Università degli Studi di Pavia

IDEA Dual-Readout calorimeter:
Full simulation chain and neural network on particle ID

Candidato

Alessandro Villa

Matricola 462495

Supervisore

Dott. Roberto Ferrari

Co-Supervisore

Dott. Lorenzo Pezzotti

Contents

Introduction

aaa

Chapter 1

Future colliders

aaa

1.1 Physics goals

aaa

1.2 Leptonic colliders

aaa

1.3 Hadronic colliders

aaa

1.4 IDEA Detector concept

aaa

1.4.1 Vertex detector

aaa

1.4.2 Drift chamber

aaa

1.4.3 Magnet system

aaa

1.4.4 Dual-readout calorimeter

aaa

1.4.5 Preshower and muon chambers

aaa

Chapter 2

Calorimetry and dual-readout

Calorimetry is an important detection principle in particle physics. Originally developed with astrophysical purpose for cosmic-ray studies, this method refers to the detection of particles and the measurement of their properties, using blocks of instrumented material. It was developed and perfected for accelerator-based particle physics experimentation primarily in order to measure the energy of particles. In these blocks, particles are fully absorbed and their energy transformed into a measurable quantity.
The incident particle interact with the detector (through electromagnetic or strong processes) producing a shower of secondary particles with progressively degraded energy. The energy deposited by the charged particles of the shower in the active material of the calorimeter, which can be detected in the form of charge or light, is used to measure the energy of the incident particle. Typical processes suitable to detect this energy are: ionization of the medium, scintillation light and the Cherenkov light produced by relativistic particles.

Calorimeters can be divided into two categories depending on the type of shower they are optimized to detect: electromagnetic calorimeters, used mainly to measure electrons and photons through their electromagnetic interactions (e.g. bremsstrahlung, pair production), and hadronic calorimeters, used to measure mainly hadrons through their strong and electromagnetic interactions.
Another classification can be made according to their construction technique defining sampling calorimeters and homogeneous calorimeters.
Homogeneous calorimeters are built of one type of material that performs both the main tasks: degrade the energy of the incident particles and provide the detectable signal.
Sampling calorimeters, instead, consist of alternating layers of an absorber,

a dense material used to perform energy degradation, and an active medium that generate the signal.

This chapter describes the physics behind both the electromagnetic and hadronic shower developments, provides a basic description of the energy response of these detectors and introduces the particular technique of the dual-readout, a modern concept of calorimeter that has the quality of overcome the non-compensating problem measuring both electromagnetic and hadronic showers through two different type of signals simultaneously (Cherenkov and scintillation light).

A more comprehensive descriptions of the topic can be found in the references [1] and [2].

2.1   Physics of shower development

A particle interact and lose part or whole of its energy traversing matter. During this process the medium get excited and heated up, is excited in this process, or heated up. From this feature the term calorimetry, literally meaning "heat measurement", was introduced.
The groundwork for the calorimetry is the interaction processes between particle and matter. They are the manifestation of the electromagnetic, the strong and, more rarely, the weak forces and they strongly depend on the energy and the nature of the incident particle, in addition to medium features.
The term particle shower refers to the production of a group of particles generated by the interaction of a primary particle with the matter. The processes and the consequent shower effects are the keys to deeply understand this topic.

2.1.1   Electromagnetic showers

Despite the complex mechanisms of particle-matter interaction, electromagnetic showers are produced via a small number of well understood QED processes. Charged particles (electrons and positrons) lose energy by ionization and by radiation, instead neutral ones (photons) are characterized by photoelectric effect, Compton scattering and pair production.

The first type of particle in its path through the medium ionize it under the condition of having an energy at least sufficient to release the atomic electrons from the Coulomb fields generated by the atomic nuclei (few of eV). The amount of energy released (in unit of path) by these particle is

predictable through the semi-empirical Bethe-Block formula restricted to electrons (and positrons) [3]:

$$-\frac{dE}{dx} = 2\pi N_a r_e^2 m_e c^2 \rho \frac{Z}{A} \frac{1}{\beta^2} \left[\ln \frac{\tau^2(\tau+2)}{2(I/m_e c^2)^2} - F(\tau) - \delta - 2\frac{C}{Z} \right] \qquad (2.1)$$

the stopping power (i.e. the quantity described by the formula) decreases as the particle energy increases ($\propto \beta^2$). Hence the ionization process is the greatest source of energy loss for particles with small energy.

The other energy loss process known as *bremsstrahlung* is the dominant source of energy loss by electrons and positrons at energies above $100\ MeV$. Relativistic electrons and positrons radiate photons as a result of the interaction between Coulomb and the atomic electric fields. The energy spectrum of these photons falls off as $1/E$ ranging till the primary particle energy, but in general most of the photons carry a small part of it. The process produces (usually small) changes in electron (or positron) direction. This is called Coulomb or multiple scattering.

At a fixed energy the relative importance of ionization and radiation losses depends on the medium and in particular on the electron density of the medium in which the shower develops. This density is in first approximation proportional to the (average) Z of the medium. The critical energy, i.e. the energy value at which the two processes have equal impact, is roughly inversely proportional to the Z value of the material:

$$\varepsilon_c = \frac{160\ MeV}{Z + 1.24}. \qquad (2.2)$$

An example of energy loss in copper by electron is sketched in figure 2.1, where the ionization and radiation contribution are separated.

The other particles that produce electromagnetic showers are photons. The interaction between photons and matter is mainly affected by three different processes: the photoelectric effect, Compton scattering and electron–positron pair production.

The photoelectric effect is the process that most likely occurs at low energies. It is characterized by an atom absorbing the photon and emitting an electron. Eventually the atom, left in an excited state, emits an Auger electrons or X-rays returning to the ground state. The photoelectric cross section strongly depends on the available number of electrons, and thus on the Z value of the absorber material. In particular it scales with Z^n, with the power n between 4 and 5. Meanwhile the photoelectric cross section rapidly decrease with greater energies, varying as E^{-3}. In this way the process rapidly loses its impact as the energy increases.

The Compton process is a scattering process where an impinging photon interact with an atomic electron transferring enough momentum and energy
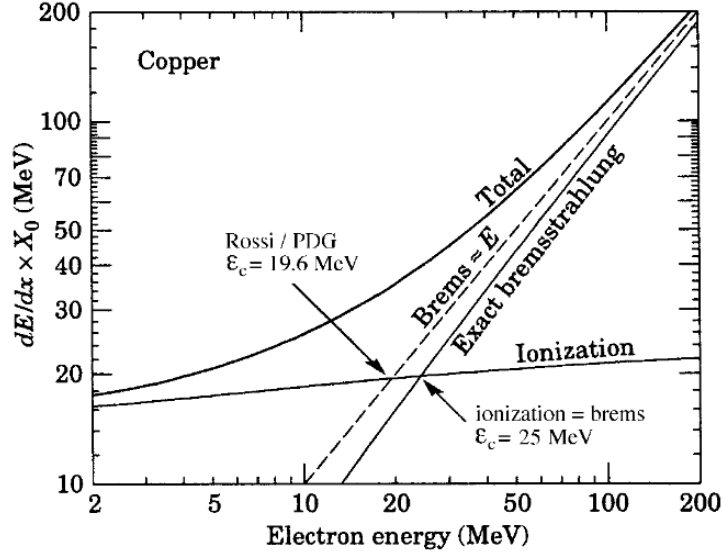
Figure 2.1: Energy losses through ionization and bremsstrahlung by electrons in copper. From [PDG 98].

to the struck electron to escape from the atomic Coulomb field. Kinematic variables such as energy transfer and scattering angles can be easily obtained applying the laws of energy and momentum conservation.

Photons in the MeV energy range are absorbed by photoelectric effect only after a sequence of Compton scattering processes, in which the photon energy is reduced step by step in each collision until it is low enough to permit the photoelectric occurrence. In each step, the amount of loss energy is:

$$T = E_\gamma \frac{\xi(1 - \cos\theta)}{1 + \xi(1 - \cos\theta)} \qquad (2.3)$$

where $\xi = E_\gamma/m_e c^2$. The cross section linked to the Compton scattering is much less dependent on the Z value than the photoelectric one. It is almost proportional to the number of target electrons in the nuclei. Also in this process the cross section decreases with increasing photon energy, but only with the first power of E. Therefore Compton scattering has more impact than photoelectric absorption above a certain threshold energy.

The pair production process, differently from the previous ones, has a binding threshold under which the effect can not occur. This threshold is the twice of the electron rest mass (2×511 MeV). If the photon has a higher energy then it can produce an electron-positron pair that can continue the path in the medium producing bremsstrahlung radiation as well as ionization.

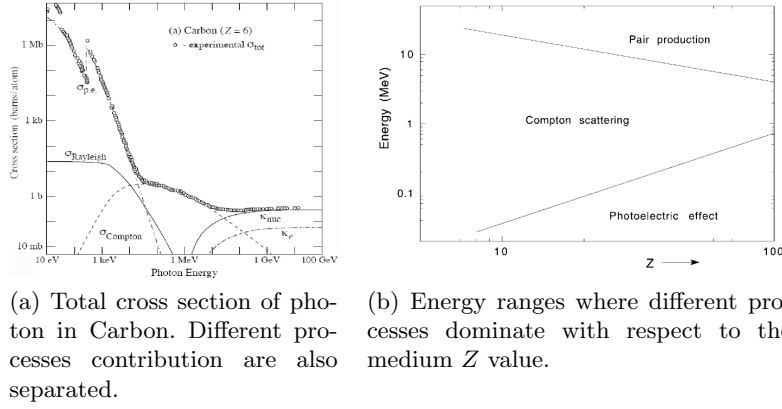The cross section for pair production rises with the energy reaching an

(a) Total cross section of photon in Carbon. Different processes contribution are also separated.

(b) Energy ranges where different processes dominate with respect to the medium Z value.

Figure 2.2: Images from [3].

asymptotic value at higher then 1 GeV. For this reason, at high energies, pair production is the most likely process to occur. Meanwhile the dependence from the medium goes, in first approximation, as Z^2.

Comparing the cross section of the three processes and its dependence with respect to the photon energy it is clear that the photoelectric effect dominates at lower energies, meanwhile at the intermediate values the Compton scattering gives the greatest contribute, at the end at higher energies almost every photon loses its energy through pair production. An example of these contributions is shown in figure 2.2a.
Knowing the dependence of the cross sections with respect to the Z of the material, ranges of energies where each process dominates can be found and parametrized with the Z value. A representation is sketched in figure 2.2b.

