Fractal dimension analysis in a highly granular calorimeter

M Ruan, V Boudry, J-C Brient, D Jeans, H Videau

Laboratoire Leprince-Ringuet,
École polytechnique, CNRS/IN2P3, Palaiseau, France

E-mail: `Manqi.Ruan@llr.in2p3.fr`

Abstract.
The concept of "particle flow" has been developed to optimise the jet energy resolution by distinguishing the different jet components. A highly granular calorimeter designed for the particle flow algorithm provides an unprecedented level of detail for the reconstruction of calorimeter showers and enables new approaches to shower analysis. In this paper the measurement and use of the fractal dimension of showers is described. The fractal dimension is a characteristic number that measures the global compactness of the shower. It is highly dependent on the primary particle type and energy. Its application in identifying particles and estimating their energy is described in the context of a calorimeter designed for the International Linear Collider.

1. Introduction

1.1. Linear Collider and a highly granular calorimeter

In recent years, the concept of a highly granular calorimeter has been studied in the context of a proposed e^+e^- collider, either the International Linear Collider (ILC)[1] or the Compact Linear Collider (CLIC)[2]. At these colliders, many interesting physics channels will involve multi-boson production, and a good separation between the Z, W and Higgs bosons will be crucial. As many of their decay channels contain jets in the final states, a detector providing good jet energy resolution is required. For example, to separate different hadronic final states when looking at WW scattering (i.e, the $ZZ\nu\nu$ and $WW\nu\nu$ events), a jet energy resolution better than 3% (sometimes referred to as $30\%/\sqrt{E_J/\mathrm{GeV}}$) is required. This is roughly a factor of two better than previous experiments (e.g, ALEPH), and is one of the biggest challenges for the detector design. This goal has now been achieved at the full simulation level by using a Particle Flow Algorithm (PFA)[3]. The key idea is to measure the energy of each jet particle individually in the best suited subdetector: charged particles with tracker devices, photons with the electromagnetic calorimeter (ECAL) and neutral hadrons with the ECAL and hadron calorimeter (HCAL).

The most crucial task for a PFA is to separate showers from different particles in the calorimeter, in order to avoid losses and double counting. On average, 65% of the jet energy is in charged particles, 26% in photons (γ) and 9% in neutral hadrons (h^0). Charged particles and photons can be measured to much higher precision in the detector than neutral hadrons. For example, at the International Large Detector (ILD, one of the proposed ILC detectors), the momenta of charged particles can be measured to a precision of $\delta(1/P_T) = 2.0 \times 10^{-5}\mathrm{GeV}^{-1}$, the energy of photon measured to $\delta(E_\gamma)/E_\gamma \sim 15\%/\sqrt{E_\gamma/\mathrm{GeV}}$, and neutral hadrons measured to $\delta(E_{h^0})/E_{h^0} \sim 50\%/\sqrt{E_{h^0}/\mathrm{GeV}}$ [4]. Therefore, for a perfect PFA with no confusion between these components, the ultimate jet energy resolution can reach $\delta E_{Jet}/E_{Jet} \sim 14\%/\sqrt{E_{Jet}/\mathrm{GeV}}$ at typical precision of subdetectors [5].

The confusion contribution dominates the jet energy resolution, and to reduce it, the calorimeter granularity is more important than the intrinsic single particle energy resolution. The calorimeters designed for PFA typically have of order 10^8 readout channels with a volume of $1\,\mathrm{cm}^3$ each. A sampling structure is used to provide a high granularity in the longitudinal direction, and each transverse sensor layer is divided into numerous cells. Calorimeters with such high granularity are referred to as PFA oriented or imaging calorimeters.

1.2. Shower Fractal Dimension

Many objects in nature have a self-similar pattern, or in other words "a rough or fragmented geometric shape that can be split into parts, each of which is, or at least approximately, a reduced-size copy of the whole" [6]. A self-similar set is also known as a fractal. The Fractal Dimension (FD) is a characteristic property of a fractal. It is defined as:

$$FD = -\lim_{\epsilon \to 0} \frac{\log N_\epsilon}{\log(\epsilon)}, \tag{1}$$

where N_ϵ represents the number of subsets at a scale ϵ.

