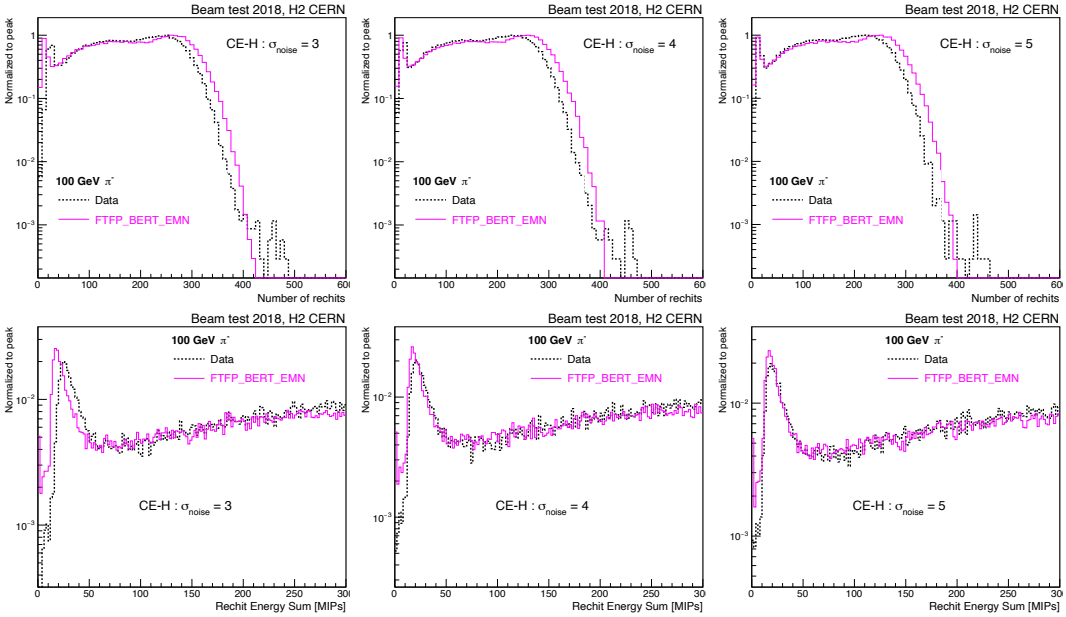

FIGURE E.2: Data-simulation comparison of hit multiplicity (top row) and reconstructed energy (bottom row) distributions for 100 GeV pions in CE-H. The left, middle and right column corresponds to noise-rejection of 3, 4 and 5 σ_{noise} respectively.

ADC counts which shows reasonable performance, for CE-H.

In CE-E (Figure E.1), the MIP-peaks lie almost on top of each other for all thresholds. Therefore, *the standard 3 σ_{noise} HG ADC counts is chosen as noise-rejection threshold for CE-E.*

Fetching the noise value for each cell for each event is a very computationally intensive task. Therefore, taking the advantage of the the observation that the noise level is stable within a chip, a chip average value of σ_{noise} is used as noise-rejection threshold. This threshold is applied on cell-by-cell basis, thereby saving memory and CPU time.

Appendix F

Tables of energy response and resolution plots

TABLE F.1: Data **response** values with **fixed weights** corresponding to Figure 7.27.

Energy [GeV]	CE-E pions	CE-H pions	All pions
20	0.672 ± 0.002	0.922 ± 0.003	0.732 ± 0.001
50	0.786 ± 0.001	1.003 ± 0.002	0.840 ± 0.001
80	0.832 ± 0.001	1.039 ± 0.001	0.889 ± 0.001
100	0.851 ± 0.001	1.053 ± 0.002	0.905 ± 0.001
120	0.868 ± 0.001	1.059 ± 0.002	0.924 ± 0.001
200	0.901 ± 0.001	1.076 ± 0.002	0.951 ± 0.001
250	0.917 ± 0.001	1.087 ± 0.001	0.964 ± 0.001
300	0.924 ± 0.001	1.089 ± 0.001	0.969 ± 0.001

TABLE F.2: Data **resolution** values with **fixed weights** corresponding to Figure 7.27.

Energy [GeV]	CE-E pions	CE-H pions	All pions
20	0.308 ± 0.003	0.279 ± 0.004	0.326 ± 0.003
50	0.217 ± 0.002	0.196 ± 0.003	0.240 ± 0.002
80	0.180 ± 0.001	0.164 ± 0.002	0.207 ± 0.001
100	0.162 ± 0.001	0.158 ± 0.002	0.193 ± 0.001
120	0.154 ± 0.002	0.143 ± 0.002	0.184 ± 0.002
200	0.130 ± 0.001	0.125 ± 0.002	0.160 ± 0.001
250	0.121 ± 0.001	0.118 ± 0.001	0.149 ± 0.001
300	0.114 ± 0.001	0.113 ± 0.001	0.142 ± 0.001

TABLE F.3: Data-MC **response** values with **fixed weights** and MC **unrescaled** corresponding to Figure 7.29.

Energy [GeV]		CE-E pions	CE-H pions	All pions
20	Data	0.672 ± 0.002	0.922 ± 0.003	0.732 ± 0.001
	FTF	0.723 ± 0.001	1.002 ± 0.003	0.781 ± 0.001
	QGS	0.734 ± 0.001	1.023 ± 0.002	0.791 ± 0.001
50	Data	0.786 ± 0.001	1.003 ± 0.002	0.840 ± 0.001
	FTF	0.826 ± 0.001	1.088 ± 0.002	0.887 ± 0.001
	QGS	0.830 ± 0.001	1.103 ± 0.002	0.893 ± 0.001
80	Data	0.832 ± 0.001	1.039 ± 0.001	0.889 ± 0.001
	FTF	0.870 ± 0.001	1.126 ± 0.002	0.929 ± 0.001
	QGS	0.872 ± 0.001	1.126 ± 0.002	0.937 ± 0.001
100	Data	0.851 ± 0.001	1.053 ± 0.002	0.905 ± 0.001
	FTF	0.890 ± 0.001	1.141 ± 0.002	0.954 ± 0.001
	QGS	0.892 ± 0.001	1.140 ± 0.001	0.957 ± 0.001
120	Data	0.868 ± 0.001	1.059 ± 0.002	0.924 ± 0.001
	FTF	0.907 ± 0.001	1.154 ± 0.001	0.970 ± 0.001
	QGS	0.908 ± 0.001	1.151 ± 0.001	0.972 ± 0.001
200	Data	0.901 ± 0.001	1.076 ± 0.002	0.951 ± 0.001
	FTF	0.944 ± 0.001	1.181 ± 0.001	1.009 ± 0.001
	QGS	0.947 ± 0.001	1.172 ± 0.001	1.010 ± 0.001
250	Data	0.917 ± 0.001	1.087 ± 0.001	0.964 ± 0.001
	FTF	0.964 ± 0.001	1.198 ± 0.001	1.028 ± 0.001
	QGS	0.966 ± 0.001	1.181 ± 0.001	1.026 ± 0.001
300	Data	0.924 ± 0.001	1.089 ± 0.001	0.969 ± 0.001
	FTF	0.976 ± 0.001	1.209 ± 0.001	1.041 ± 0.001
	QGS	0.979 ± 0.001	1.192 ± 0.001	1.037 ± 0.001

TABLE F.4: Data-MC **resolution** values with **fixed weights** and MC **unrescaled** corresponding to Figure 7.29.

Energy [GeV]		CE-E pions	CE-H pions	All pions
20	Data	0.308 ± 0.003	0.279 ± 0.004	0.326 ± 0.003
	FTF	0.282 ± 0.002	0.277 ± 0.004	0.302 ± 0.002
	QGS	0.259 ± 0.002	0.265 ± 0.003	0.290 ± 0.002
50	Data	0.217 ± 0.002	0.196 ± 0.003	0.240 ± 0.002
	FTF	0.205 ± 0.002	0.193 ± 0.002	0.237 ± 0.002
	QGS	0.191 ± 0.002	0.189 ± 0.002	0.226 ± 0.001
80	Data	0.180 ± 0.001	0.164 ± 0.002	0.207 ± 0.001
	FTF	0.170 ± 0.001	0.164 ± 0.002	0.207 ± 0.001
	QGS	0.166 ± 0.001	0.160 ± 0.002	0.209 ± 0.001
100	Data	0.162 ± 0.001	0.158 ± 0.002	0.193 ± 0.001
	FTF	0.158 ± 0.001	0.152 ± 0.002	0.198 ± 0.001
	QGS	0.156 ± 0.001	0.148 ± 0.002	0.200 ± 0.001
120	Data	0.154 ± 0.002	0.143 ± 0.002	0.184 ± 0.002
	FTF	0.148 ± 0.001	0.139 ± 0.002	0.190 ± 0.001
	QGS	0.147 ± 0.001	0.141 ± 0.002	0.188 ± 0.001
200	Data	0.130 ± 0.001	0.125 ± 0.002	0.160 ± 0.001
	FTF	0.124 ± 0.001	0.117 ± 0.001	0.174 ± 0.001
	QGS	0.126 ± 0.001	0.118 ± 0.001	0.169 ± 0.001
250	Data	0.121 ± 0.001	0.118 ± 0.001	0.149 ± 0.001
	FTF	0.116 ± 0.001	0.109 ± 0.001	0.171 ± 0.001
	QGS	0.118 ± 0.001	0.109 ± 0.001	0.162 ± 0.001
300	Data	0.114 ± 0.001	0.113 ± 0.001	0.142 ± 0.001
	FTF	0.109 ± 0.001	0.102 ± 0.001	0.170 ± 0.001
	QGS	0.113 ± 0.001	0.104 ± 0.001	0.157 ± 0.001

TABLE F.5: Data-MC **response** values with **fixed weights** and MC **rescaled** corresponding to Figure 7.32.

Energy [GeV]		CE-E pions	CE-H pions	All pions
20	Data	0.672 ± 0.002	0.922 ± 0.003	0.732 ± 0.001
	FTF	0.680 ± 0.001	0.918 ± 0.002	0.731 ± 0.001
	QGS	0.697 ± 0.001	0.935 ± 0.002	0.746 ± 0.001
50	Data	0.786 ± 0.001	1.003 ± 0.002	0.840 ± 0.001
	FTF	0.779 ± 0.001	0.996 ± 0.001	0.832 ± 0.001
	QGS	0.782 ± 0.001	1.007 ± 0.002	0.833 ± 0.001
80	Data	0.832 ± 0.001	1.039 ± 0.001	0.889 ± 0.001
	FTF	0.818 ± 0.001	1.029 ± 0.001	0.868 ± 0.001
	QGS	0.820 ± 0.001	1.028 ± 0.001	0.872 ± 0.001
100	Data	0.851 ± 0.001	1.053 ± 0.002	0.905 ± 0.001
	FTF	0.837 ± 0.001	1.043 ± 0.001	0.887 ± 0.001
	QGS	0.837 ± 0.001	1.041 ± 0.001	0.889 ± 0.001
120	Data	0.868 ± 0.001	1.059 ± 0.002	0.924 ± 0.001
	FTF	0.852 ± 0.001	1.054 ± 0.001	0.902 ± 0.001
	QGS	0.852 ± 0.001	1.052 ± 0.001	0.903 ± 0.001
200	Data	0.901 ± 0.001	1.076 ± 0.002	0.951 ± 0.001
	FTF	0.886 ± 0.001	1.078 ± 0.001	0.935 ± 0.001
	QGS	0.886 ± 0.001	1.071 ± 0.001	0.934 ± 0.001
250	Data	0.917 ± 0.001	1.087 ± 0.001	0.964 ± 0.001
	FTF	0.903 ± 0.001	1.094 ± 0.001	0.953 ± 0.001
	QGS	0.902 ± 0.001	1.078 ± 0.001	0.949 ± 0.001
300	Data	0.924 ± 0.001	1.089 ± 0.001	0.969 ± 0.001
	FTF	0.914 ± 0.001	1.105 ± 0.001	0.964 ± 0.001
	QGS	0.914 ± 0.001	1.088 ± 0.001	0.959 ± 0.001

TABLE F.6: Data-MC **resolution** values with **fixed weights** and MC **rescaled** corresponding to Figure 7.32.

Energy [GeV]		CE-E pions	CE-H pions	All pions
20	Data	0.308 ± 0.003	0.279 ± 0.004	0.326 ± 0.003
	FTF	0.283 ± 0.003	0.278 ± 0.003	0.299 ± 0.002
	QGS	0.253 ± 0.002	0.263 ± 0.003	0.272 ± 0.002
50	Data	0.217 ± 0.002	0.196 ± 0.003	0.240 ± 0.002
	FTF	0.203 ± 0.002	0.190 ± 0.002	0.226 ± 0.001
	QGS	0.190 ± 0.001	0.190 ± 0.002	0.216 ± 0.001
80	Data	0.180 ± 0.001	0.164 ± 0.002	0.207 ± 0.001
	FTF	0.171 ± 0.001	0.163 ± 0.002	0.197 ± 0.001
	QGS	0.164 ± 0.001	0.160 ± 0.002	0.194 ± 0.001
100	Data	0.162 ± 0.001	0.158 ± 0.002	0.193 ± 0.001
	FTF	0.158 ± 0.001	0.151 ± 0.002	0.186 ± 0.001
	QGS	0.153 ± 0.001	0.148 ± 0.002	0.183 ± 0.001
120	Data	0.154 ± 0.002	0.143 ± 0.002	0.184 ± 0.002
	FTF	0.147 ± 0.001	0.140 ± 0.002	0.179 ± 0.001
	QGS	0.143 ± 0.001	0.141 ± 0.002	0.176 ± 0.001
200	Data	0.130 ± 0.001	0.125 ± 0.002	0.160 ± 0.001
	FTF	0.122 ± 0.001	0.117 ± 0.001	0.158 ± 0.001
	QGS	0.122 ± 0.001	0.118 ± 0.001	0.156 ± 0.001
250	Data	0.121 ± 0.001	0.118 ± 0.001	0.149 ± 0.001
	FTF	0.114 ± 0.001	0.109 ± 0.001	0.151 ± 0.001
	QGS	0.113 ± 0.001	0.108 ± 0.001	0.147 ± 0.001
300	Data	0.114 ± 0.001	0.113 ± 0.001	0.142 ± 0.001
	FTF	0.106 ± 0.001	0.102 ± 0.001	0.148 ± 0.001
	QGS	0.108 ± 0.001	0.104 ± 0.001	0.141 ± 0.001

TABLE F.7: Data **response** values with **energy dependent weights** corresponding to Figure 7.36.

Energy [GeV]	CE-E pions	CE-H pions	All pions
20	0.986 ± 0.002	0.993 ± 0.003	0.989 ± 0.002
50	0.999 ± 0.002	0.998 ± 0.002	0.999 ± 0.001
80	1.002 ± 0.001	1.004 ± 0.001	1.003 ± 0.001
100	1.004 ± 0.001	1.006 ± 0.001	1.004 ± 0.001
120	1.004 ± 0.001	1.005 ± 0.001	1.004 ± 0.001
200	1.006 ± 0.001	1.004 ± 0.002	1.006 ± 0.001
250	1.005 ± 0.001	1.005 ± 0.001	1.005 ± 0.001
300	1.007 ± 0.001	1.005 ± 0.001	1.006 ± 0.001

TABLE F.8: Data **resolution** values with **energy dependent weights** corresponding to Figure 7.36.

Energy [GeV]	CE-E pions	CE-H pions	All pions
20	0.290 ± 0.003	0.275 ± 0.004	0.285 ± 0.002
50	0.204 ± 0.002	0.194 ± 0.003	0.203 ± 0.002
80	0.175 ± 0.001	0.166 ± 0.002	0.172 ± 0.001
100	0.159 ± 0.001	0.160 ± 0.002	0.159 ± 0.001
120	0.150 ± 0.001	0.142 ± 0.002	0.148 ± 0.001
200	0.129 ± 0.001	0.126 ± 0.002	0.128 ± 0.001
250	0.117 ± 0.001	0.117 ± 0.001	0.117 ± 0.001
300	0.112 ± 0.001	0.112 ± 0.001	0.112 ± 0.001

TABLE F.9: Data-MC **response** values with **energy dependent weights** and MC **rescaled** corresponding to Figure 7.38.

Energy [GeV]		CE-E pions	CE-H pions	All pions
20	Data	0.986 ± 0.002	0.993 ± 0.003	0.989 ± 0.002
	FTF	1.004 ± 0.002	0.988 ± 0.002	0.998 ± 0.001
	QGS	1.037 ± 0.002	1.003 ± 0.002	1.030 ± 0.001
50	Data	0.999 ± 0.002	0.998 ± 0.002	0.999 ± 0.001
	FTF	0.996 ± 0.001	0.995 ± 0.002	0.995 ± 0.001
	QGS	1.007 ± 0.001	1.004 ± 0.002	1.006 ± 0.001
80	Data	1.002 ± 0.001	1.004 ± 0.001	1.003 ± 0.001
	FTF	0.989 ± 0.001	0.996 ± 0.001	0.992 ± 0.001
	QGS	0.995 ± 0.001	0.996 ± 0.001	0.996 ± 0.001
100	Data	1.004 ± 0.001	1.006 ± 0.001	1.004 ± 0.001
	FTF	0.992 ± 0.001	0.995 ± 0.001	0.993 ± 0.001
	QGS	0.992 ± 0.001	0.994 ± 0.001	0.993 ± 0.001
120	Data	1.004 ± 0.001	1.005 ± 0.001	1.004 ± 0.001
	FTF	0.992 ± 0.001	1.001 ± 0.001	0.995 ± 0.001
	QGS	0.992 ± 0.001	1.001 ± 0.001	0.995 ± 0.001
200	Data	1.006 ± 0.001	1.004 ± 0.002	1.006 ± 0.001
	FTF	0.996 ± 0.001	1.008 ± 0.001	0.999 ± 0.001
	QGS	0.992 ± 0.001	1.000 ± 0.001	0.995 ± 0.001
250	Data	1.005 ± 0.001	1.005 ± 0.001	1.005 ± 0.001
	FTF	1.000 ± 0.001	1.011 ± 0.001	1.004 ± 0.001
	QGS	0.994 ± 0.001	1.001 ± 0.001	0.997 ± 0.001
300	Data	1.007 ± 0.001	1.005 ± 0.001	1.006 ± 0.001
	FTF	1.004 ± 0.001	1.019 ± 0.001	1.009 ± 0.001
	QGS	0.999 ± 0.001	1.008 ± 0.001	1.002 ± 0.001

TABLE F.10: Data-MC **resolution** values with **energy dependent weights** and MC **rescaled** corresponding to Figure 7.38.