Shower principle

Minimal showers may also develop at vary low energy of primary particle. Starting for example from a photon of tens of MeV, it can eventually produce a electron-positron pair in the calorimeter. The charged particles lose their energy in the matter through ionization. When the positron loses all the kinetic energy, it annihilates with an electron producing two 511 keV γs. These photons are absorbed through the photoelectric effect after a sequence of Compton scattering. During the process, the energy of the primary particle is released to the material by charged particle in ionization processes. At energy values of 1 GeV and higher, electrons, protons and photons initiate actual electromagnetic showers in the materials in which they penetrate. At these energies charged particles lose their energy mostly by brehemstralung, the majority of these photons are very soft, and interact with Compton scattering until their absorption through photoelectric effect. Meanwhile the
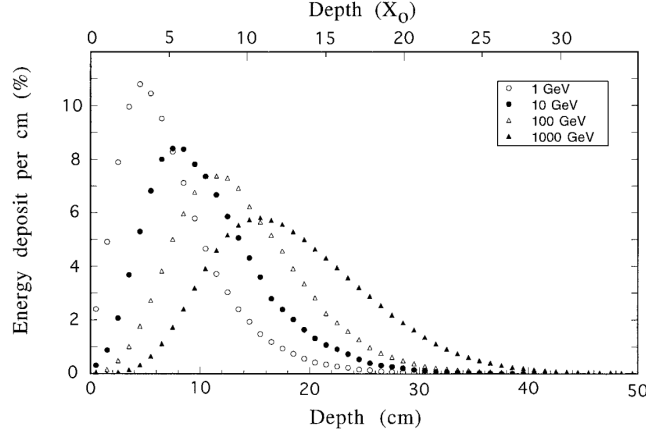
Figure 2.3: The energy deposit as a function of depth, for 1, 10, 100 and 1000 GeV electron showers developing in a block of copper [3].

photons with energy more than 5–10 MeV produce $e^+ - e^-$ pairs, which eventually radiate more γs. The process continues till the secondary particles energy is higher enough to sustain it. The shower maximum is defined as the point at which the number of shower particles produced in this particle multiplication process reaches a maximum. The depth inside the absorber associated to the shower maximum increases logarithmically with the energy of the primary particle (see figure 2.3). The longitudinal shower development is described by the radiation length (X_0), it is defined as distance at which the electron (or positron) loses on average 63% $(1 - e^{-1})$ of its energy by radiation. Expressing the shower containment in term of X_0 helps to mitigate the material-dependence effects.

Another quantity useful to describe the spatial shower development, in particular the transverse one, is the Molière radius. It is defined in terms of the radiation length and the critical energy:

$$\rho_M = E_s \frac{X_0}{\varepsilon_c} \tag{2.4}$$

where E_s is defined as $m_c^2 \sqrt{4\pi/\alpha} \simeq 21.2 \ MeV$. This quantity is almost material-independent and, on average, a cylindrical volume with this radius around the shower axis contains 90% of the shower energy. The lateral spread is mainly due to two effects: at high energy, electrons and positrons are moved away from the shower axis because of the deviation occurring in Compton scattering; photons and electrons are also produced in isotropic processes moving them away from the axis (spread more important in lower energy particles). Also brehemstralung process produces photons with a certain angle, contributing to the shower spread. Figure 2.4 give an idea of the radial distribution of deposited energy.
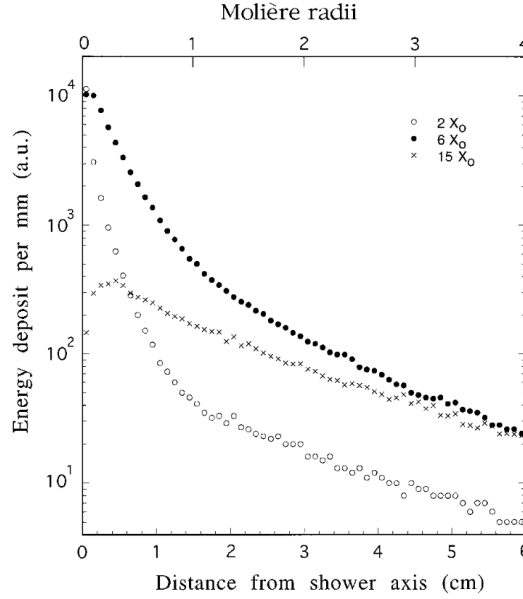
Figure 2.4: The radial distributions of the energy deposited by 10 *GeV* electron showers in copper, at various depths [3].

The lateral and longitudinal shower development generated by charged particles and by neutral ones are basically identical except for the initial stages. Electrons start radiating as soon as they enter the calorimeter, instead photons must convert before releasing any energy. Once they start producing electrons and positrons, they can release even more energy than electron induced showers. This behaviour is shown in Figure 2.5, where the distribution of the energy fraction deposited in the first 5 X_0 by 10 *GeV* electrons and photons in lead is plotted.

2.1.2   Hadronic showers

The hadronic showers, a new degree of complexity arises. The showers produced by hadrons are affected by the strong interaction. This interaction is responsible for:

- The production of hadronic secondary particles, most of which, $\simeq 90\%$, are pions. Neutral pions mainly decay in two photons, which develop electromagnetic showers.

- The occurrence of nuclear reactions where atomic nuclei release neutrons and protons. The fraction of the shower energy needed to unbind the nucleons does not contribute to the energy released to the calorimeter (invisible energy phenomenon).
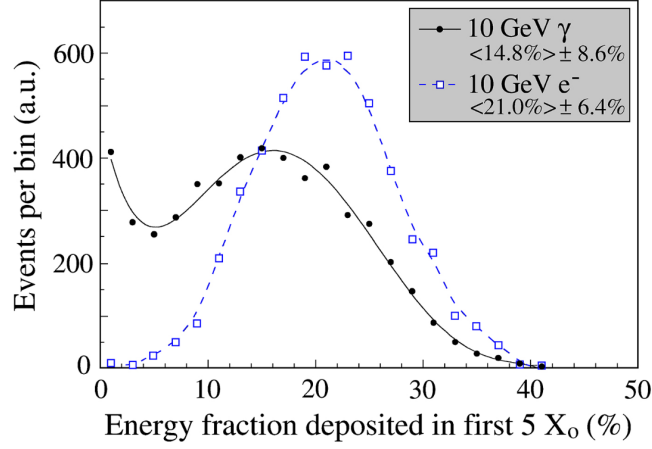
Figure 2.5: Distribution of the energy fraction deposited in the first 5 X_0 by 10 GeV electrons and photons showering in lead. Image from [4].

The part of energy released via electromagnetic showers is commonly named em fraction (f_{em}), and once the energy is converted to em showers it can not transformed back to hadronic energy. Another contribution to the complexity is given by the variability of this em fraction event by event and its energy dependence. On average, f_{em} increases with the primary particle energy, since πs may also be produced by secondary and higher-order particles: the higher the energy,the more generations of shower particles, the larger em fraction. The average electromagnetic fraction has been evaluated to increase with the energy following the power law:

$$f_{em} = 1 - 1 \left(\frac{E}{E_0} \right)^{k-1} \tag{2.5}$$

where $K \simeq 0.82$ and E_0 is a matter dependent value of the order of GeV. The behaviour in lead and copper is shown in figure 2.6.

In the calorimetry context, these characteristics have important consequences:

- *Non-compensation*; the calorimeter signals for hadrons and for electrons of the same energy are in general different a result of the invisible-energy phenomenon, in particular the one for hadrons is generally smaller.

- *Non-linearity*; a calorimeter for hadron detection does not present a linear response with respect to the primary particle due tio the energy dependence of the em fraction.

Another important difference between em and hadronic showers is the greater spatial profiles of released energy. Analyzing for example several
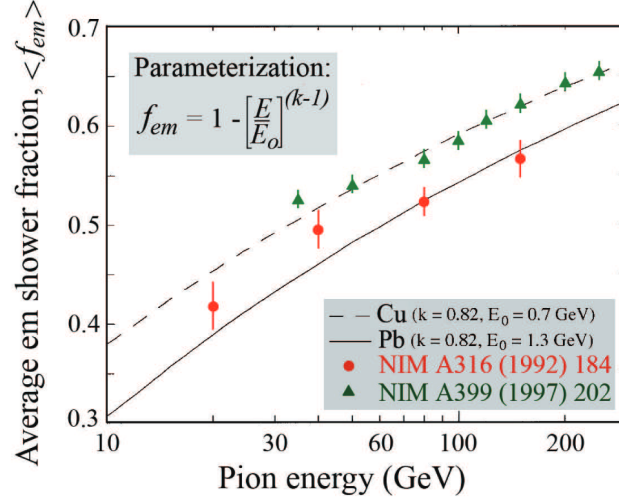
Figure 2.6: Comparison between the experimental results on the em fraction of pion-induced showers in copper-based and lead-based calorimeters. Image from [4].

pions traces, peaks of signals are produced at a depth near the one where a π^0 is generated. However, the neutral pion production can occur in the second or third generation of the shower development releasing energy at different depth. An example of 4 different showers from 270 GeV pions is illustrated in figure 2.7.

The depth of the calorimeter required to contain hadronic showers increases logarithmically with energy, as already seen for em showers, but the large longitudinal fluctuations make the leakage effect a point of extreme interest also in configuration that would contain, on average, 99% of the shower.

Laterally, an hadronic shower is easier contained if the primary particle has higher energy. This is due to the fact that the electromagnetic shower fraction increases with energy. The em showers produced tend to develop laterally close to the shower axis.

2.2 Energy response of calorimeters

Starting from the definition of the term *response*, it is the average calorimeter signal per unit of deposited energy. The response is usually expressed in terms, for example, of number of photoelectrons per GeV or charge (pC) per MeV. A straight forward definition is the linearity of a calorimeter, it is defined *linear* if its response is constant.
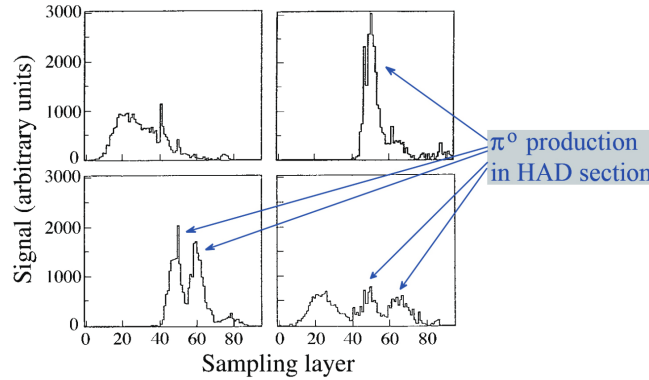
Figure 2.7: Longitudinal profiles for 4 different showers induced by 270 GeV pions in a lead/iron/plastic-scintillator calorimeter [4].