A particle shower is a cascade of particles produced when an energetic primary particle interacts with dense matter, producing multiple daughter particles with less energy. Each of these then interacts in a similar way. This process stops when particles are absorbed in matter. Because each shower particle with sufficient energy can produce multiple daughter particles, the shower has a self-similar pattern, and can be regarded as a fractal. Thus a shower can be described by its fractal dimension.

Based on the type of interactions, showers can be classified into three different kinds: the minimum ionization particle (MIP) shower/track, the electromagnetic shower and the hadronic shower. Electromagnetic showers (composed of e^+, e^- and photons) are developed by pair-production and bremsstrahlung, whose large cross sections ensure the compactness of electromagnetic showers. Hadronic showers are governed by the production of hadrons (pions, kaons, protons, neutrons, etc.) and ions in the hard collisions with nuclei. Therefore they are composed of long travelling hadrons and compact clusters from nuclear fragments and electromagnetic showers stemming from π^0 decays.

2. Measurement of shower fractal dimension in a highly granular calorimeter

Showers of different particle types (π^+, μ^+, e^+, K_L^0, proton and neutron, see Fig. 1) with energies from 2 GeV to 40 GeV were simulated with Mokka[7], the official GEANT4 [8] based full simulation software for ILD. The simulated detector was only the Digital Hadron Calorimeter (DHCAL) in a solenoidal magnetic field of 3.5 T. The DHCAL is a sampling Resistive Plate Chamber(RPC)-iron HCAL with 48 longitudinal layers, each 26.5 mm thick with a 20 mm iron absorber and a 6.5 mm RPC sensor layer. In the simulation, each RPC layer is divided into $1 \times 1\,\mathrm{mm}^2$ cells (modified from the original design of $10 \times 10\,\mathrm{mm}^2$). This cell size is chosen to record detailed information about shower development information. In the following analysis, only digital information has been used, meaning only one bit is used for each channel to record whether this channel was hit or not.

The effective cell size can be varied by grouping $\alpha \times \alpha$ nearby cells. α then defines the scale at which the shower is analyzed, see Fig. 2. The shower fractal dimension can be estimated by

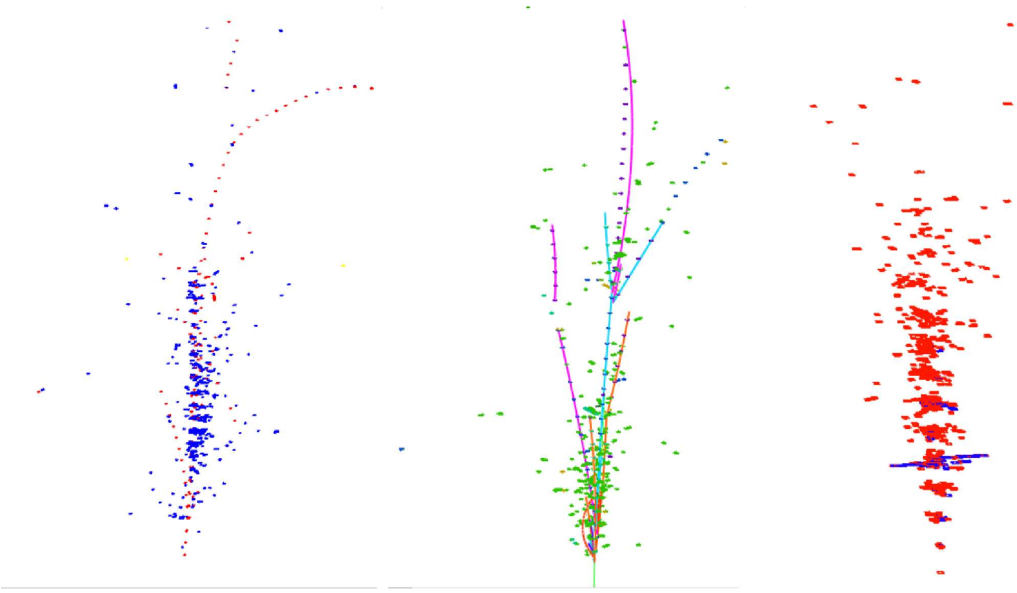


Figure 1. From left to right: simulated shower of 40 GeV π^+, K_L^0 and e^+.