Energy [GeV]		CE-E pions	CE-H pions	All pions
20	Data	0.290 ± 0.003	0.275 ± 0.004	0.285 ± 0.002
	FTF	0.271 ± 0.002	0.276 ± 0.003	0.275 ± 0.002
	QGS	0.261 ± 0.002	0.267 ± 0.003	0.258 ± 0.001
50	Data	0.204 ± 0.002	0.194 ± 0.003	0.203 ± 0.002
	FTF	0.194 ± 0.002	0.196 ± 0.002	0.195 ± 0.001
	QGS	0.185 ± 0.001	0.192 ± 0.002	0.188 ± 0.001
80	Data	0.175 ± 0.001	0.166 ± 0.002	0.172 ± 0.001
	FTF	0.163 ± 0.001	0.168 ± 0.002	0.165 ± 0.001
	QGS	0.161 ± 0.001	0.161 ± 0.002	0.161 ± 0.001
100	Data	0.159 ± 0.001	0.160 ± 0.002	0.159 ± 0.001
	FTF	0.153 ± 0.001	0.154 ± 0.002	0.153 ± 0.001
	QGS	0.149 ± 0.001	0.149 ± 0.002	0.150 ± 0.001
120	Data	0.150 ± 0.001	0.142 ± 0.002	0.148 ± 0.001
	FTF	0.141 ± 0.001	0.140 ± 0.002	0.142 ± 0.001
	QGS	0.141 ± 0.001	0.141 ± 0.001	0.142 ± 0.001
200	Data	0.129 ± 0.001	0.126 ± 0.002	0.128 ± 0.001
	FTF	0.117 ± 0.001	0.119 ± 0.001	0.118 ± 0.001
	QGS	0.120 ± 0.001	0.118 ± 0.001	0.119 ± 0.001
250	Data	0.117 ± 0.001	0.117 ± 0.001	0.117 ± 0.001
	FTF	0.109 ± 0.001	0.109 ± 0.001	0.110 ± 0.001
	QGS	0.111 ± 0.001	0.111 ± 0.001	0.111 ± 0.001
300	Data	0.112 ± 0.001	0.112 ± 0.001	0.112 ± 0.001
	FTF	0.102 ± 0.001	0.100 ± 0.001	0.102 ± 0.001
	QGS	0.105 ± 0.001	0.105 ± 0.001	0.105 ± 0.001

TABLE F.11: Data **response** values with E_{beam} **and** E_{reco} as reference corresponding to Figure 7.40.

Energy [GeV]		CE-E pions	CE-H pions	All pions
20	Ebeam	0.986 ± 0.002	0.993 ± 0.003	0.989 ± 0.002
	Ereco	1.067 ± 0.002	0.999 ± 0.003	1.047 ± 0.002
50	Ebeam	0.999 ± 0.002	0.998 ± 0.002	0.999 ± 0.001
	Ereco	1.026 ± 0.002	0.997 ± 0.002	1.018 ± 0.001
80	Ebeam	1.002 ± 0.001	1.004 ± 0.001	1.003 ± 0.001
	Ereco	1.020 ± 0.001	1.005 ± 0.001	1.014 ± 0.001
100	Ebeam	1.004 ± 0.001	1.006 ± 0.001	1.004 ± 0.001
	Ereco	1.017 ± 0.001	1.008 ± 0.001	1.014 ± 0.001
120	Ebeam	1.004 ± 0.001	1.005 ± 0.001	1.004 ± 0.001
	Ereco	1.016 ± 0.001	1.004 ± 0.001	1.012 ± 0.001
200	Ebeam	1.006 ± 0.001	1.004 ± 0.002	1.006 ± 0.001
	Ereco	1.012 ± 0.001	1.002 ± 0.002	1.009 ± 0.001
250	Ebeam	1.005 ± 0.001	1.005 ± 0.001	1.005 ± 0.001
	Ereco	1.013 ± 0.001	1.004 ± 0.001	1.010 ± 0.001
300	Ebeam	1.007 ± 0.001	1.005 ± 0.001	1.006 ± 0.001
	Ereco	1.010 ± 0.001	1.001 ± 0.001	1.007 ± 0.001

TABLE F.12: Data **resolution** values with E_{beam} **and** E_{reco} as reference corresponding to Figure 7.40.

Energy [GeV]		CE-E pions	CE-H pions	All pions
20	Ebeam	0.290 ± 0.003	0.275 ± 0.004	0.285 ± 0.002
	Ereco	0.235 ± 0.002	0.246 ± 0.004	0.244 ± 0.002
50	Ebeam	0.204 ± 0.002	0.194 ± 0.003	0.203 ± 0.002
	Ereco	0.180 ± 0.002	0.181 ± 0.003	0.183 ± 0.002
80	Ebeam	0.175 ± 0.001	0.166 ± 0.002	0.172 ± 0.001
	Ereco	0.158 ± 0.001	0.157 ± 0.002	0.158 ± 0.001
100	Ebeam	0.159 ± 0.001	0.160 ± 0.002	0.159 ± 0.001
	Ereco	0.146 ± 0.001	0.154 ± 0.002	0.148 ± 0.001
120	Ebeam	0.150 ± 0.001	0.142 ± 0.002	0.148 ± 0.001
	Ereco	0.137 ± 0.001	0.135 ± 0.002	0.137 ± 0.001
200	Ebeam	0.129 ± 0.001	0.126 ± 0.002	0.128 ± 0.001
	Ereco	0.121 ± 0.001	0.122 ± 0.002	0.121 ± 0.001
250	Ebeam	0.117 ± 0.001	0.117 ± 0.001	0.117 ± 0.001
	Ereco	0.111 ± 0.001	0.113 ± 0.001	0.112 ± 0.001
300	Ebeam	0.112 ± 0.001	0.112 ± 0.001	0.112 ± 0.001
	Ereco	0.106 ± 0.001	0.109 ± 0.001	0.107 ± 0.001

TABLE F.13: Resolution values for **default and larger TB geometry** implemented in simulation (FTFP_BERT_EMN) corresponding to Figure 7.42.

Energy [GeV]		CE-E pions	CE-H pions	All pions
20	DefaultGeom	0.271 ± 0.002	0.276 ± 0.003	0.275 ± 0.002
	LargerGeom	0.234 ± 0.002	0.259 ± 0.003	0.243 ± 0.002
50	DefaultGeom	0.194 ± 0.002	0.196 ± 0.002	0.195 ± 0.001
	LargerGeom	0.169 ± 0.001	0.186 ± 0.002	0.174 ± 0.001
80	DefaultGeom	0.163 ± 0.001	0.168 ± 0.002	0.165 ± 0.001
	LargerGeom	0.140 ± 0.001	0.146 ± 0.002	0.144 ± 0.001
100	DefaultGeom	0.153 ± 0.001	0.154 ± 0.002	0.153 ± 0.001
	LargerGeom	0.130 ± 0.001	0.137 ± 0.002	0.133 ± 0.001
120	DefaultGeom	0.141 ± 0.001	0.140 ± 0.002	0.142 ± 0.001
	LargerGeom	0.120 ± 0.001	0.129 ± 0.002	0.124 ± 0.001
200	DefaultGeom	0.117 ± 0.001	0.119 ± 0.001	0.118 ± 0.001
	LargerGeom	0.103 ± 0.001	0.107 ± 0.001	0.104 ± 0.001
250	DefaultGeom	0.109 ± 0.001	0.109 ± 0.001	0.110 ± 0.001
	LargerGeom	0.093 ± 0.001	0.101 ± 0.001	0.096 ± 0.001
300	DefaultGeom	0.102 ± 0.001	0.100 ± 0.001	0.102 ± 0.001
	LargerGeom	0.089 ± 0.001	0.093 ± 0.001	0.090 ± 0.001

Appendix G

Energy resolution plots as a function of $1/\sqrt{E}$

For the energy region of interest (a few GeV to hundred of GeV) in the high energy experimental particle physics, the resolution of calorimetric energy measurement scales as $1/\sqrt{E}$. Therefore, a linearity in resolution versus $1/\sqrt{E}$ plot is another indicator of good calorimetric performance. This appendix shows such plots corresponding to resolution plots presented in chapter 7, and are found to be reasonably linear as a function of $1/\sqrt{E}$.

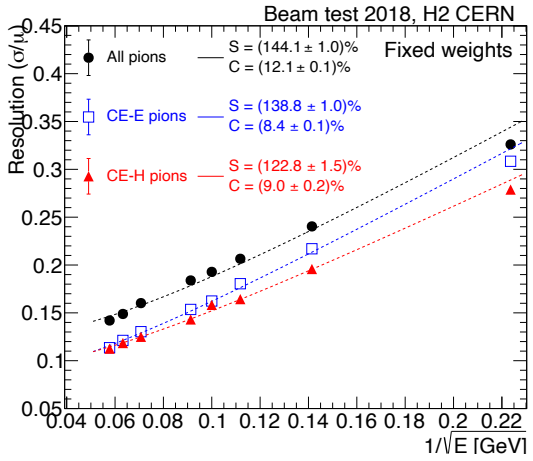

FIGURE G.1: Energy resolution as a function of $1/\sqrt{E_{beam}}$ for CE-E pions (blue open square), CE-H pions (red full triangles) and inclusive in both categories (black full circles) corresponding to Figure 7.27.

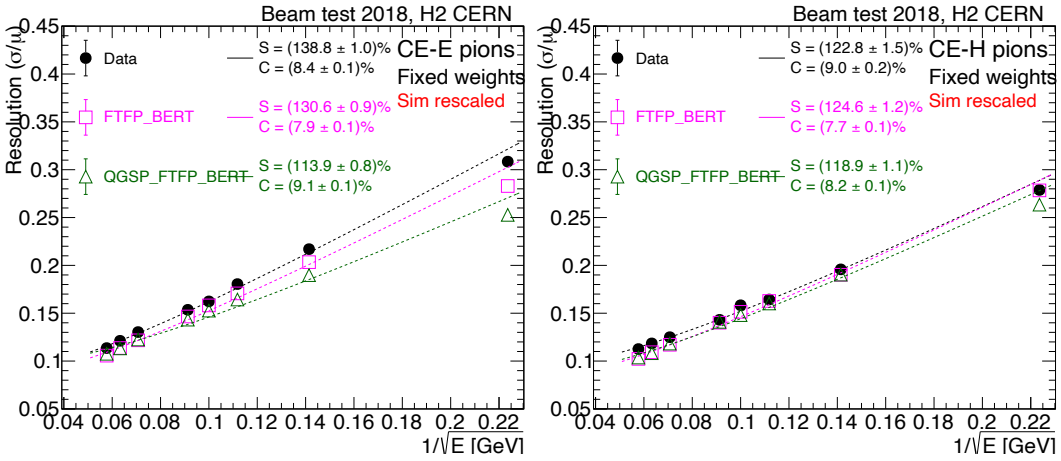

FIGURE G.2: Energy resolution comparison between data and simulation (with two physics lists) as a function of $1/\sqrt{E_{beam}}$ for CE-E pions (left) and CE-H pions (right) after applying global energy correction factor on simulation, corresponding to Figure 7.32.

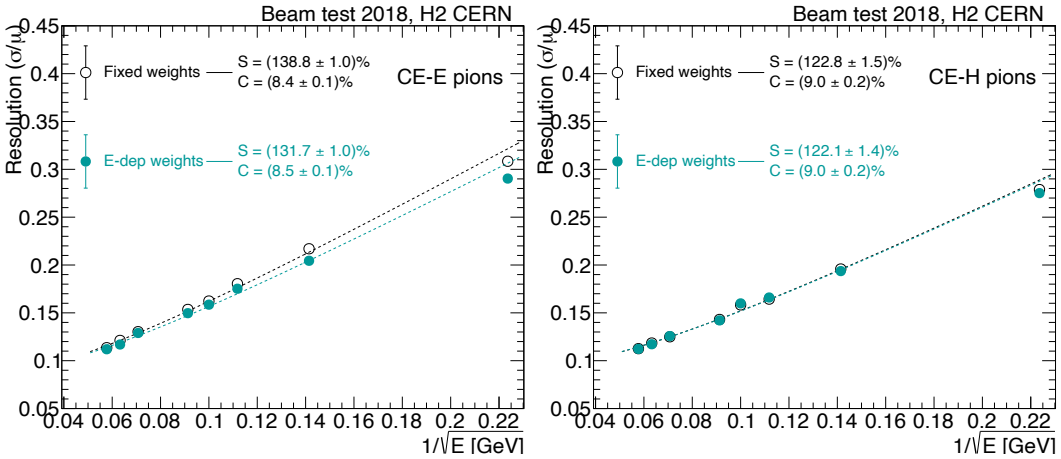

FIGURE G.3: Resolution comparison after applying fixed weights (black open circles) and energy dependent weights (cyan full circles) for CE-E pions (left) and CE-H pions (right) as a function of $1/\sqrt{E_{beam}}$ in data, corresponding to Figure 7.35.

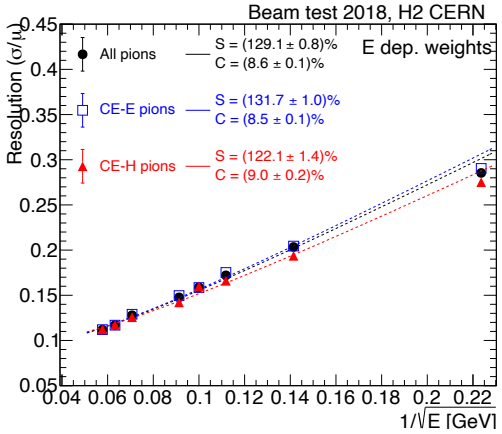

FIGURE G.4: Resolution (right) as a function of $1/\sqrt{E_{beam}}$ for different categories of pions after applying energy dependent weights, corresonding to Figure 7.36.

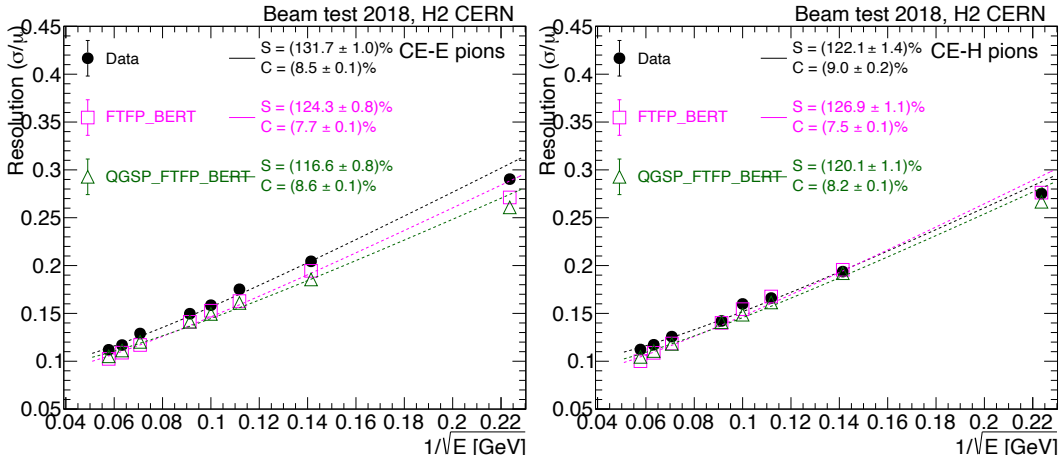

FIGURE G.5: Energy Resolution comparison between data and simulation (with two physics lists) as a function of $1/\sqrt{E_{beam}}$ for CE-E pions (left) and CE-H pions (right) after the application of energy dependent weights, corresponding to Figure 7.38.

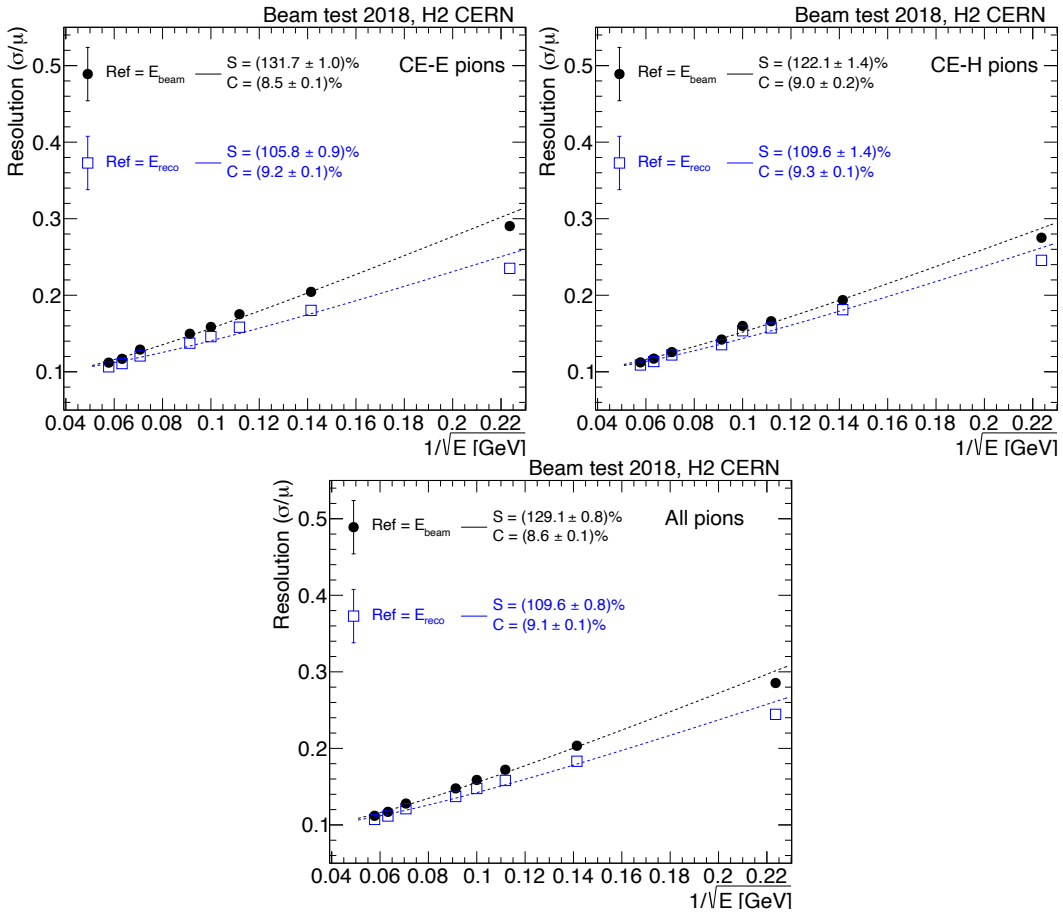

FIGURE G.6: Energy resolution comparison between two references used to extract the weights (black full circle: beam energy, blue open rectangle: reconstructed calorimeter energy) for CE-E pions (left top), CE-H pions (right top) and all pions (bottom), corresponding to Figure 7.40.