Electromagnetic calorimeters

Electromagnetic calorimeters are detector optimized to induce em showers. In these showers, all the energy carried by the primary particle is released by a very large number of shower particles through processes that may generate signals (excitation or ionization of the absorbing medium). The number of particles is on average proportional to the primary energy and the signal generated by each one of these is established by the detector properties. The direct consequence is that em calorimeters are in general linear detectors.
A non-linearity response is usually an indication of instrumental problems. The most common ones are:

- leakage effects, caused by too small dimensions of the calorimeter;

- saturation effects of the active components, originated by possible high localized energy depositions;

- recombination of electrons and ions, making the energy carried by the electron undetected.

At the same time, the small lateral and longitudinal em shower development allows to design relatively small calorimeters, often used as electromagnetic components in more complex systems.

A structural distinction can be made between *homogeneous* and *sampling* calorimeter.
The main advantage of homogeneous detectors is their excellent energy resolution, achievable because the whole energy of the primary particle is released in the active medium. At the same time, if a well position measurement and particle identification are required, homogeneous calorimeters are

not the best choice because of the challenging lateral and longitudinal segmentation task. Homogeneous calorimeter can be classified in four groups:

- semiconductor calorimeters;

- Cherenkov calorimeters;

- scintillator calorimeters;

- noble-liquid calorimeters.

On the other hand, sampling calorimeters have in general worst energy resolution due to the presence of an absorber material (copper, iron, lead and uranium) that does not contribute to the signal and the sampling fluctuations. For electromagnetic sampling calorimeters, typical resolution values are in the range of $5 - 20\%/\sqrt{E(GeV)}$. Meanwhile they are relatively simple to segment longitudinally and laterally due to their sampled structure contributing to achieve a better space resolution. Although this structure can be used to reduce the em calorimeter sizes, they are universally used in hadronic calorimeters.

Hadronic calorimeters

Hadronic calorimeters are detector optimized to induce hadronic showers, are more complex because both em and hadronic shower are produced in them. The calorimeter response is, therefore, divided in em (e) and non-em (h) components. The ratio e/h classifies calorimeters in three categories: *compensating*, if $e/h = 1$; *undercompensating*, if $e/h > 1$; *overcompensating*, if $e/h < 1$. Hence, the total calorimeter response (to charged pions) is a combination of the two:

$$\pi = f_{em} \cdot e + (1 - f_{em}) \cdot h. \tag{2.6}$$

Since the average f_{em} value, as already seen, increases with the energy, a non-compensating calorimeter response ($e/h \neq 1$) is not constant and hence non-compensating calorimeters are intrinsically non-linear detectors. The e/h value cannot be directly measured. However, it can be derived from the e/π ratios, measured at various energies. The relationship between e/π and e/h is as follows, where f_{em} is the energy dependent average em fraction:

$$\frac{e}{\pi} = \frac{e/h}{1 - f_{em}(1 - e/h)} \tag{2.7}$$

The relation is represented in figure 2.8.

The basic reason for different response for em and hadronic showers lies in the fact that in the absorption of hadronic showers, a significant fraction of the energy is *invisible* not contributing to the calorimeter signal. The main
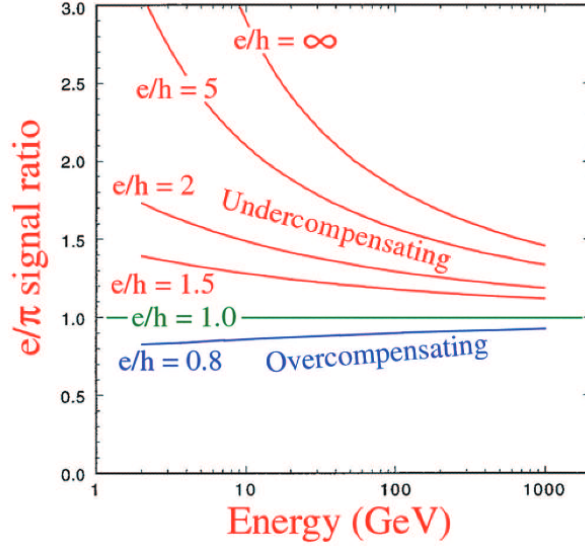
Figure 2.8: Relation between the calorimeter response ratio to em and non-em energy deposits, e/h, and the measured e/π signal ratios [4].

source of this energy is the energy used to release nucleons from nuclei.

To design a linear calorimeter various compensation methods have been developed. There are two main approach to obtain $e/h = 1$: one is reducing the electromagnetic response and the other one is to increase the non-em response.

The most effective way to reduce the em response of a sampling calorimeter is to use as absorber an high-Z material. The concept is based on the way that low energy photons release their energy. As seen, the main process at low energies is the photoelectric effect that is highly Z dependent. For this reason, using high-Z absorber materials causes an energy deposition in the absorber making this energy not detected.

The other strategy is known as compensation by neutrons signal boosting or by Signal Amplification through Neutron Detection (SAND). It consist in increasing the non-em response taking advantage of the kinetic energy transported by neutrons. The good correlation between the invisible energy and the kinetic energy of neurons opens the possibility of an event-by-event correction boosting the signals generated by the neurons. The most likely process to occur at low energies for a neutron is elastic scattering on a nucleus target. The fraction of energy transferred in this process is on average $f_{\text{elastic}} = 2A/(A+1)^2$ (where A is the mass number of the target). Considering that hydrogen maximize this fraction it is the element that must be present in the active medium to produce the non-em response increment. With this consideration and fine tuning the sampling fraction, a e/h ra-

tio can be forced to be 1 obtaining a linear hadronic calorimeter. In short, compensation through SAND can be achieved in sampling calorimeters with active materials containing hydrogen and a precisely tuned sampling fraction.

In the IDEA dual-readout calorimeter, a more recent compensation strategy is designed. The dual-readout compensation used will be described in section 2.3.

2.2.1   Fluctuations

The energy response of a calorimeter represent an average value, but the energy loss is a statistical phenomenon therefore the precision of these detectors is affected and limited by fluctuations. The effect is quantified by the relative resolution defined energy distribution standard deviation divided by the mean value (σ/E).

Considering the simpler case of em calorimeters, four are the main sources of fluctuation:

- sampling fluctuations;

- signal quantum fluctuations;

- shower leakage fluctuations;

- instrumental fluctuations.

Sampling effects are affected both by the sampling fraction (i.e. the ratio of the active material on the total) and the sampling frequency (the thickness of the layers). The sampling fraction is defined as:

$$f_{samp} = \frac{E_{active}}{E_{passive} + E_{active}} \tag{2.8}$$

where E_{active} and $E_{passive}$ are the energies deposited in the active and passive part by an incident minimum ionizing particle (mip). This type of fluctuation is dominated by the Poisson statistics, hence it contribute with a term proportional to $1/\sqrt{E}$ to the energy resolution. In em calorimeter with non-gaseous active material the sampling contribution follow the empirical law:

$$\frac{\sigma}{E} = \frac{2.7\%\sqrt{d/f_{samp}}}{\sqrt{E}} \tag{2.9}$$

where d is the thickness of the active material layer measured in mm.

Under the label of signal quantum fluctuations are grouped fluctuation effects due, for example, to photon statistics in light emitting active materials. Also these fluctuations follow the rules of Poisson statistics, with the constrain of uncorrelated separate contributions. Another contribution that

scales with $E^{-1/2}$ is added to the relative energy resolution expression.

In leakage fluctuations there are three possibilities: longitudinal, lateral and albedo. These fluctuation are highly dependent to the geometry and the material of the calorimeter. For this reason their contribution to the resolution does not have a precise energy dependent form. The typical solution to numerically evaluate their effect is to produce Monte Carlo simulation for the calorimeter and study the obtained data.

Finally, instrumental fluctuations are the ones provided, for example, by electronic noise or geometrical inhomogeneities. The electronic noise contribution is usually described with a term that scales with $1/E$ because it is largely independent of the shower energy. Meanwhile fluctuations induced by structural inhomogeneities depend on the shower position and might be energy dependent.

Typically, these contribution are uncorrelated and have to be added in quadrature to obtain the total resolution. As seen, different contribution with different energy dependence have to be considered. The consequence is that different effects dominate the energy resolution in different energy ranges. These behaviour have to be considered in the design of the calorimeter in order to optimize the resources.

All the sources described affect also hadronic calorimeters. For example, sampling fluctuation has a higher impact in hadronic showers due mainly to the lower average number of particle that release energy in the hadronic ones. At the same time, in hadronic detectors, resolution is affected also by other fluctuation effects.

The introduction of quantities such as f_{em} in hadronic calorimeters produces new sources of fluctuations. In non-compensating calorimeter, event-by-event fluctuation in em component affect the detector response. It contributes to the relative energy resolution through a term scaling as $cE^{-0.28}$.

$$\frac{\sigma}{E} = \frac{a_1}{\sqrt{E}} \oplus cE^{-0.28} \tag{2.10}$$

Up to $400 GeV$, this law runs almost parallel to an energy resolution in which only the stochastic term is included ($\sigma/E = a/\sqrt{E}$). For this reason, the resolution of non-compensating calorimeter is described as:

$$\frac{\sigma}{E} = \frac{a_2}{\sqrt{E}} + b. \tag{2.11}$$

l A useful solution that bring remove these fluctuations is to design a compensating calorimeter, greatly improving the energy resolution.

Finally, the calorimeter resolution is also affected by invisible energy fluctuations. To evaluate the impact of these fluctuations the correlation between f_{em} and f_{inv} has to be known, and in particular if they are correlated or not. If they are uncorrelated the contribution to the resolution has to be added

in quadrature. Moreover, in compensating calorimeter, the contribution of invisible energy fluctuations are different depending on the techniques used to achieve compensation.

2.3 Dual-readout compensation

All the benefits obtained with compensating calorimeters are obtaining by sampling em and non-em components in hadronic showers with the same response ($e = h$). Dual-readout is a method that perform this correction event-by-event, constantly measuring the electromagnetic fraction [cite?].

2.3.1 Working principles

This technique is based on the concept that the em component is mostly provided by relativistic electrons and positrons, while the majority of non-em one is carried by non-relativistic particles. Hence, collecting the Cherenkov signal produced by an hadronic shower is almost equivalent to sampling the em component, estimating the electromagnetic fraction. A second simultaneous signal, typically from scintillators, is recorded. Thanks to this second signal, the non-em component can be evaluated starting from the f_{em} already obtained.