exploring the variation of the number of hits with the scale. Defining N_α as the number of hits at a scale α, the ratio of hit counts at different scales can be written as:

$$R_{\alpha,\beta} = N_\beta / N_\alpha. \tag{2}$$

At a fixed β, the ratio $R_{\alpha,\beta}$ is a function of α. In the following discussing β will be called the initial scale. To make full use of the recorded information, β is usually fixed at the initial cell size. Fig. 3 shows the correlation between $R_{\alpha,1}$ and the scale α for different samples. An approximate linear correlation with α is observed in a double logarithmic scale (especially for the e^+ and π^+ samples), thus we define the shower fractal dimension as the average slope + 1, where +1 is added for the longitudinal degree of freedom:

$$FD_\beta = \left\langle \frac{\log(R_{\alpha,\beta})}{\log(\alpha)} \right\rangle + 1. \tag{3}$$

The shower fractal dimension depends on the type of the primary particle. A MIP has a trajectory close to a helix, and its fractal dimension is close to 1 (in the ideal case, a MIP track has only one hit per layer). On the other hand, electromagnetic showers (e^+ samples) have the largest fractal dimension because of their compactness. Hadronic showers (for example π^+ samples) are composed of MIP tracks from long travelling charged hadrons and compact clusters from electromagnetic interactions and heavy nuclear fragments, thus their fractal dimension lies in between.

Typical values for the shower fractal dimension for 40 GeV particles are (see Fig. 6):

$$FD_{1\mathrm{mm}}(\mu) = 1.2,$$

$$FD_{1\mathrm{mm}}(\pi^+) < FD_{1\mathrm{mm}}(K_L^0) = 1.5,$$
$$FD_{1\mathrm{mm}}(e^+) = 1.75.$$

At the same particle energy, the fractal dimension of the K_L^0 showers is slightly larger than that of the π^+ showers, as the K_L^0 showers don't have an initial MIP track before the first interaction.

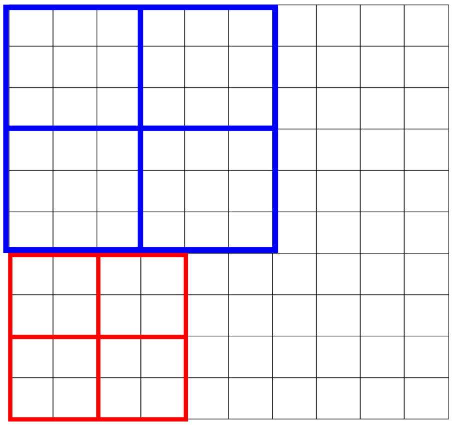


Figure 2. Basic method: group nearby cells.

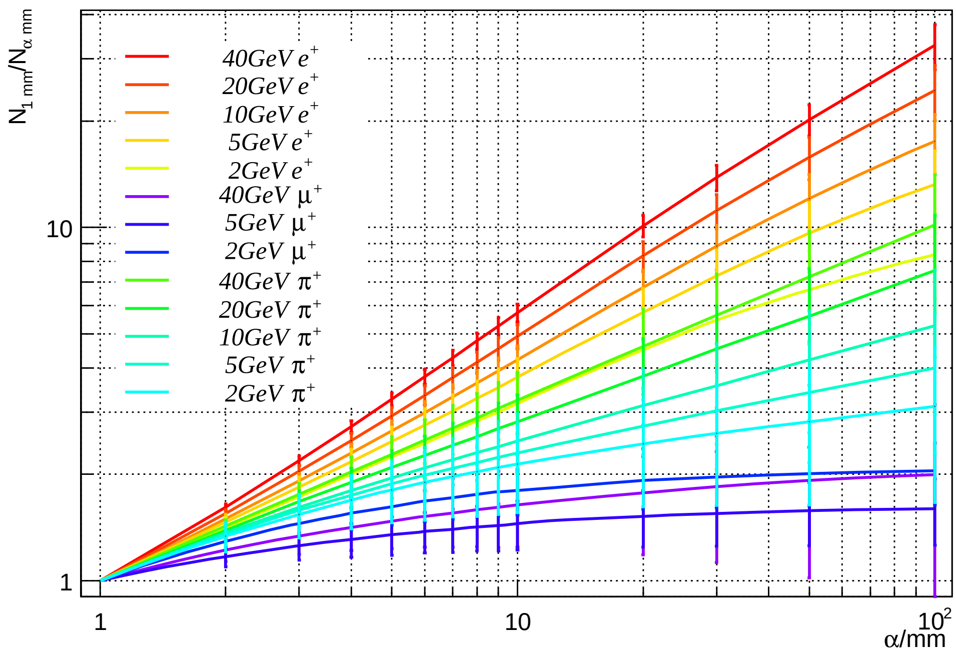


Figure 3. Dependence of hit count ratio and scale for different samples (the error bar scaled to rms, 1k events for each sample).