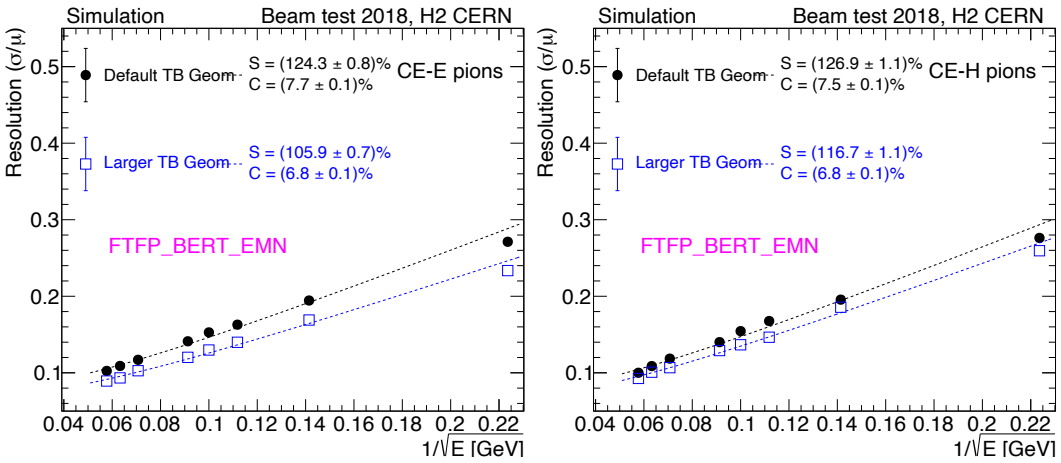

FIGURE G.7: Comparison of energy resolution with default (black full circle) and larger TB geometry (blue open squares) in simulation, corresponding to Figure 7.42.

Appendix H

Energy integrated in circles of different radii in beam test data

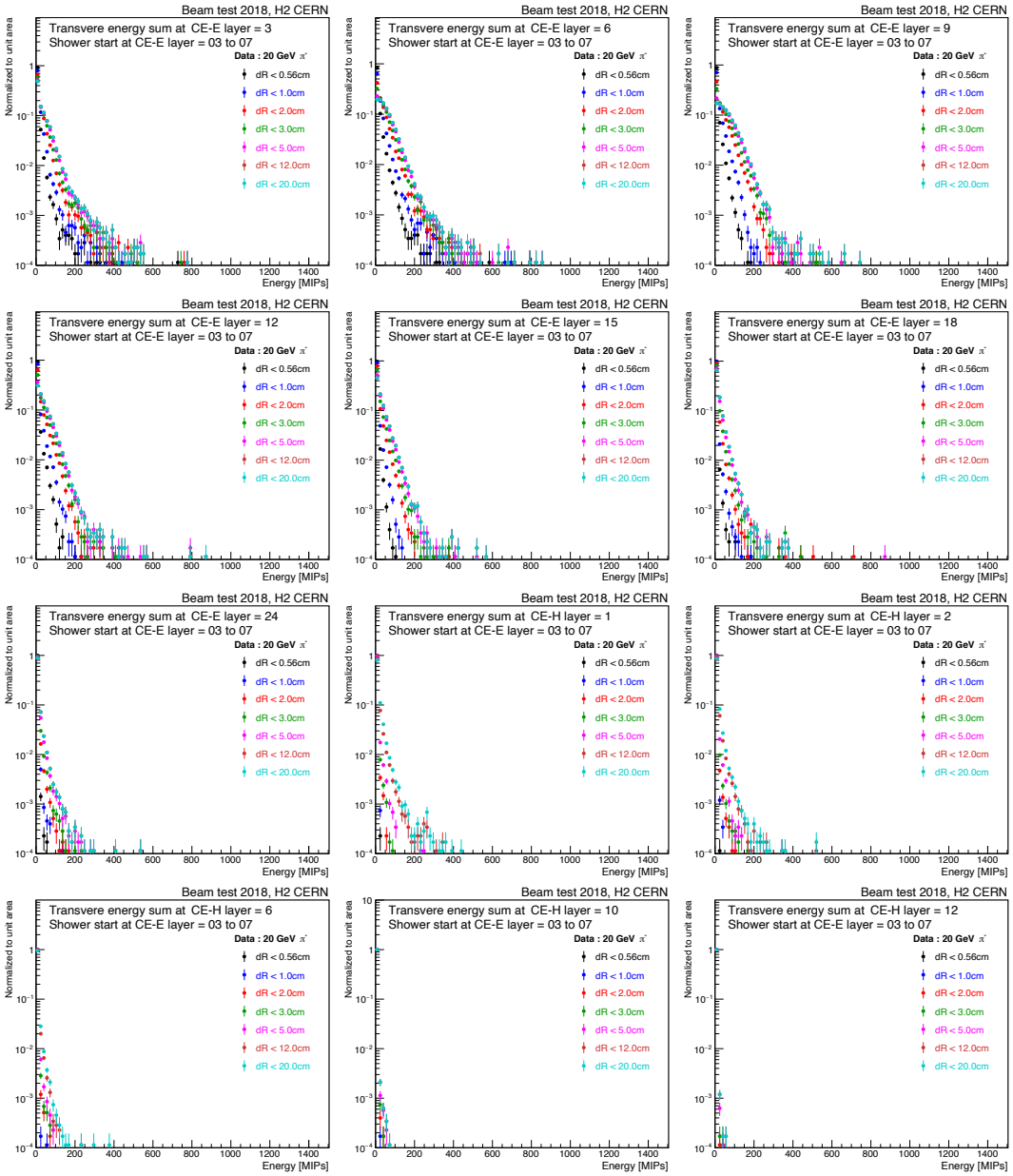

FIGURE H.1: Energy distributions in the circle of different radii (shown in different colour points) from the track impact position at layers downstream to shower start location from **CE-E layer 3 to 7** for **20 GeV/c** pion beam.

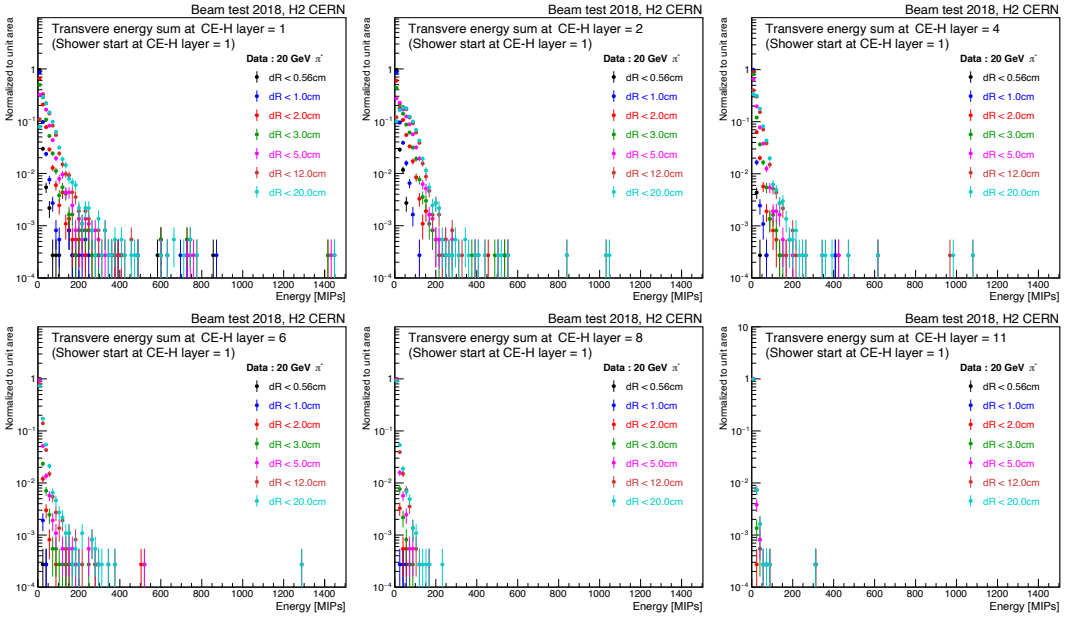

FIGURE H.2: Energy distributions in the circle of different radii (shown in different colour points) from the track impact position at layers downstream to shower start location from **CE-H layer 1** for **20 GeV/c** pion beam.

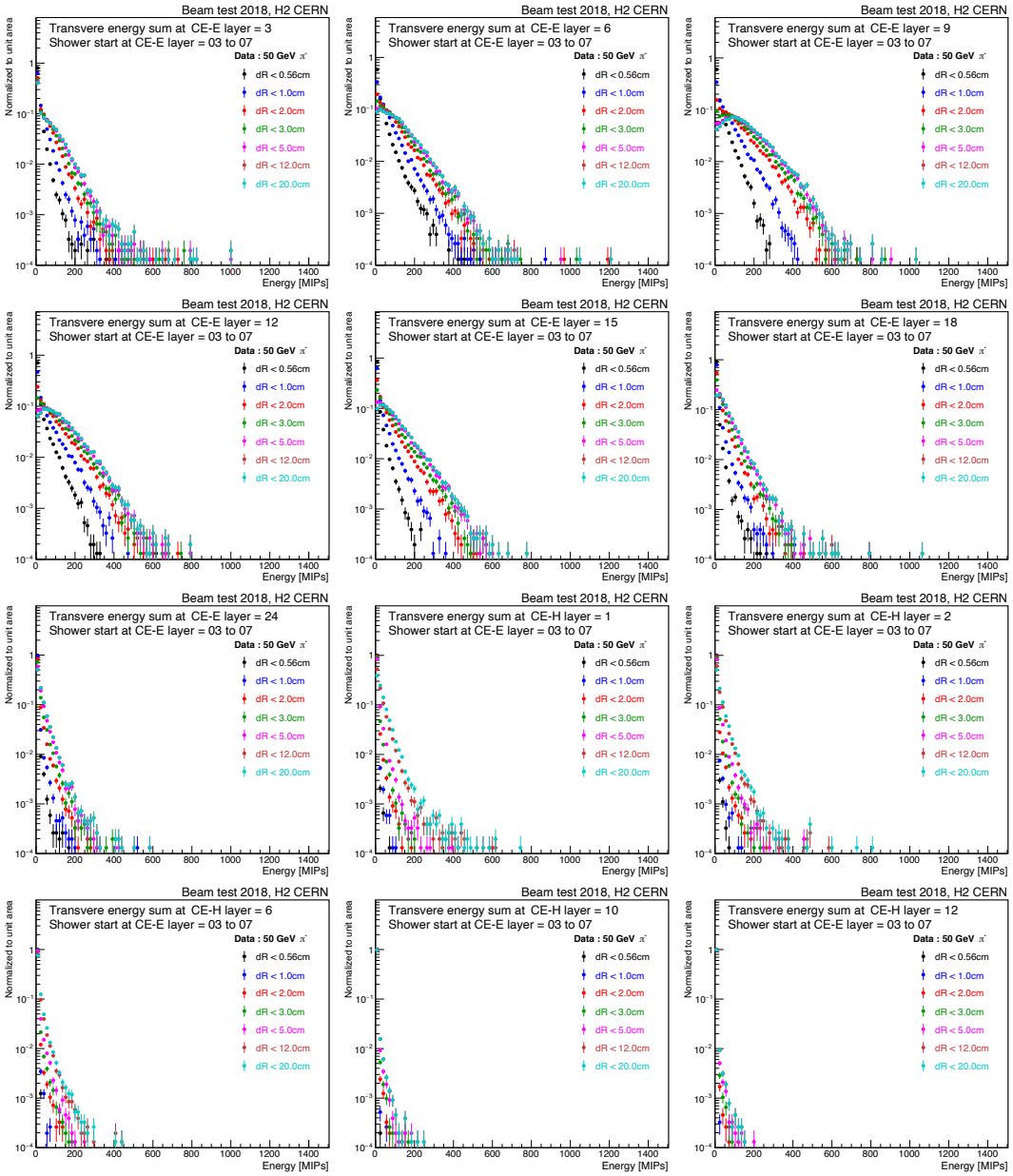

FIGURE H.3: Energy distributions in the circle of different radii (shown in different colour points) from the track impact position at layers downstream to shower start location from **CE-E layer 3 to 7** for **50 GeV/c** pion beam.

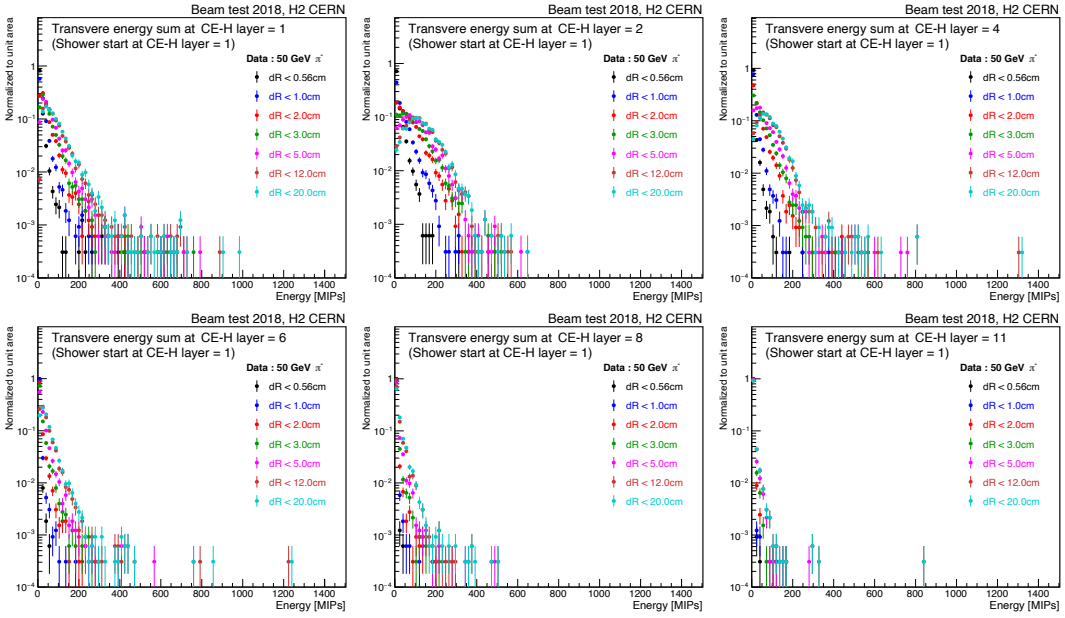

FIGURE H.4: Energy distributions in the circle of different radii (shown in different colour points) from the track impact position at layers downstream to shower start location from **CE-H layer 1** for **50 GeV/c** pion beam.