Entering the details of the process, let's consider scintillation (S) and Cherenkov (C) signals produced by an hadronic shower. The mean values of these signals are calibrated with an electron beam of known energy E so that, for em showers, $\langle S \rangle = \langle C \rangle = E$. The hadronic signals can be written (for each event) as:

$$S = E\left[f_{em} + (h/e)_S(1 - f_{em})\right], \tag{2.12}$$

$$C = E\left[f_{em} + (h/e)_C(1 - f_{em})\right], \tag{2.13}$$

where h/e quantify the two different degree of non-compensation for the two signals. Considering that h/e are measurable quantities and are assumed to be constants, the ratio C/S is also energy independent:

$$\frac{S}{C} = \frac{f_{em} + (h/e)_S(1 - f_{em})}{f_{em} + (h/e)_C.(1 - f_{em})} \tag{2.14}$$

This expression can straightforwardly give an evaluation of the em fraction:

$$f_{em} = \frac{(h/e)_C - (C/S)(h/e)_S}{(C/S)\left[1 - (h/e)_S\right] - \left[1 - (h/e)_C\right]} \tag{2.15}$$

allowing the rewriting of 2.12 and 2.13 as

$$S/E = (h/e)_S + f_{em}\left[1 - (h/e)_S\right], \tag{2.16}$$

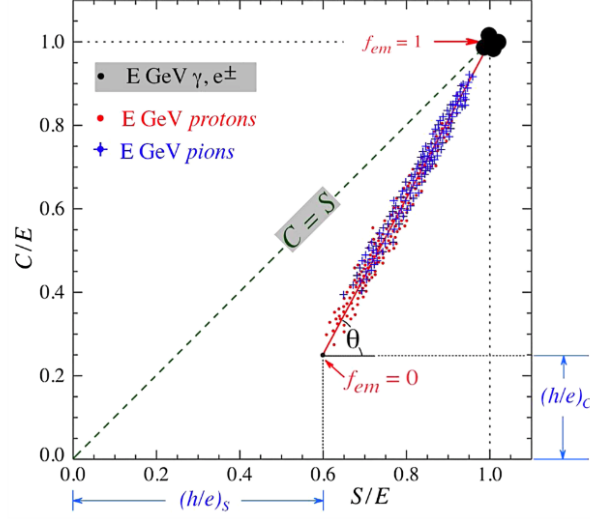$$C/E = (h/e)_C + f_{em}\left[1 - (h/e)_C\right]. \tag{2.17}$$

Figure 2.9: Relation between the calorimeter response ratio to em and non-em energy deposits, e/h, and the measured e/π signal ratios [4].

A C/E vs S/E scatter plot can be sketched using signals from a generic DR calorimeter for electrons, pions and protons showers. An example is shown in figure 2.9. As expected, data generated from electrons are always around the point $[1,1]$ (where $f_{em} = 1$). On the other hand an hypothetical signals with only non-em component would be at $[(h/e)_S, (h/e)_C]$ where f_{em} is 0. Signals from pions and protons showers are clustered along a straight line linking these two points. From this geometric construction it is clear that the slope of this line (the θ angle) is energy independent because it is determined only by the two e/h values. Thanks to the fact that θ is energy and particle independent, it is possible to define the parameter:

$$\chi = \frac{1 - (h/e)_S}{1 - (h/e)_C} = \cot\theta, \tag{2.18}$$

this parameter can also easily estimated with test beam data depending only on calorimeter material and geometry. Finally, event-by-event, the energy of each hadron shower can reconstructed using the S and C signals in an unique way:

$$E = \frac{S - \chi C}{1 - \chi}. \tag{2.19}$$

The ideal DR calorimeter is able to identify the Cherenkov signal as the exact electromagnetic component ($h/e = 0$) being a direct measurement of f_{em}. The worst DR calorimeter instead is designed when $(h/e)_C = (h/e)_S$, in this case the two signals would sample the two components (em and non-em) with the same response giving no information about the f_{em} from
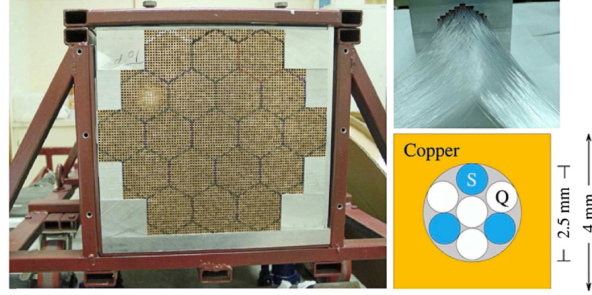
Figure 2.10: The DREAM calorimeter layout cite?.

the S/C ratio. Therefore, the best real DR calorimeter is the one with the lower χ parameter found with the lower $(h/e)_C$ and the higher $(h/e)_S$ values possible.

2.3.2 Experiments

At present a few dual-readout calorimeter have been built and tested, and more project are under study.

The first practical usefulness of the dual-readout method was tested in a R&D study for the Advanced Cosmic Composition Experiment at the Space Station (ACCESS) cite?. The prototype had a depth of $1.4\lambda_{int}$ and was equipped with both plastic-scintillator (S) and quartz (Q) optical fibres to measure and collect the scintillation and Cherenkov light, respectively. With this detector, the Q/S signals ratio was found to represent a good event-by-event measurement of the shower energy fraction carried by π^0.

Later, a $10\lambda_{int}$ deep calorimeter known as Dual-REAdout Method (DREAM) was built and tested. The detector is composed by 5580 basic element represented in figure 2.10. These are 200 cm long, hollow, extruded copper rod of $4 \times 4mm^2$ cross section in which 3 scintillating and 4 Cherenkov fibres are inserted.

Figure 2.11 shows the $C - S$ scatter plot of the calibrated signals, where the correlation between them its clear. The two h/e ratios for Cherenkov and scintillator DREAM structures are 0.21 and 0.77, respectively. The expression of the em fraction 2.14 becomes:

$$f_{em} = \frac{0.21 - 0.77(C/S)}{(C/S)\left[1 - 0.77\right] - \left[1 - 0.21\right]} \qquad (2.20)$$

The DREAM results, together with other studies (e.g. SPACAL and RD52) confirmed the potentiality and the good concept of the dual-readout compensation method, bringing to the IDEA dual-readout calorimetry project.
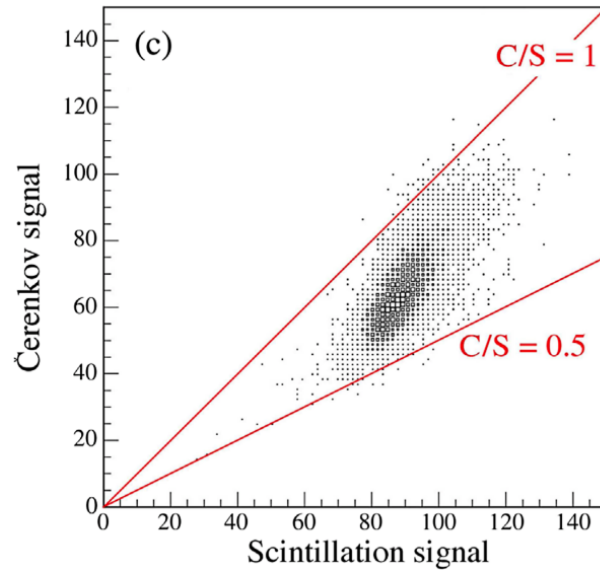
Figure 2.11: The $C - S$ scatter plot showing the correlation between both types of signals. The signals are expressed in the same units used to calibrate the calorimeter (em GeV) cite?.

Chapter 3

Silicon Photomultipliers

aaa

3.1 Working principles

aaa

Single photon avalanche diode

aaa

Electrical model

aaa

Gain and amplitude

aaa

3.2 SiPM efficiency

aaa

Geometrical fill factor

aaa

Quantum efficiency

aaa

Avalanche breakdown trigger probability

aaa

3.2.1   Occupancy effect

aaa

3.3   Noise effects

aaa

3.3.1   Dark Count Rate

aaa

3.3.2   After-Pulse

aaa

3.3.3   Optical Cross-Talk

aaa

3.4   Timing properties

aaa

Chapter 4

IDEA DR calorimeter full simulation

As already said, the concept described in chapter 1.4 is an on-going project and it has to be supported by simulation. With this goal, a dual-readout calorimeter full simulation has been developed allowing to generate data and monitor the whole process from the collision on the interaction point to the digitized signal produced by SiPMs.

The chapter presents a description of the simulation structure. The section 4.1 describes in details the simulation dividing it in two main Monte Carlo processes:

- the calorimeter simulation, coded in GEANT4;

- the SiPM response digitization ("pySIPM"), coded in Python.

Later, the performances obtained will be shown. The temporal behavior, the SiPM saturation effect and the energy resolution will be described in section 4.2.

Then the chapter treats of the possibility of simple particle identification using neural network structures.
The last section (sec.5) exposes a neural network that has the purpose of identify if signal are generated from photons (γ) or neutral pions (π^0). This aim is achieved analyzing the spatial pattern of energy released in the calorimeter.
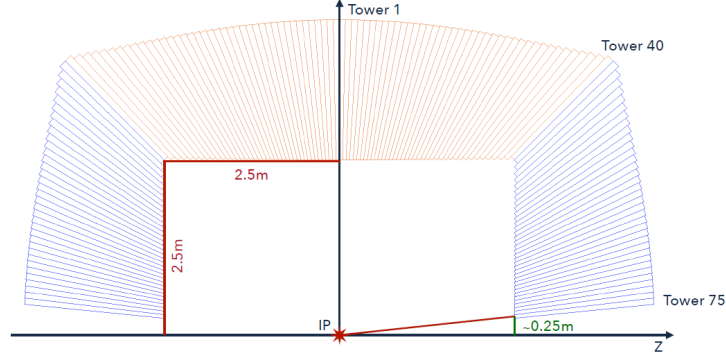
Figure 4.1: Calorimeter single slice.

4.1 Simulation structure

4.1.1 Calorimeter simulation

The calorimeter simulated follows the idea show in chapter 1.4. It has a cylindrical symmetry characterized by a barrel and two endcaps. This 4π structure is obtained rotating 36 simpler component, called slices, around the z axis. The dimensions of the slices are shown in figure 4.1, therefore the inner diameter and the inner length are both 5 m meanwhile the overall outer diameter and length are 9 m.

Each half slice is composed by 75 2 m long towers (40 part of the barrel and 35 part of the endcap), 5400 of this element are used to set up the whole calorimeter. To correctly cover an almost 4π solid angle each tower has different trapezoidal inner face with dimensions that can vary from ~ 5 cm to ~ 8 cm. A small circular area, with 0.25 m of radius, centered along the z axis is not covered by the calorimeter to permit the beam to reach the interaction point (IP).

The towers are copper based and play the role of absorber. To have a sensitive element they are filled by optical fibres. The idea of a projective calorimeter make the absorber volume greater increasing the distances from the IP. New fibres at different depth have to be placed inside the calorimeter to keep constant the sampling fraction.