The shower fractal dimension also depends on the energy of the primary particle. The volume of a shower increases logarithmically with the primary particle energy while, to first order, the number of hits is proportional to the particle energy (ignoring the effects of saturation). A higher energy shower is therefore more compact, and has a larger fractal dimension (except for a MIP). The fractal dimension of electromagnetic and hadronic showers follows approximately (see Fig. 4):

$$FD_{em,1\mathrm{mm}}(E) = 1.41 + 0.21 \times \log_{10}(E/\mathrm{GeV}), \tag{4}$$

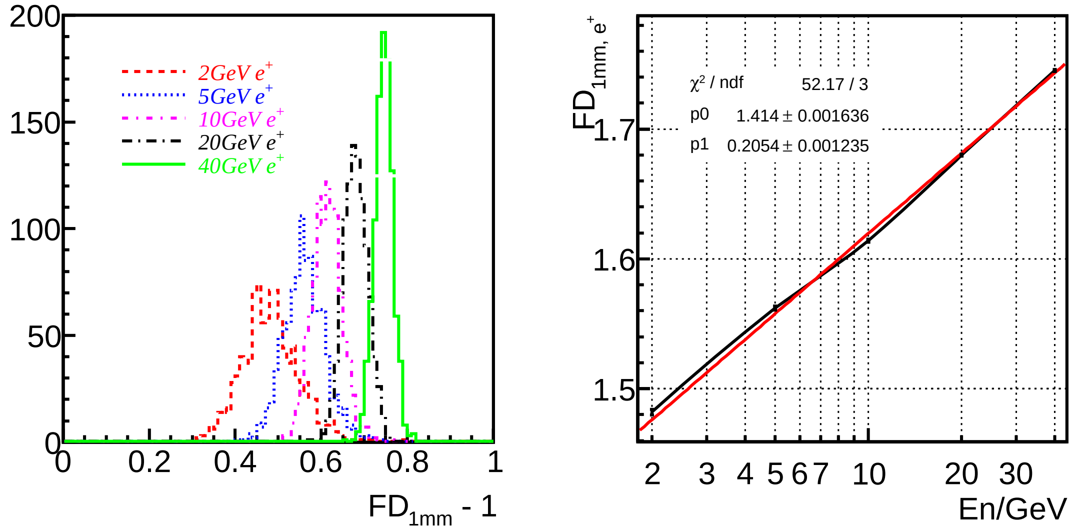


Figure 4. Left: Fractal dimension for e^+ samples at different energies;
Right: Correlation between fractal dimension and particle energy for e^+ samples.

$$FD_{had,1\mathrm{mm}}(E) = 1.24 + 0.15 \times \log_{10}(E/\mathrm{GeV}). \tag{5}$$

3. Application of shower fractal dimension

The fact that the shower fractal dimension is sensitive to the particle type and energy makes it a potential tool for particle identification and shower energy estimation.