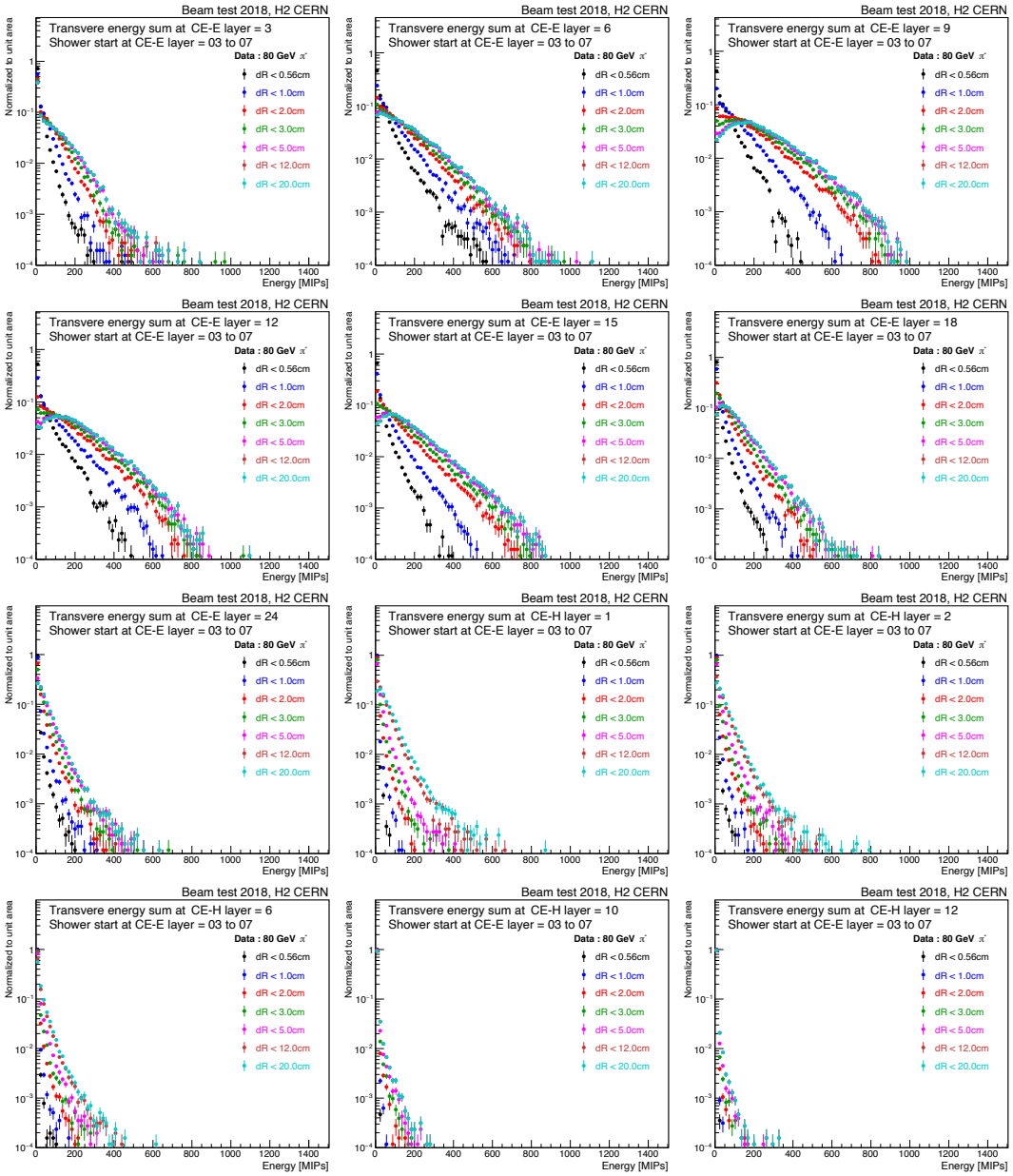

FIGURE H.5: Energy distributions in the circle of different radii (shown in different colour points) from the track impact position at layers downstream to shower start location from **CE-E layer 3 to 7** for **80 GeV/c** pion beam.

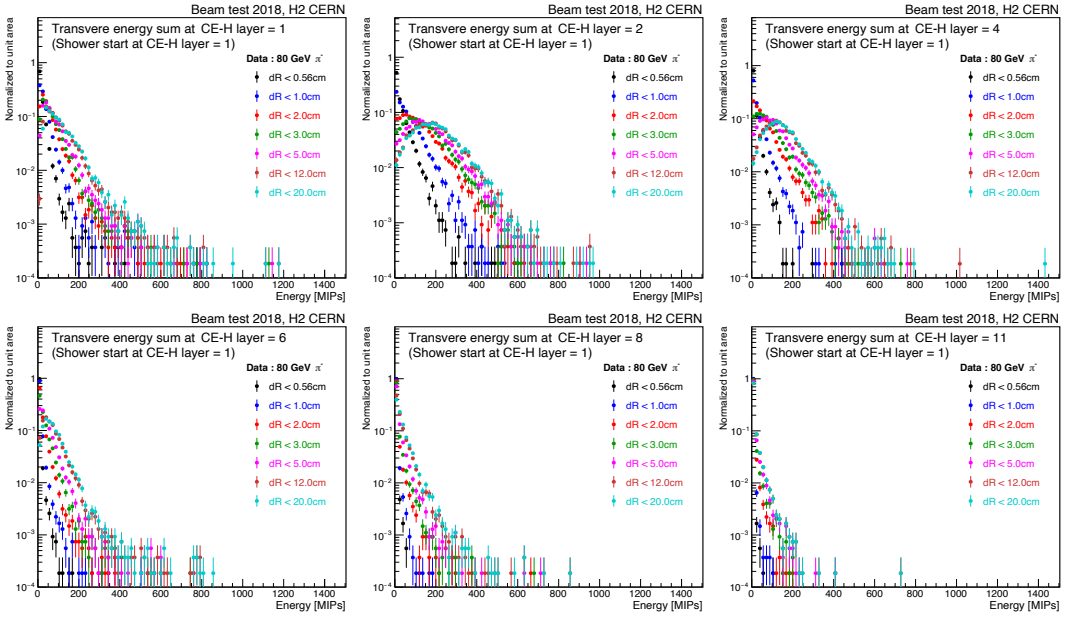

FIGURE H.6: Energy distributions in the circle of different radii (shown in different colour points) from the track impact position at layers downstream to shower start location from **CE-H layer 1** for **80 GeV/c** pion beam.

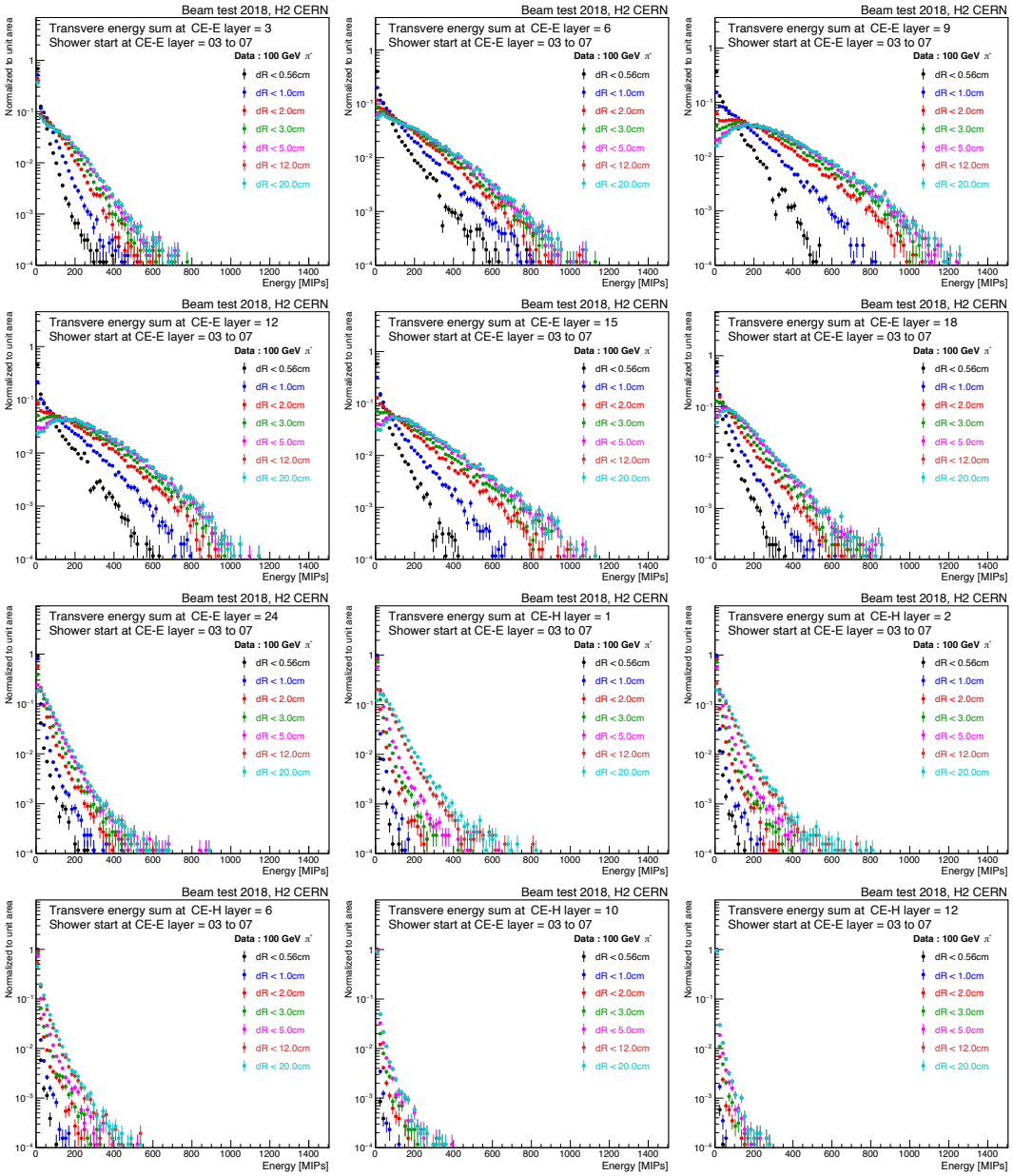

FIGURE H.7: Energy distributions in the circle of different radii (shown in different colour points) from the track impact position at layers downstream to shower start location from **CE-E layer 3 to 7** for **100 GeV/c** pion beam.

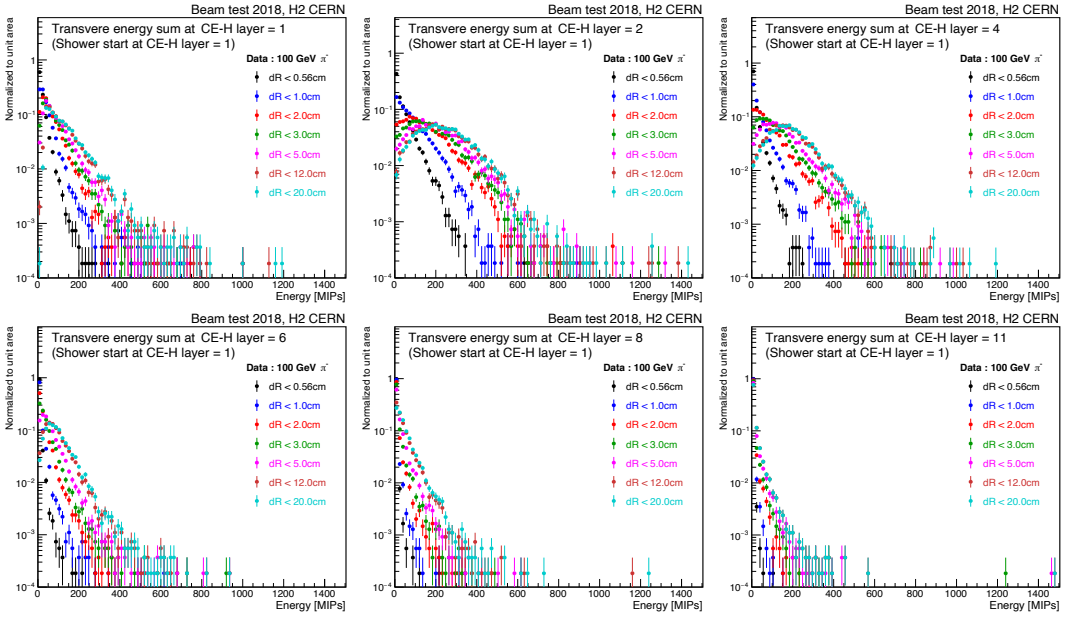

FIGURE H.8: Energy distributions in the circle of different radii (shown in different colour points) from the track impact position at layers downstream to shower start location from **CE-H layer 1** for **100 GeV/c** pion beam.

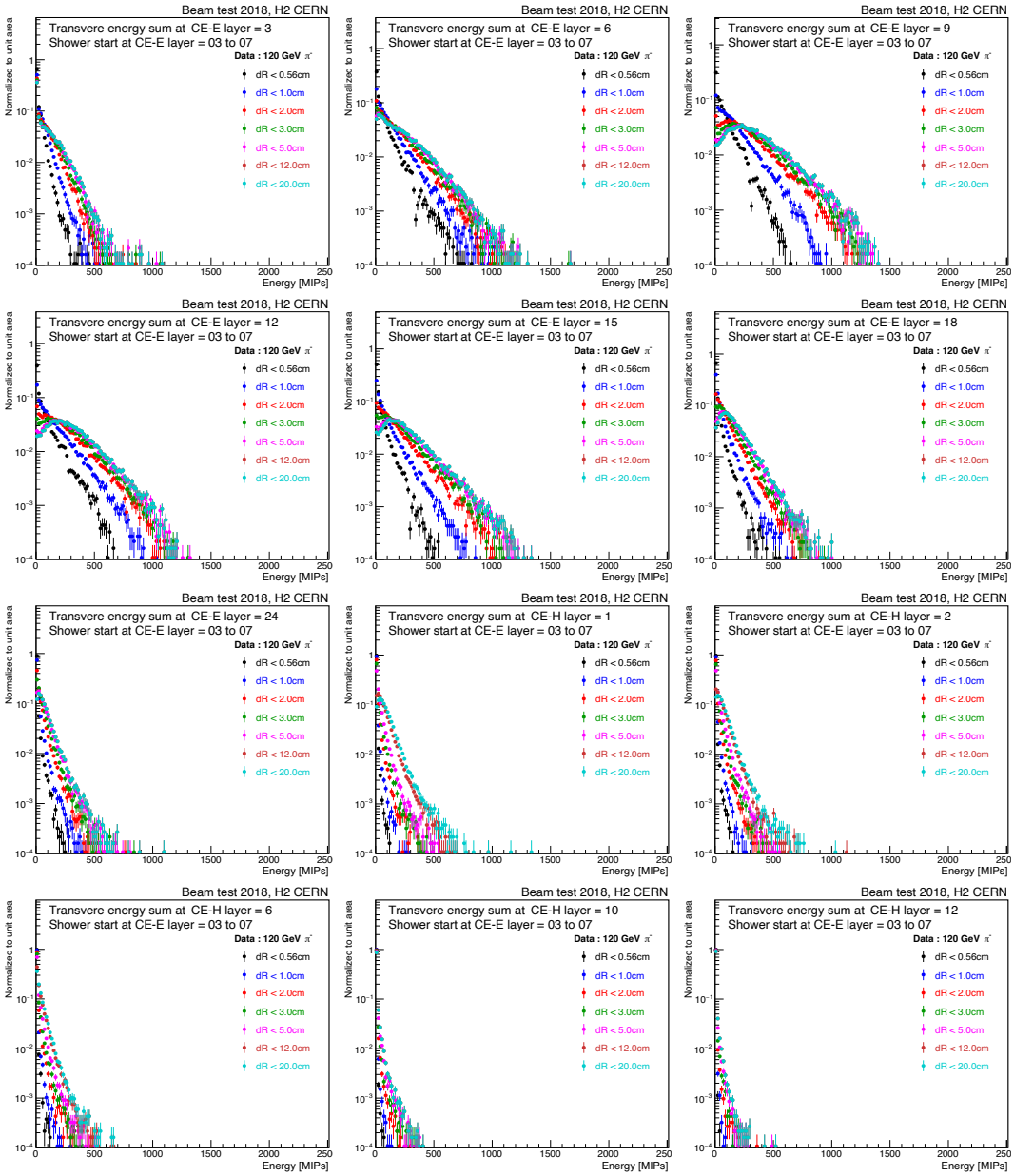

FIGURE H.9: Energy distributions in the circle of different radii (shown in different colour points) from the track impact position at layers downstream to shower start location from **CE-E layer 3 to 7** for **120 GeV/c** pion beam.

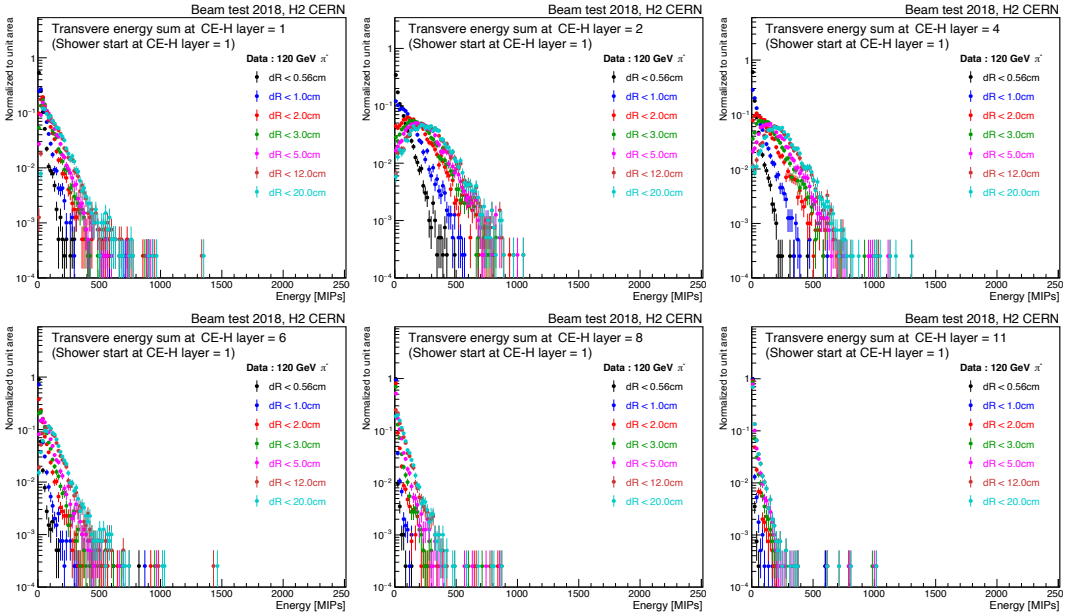

FIGURE H.10: Energy distributions in the circle of different radii (shown in different colour points) from the track impact position at layers downstream to shower start location from **CE-H layer 1** for **120 GeV/c** pion beam.