As the dual-readout technique needs to distinguish Scintillating (S) and Cherenkov (C) signal, two types of fibres are used (fig. 4.2). Their characteristic are shown in tab. 4.1.

The fibre refractive indices determine the light transport (as consequence of the Snell's law [5]). The signal from the scintillating fibres is parametrised by the deposited energy while the Cherenkov photons are produced accordingly to the Cherenkov emission process.

Kuraray SCSF-78 (S)	
Core:	$r = 0.485$ mm, Polystyrene (C_5H_5), $\rho = 1.95$ g/cm^3, $n = 1.59$
Cladding:	Thickness = 2% of r, PMMA ($C_5H_8o_2$), $\rho = 1.19$ g/cm^3, $n = 1.49$
Main properties:	Emission constant = 2.8 ns, LY = 10^4 photons/MeV, $\lambda_{att} = 4$ m
Mitsubishi SK-40 (C)	
Core:	$r = 0.485$ mm, PMMA ($C_5H_8o_2$), $\rho = 1.19$ g/cm^3, $n = 1.49$
Cladding:	Thickness = 2% of r, Fluorinated Polymer (C_2F_2), $\rho = 1.43$ g/cm^3, $n = 1.42$
Main properties:	$\lambda_{att} = 8.9$ m

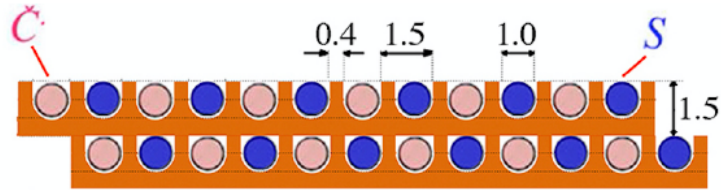Table 4.1: Technical characteristics of the two type of fibres used.

Figure 4.2: Chess-like configuration to dispose the fibres in the copper structure.

For each event, the simulation gives as output useful information:

- Event ID;

- Fibre Type;

- Fibre ID;

- the position of the fibre end closer to the IP;

- the number of photons reaching the fibre further end;

- the list of photons time of arrival to the fibre end.

The computation of light propagation is extremely time consuming, so that it has to be fine tuned to optimize the process. In particular, the propagation of C photons is tracked until the single photon reach the core-cladding boundary (i.e. at the distance R from the further end of the fibre and at the time t_0). If the emission angle is inside the range of the fibre numerical aperture, the photon is added to the final number of photons (after a Poissonian smearing on their number). The time of arrival on the sensor for each photon is estimated as:

$$t_C = t_0 + R\frac{n_C}{c} \tag{4.1}$$

where $n_C = 1.49$ is the fibre refractive index and c is the speed of light.

The S fibres, instead, carry scintillating photons produced considering the light yield of the fibres and the energy deposited by the interacting particle. The number of photons is smeared with as Poissonian law and de time of arrival on the sensor is obtained as:

$$t_S = t_0 + R\frac{n_S}{c \cdot \cos(\vartheta)} + t^* \tag{4.2}$$

where $n_S = 1.59$ is the refractive index and t^* is a random time that considers the fibres decay time, it is chosen from an exponential distribution with 2.8 ns as mean value. Considering the internal reflection, the photon path depends on the ϑ angle (i.e. the angle between the photon direction and the fibre axis). It is chosen randomly in the range $[\cos(\alpha), \cos(0)]$, where $\alpha = 20.4°$ is the fibre critical angle.
Eventually, the light produced is smeared by two Poissonian distribution, one for scintillation photo-electrons and the other one for Cherenkov photo-electrons. This procedure correctly reproduces statistic fluctuations in scintillation and Cherenkov light production and allows to reproduce within simulations the scintillating light yield (p.e./GeV) desired. To make the full simulation lighter, this process also includes the attenuation due to the PDE of the simulated SiPMs. The simulation is tuned to produce $\sim 400\ Spe/GeV$ and $\sim 100\ Cpe/GeV$ at the electromagnetic scale.

4.1.2 SiPM response digitization

The results obtained are the input of the second part of the simulation: *pySiPM*, a Monte Carlo simulation, performed mostly in Python, able to reproduce the SiPM response to a light source and replicate the waveforms recorded with a digitizer [6].

The importance of this software goes beyond our context, but perfectly fits our needs. In particular each fibre from the calorimeter simulation is considered coupled to a single SiPM, which digitized response is simulated through *pySiPM*.

The simulation allows to set most of the SiPM parameters:

- **Geometrical parameters**: the sensor dimensions and the pixel pitch.

- **Sensor parameters**: Photon Detection Efficiency, Dark Count Rate, After-Pulse probability, Optical Cross-Talk probability.

- **Signal parameters**: rise time constant, decay time constant.

- **Waveform parameters**: time window, sampling time, integration window.

For each event and fibre, random parameters determine the photon position inside the sensor. Meanwhile the sensor PDE is set at 100% to be consistent with the smearing applied at the calorimeter simulation level. A control stop the count of impinging photons on the same cell to a maximum one, then each element of noise is generated with the set probability.
The pulse generated is a combination of two exponentials characterized by the rise time constant (τ_{rise}) and the decay time constant (τ_{fall}), considering the different photon time of arrival (t_S and t_C):

$$y(t) = A \cdot \left(e^{-\frac{t}{\tau_{fall}}} - e^{-\frac{t}{\tau_{rise}}} \right).$$

(4.3)

The total signal of each SiPM is the sum of all the signals generated from the activated cells.

The information given as output of the simulation are:

- **Data reported from GEANT4 simulation**: event ID, type of fibre, fibre ID, fibre position;

- **Computed quantities**: integral, peak height, time of arrival, time over threshold, time of peak;

- **Digitized waveform**.

Geometrical Parameter	
SiPM area	$1 \times 1 \ mm^2$
Sensor Parameters	
DCR	$200 \ kHz$
After-Pulse	3%
Cross-Talk	1%
Signal Parameter	
Rise time	$1 \ ns$
Waveform Parameters	
Time window	$500 \ ns$
Integration window	$300 \ ns$
Sampling frequency	$10 \ GHz$

Table 4.2: Fixed SiPM parameters used in the simulation configuration file. To know also variable parameters see text.

All this features can be rejected through a trigger, in particular a threshold can be set to establish the informations that have to be recorded. The threshold is defined as a scale of the maximum value of a waveform generated by a single photoelectron (neglecting the electrical noise). In the results showed later a one-suppression has been applied using a threshold factor of 1.5, a typical setup to filter isolated signals of DCR.

4.2   Simulation performances

4.2.1   Different configurations

The results shown in this last chapter are obtained considering different SiPM parameters configurations.
They have been chosen in a common parameter space identified checking the lineup of SiPMs produced by Hamamatsu [7]. Two are the parameters that has been changed in our studies:

- the decay time constant of the signal, the chosen values are 10 ns and 50 ns;

- the pixel size, the chosen values are 10 μm, 15 μm and 25 μm.

The other fixed values of parameters are listed in the table 4.2.

An example of waveform generated is plotted in figure 4.3 where is clear the difference produced using two different decay time constant.
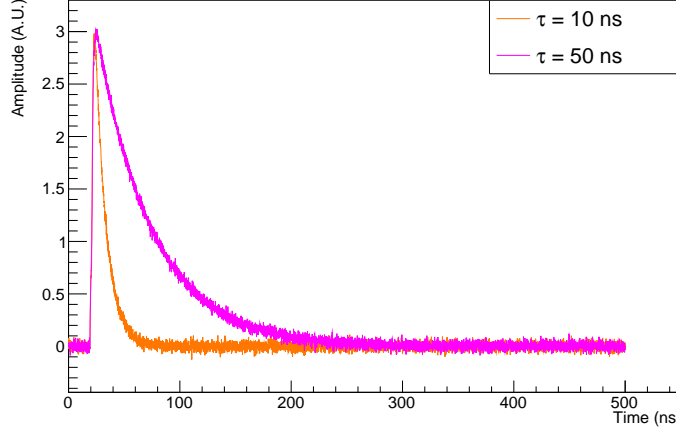
Figure 4.3: Single waveforms generated in two identical condition except for the decay time constant in the SiPM configuration file.

4.2.2 Time studies

An important aspect that has to be studied is the temporal one. In order to do this, data of 1000 events are produced. In each event a 20 GeV electron is produced from the interaction point.

A first step is to analyze the distribution of the time of arrival of the photons converted at the SiPMs (i.e. the time recorded in the GEANT4 simulation output).

The distribution obtained from C and S photons are plotted in figure 4.4. As expected, the distribution of C photons time extremely narrow due to the instantaneous production of photons at the passage of relativistic charged particle in the fibres, instead the S photons time distribution shows an exponential tail that is a direct consequence of the emission time constant of the Polystyrene ($\tau = 2.8\ ns$).

Now a step forward can be done using this data as input for *pySiPM*. The SiPM parameters are chosen as described in paragraph 4.2.1. In this context the most interesting editable parameter is the decay time constant. Figure 4.5b and 4.5a are in analogy with respect to the last described and presents clearly a widening of the distributions, the cause of this phenomenon has to be associated to the characteristic response function of the sensors 4.3.

These data can be compared looking for differences in changing SiPMs configurations. As we can see in figure 4.6, the same C and S photons produce narrower time of peak distribution due to the less impact of electronic noise on a sharper response function.
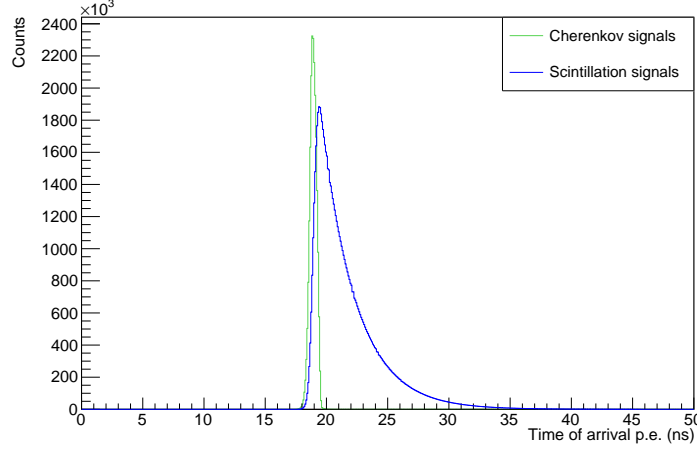
Figure 4.4: The distributions of the time of arrivals of photons on the SiPMs surface, distinguishing the ones transported from Cherenkov fibres from the ones transported from scintillation fibres.