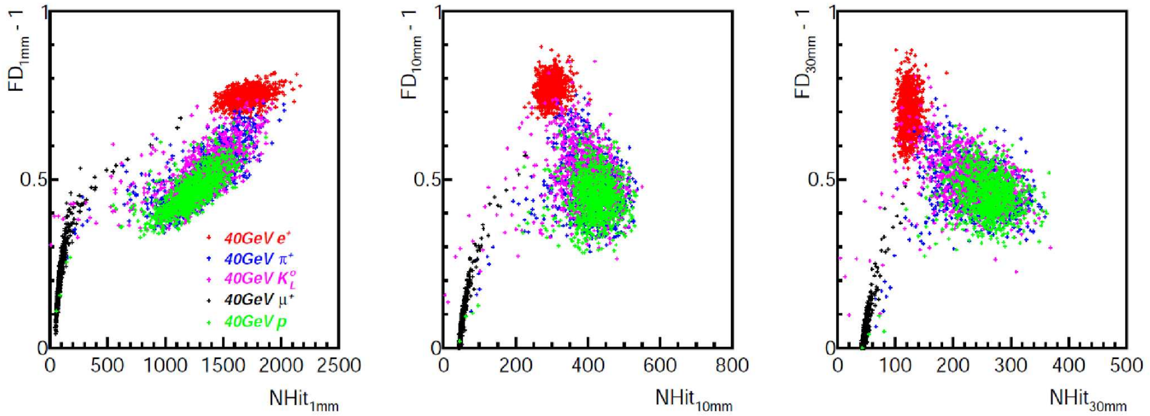
3.1. Particle identification

A particle identification technique purely based on the calorimeter information has been developed. It makes use of two input variables: the measured fractal dimension of the shower and the number of hits. This technique has been tested on different data samples at an energy of 40 GeV, see Fig. 5. To study its performance at different cell size, these two observables have been measured at three different initial scales: 1 mm, 10 mm and 30 mm.

A clear separation between muons, hadrons and positrons is observed at each initial scale. Using some simple cuts (for example, at initial scale of 1 mm, showers with $FD \geqslant 0.68$ are identified as electromagnetic showers, showers with $FD < 0.68$ and $N_{hits} > 500$ are identified as hadronic showers, and all the rest are identified as MIPs), the tables in Fig. 5 show the output of this particle identification, where the rows present input samples and columns present the output for 1000 events. In fact, the separation is clear at each initial scale and similar performance is achieved at different scales.

These results are achieved with single particles at a fixed energy, thus in a full event, the performance of this method will be affected by the performance of the PFA to identify individual showers and the measurement of the shower energy. However, for well isolated particles, especially charged particles whose track momentum can be taken as a measurement of their energy, this performance should be easily obtained.

The evolution of the fractal dimension with the initial scale can provide more information. Fig. 6 shows the measured fractal dimension for the same data samples at three different initial scales. Because of the compactness of electromagnetic showers, the e^+ samples have a narrower



1mm	e+	u	h
e+	998	0	2
u	1	994	5
h	15	14	971

10mm	e+	u	h
e+	1000	0	0
u	0	995	5
h	17	14	969

30mm	e+	u	h
e+	1000	0	0
u	0	996	4
h	18	11	971

Figure 5. Particle identification with the fractal dimension of shower and the number of hits.

fractal dimension distribution at smaller initial scale, and get significantly smeared at larger initial scales. On the other hand, with increasing cell size, the fractal dimension distribution of the μ^+ sample becomes narrower, providing a better muon-hadron separation. The reason is that at large cell size, the fractal dimension is less sensitive to soft bremsstrahlung photons and noise (especially the noise caused by low momentum charged particles trapped in the gas gap by the magnetic field), see Fig. 7.

At any scale, the fractal dimension of π^+ samples is a continuous distribution which lies between those of μ^+ and e^+ of the same energy. The pion events with comparable fractal dimension to muons are due to pion decay ($\pi^+ \to \mu^+ + \nu_\mu$) before reaching the calorimeter, resulting in a MIP cluster with small fractal dimension. Events with large fractal dimension are cases in which the majority of energy is deposited electromagnetically: for example, a charged pion can convert into π^0 through isospin exchange ($\pi^+ + n \to \pi^0 + p$). This can be distinguished from e^+, e^- or γ by tagging the position of the first interaction point, or by measuring the fractal dimension at different depths in the calorimeter, see Fig. 7.

In the μ^+ sample, the fractal dimension of the shower can be as large as that of the electromagnetic shower through hard bremsstrahlung. More interestingly, in viewing the fractal dimension measured at different scales, some of the muon events have small fractal dimension at large scale, while having a large fractal dimension at small scale. As explained above, these showers have some micro structure, which is not resolved at large scale, see Fig. 7.

The fractal dimension depends strongly on the type of interaction, and is therefore a promising tool for particle identification. It also has the capability to tag some of the details of shower development.