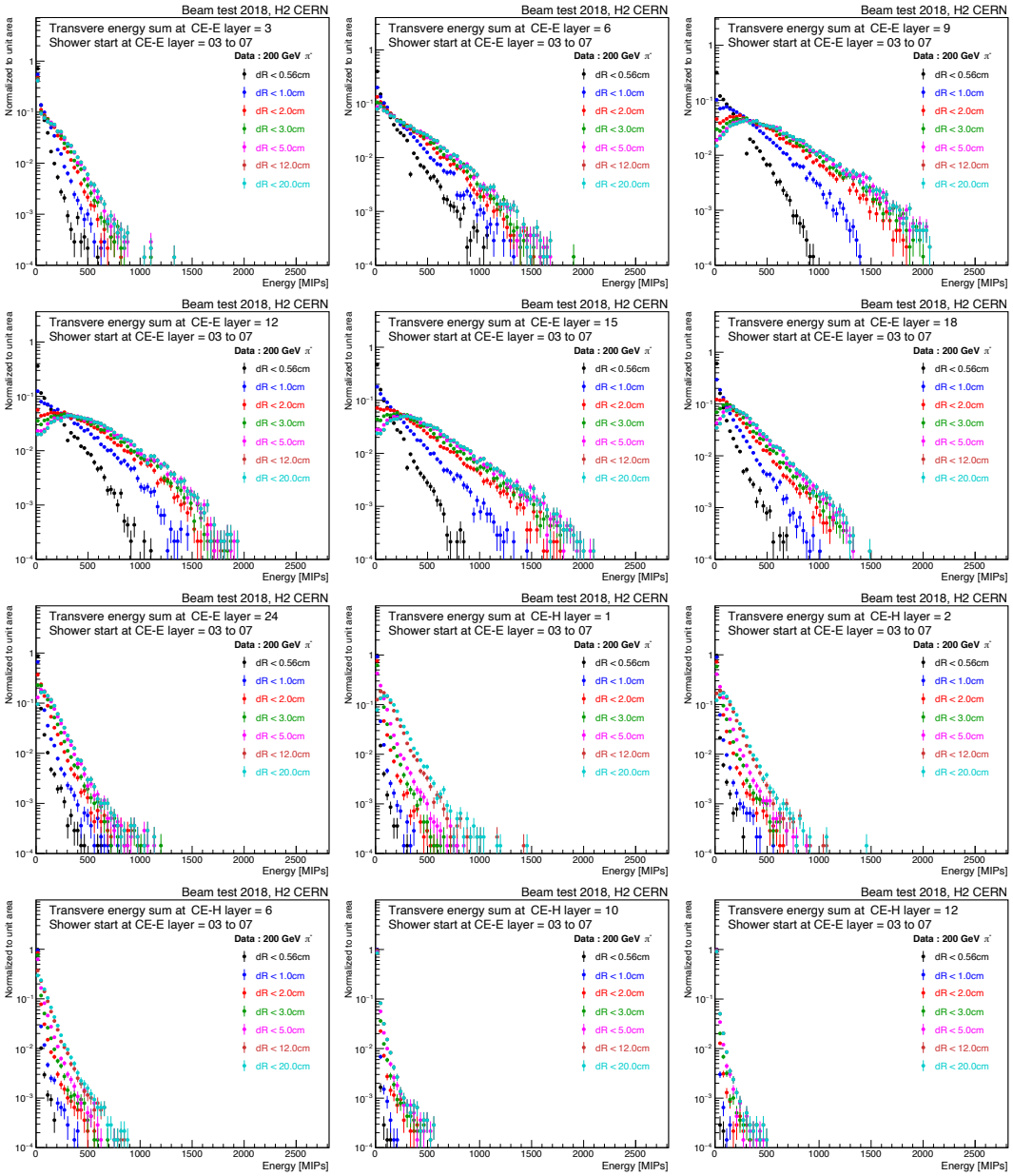

FIGURE H.11: Energy distributions in the circle of different radii (shown in different colour points) from the track impact position at layers downstream to shower start location from **CE-E layer 3 to 7** for **200 GeV/c** pion beam.

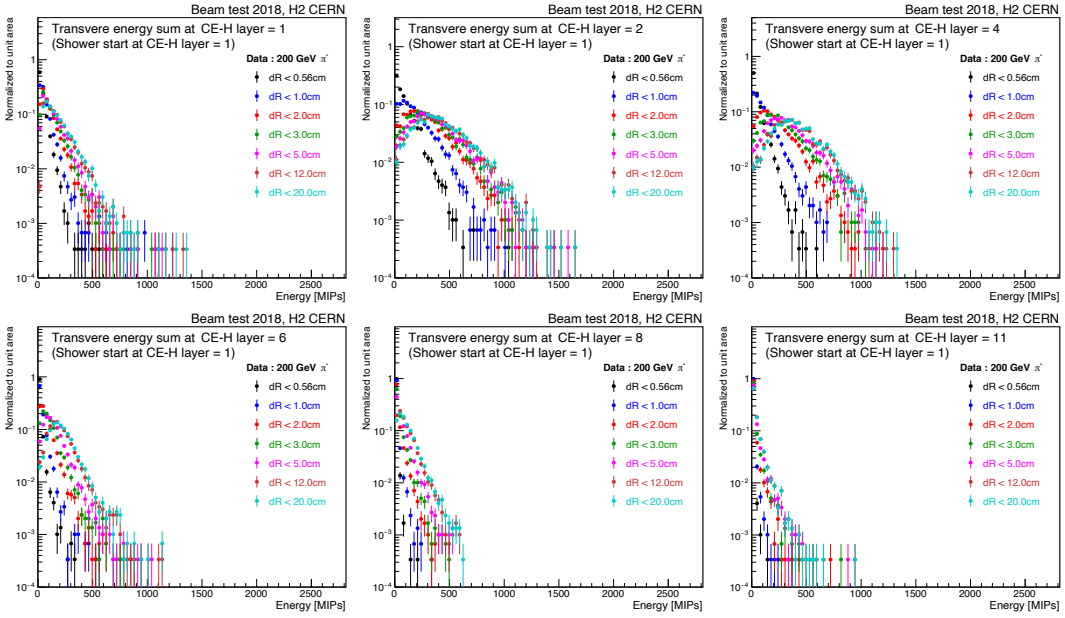

FIGURE H.12: Energy distributions in the circle of different radii (shown in different colour points) from the track impact position at layers downstream to shower start location from **CE-H layer 1** for **200 GeV/c** pion beam.

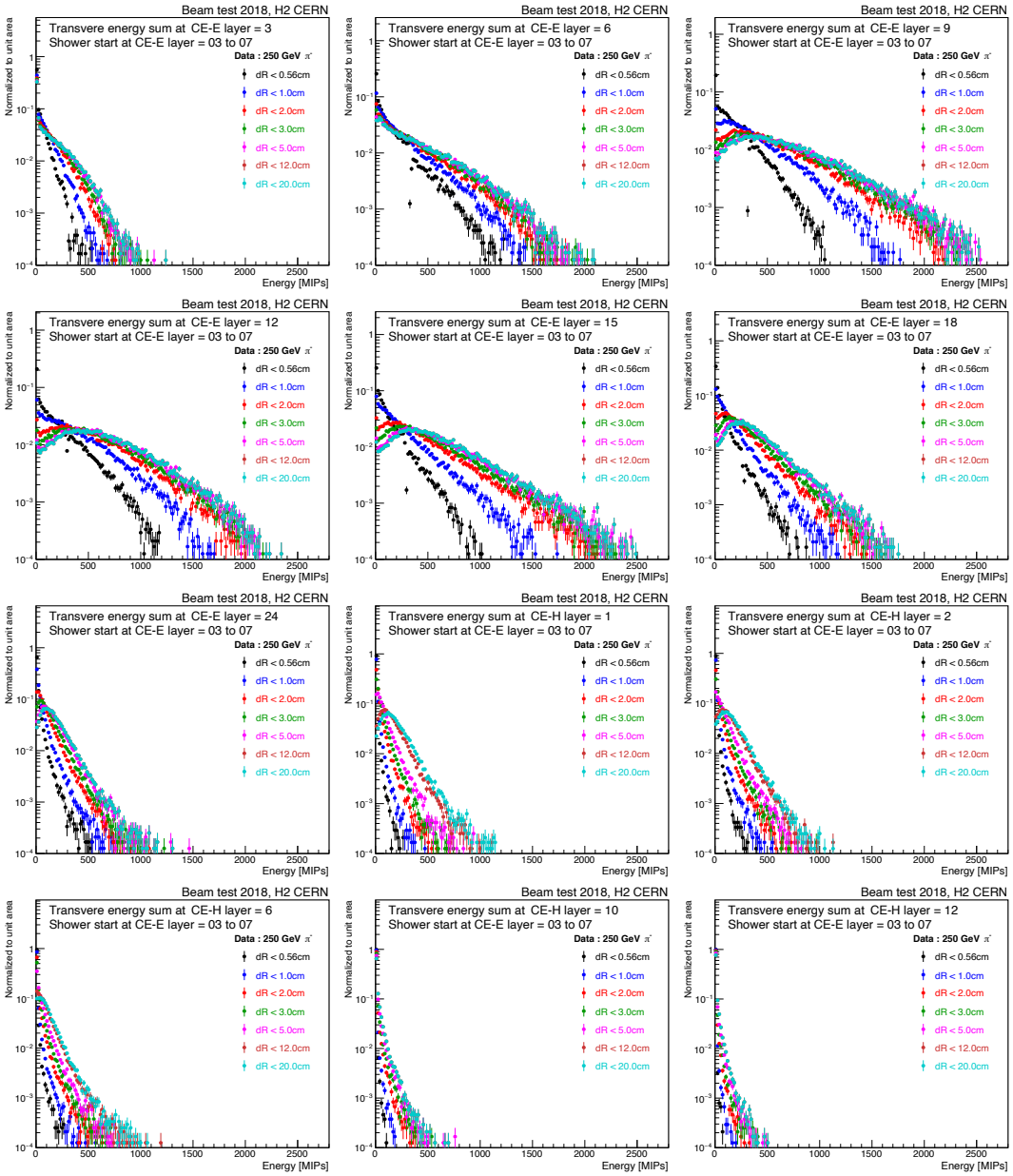

FIGURE H.13: Energy distributions in the circle of different radii (shown in different colour points) from the track impact position at layers downstream to shower start location from **CE-E layer 3 to 7** for **250 GeV/c** pion beam.

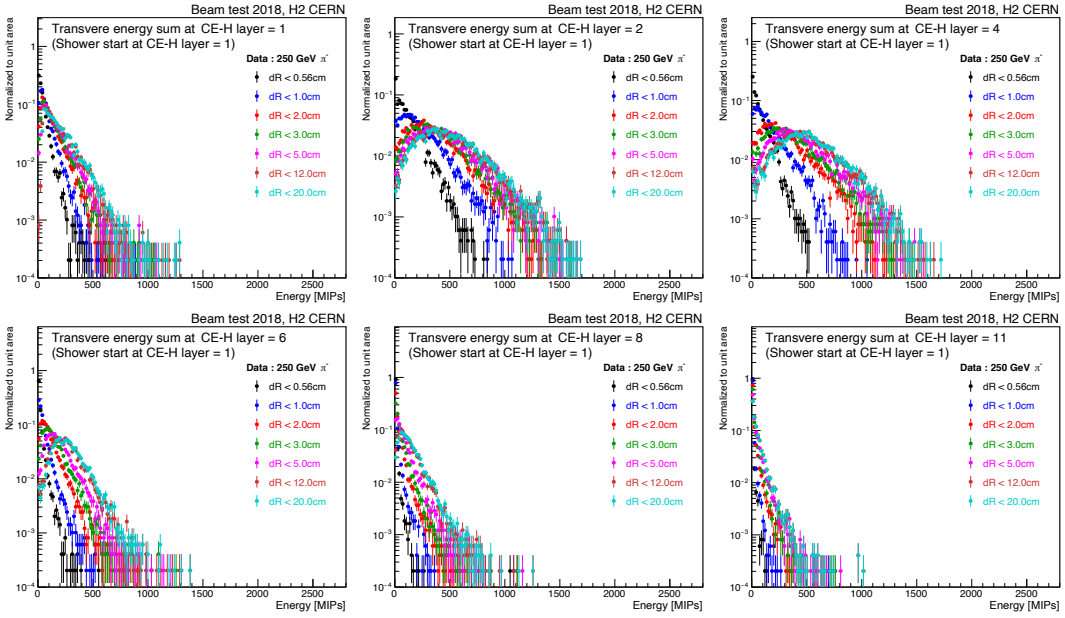

FIGURE H.14: Energy distributions in the circle of different radii (shown in different colour points) from the track impact position at layers downstream to shower start location from **CE-H layer 1** for **250 GeV/c** pion beam.

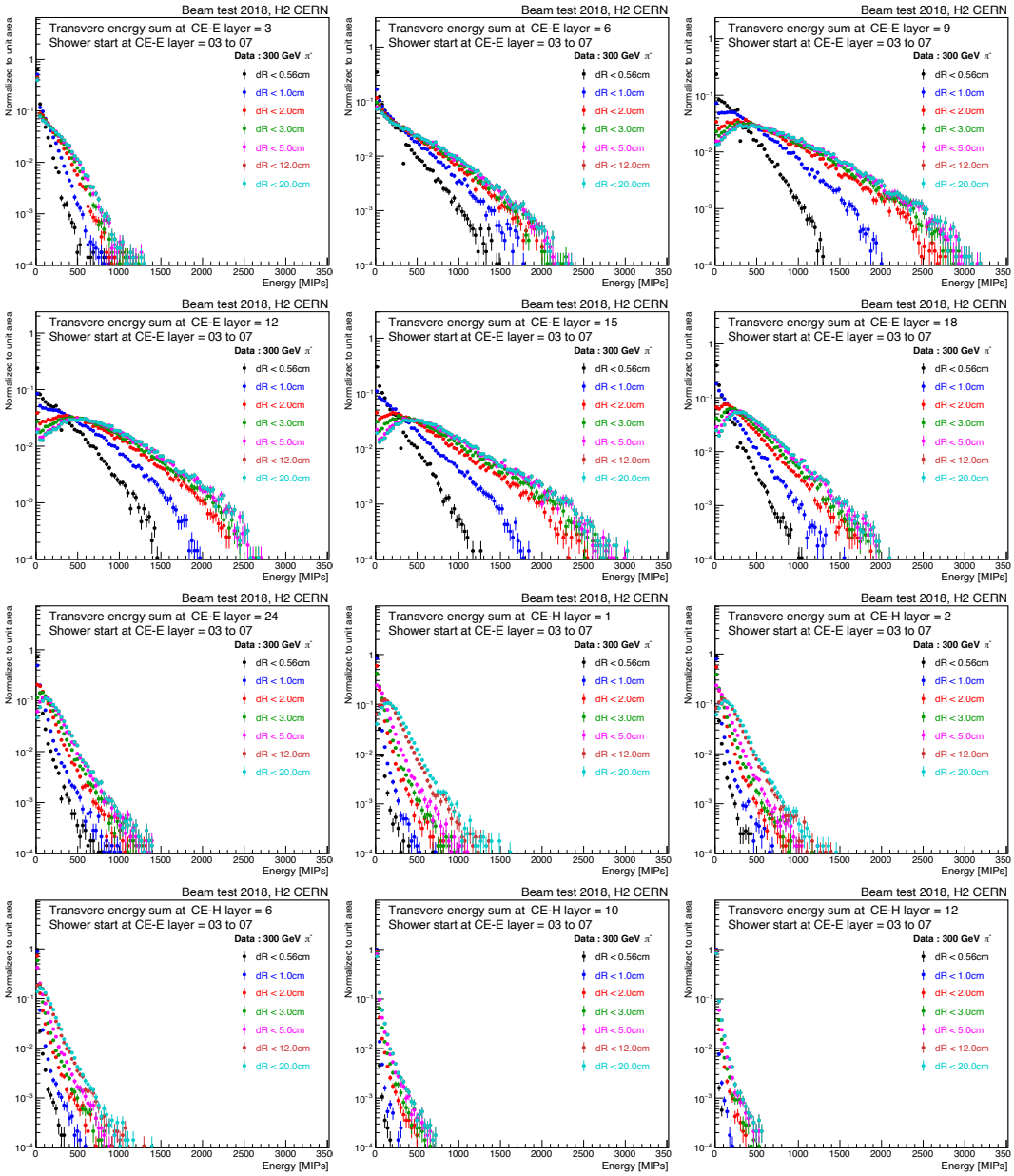

FIGURE H.15: Energy distributions in the circle of different radii (shown in different colour points) from the track impact position at layers downstream to shower start location from **CE-E layer 3 to 7** for **300 GeV/c** pion beam.

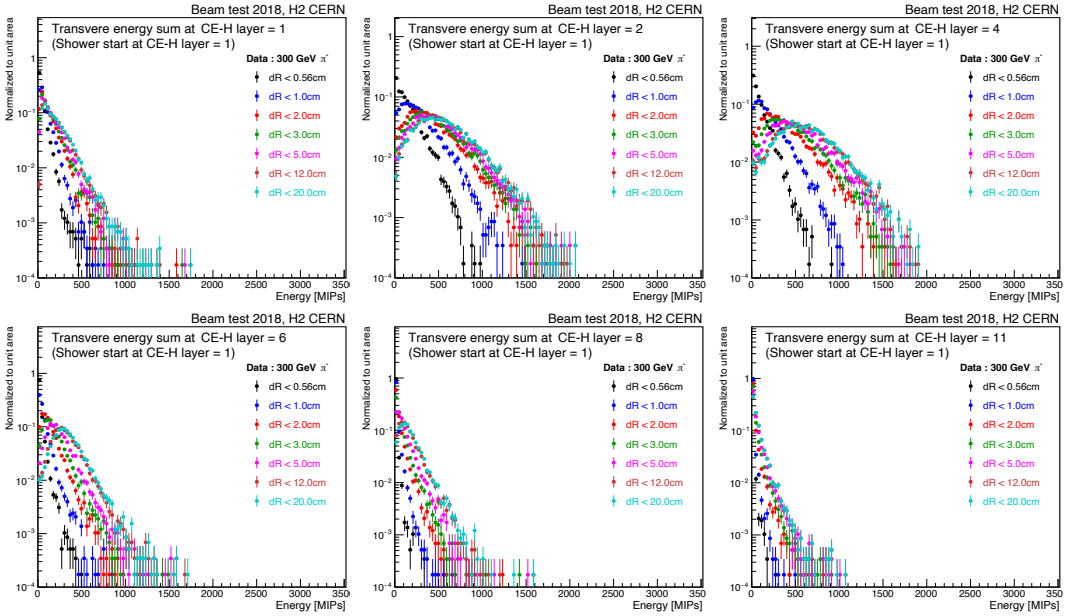

FIGURE H.16: Energy distributions in the circle of different radii (shown in different colour points) from the track impact position at layers downstream to shower start location from **CE-H layer 1** for **300 GeV/c** pion beam.