The impact of noise on time of peak precision is also dependent on the number of photons impinging the same SiPM, in particular the peak precision increase with the number of photons.
To prove this 10000 SiPMs have been fired with an increasing number of simultaneous photons. For each fixed number of photons, the time of peak has been recorded, plotted in an histogram and fitted with a Gaussian function. An example of these histograms is shown in figure 4.7.

The standard deviation of these Gaussian fit is the interested quantity that has been recorded and reported in the table 4.3.
It is interesting to plot these data and study the behavior of the standard deviation in fuction of the number of photons. Figure 4.8 shows graphically the data, which are well fitted with a function of the form:

$$\sigma = \frac{A}{\sqrt{n}} + B. \tag{4.4}$$

The parameters obtained are $A = 0.8712 \, ns$ and $B = 0.08734 \, ns$ for data associated to SiPMs with $\tau_{fall} = 10 \, ns$, and $A = 1.949 \, ns$ and $B = 0.008217 \, ns$ for data associated to SiPMs with $\tau_{fall} = 50 \, ns$.

4.2.3   Occupancy effect

The occupancy effect, as shown in the paragraph 3.1, is an important characteristic that has to be deeply studied to know the behaviour of the SiPMs
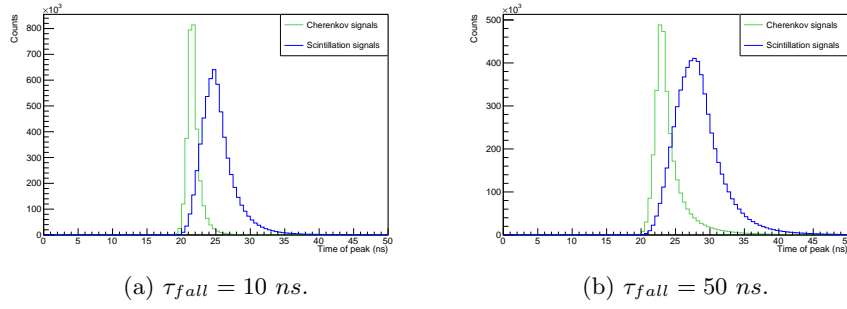
(a) $\tau_{fall} = 10\ ns$.   (b) $\tau_{fall} = 50\ ns$.

Figure 4.5: Time of peak distributions comparing in each histogram signals from Chernkov and from scintillation fibres. Different figures correspond to different decay time constants.
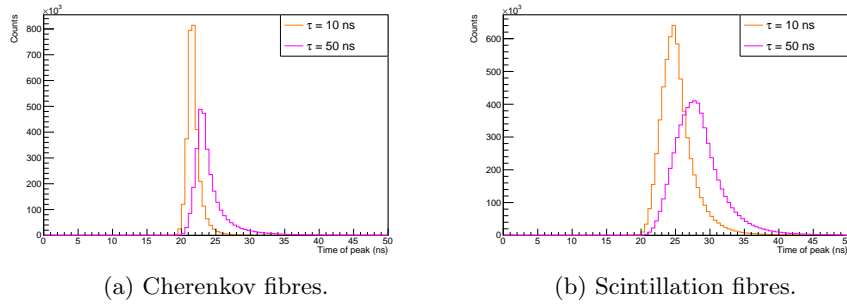

(a) Cherenkov fibres.   (b) Scintillation fibres.

Figure 4.6: Time of peak distributions comparing in each histogram signals with different decay time constant of SiPMs. Different figures correspond to different photon emission processes.

Number of photons	σ with $\tau_{fall} = 10\ ns$ (ns)	σ with $\tau_{fall} = 50\ ns$ (ns)
1	1.4150	7.0680
2	0.8717	2.6420
3	0.6738	1.7370
4	0.5742	1.3770
5	0.5146	1.1230
6	0.4624	0.9719
7	0.4314	0.9148
8	0.3998	0.8508
9	0.3811	0.7717
10	0.3605	0.7169
25	0.2339	0.4481
50	0.1679	0.3112
100	0.1229	0.2297

Table 4.3: Standard deviations obtained through the Gaussian fit applied on the time of peak distributions under different conditions.
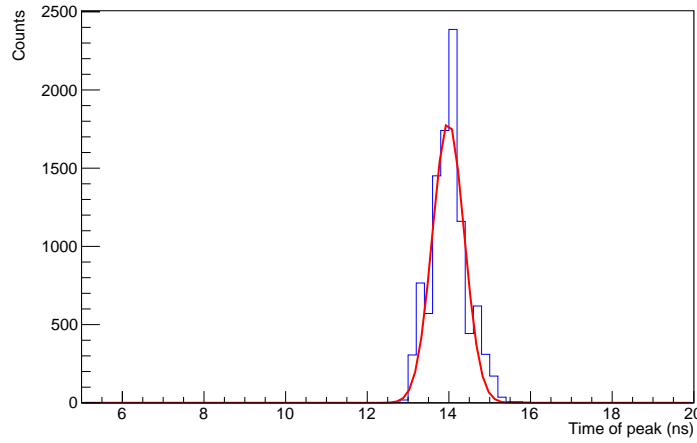

Figure 4.7: Time of peak distribution obtained firing 10000 SiPMs with 11 photoelectrons each, all at the same time (10ns). Over this distribution a Gaussian fit has been applied.
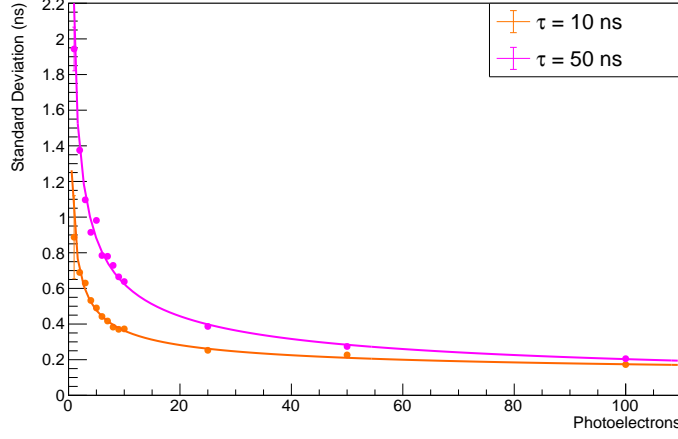
Figure 4.8: Standard deviation behaviour with respect to the number of simultaneous photoelectrons. Study performed with two different decay time constant in SiPMs configuration.

under high number of impinging photons and to correctly reconstruct the released energy.

As can be seen in figure 4.9, this effect is reproduced correctly in our simulations. The example shows the charge integral in dependence to the number of p.e. and follows the concepts already seen in paragraph 3.2.1.

The first step to perform these studies is to reconstruct a calibration law that reproduces the charge integral with respect to the number of p.e. firing the same SiPM.
Starting by assuming that in the range of photoelectrons from 2 to 10 the saturation effect does not occur in our configurations (i.e. 10000, 4356 and 1600 cells in each SiPM), 10000 SiPM has been fired 9 times with an increasing number of simultaneous p.e. in the range considered.
To assume that no saturation effect occurs in these data the charge integral distribution from the three different configurations has been compared finding no bias as seen in figure 4.10. The mean and the RMS values has been recorded and fitted with a strait line corresponding to the calibration law desired:

$$I(n) = A \cdot n + B \tag{4.5}$$

with $A = 49.43 \pm 0.03852$ and $B = 2.491 \pm 0.2516$ (fig 4.11).
The parameter A represents the contribution to the charge integral associated to a single photoelectron. Meanwhile B is the pedestal that is originated mostly by the DCR. This contribution can be evaluated considering
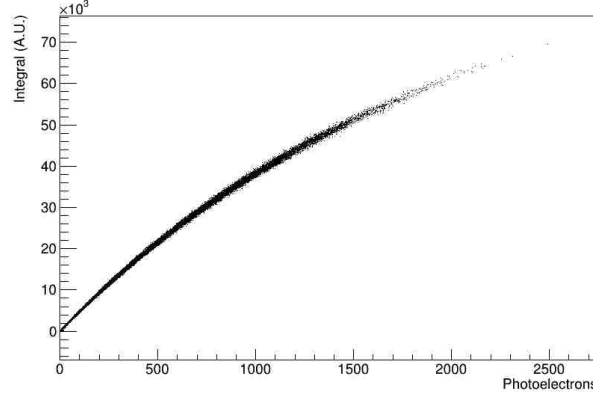
Figure 4.9: Example of saturation effect produced with the simulation chain. Each point correspond to a SiPM with a cell size of 25 μm.

our parameters of DCR $= 200 \ kH$ and the integration window of 300 ns.
$B_{DCR} = 2 \cdot 10^5 \cdot 3 \cdot 10^{-7} p.e. = 6 \cdot 10^{-2} p.e. = 6 \cdot 10^{-2} \cdot 49.43 = 2.96$.
In the following result, the value of the pedestal has been subtracted to the integral value of each SiPM.

The occupancy effect has been studied in electromagnetic shower produced by single electrons of different energies at each event with discrete values of 20, 40, 60, 80 GeV.
The impact of the saturation has been quantified using the line $I = A \cdot n$ as no saturation reference. The result obtained from 10000 events of single 40 GeV e^- with different SiPM configurations are separated in Cherenkov and Scintillation signals. They are shown in figure 4.12 where each point correspond to a single SiPM. The smaller is the number of cells, the greater is the occupancy effect.
Moreover, as expected, the scintillation fibres transport ~ 4 times the p.e. from Cherenkov ones on average, therefore they are more affected to the saturation.

This process can be extended considering one event at the time and adding the charge integral and the corresponding number of photoelectrons. The effect produced is represented in figure 4.15 where each point correspond to a single event.