3.2. Energy Estimation

A strong anti-correlation is observed between the fractal dimension measured at small cell size and the number of hits at a large scale. For example, the left plot in Fig. 8 shows the distribution of the fractal dimension measured from 10 mm cells and the number of hits at 30 mm cells for

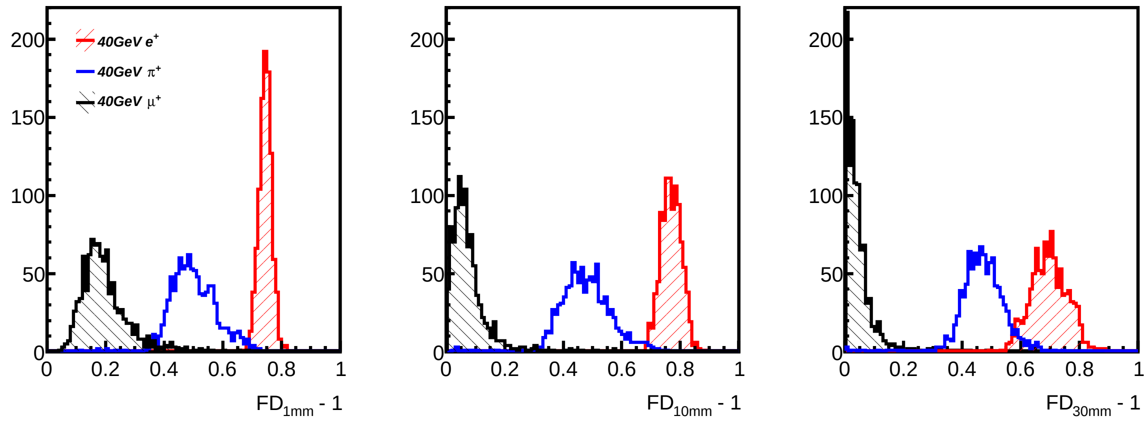


Figure 6. Fractal dimension measured from 1, 10 and 30 mm cell for μ^+, e^+ and π^+ samples.

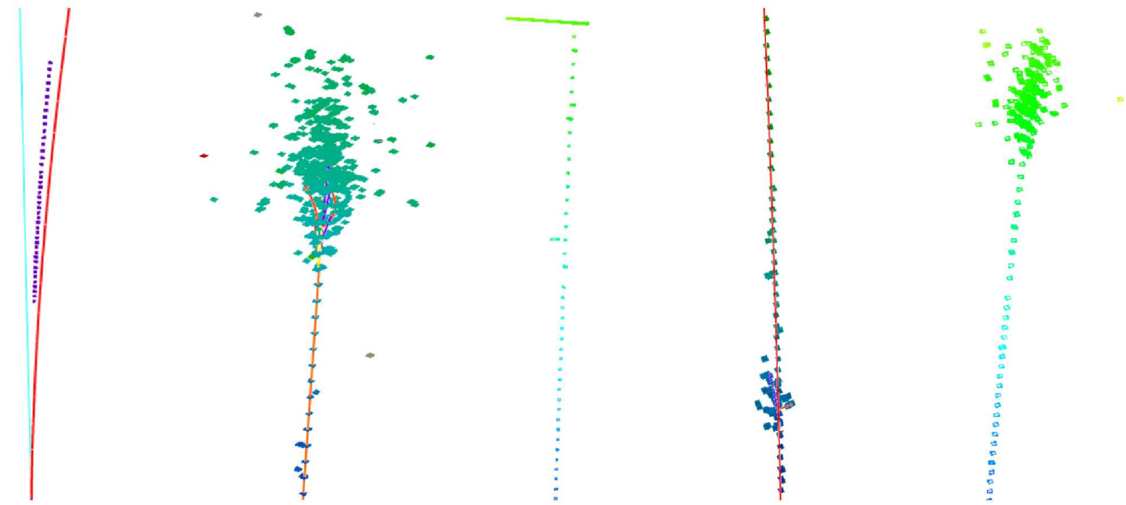


Figure 7. Extreme cases for π^+ and μ^+ shower, from left to right: pion decay, isospin exchange, noise, soft and hard bremsstrahlung radiation of μ^+.

the 40 GeV μ^+, π^+ and e^+ samples. This correlation allows a linear combination of these two observables to be used as an energy estimator:

$$E = N_1 \times (Nhit_{30\mathrm{mm}} + c_1 \times FD_{\beta\mathrm{mm}}), \beta < 30, \tag{6}$$

where N_1 is a calibration constant, $Nhit_{30\mathrm{mm}}$ is the number of hits at 30 mm cell size, and $FD_{\beta\mathrm{mm}}$ is the shower fractal dimension measured at the cell size of β mm.