Appendix I

Data simulation comparison of energy integrated in circles of different radii

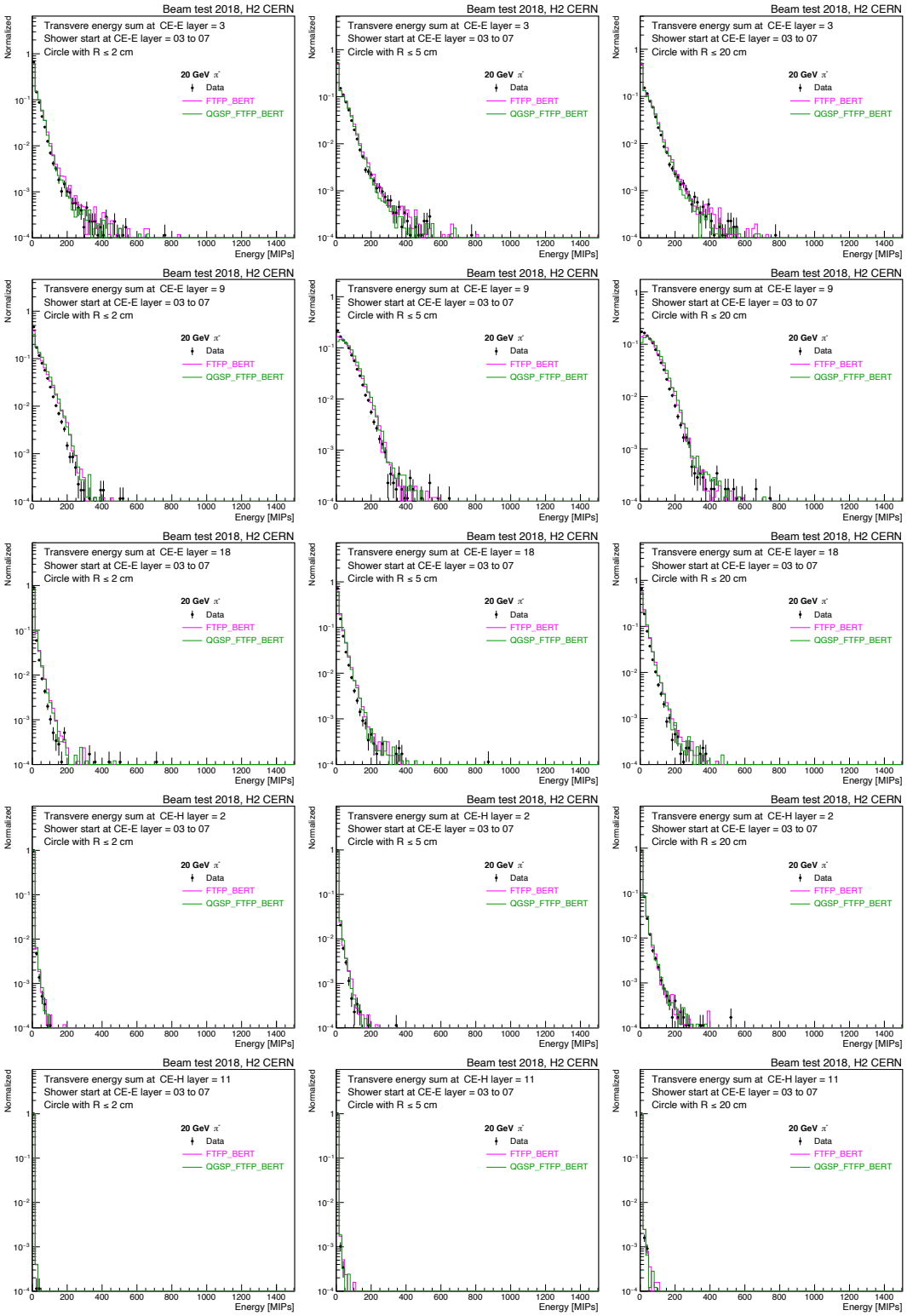

FIGURE I.1: Data-simulation comparison of energy distributions in the circle of radius 2 cm (left column), 5 cm (middle column) and 20 cm (right column) from the track impact position at layers downstream to shower start location from CE-E layer 3 to 7 for **20 GeV/c** pion beam.

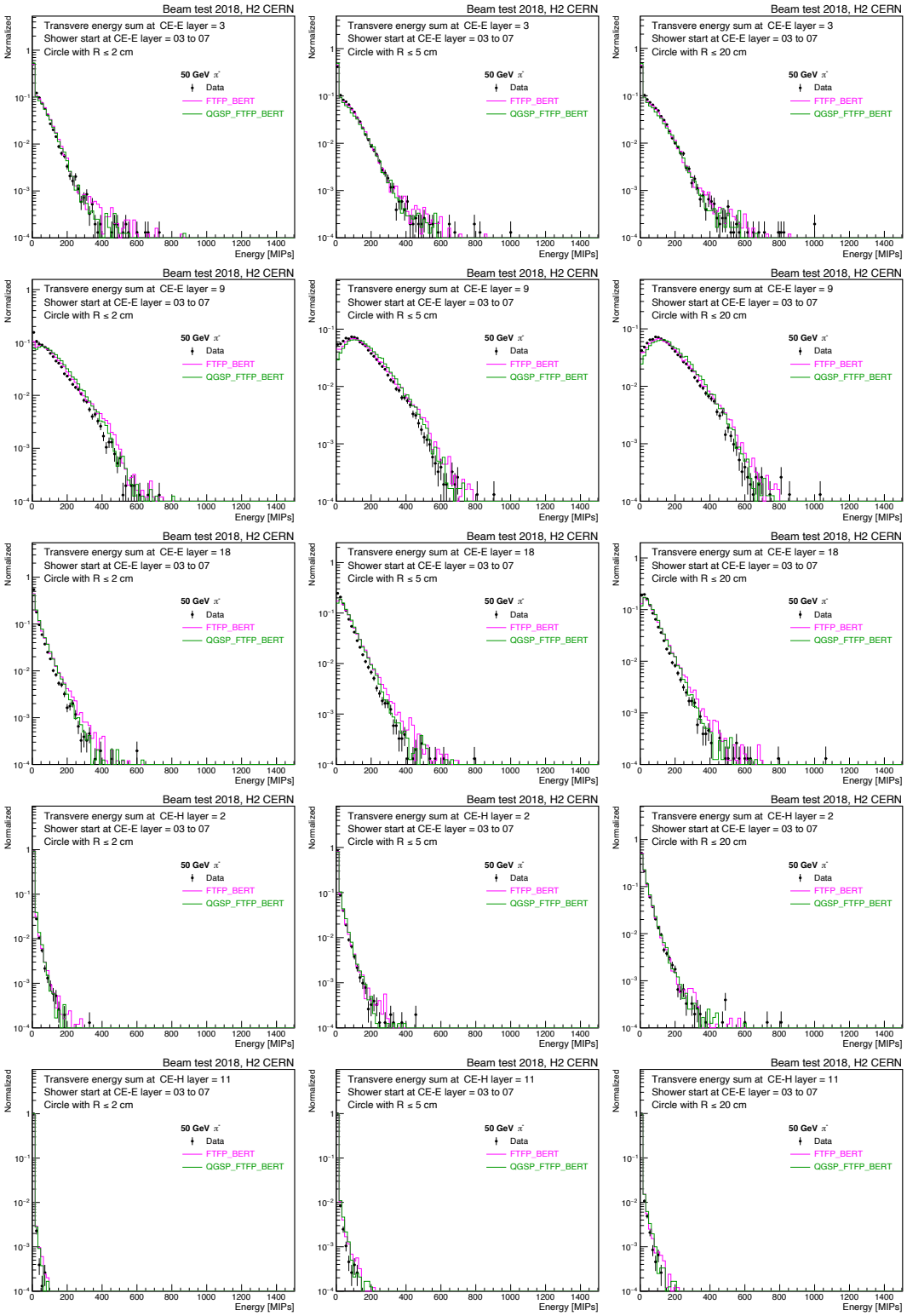

FIGURE I.2: Data-simulation comparison of energy distributions in the circle of radius 2 cm (left column), 5 cm (middle column) and 20 cm (right column) from the track impact position at layers downstream to shower start location from CE-E layer 3 to 7 for **50 GeV/c** pion beam.

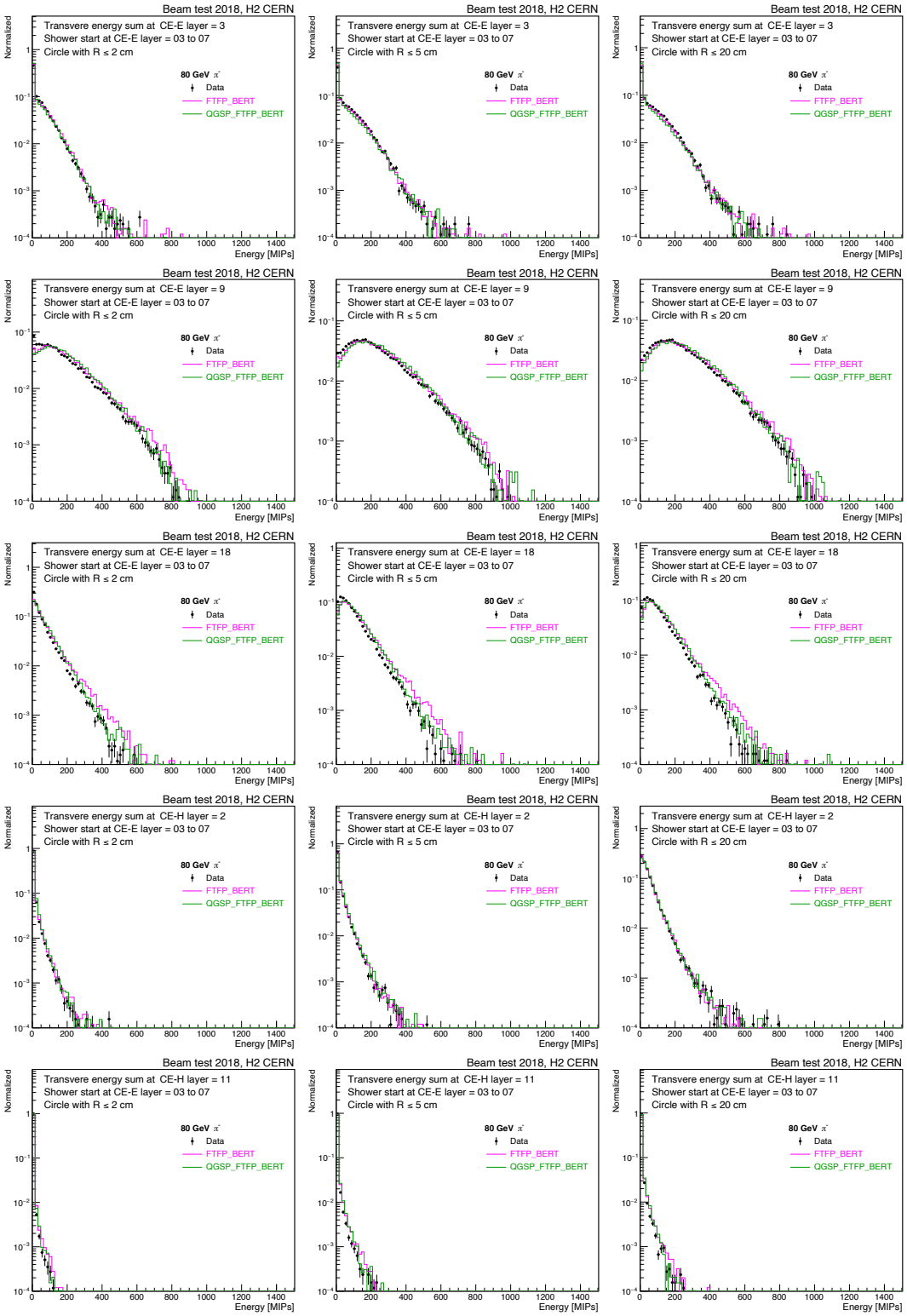

FIGURE I.3: Data-simulation comparison of energy distributions in the circle of radius 2 cm (left column), 5 cm (middle column) and 20 cm (right column) from the track impact position at layers downstream to shower start location from CE-E layer 3 to 7 for **80 GeV/c** pion beam.

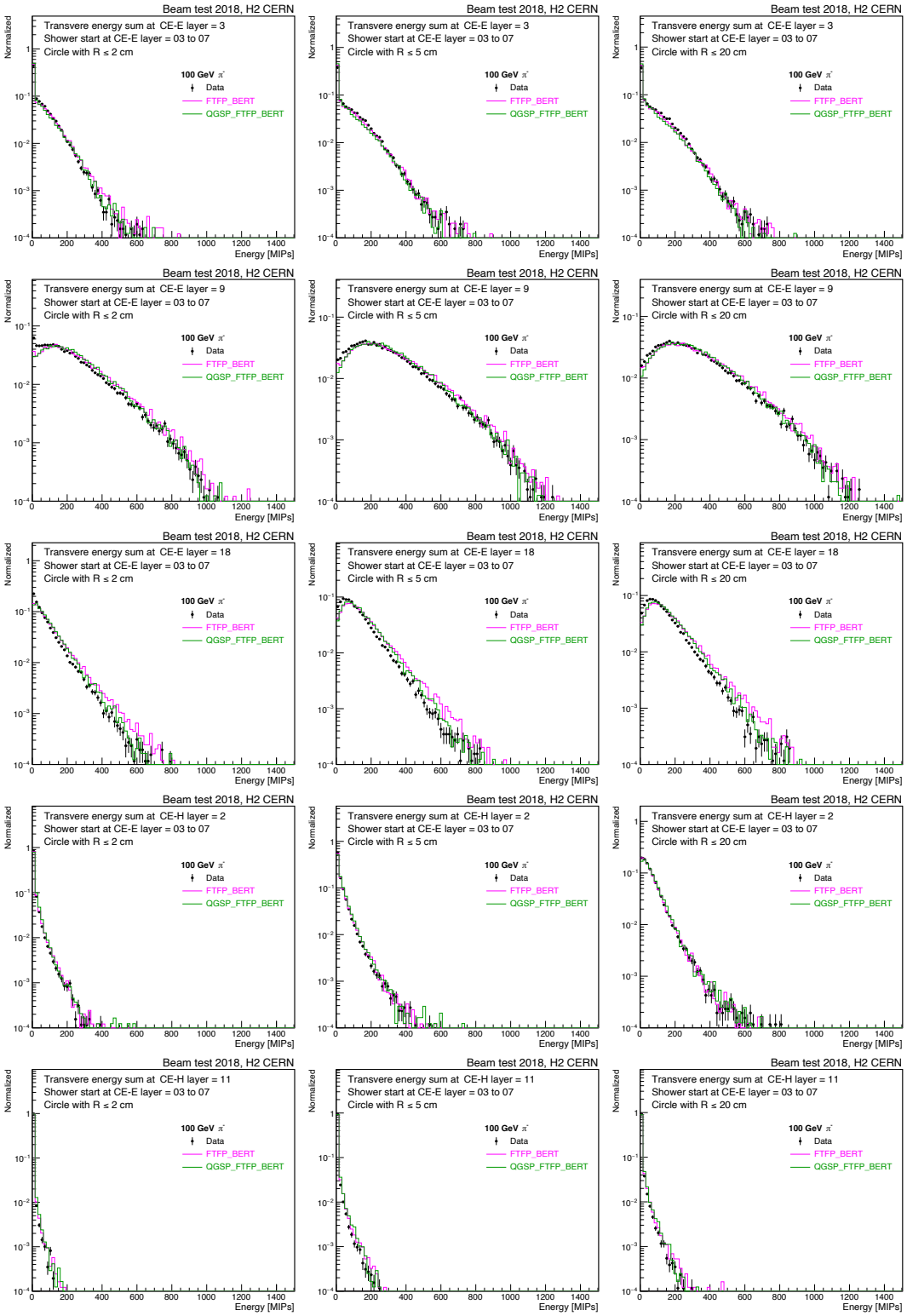

FIGURE I.4: Data-simulation comparison of energy distributions in the circle of radius 2 cm (left column), 5 cm (middle column) and 20 cm (right column) from the track impact position at layers downstream to shower start location from CE-E layer 3 to 7 for **100 GeV/c** pion beam.