As can be seen, the occupancy effect in our conditions is consistent. To mitigate this problem an analytical correction can be performed through the formula:

$$N_{fired} = N_{cells} \cdot \left[1 - \exp \left(-\frac{N_{p.e.}}{N_{cells}} \right) \right] \qquad (4.6)$$
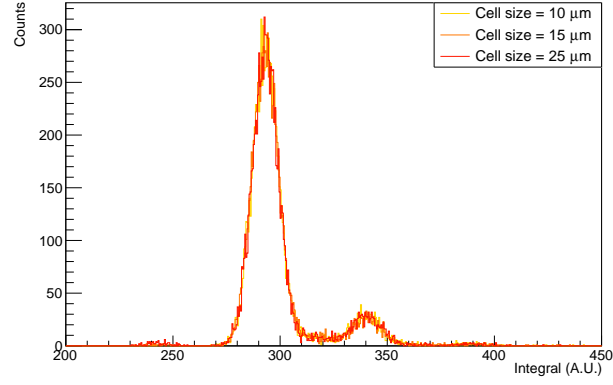
Figure 4.10: Charge integral distributions with different SiPM cell size considering 6 simultaneous photoelectrons. The distribution shape with less then 10 simultaneous p.e. is not dependent to the number of cells.
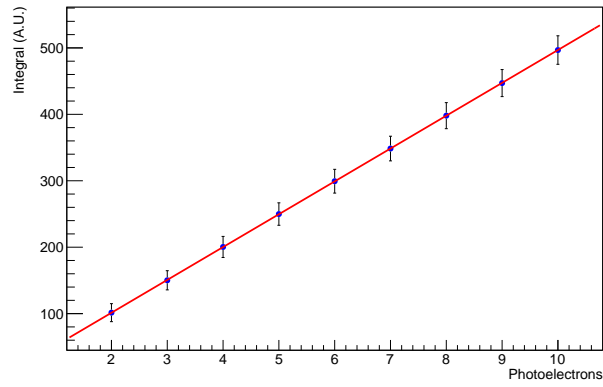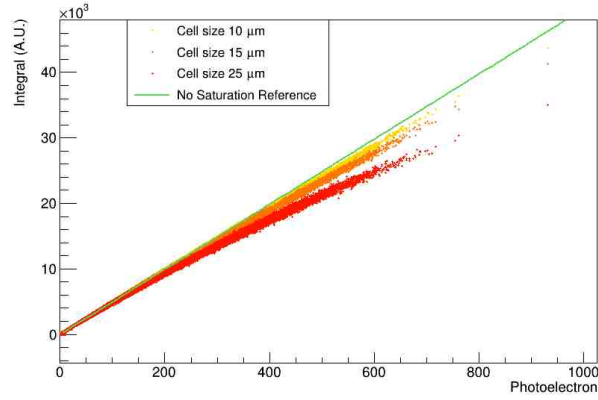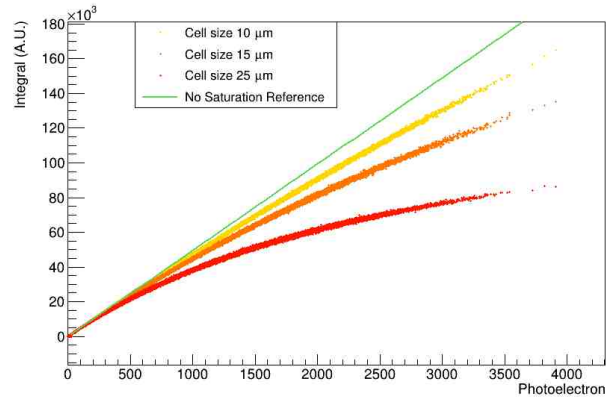

Figure 4.11: Study of the proportional relation between integral and number of photoelectrons in the range from 2 to 10. A linear fit has been applied to find calibration law without saturation.
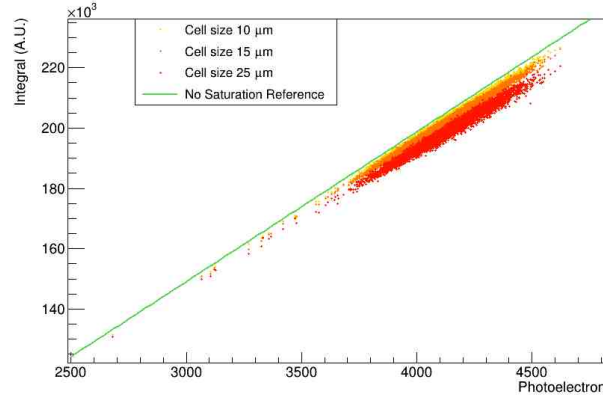
(a) Cherenkov signals.

(b) Scintillation signals.

Figure 4.12: The plots, divided considering the two different photon production processes, show the occurrence of the occupancy effect studying single 40 *GeV* electrons. Each point correspond to a SiPM with different values of cell size. The no saturation line shown in figure 4.11 has been added as reference.
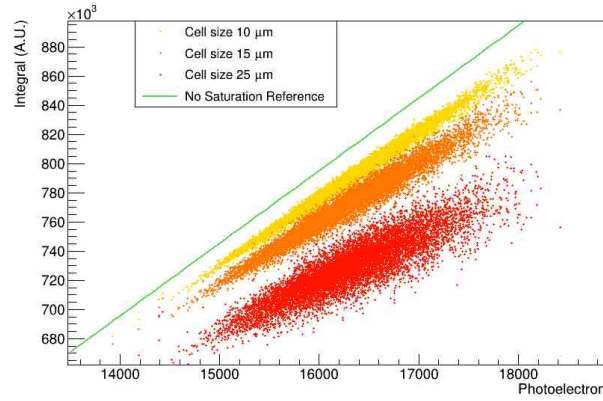
(a) Cherenkov signals.

(b) Scintillation signals.

Figure 4.13: The plots, divided considering the two different photon production processes, show the occurrence of the occupancy effect studying single 40 GeV electrons. Each of the 10000 points correspond to a single event where integral and number of photoelectrons have been added over the fired SiPMs. Different values of cell size are represented with different colors. The no saturation line shown in figure 4.11 has been added as reference.

therefore the correction has been applied modifying the integral values such as:

$$I_{corr} = -AN_{cells}\left[\ln\left(1 - \frac{I}{AN_{cells}}\right)\right] \tag{4.7}$$

The results obtained can be visualized in figure 4.14 where are compared the data from SiPM with cell size of 10 μm with and without analytical correction.

The discrepancy from the no saturation reference quantifies the effect of the occupancy when performing the energy reconstruction task. The percentage difference has been evaluated through the formula $\frac{E_{NoSat}-E}{E_{NoSat}}$, and the value obtained fill the histograms in figure 4.13.
After applying the analytical correction a clear improvement is show in figures 4.16.

This whole process has been performed simulating electrons with energies of 20, 40, 60, 80 GeV. Mean and standard deviation of the gaussian fit applied on the percentage discrepancy have been recorded and the obtained plots are shown in figure 4.17.

4.2.4 Energy resolution

The first step to study the energy resolution is to calibrate the full simulation.
Dual-readout calorimeters are typically calibrated at the electromagnetic scale. This is specially useful in a leptonic collider because, having easily access to electrons and positrons, the calorimeter can be calibrated precisely during all the life of the experiment.

The calibration has been performed on 40 GeV electrons applying the one-suppression already introduce un paragraph 4.1.2. 10000 events with single 40 GeV electrons have been fired from the interaction point obtaining the charge integral distributions for scintillation and Cherenkov signals.The calibration constants obtained to transform these data in energy distributions centered around 40 GeV are: $k_S = 4.998 \times 10^{-5}$ and $k_C = 2.023 \times 10^{-4}$.
Two analogue calibration constants have been obtained starting from the number of photoelectrons distribution with values of: $k_{pe,S} = 2.48 \times 10^{-3}$ and $k_{pe,C} = 1.00 \times 10^{-2}$.

Starting from these calibration constant, the energy distributions can be obtained from the number of p.e. (pre SiPM digitization simulation) of from the charge integral (post SiPM digitization simulation). The two types of distribution are compared in figure 4.18.
As expected, the energy distributions obtained from the charge integrals are
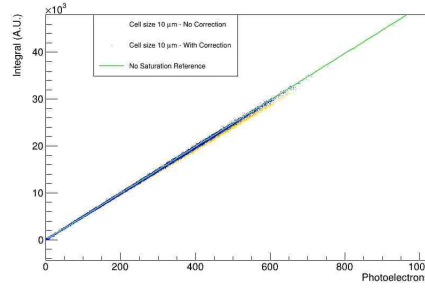
(a) Cherenkov signals. Points are single SiPMs.
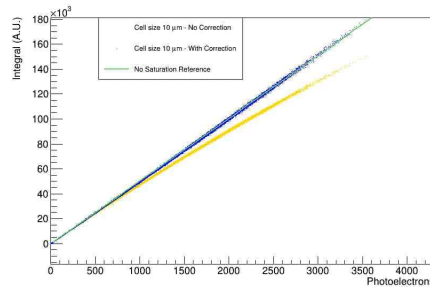
(b) Scintillation signals. Points are single SiPMs.
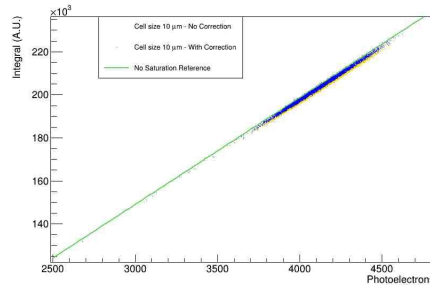
(c) Cherenkov signals. Points are single events.
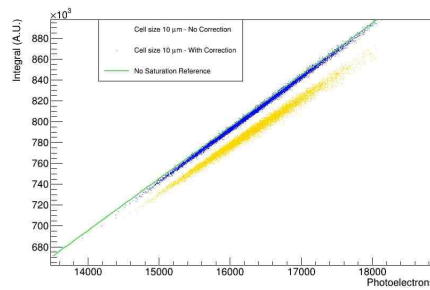
(d) Scintillation signals. Points are single events.

Figure 4.14: Similar plots to figures 4.12 and 4.15. Two series of data are presented to show the effectiveness of the analytical correction. The no saturation line shown in figure 4.11 has been added as reference.
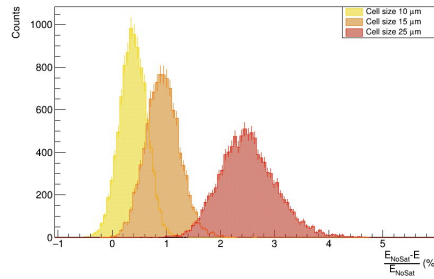
(a) Cherenkov signals.
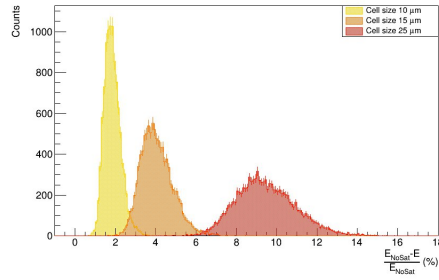
(b) Scintillation signals.

Figure 4.15: Percentage discrepancy distribution considering events with single 40GeV electrons. Different colors correspond to different cell size values.
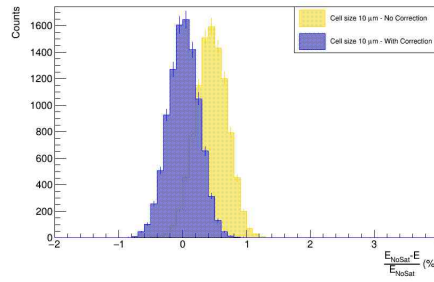
(a) Cherenkov signals.
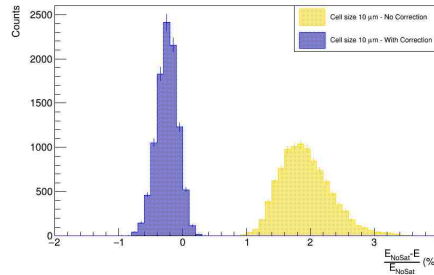
(b) Scintillation signals.