The right plot of Fig. 8 shows the distribution of measured energy with different estimators for 40 GeV π^+ sample (with mean value of the distribution normalized to 1). Using the energy estimator shown in Eq. 6, the shower energy resolution can be improved by roughly a factor of two comparing to hit counting at any scale. Unfortunately, the correlation coefficient c_1 depends on the particle energy. This energy resolution can only be obtained if the energy of the primary particle is known. Therefore, the application of the fractal dimension on energy estimation is two-fold. First, for charged particles, this phenomenon can be used to improve the track-cluster

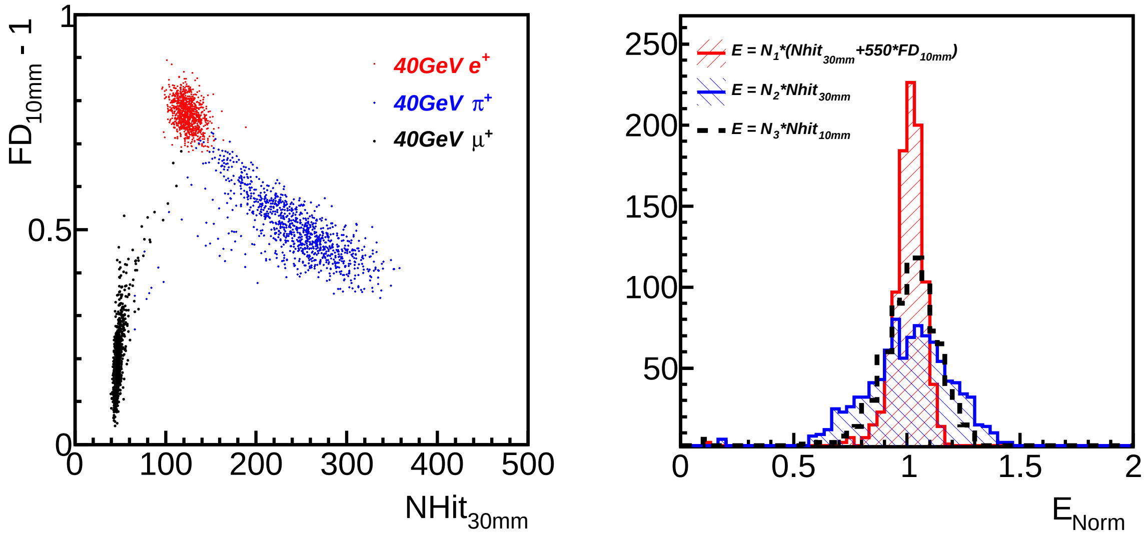


Figure 8. Left: Correlation between $FD_{10\,\mathrm{mm}} - 1$ and $Nhit_{30\mathrm{mm}}$ for different samples. **Right:** Energy measurement with different estimators for 40 GeV π^+ sample.

matching in the PFA, which means the comparison of the position and momentum of tracks to a given cluster, to check if the cluster should be linked to a track and reconstructed as a charged particle. In this case, the track momentum can be taken as a reference of the injected particle energy, with which the correlation coefficient can be determined, and a better energy estimation can be achieved. Secondly, to first order, the correlation coefficient c_1 is proportional to the particle energy. Eq. 6 can be written as:

$$E = N_1 \times (Nhit_{30\mathrm{mm}} + \delta \times E \times FD_{\beta\mathrm{mm}}), c_1 = \delta \times E.$$

From which the energy of the shower can be resolved:

$$\begin{aligned} E &= N_1 \times (Nhit_{30\mathrm{mm}}/(1 - \delta \times N_1 \times FD_{10\mathrm{mm}})) \\ &= N_1 \times (Nhit_{30\mathrm{mm}}/(1 - c_2 \times FD_{10\mathrm{mm}})), c_2 = \delta \times N_1, \end{aligned} \quad (7)$$

where c_2 is a constant. The right hand side of Eq. 7 is independent of the primary particle energy, therefore it can be used for the neutral particle energy measurement.

The performances of the different estimators tested on K_L^0 samples have been compared. The distributions of the measured energy have been fitted with a Gaussian function. On the left plot of Fig. 9, the resolution is expressed as the ratio between the width and mean of the Gaussian distributions (σ/M). The right plot shows the linearity in terms of the ratio between the measured energy and true particle energy (E/E_{exp}), fixing this ratio to 1 at 80 GeV.