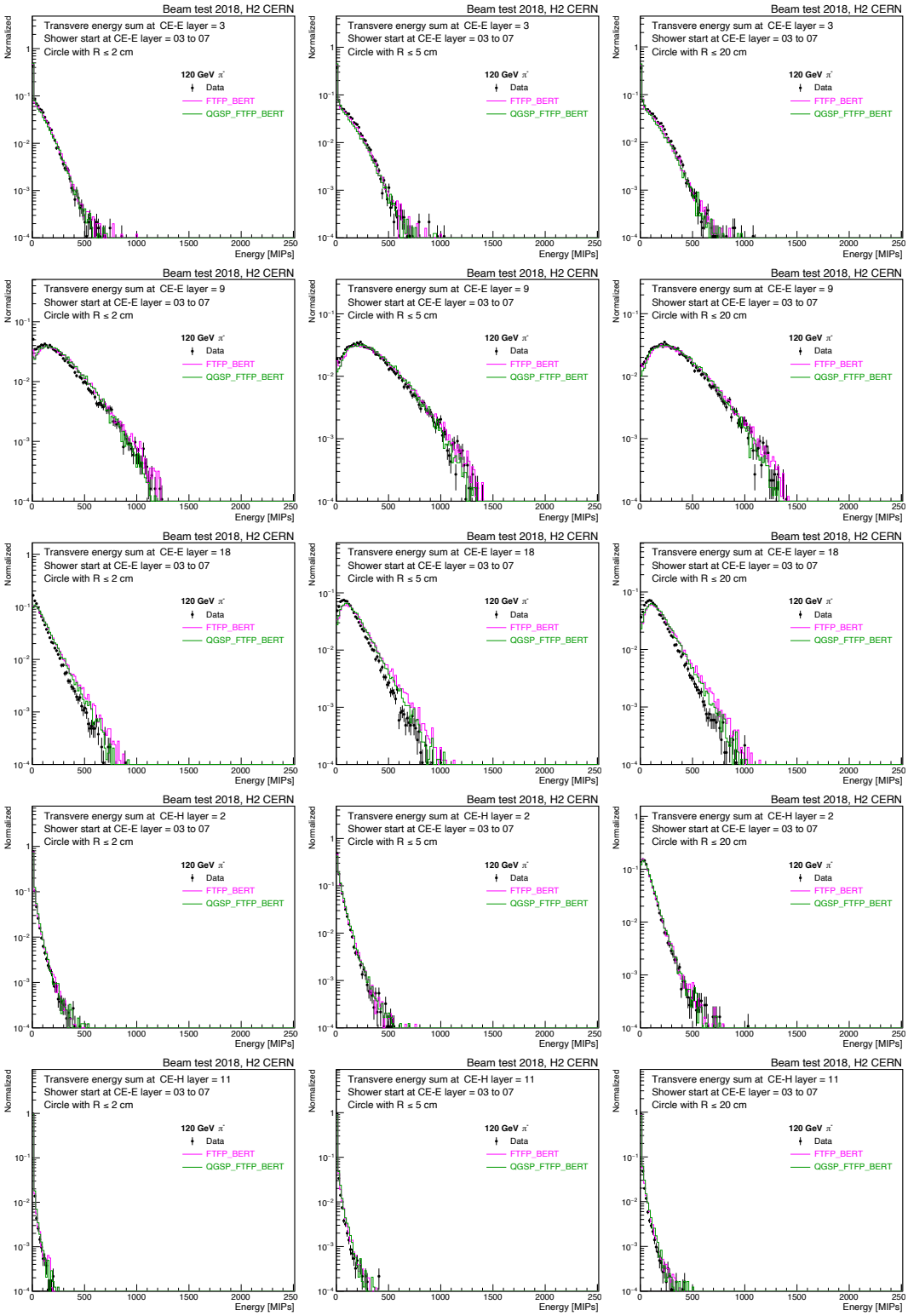

FIGURE I.5: Data-simulation comparison of energy distributions in the circle of radius 2 cm (left column), 5 cm (middle column) and 20 cm (right column) from the track impact position at layers downstream to shower start location from CE-E layer 3 to 7 for **120 GeV/c** pion beam.

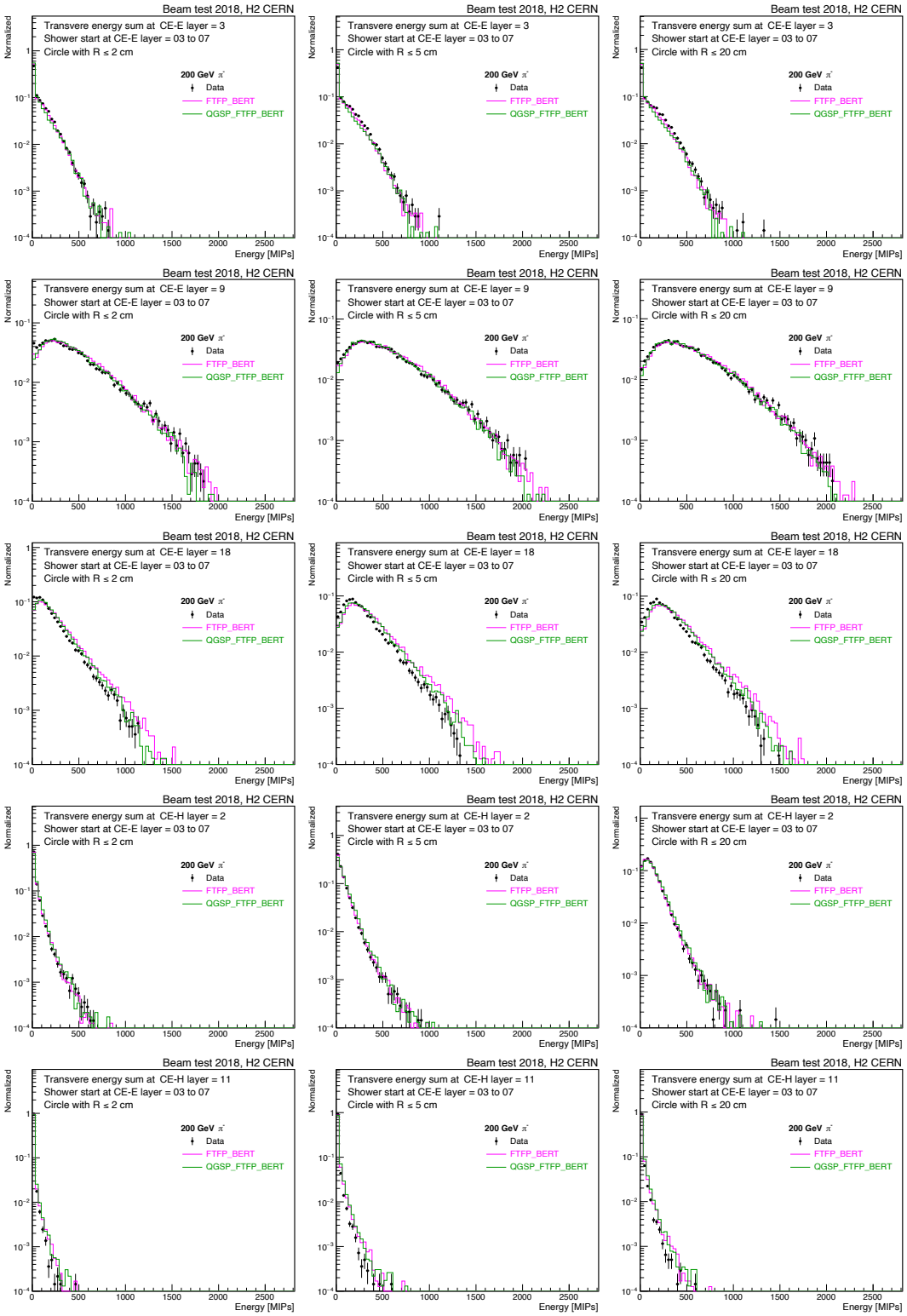

FIGURE I.6: Data-simulation comparison of energy distributions in the circle of radius 2 cm (left column), 5 cm (middle column) and 20 cm (right column) from the track impact position at layers downstream to shower start location from CE-E layer 3 to 7 for **200 GeV/c** pion beam.

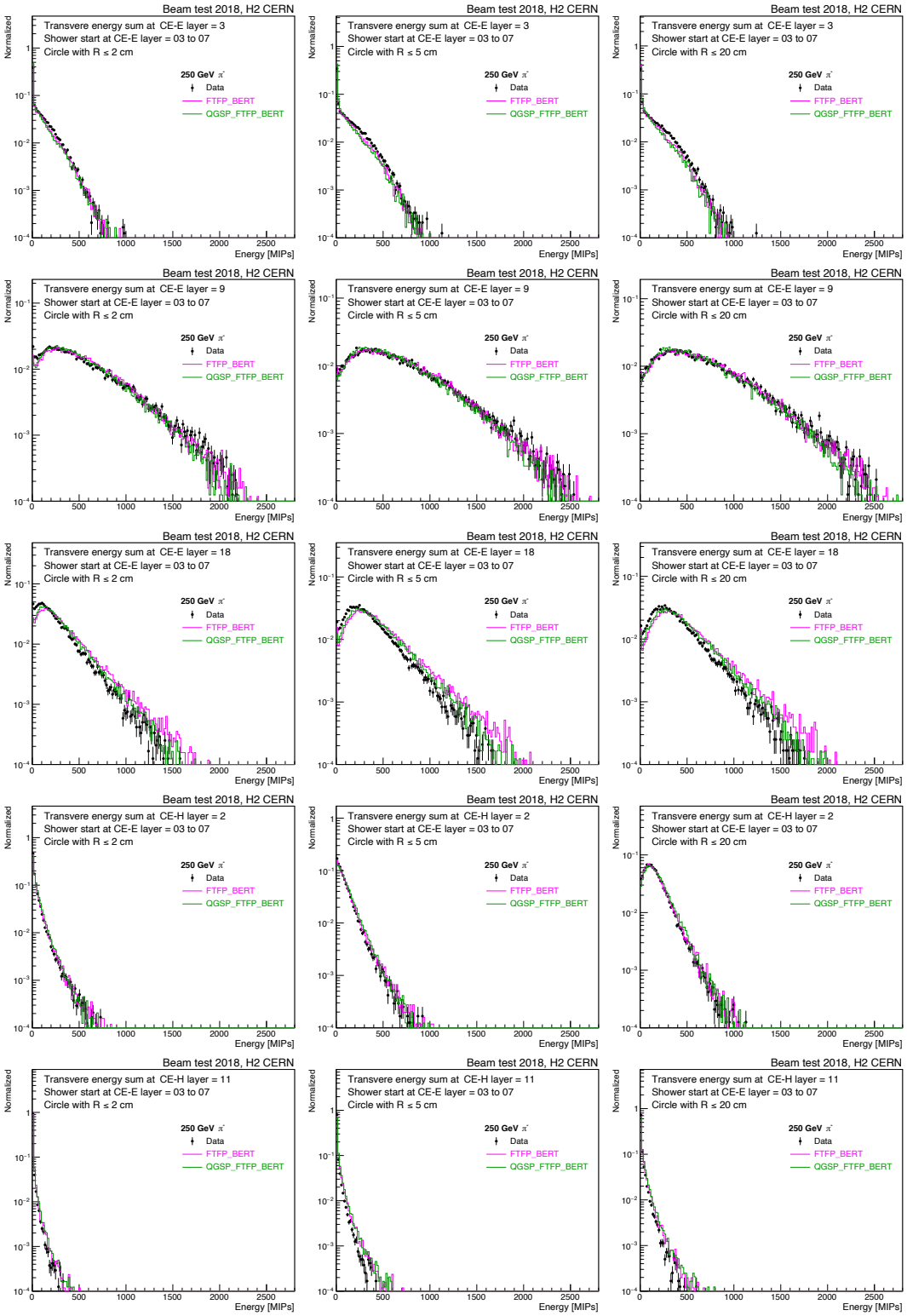

FIGURE I.7: Data-simulation comparison of energy distributions in the circle of radius 2 cm (left column), 5 cm (middle column) and 20 cm (right column) from the track impact position at layers downstream to shower start location from CE-E layer 3 to 7 for **250 GeV/c** pion beam.

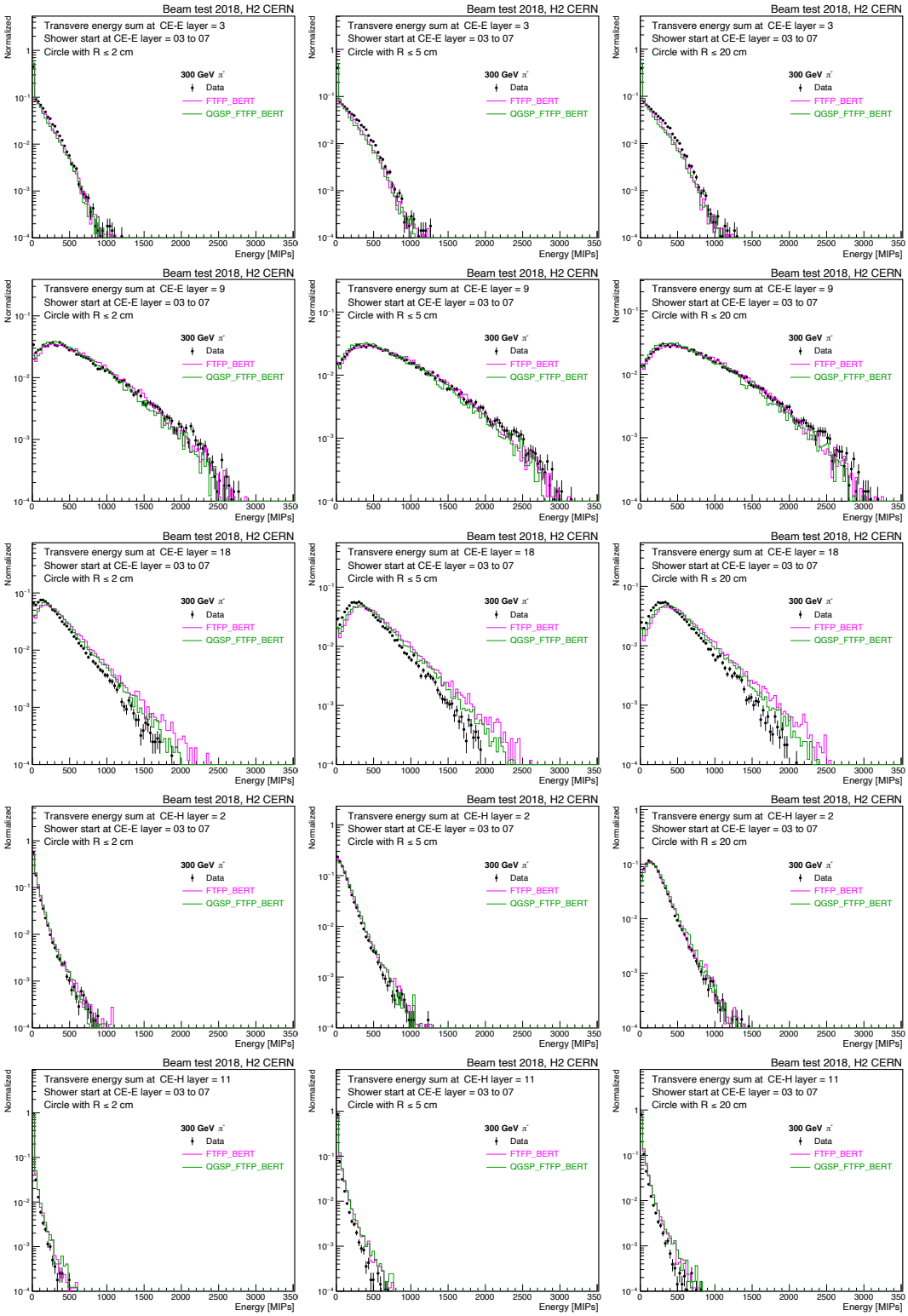

FIGURE I.8: Data-simulation comparison of energy distributions in the circle of radius 2 cm (left column), 5 cm (middle column) and 20 cm (right column) from the track impact position at layers downstream to shower start location from CE-E layer 3 to 7 for **300 GeV/c** pion beam.

Bibliography

[1] Stanford Encyclopedia of Philosophy. *Leucippus*.

[2] Stanford Encyclopedia of Philosophy. *Democritus*.

[3] Chandradhar Sharma. *A Critical Survey of Indian Philosophy*. Motilal Banarsidass Publ., 2000. ISBN: 9788120803657.

[4] Science History Institute. *John Dalton*.

[5] Science History Institute. *Joseph John Thompson*.

[6] Professor E. Rutherford F.R.S. "LXXIX. The scattering of alpha and beta particles by matter and the structure of the atom". *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 21.125 (1911), pp. 669–688. DOI: 10.1080/14786440508637080.

[7] Louis de Broglie. "XXXV. A tentative theory of light quanta". *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 47.278 (1924), pp. 446–458. DOI: 10.1080/14786442408634378.

[8] WK Panofsky. "Evolution of particle accelerators". *SLAC Beam Line* 27.1 (1997), pp. 36–44.

[9] Guido Altarelli. "The Standard model of particle physics". *arXiv preprint hep-ph/0510281* (2005).

[10] Brian R Martin and Graham Shaw. *Particle physics*. John Wiley & Sons, 2017.

[11] B. Odom et al. "New Measurement of the Electron Magnetic Moment Using a One-Electron Quantum Cyclotron". *Phys. Rev. Lett.* 97 (3 2006), p. 030801. DOI: 10.1103/PhysRevLett.97.030801.

[12] John D Dowell. *The UA1 experiment with ACOL*. Tech. rep. 198612094, 1986.

[13] CN Booth. *The UA2 experiment at ACOL*. Tech. rep. 198706279, 1987.

[14] W. de Boer. "Precision Experiments at LEP". *60 Years of CERN Experiments and Discoveries* (2015), 107–136. ISSN: 1793-1339. DOI: 10.1142/9789814644150_0005.

[15] Arnd Meyer. "The CDF experiment at the Tevatron: The first two years of Run II". *Modern Physics Letters A* 18.24 (2003), pp. 1643–1660.

[16] Martin Grunewald et al. "Testing the standard model at the Fermilab Tevatron". *arXiv preprint hep-ph/0111217* (2001).

[17] B Foster. "ZEUS AT HERA II". *Deep Inelastic Scattering: DIS 2001*. World Scientific, 2002, pp. 145–152.

[18] Lyndon Evans and Philip Bryant. "LHC Machine". *Journal of Instrumentation* 3.08 (2008), S08001–S08001. DOI: 10.1088/1748-0221/3/08/s08001.

[19] The CMS collaboration. "The CMS experiment at the CERN LHC". *Journal of instrumentation* 3 (2008), S08004.

[20] Stephen P Martin. "A supersymmetry primer". *Perspectives on supersymmetry II*. World Scientific, 2010, pp. 1–153.

[21] Giorgio Apollinari et al. "High luminosity large hadron collider HL-LHC". *arXiv preprint arXiv:1705.08830* (2017).

[22] D Contardo et al. *Technical Proposal for the Phase-II Upgrade of the CMS Detector*. Tech. rep. Geneva, 2015.

[23] *The Phase-2 Upgrade of the CMS Endcap Calorimeter*. Tech. rep. Geneva: CERN, 2017.

[24] S. Pandey, S. Sharma, and A. Lobanov. "Performance of CMS High Granularity Calorimeter prototypes in beam test experiments". *Journal of Instrumentation* 15.06 (2020), pp. C06028–C06028. DOI: 10.1088/1748-0221/15/06/c06028.