As the digital readout can not distinguish if a hit is induced by a single or multiple particles, at higher energy, the curve for number of hits at 10 mm has a significant saturation effect and a degraded resolution. However, this can be corrected using the fractal dimension. With a roughly tuned coefficient c_2, the resolution significantly improves at energies above 40 GeV and below 2 GeV, while it is worse in the range $10 - 20$ GeV. The linearity simultaneously improves to a level comparable to that of 5 mm.

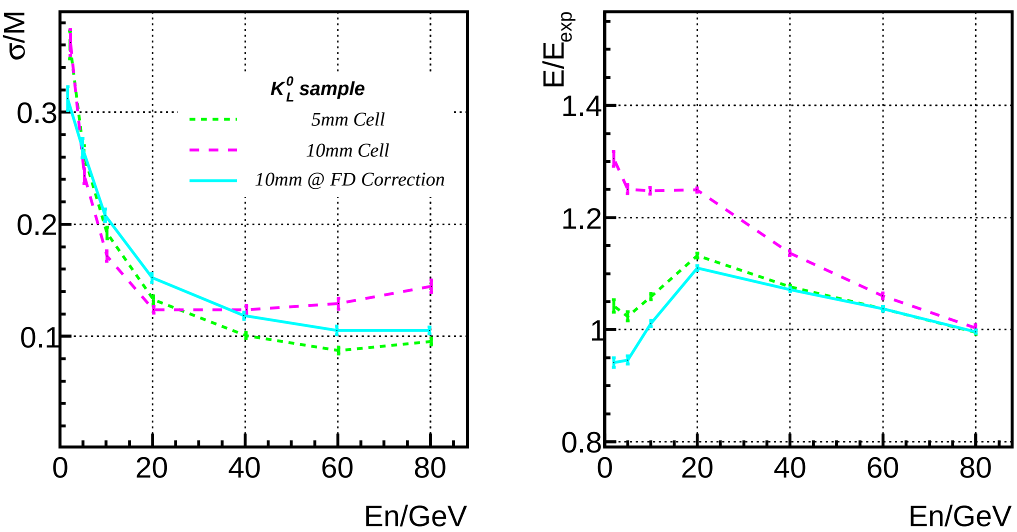


Figure 9. σ/M and linearity of different energy estimators. Pink curve: hit counts at 10 mm. Light blue curve: hit counts at 5 mm. Green curve: estimator of Eq. 7.

4. Summary

Showers created by particle interactions in material are fractal objects, to which a fractal dimension can be assigned. This shower fractal dimension can be measured with a highly granular calorimeter such as the ones designed for a future linear collider. The shower fractal dimension depends strongly on the type and energy of the incident particle, making it a potential tool for particle identification and shower energy estimation. Preliminary results from full simulation of the detector show that the fractal dimension can provide particle identification with excellent efficiency and purity on positrons, muons and pions at 40 GeV using only calorimetric information. In using the fractal dimension measured at different scales, this method can tag more detailed information of the shower development.

A strong anti-correlation between the fractal dimension at small scale and the number of hits at large scale is observed. For charged particles with a known track momentum, this correlation enables an energy resolution improved by a factor of up to two compared to hit counting, and can be used to enhance the track-cluster linking in the PFA. For neutral showers, the fractal dimension can improve the energy measurement in terms of linearity and resolution, especially at large energy where the saturation effects limit the performance of a digitally readout calorimeter with large or medium cell sizes.

5. Acknowledgments

The research leading to these results has received funding from the European Commission under the FP7 Research Infrastructures project AIDA, grant agreement no. 262025. We are grateful to the CALICE collaboration for their constructive comments, and to the linear collider software team, especially to G. Musat from LLR, for the development of many tools used in this study.

Reference

[1] http://www.linearcollider.org
[2] http://clic-study.org

[3] Thomson M A 2009 *Particle Flow Calorimetry and the PandoraPFA Algorithm* Nucl. Instrum. Meth. A 610, 25-40
[4] Abe T et al 2010 *The International Large Detector: Letter of Intent* DESY-2009-87
[5] Brient J-C 2009 *Particle Flow Algorithm and calorimeter design J. Phys.: Conf. Series* **160** 012025
[6] Mandelbrot B B 1982 *The Fractal Geometry of Nature. W.H. Freeman and Company* ISBN 0-7167-1186-9
[7] http://polzope.in2p3.fr:8081/MOKKA
[8] Agostinelli S 2003 *Geant4 a simulation toolkit* Nucl. Instrum. Meth. A 506, 250-303