[25] Felix Sefkow and Frank Simon. "A highly granular SiPM-on-tile calorimeter prototype". *Journal of Physics: Conference Series* 1162 (Jan. 2019), p. 012012. DOI: 10.1088/1742-6596/1162/1/012012.

[26] Sea Agostinelli et al. "GEANT4—a simulation toolkit". *Nuclear instruments and methods in physics research section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 506.3 (2003), pp. 250–303.

[27] Michael E Peskin. *An introduction to quantum field theory*. CRC press, 2018.

[28] Emmy Noether. "Invariant variation problems". *Transport theory and statistical physics* 1.3 (1971), pp. 186–207.

[29] David Griffiths. *Introduction to elementary particles*. John Wiley & Sons, 2020.

[30] Rohini M. Godbole. *Lectures on EW Standard Model*. 2017. arXiv: 1703.04978 [hep-ph].

[31] Gautam Bhattacharyya. "Hierarchy problem and BSM physics". *Pramana* 89.4 (2017), pp. 1–3.

[32] Matts Roos. "Dark Matter: The evidence from astronomy, astrophysics and cosmology". *arXiv preprint arXiv:1001.0316* (2010).

[33] Abdel Perez-Lorenzana. "An introduction to extra dimensions". *Journal of Physics: Conference Series*. Vol. 18. 1. IOP Publishing. 2005, p. 006.

[34] Zurab G Berezhiani and Rabindra N Mohapatra. "Reconciling present neutrino puzzles: Sterile neutrinos as mirror neutrinos". *Physical Review D* 52.11 (1995), p. 6607.

[35] H. Bethe. "Zur Theorie des Durchgangs schneller Korpuskularstrahlen durch Materie". *Annalen der Physik* 397.3 (1930), pp. 325–400. DOI: https://doi.org/10.1002/andp.19303970303.

[36] P.A. Zyla et al. "Review of Particle Physics". *PTEP* 2020.8 (2020), p. 083C01. DOI: 10.1093/ptep/ptaa104.

[37] R. Wigmans and R. Wigmans. *Calorimetry: Energy Measurement in Particle Physics*. International series of monographs on physics. Oxford University Press, 2017. ISBN: 9780198786351.

[38] Tejinder S Virdee. *Experimental techniques*. Tech. rep. CERN, 1999.

[39] Donald E. Groom. "Energy flow in a hadronic cascade: Application to hadron calorimetry". *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 572.2 (2007), pp. 633–653. ISSN: 0168-9002. DOI: https://doi.org/10.1016/j.nima.2006.11.070.

[40] G. Grindhammer and S. Peters. *The Parameterized Simulation of Electromagnetic Showers in Homogeneous and Sampling Calorimeters*. 2000. arXiv: hep-ex/0001020 [hep-ex].

[41] Larry C Andrews et al. "Theory of optical scintillation". *JOSA A* 16.6 (1999), pp. 1417–1429.

[42] Boris M Bolotovskii. "Vavilov–Cherenkov radiation: its discovery and application". *Physics-Uspekhi* 52.11 (2009), p. 1099.

[43] E Armengaud et al. "Performance of the EDELWEISS-III experiment for direct dark matter searches". *Journal of Instrumentation* 12.08 (2017), P08010.

[44] Richard W Schnee et al. "The superCDMS experiment". *Dark matter in astro-and particle physics*. Springer, 2006, pp. 259–268.

[45] Leonardo Rossi et al. *Pixel Detectors*. Springer, 2006. ISBN: 9783540283324. DOI: https://doi.org/10.1007/3-540-28333-1.

[46] Sven Menke. *Electromagnetic shower event displays*. URL: https://www.mpp.mpg.de/~menke/elss/pic1.shtml?unwrap.

[47] C D'Ambrosio et al. "Particle Detectors-Principles and Techniques". *CERN Academic Training Programming* 2005 (2004).

[48] The CALICE collaboration. "Design and electronics commissioning of the physics prototype of a Si-W electromagnetic calorimeter for the International Linear Collider". *Journal of Instrumentation* 3.08 (2008), P08001–P08001. DOI: 10.1088/1748-0221/3/08/p08001.

[49] Christopher Z Mooney. *Monte carlo simulation*. 116. Sage, 1997.

[50] Vladimir Ivanchenko et al. "Recent improvements in geant4 electromagnetic physics models and interfaces". *Progress in nuclear science and technology* 2 (2011), pp. 898–903.

[51] The GEANT4 Collaboration. *GEANT4 Physics Reference Manual*. 2017.

[52] Bo Andersson, G Gustafson, and Hong Pi. "The FRITIOF model for very high energy hadronic collisions". *Zeitschrift für Physik C Particles and Fields* 57.3 (1993), pp. 485–494.

[53] B Anderson. "The Lund string model". *Antiproton 1984. Proceedings of the VII European symposium on antiproton interactions held in Durham, 9-13 July 1984*. 1985.

[54] Torbjörn Sjöstrand. "The Pythia event generator: Past, present and future". *Computer Physics Communications* 246 (2020), p. 106910. ISSN: 0010-4655. DOI: 10.1016/j.cpc.2019.106910.

[55] GEANT4 collaboration. *Geant4 Book For Application Developers*.

[56] The ATLAS Collaboration. "The ATLAS experiment at the CERN large hadron collider". *Journal of instrumentation* 3 (2008), S08003.

[57] Kenneth Aamodt et al. "The ALICE experiment at the CERN LHC". *Journal of Instrumentation* 3.08 (2008), S08002.

[58] A Augusto Alves Jr et al. "The LHCb detector at the LHC". *Journal of instrumentation* 3.08 (2008), S08005.

[59] *The accelerator complex at CERN*. URL: https://home.cern/science/accelerators/accelerator-complex.

[60] Esma Mobs. "The CERN accelerator complex - 2019. Complexe des accélérateurs du CERN - 2019" (2019). General Photo.

[61] *CMS Luminosity - Public results*. URL: https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults.

[62] Guido Tonelli. "The silicon tracking system of CMS". *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 386.1 (1997). Proceedings of the 5th International Workshop on Vertex Detectors, pp. 129–137. ISSN: 0168-9002. DOI: https://doi.org/10.1016/S0168-9002(96)01106-0.

[63] Giovanni Franzoni, Cms Collaboration, et al. "Performance of CMS ECAL with first LHC data". *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 628.1 (2011), pp. 90–94.

[64] CMS HCAL et al. "The CMS barrel calorimeter response to particle beams from 2 to 350 GeV/c". *The European Physical Journal C* 60.3 (2009), pp. 359–373.

[65] Cms Collaboration et al. *Technical proposal for the upgrade of the CMS detector through 2020*. Tech. rep. 2011.

[66] "The CMS high level trigger". *The European Physical Journal C* 46.3 (2006), 605–667. ISSN: 1434-6052. DOI: 10.1140/epjc/s2006-02495-8.

[67] The CMS Collaboration. "Particle-flow reconstruction and global event description with the CMS detector". *Journal of Instrumentation* 12.10 (2017), P10003–P10003. DOI: 10.1088/1748-0221/12/10/p10003.

[68] The CMS Collaboration. "Description and performance of track and primary-vertex reconstruction with the CMS tracker". *Journal of Instrumentation* 9.10 (2014), P10009–P10009. ISSN: 1748-0221. DOI: 10.1088/1748-0221/9/10/p10009.

[69] Kenneth Rose. "Deterministic annealing for clustering, compression, classification, regression, and related optimization problems". *Proceedings of the IEEE* 86.11 (1998), pp. 2210–2239.

[70] R Frühwirth, Pascal Vanlaer, and Wolfgang Waltenberger. *Adaptive vertex fitting*. Tech. rep. CERN-CMS-NOTE-2007-008, 2007.

[71] CMS Collaboration. "Performance of tau-lepton reconstruction and identification in CMS". *Journal of Instrumentation* 7.01 (2012), P01001–P01001. ISSN: 1748-0221. DOI: 10.1088/1748-0221/7/01/p01001.

[72] Matteo Cacciari, Gavin P Salam, and Gregory Soyez. "The anti-ktjet clustering algorithm". *Journal of High Energy Physics* 2008.04 (2008), pp. 063–063. DOI: 10.1088/1126-6708/2008/04/063.

[73] A. M. Sirunyan et al. "Evidence for Higgs boson decay to a pair of muons". *Journal of High Energy Physics* 2021.1 (2021). ISSN: 1029-8479. DOI: 10.1007/jhep01(2021)148.

[74] A.M. Sirunyan et al. "A measurement of the Higgs boson mass in the diphoton decay channel". *Physics Letters B* 805 (2020), p. 135425. ISSN: 0370-2693. DOI: 10.1016/j.physletb.2020.135425.

[75] *CMS Publication*. URL: http://cms-results.web.cern.ch/cms-results/public-results/publications/.

[76] *CMS Supersymmetry Physics Results*. URL: https://twiki.cern.ch/twiki/bin/view/CMSPublic/PhysicsResultsSUS.

[77] ATLAS and CMS Collaborations. *Report on the Physics at the HL-LHC and Perspectives for the HE-LHC*. 2019. arXiv: 1902.10229 [hep-ex].

[78] *Projections for measurements of Higgs boson cross sections, branching ratios, coupling parameters and mass with the ATLAS detector at the HL-LHC*. Tech. rep. Geneva: CERN, 2018.

[79] *Sensitivity projections for Higgs boson properties measurements at the HL-LHC*. Tech. rep. Geneva: CERN, 2018.

[80] *Measurement prospects of the pair production and self-coupling of the Higgs boson with the ATLAS experiment at the HL-LHC*. Tech. rep. Geneva: CERN, 2018.

[81] *Prospects for HH measurements at the HL-LHC*. Tech. rep. Geneva: CERN, 2018.

[82] *Prospects for searches for staus, charginos and neutralinos at the high luminosity LHC with the ATLAS Detector*. Tech. rep. Geneva: CERN, 2018.

[83] Collaboration CMS. *A MIP timing detector for the CMS phase-2 upgrade*. Tech. rep. 2019.

[84] L Rossi. "LHC Upgrade Plans: Options and Strategy" (2011), TUYA02. 6 p.

[85] CMS collaboration et al. "Projected Performance of an Upgraded CMS Detector at the LHC and HL-LHC: Contribution to the Snowmass Process". *arXiv preprint arXiv:1307.7135* (2013).

[86] Wolfgang Adam et al. "The CMS Phase-1 pixel detector upgrade". *Journal of Instrumentation* 16.02 (2021), P02027.

[87] J Man et al. *CMS technical design report for the phase 1 upgrade of the hadron calorimeter*. Tech. rep. Fermi National Accelerator Lab.(FNAL), Batavia, IL (United States), 2012.

[88] A Colaleo et al. *CMS technical design report for the muon endcap GEM upgrade*. Tech. rep. 2015.

[89] Manfred Jeitler, CMS collaboration, et al. "The upgrade of the CMS trigger system". *Journal of Instrumentation* 9.08 (2014), p. C08002.

[90] Michael Moll. "Radiation damage in silicon particle detectors. Microscopic defects and macroscopic properties" (1999).

[91] Ren-yuan Zhu. "Radiation damage in scintillating crystals". *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 413.2-3 (1998), pp. 297–311.

[92] Francesca Cavallari and Chiara Rovelli. "Calibration and Performance of the CMS Electromagnetic Calorimeter in LHC Run2". *EPJ Web of Conferences*. Vol. 245. EDP Sciences. 2020, p. 02027.

[93] *The Phase-2 Upgrade of the CMS Tracker*. Tech. rep. Geneva: CERN, 2017.

[94] G. Kramberger et al. "Performance of silicon pad detectors after mixed irradiations with neutrons and fast charged hadrons". *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 609.2 (2009), pp. 142–148. ISSN: 0168-9002. DOI: https://doi.org/10.1016/j.nima.2009.08.030.

[95] Riccardo Paramatti, CMS Collaboration, et al. "Design options for the upgrade of the CMS electromagnetic calorimeter". *Nuclear and particle physics proceedings* 273 (2016), pp. 995–1001.

[96] Rihua Mao et al. "Crystal growth and scintillation properties of LSO and LYSO crystals". *Journal of Crystal Growth* 368 (2013), pp. 97–100. ISSN: 0022-0248. DOI: https://doi.org/10.1016/j.jcrysgro.2013.01.038.

[97] The CALICE collaboration. "Design and electronics commissioning of the physics prototype of a Si-W electromagnetic calorimeter for the International Linear Collider". *Journal of Instrumentation* 3.08 (2008), P08001–P08001. DOI: 10.1088/1748-0221/3/08/p08001.

[98] G. Pellegrini et al. "Technology developments and first measurements of Low Gain Avalanche Detectors (LGAD) for high energy physics applications". *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 765 (2014). HSTD-9 2013 - Proceedings of the 9th International "Hiroshima" Symposium on Development and Application of Semiconductor Tracking Detectors, pp. 12–16. ISSN: 0168-9002. DOI: https://doi.org/10.1016/j.nima.2014.06.008.

[99] Cristián H. Peña and. "Precision Timing with the CMS MIP Timing Detector". *Journal of Physics: Conference Series* 1162 (2019), p. 012035. DOI: 10.1088/1742-6596/1162/1/012035.

[100] PICMG. *AdvancedTCA*.

[101] Gianluigi Casse. "P-type sensor results from the RD50 collaboration". *Proceedings of the 16th International Conference on Vertex Detectors (Vertex2007)*. 2007.

[102] G. Casse et al. "Measurements of charge collection efficiency with microstrip detectors made on various substrates after irradiations with neutrons and protons with different . . . " *PoSVERTEX* (Jan. 2008).

[103] G Bombardi et al. "HGCROC-Si and HGCROC-SiPM: the front-end readout ASICs for the CMS HGCAL". *2020 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC)*. IEEE. 2020, pp. 1–4.

[104] S Callier et al. "SKIROC2, front end chip designed to readout the Electromagnetic CALorimeter at the ILC". *Journal of Instrumentation* 6.12 (2011), pp. C12040–C12040. DOI: 10.1088/1748-0221/6/12/c12040.

[105] Paul Rubinov. *Development of Flexible, Scalable, Low Cost Readout for Beam Tests of High Granularity Calorimeter for the CMS Endcap*. Tech. rep. Geneva: CERN, 2016.

[106] Harold M. Lucal. "Arithmetic Operations for Digital Computers Using a Modified Reflected Binary Code". *IRE Transactions on Electronic Computers* EC-8.4 (1959), pp. 449–458. DOI: 10.1109/TEC.1959.5222057.

[107] N. Akchurin et al. 13.10 (2018), P10023–P10023. DOI: 10.1088/1748-0221/13/10/p10023.

[108] "56 - ON THE ENERGY LOSS OF FAST PARTICLES BY IONISATION". *Collected Papers of L.D. Landau*. Ed. by D. TER HAAR. Pergamon, 1965, pp. 417–424. ISBN: 978-0-08-010586-4. DOI: https://doi.org/10.1016/B978-0-08-010586-4.50061-4.

[109] J. Borg et al. "SKIROC2_CMS an ASIC for testing CMS HGCAL". *Journal of Instrumentation* 12.02 (2017), pp. C02019–C02019. DOI: 10.1088/1748-0221/12/02/c02019.

[110] B. Acar et al. "The DAQ system of the 12,000 channel CMS high granularity calorimeter prototype". *Journal of Instrumentation* 16.04 (2021), T04001. DOI: 10.1088/1748-0221/16/04/t04001.

[111] B Acar et al. "Construction and commissioning of CMS CE prototype silicon modules". *Journal of Instrumentation* 16.04 (2021), T04002.

[112] The CALICE Collaboration. *Design, Construction and Commissioning of a Technological Prototype of a Highly Granular SiPM-on-tile Scintillator-Steel Hadronic Calorimeter*. Manuscript in preparation.

[113] Howard Baer et al. "The international linear collider technical design report-volume 2: physics". *arXiv preprint arXiv:1306.6352* (2013).

[114] J Spanggaard. *Delay wire chambers-a users guide*. Tech. rep. CERN-SL-Note-98-023-BI, 1998.

[115] Thorben Quast, André David Tinoco Mendes, and Martin Erdmann. *Qualification, performance validation and fast generative modelling of beam test calorimeter prototypes for the CMS calorimeter endcap upgrade*. Tech. rep. Fachgruppe Physik, 2020.

[116] Dipanwita Banerjee et al. "The North Experimental Area at the Cern Super Proton Synchrotron" (2021).

[117] N. Charitonidis and B. Rae. *H2 beam line*. 2017.

[118] The CMS HGCAL Collaboration. *Response of a CMS HGCAL silicon-pad electromagnetic calorimeter prototype to 20-300 GeV positrons*. 2021. arXiv: 2111.06855 [physics.ins-det].

[119] Ufuk Guney Tok et al. *SPS H2 Beam Line Simulation: Energy Loss and Material Impact at HGCAL Test Beam*.

[120] *OMEGA Group*. URL: https://portail.polytechnique.edu/omega/en.

[121] W. Verkerke and D. Kirkby. *RooFIT User Manual*. 2019.

[122] VV Avakian et al. "Determination of the cross section of the pion and nucleon interaction with iron nuclei in the energy range 0.5–5.0 TeV". *Nuclear Physics B* 259.1 (1985), pp. 163–169.

[123] B Bilki et al. "Pion and proton showers in the CALICE scintillator-steel analogue hadron calorimeter". *Journal of Instrumentation* 10.04 (2015), P04014.

[124] T.A. Gabriel et al. "Energy dependence of hadronic activity". *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 338.2 (1994), pp. 336–347. ISSN: 0168-9002. DOI: https://doi.org/10.1016/0168-9002(94)91317-X.

[125] T. Åkesson et al. "Performance of the Uranium / Plastic Scintillator Calorimeter for the Helios Experiment at CERN". *Nucl. Instrum. Meth. A* 262 (1987), p. 243. DOI: 10.1016/0168-9002(87)90862-X.

[126] *TProfile implementation in ROOT framework*. URL: https://root.cern.ch/doc/master/classTProfile.html.