Figure 4.16: Percentage discrepancy distribution considering events with single 40GeV electrons to evaluate the effect of the analytical correction applied on data obtained with a cell size of 10 μm.
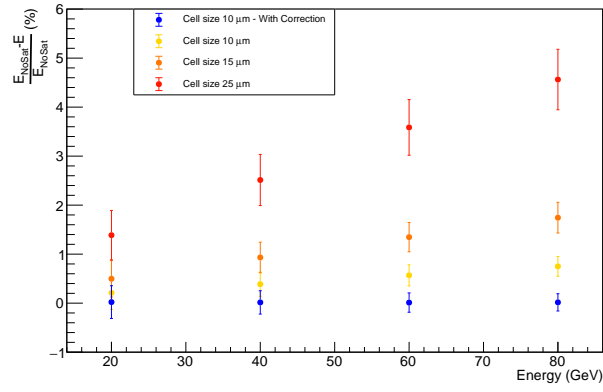
(a) Cherenkov signals.
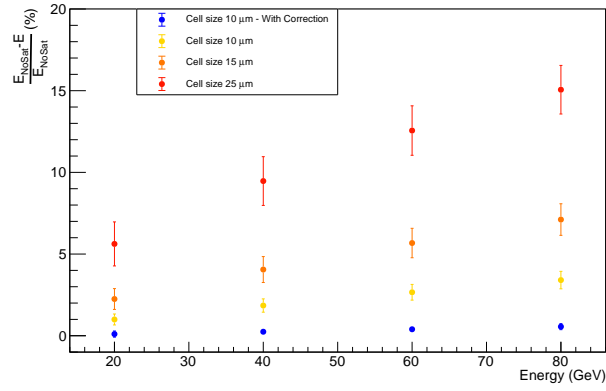
(b) Scintillation signals.

Figure 4.17: Percentage discrepancy behaviour with respect to the primary particle energy. Different cell sizes are compared also with data obtained performing the analytical correction.

	True E (GeV)	Mean E (GeV)	Standard Deviation (GeV)
Scintillation	20	19.90	0.95
	40	40.00	1.44
	60	60.65	1.85
	80	81.18	2.26
Cherenkov	20	19.88	0.95
	40	40.00	1.42
	60	60.61	1.73
	80	81.12	2.08

Table 4.4: Mean and standard deviation obtained by fitting with Gaussian function the energy distributions. The data are obtained with results from the full simulation.

wider due to the introduction of the electronic noise. However, the effect is minimal considering the fact that in each event hundreds of SiPMs are active and the white noise has mean zero.

The energy distributions have been fitted with a Gaussian function to obtain mean, standard deviation e respective errors. Doing this with different primary electron energies the energy resolution can be studied. Mean and standard deviation are listed in tables 4.4 4.5.
Then fitting the σ/E values in the energy range $20 - 80$ GeV can be seen there is a good agreement with the functions:

$$\frac{\sigma}{E} = \frac{A_1}{\sqrt{E}} + B_1 \qquad \text{or} \qquad \frac{\sigma}{E} = \frac{A_2}{\sqrt{E}} \oplus B_2 \qquad (4.8)$$

as shown in figure 4.19, with the fit parameters listed in table 4.6 and 4.7. The plots also show the loss of resolution with the introduction of the SiPM digitization software.
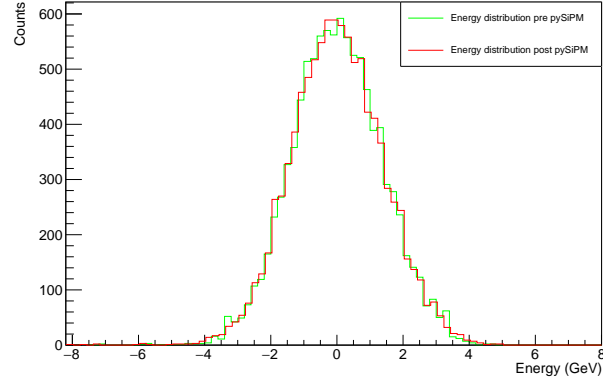
(a) Cherenkov signals.
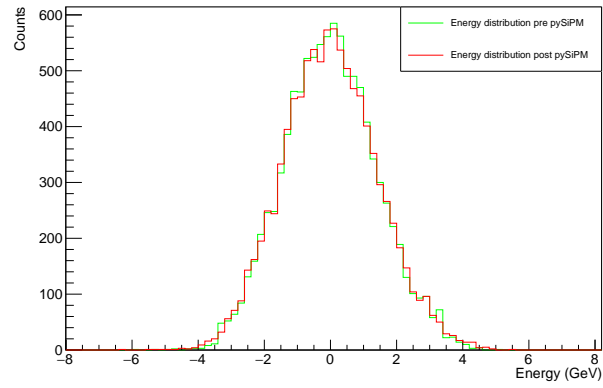
(b) Scintillation signals.

Figure 4.18: Charge integral distributions generated with data from GEANT4 (*pre pySiPM*) or from the full simulation (*post pySiPM*). The histograms are shifted by their mean values to show the minimal difference in the distributions spread.

	True E (GeV)	Mean E (GeV)	Standard Deviation (GeV)
Scintillation	20	19.82	0.92
	40	40.00	1.39
	60	60.28	1.78
	80	81.58	2.16
Cherenkov	20	19.76	0.92
	40	40.00	1.36
	60	60.40	1.66
	80	81.86	1.91

Table 4.5: Mean and standard deviation obtained by fitting with Gaussian function the energy distributions. The data are obtained with results from DR calorimeter simulation applying the calibration from the photoelectrons number.

	Scintillation	Cherenkov
pre pySiPM	$\frac{17.39\%}{\sqrt{E}} + 0.37\%$	$\frac{21.13\%}{\sqrt{E}} + 0.01\%$
post pySiPM	$\frac{17.84\%}{\sqrt{E}} + 0.77\%$	$\frac{20.01\%}{\sqrt{E}} + 0.30\%$

Table 4.6: Fit results on the energy resolution using the functional form 4.8 on the left.

	Scintillation	Cherenkov
pre pySiPM	$\frac{19.78\%}{\sqrt{E}} \oplus 1.55\%$	$\frac{21.21\%}{\sqrt{E}} \oplus 0.07\%$
post pySiPM	$\frac{20.14\%}{\sqrt{E}} \oplus 1.62\%$	$\frac{21.04\%}{\sqrt{E}} \oplus 0.98\%$

Table 4.7: Fit results on the energy resolution using the functional form 4.8 on the right.
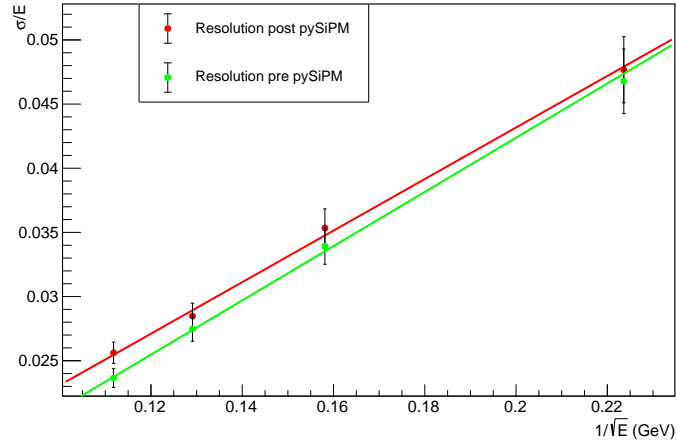
(a) C signals.
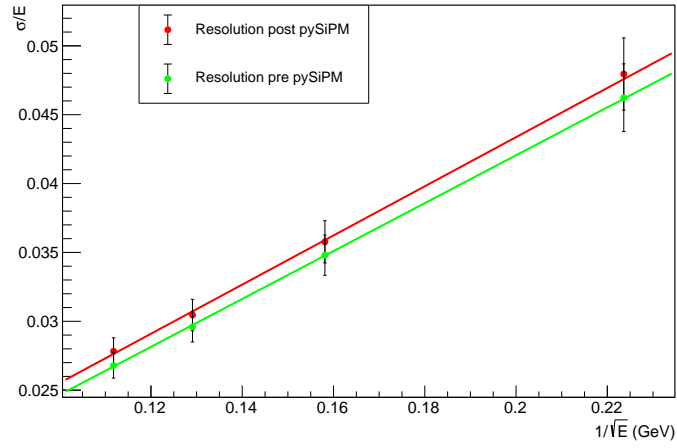
(b) S signals.

Figure 4.19: Calorimeter resolution plot where the data have been fitted with the formula on the left of line 4.8.

Chapter 5

Neural Network: Particle ID on imaging

Neural network and machine learning algorithm are powerful and versatile instruments that can improve the process of data analysis learning from datasets and making prediction with a certain degree of accuracy. Modern science, including physics, is taking more and more advantage of these techniques as the years go by. The availability of big data is a key aspect in performing these studies and experimental particle physics is an area that can provide large datasets.

The chapter describes a possible application of deep neural networks on data generated through the IDEA DR Calorimeter full simulation with the aim to obtain simple particle identification analyzing the spatial deposited energy distribution.
An introduction to the computational techniques is provided, briefly describing the role of each component and the common structures that are used to perform the study.

All the data have to be prepared to be used in a neural network structure, the section 5.3 shows which information has been used and how it has been processed to be analyzed by the deep learning algorithms. After that the structures are presented in details showing the performances obtained in training and testing phases.
Eventually the study has been performed extending the energy range, the impact of this generalization conclude the chapter.

5.1 Project goal

The chapter will explore the possibility to perform particle identification through neural network algorithms. The chosen task is to recognise and distinguish neutral pions and photons analyzing the spatial released energy distribution in a fixed area where the particles are fired.

As already introduced in paragraph 2.1.1, photons produce electromagnetic showers in their path through matter. Considering the geometry of our DR calorimeter where the fibres are oriented towards the interaction point, a photon will release most of its energy in few adjacent fibres and the remnant energy will be absorbed by the surrounding fibres in a elliptical shape (an example can be seen in figure 5.1a. On the other hand, the meson has a different behaviour. It decays in two main modes shown in figure 5.2:

$$\pi^0 \to 2\gamma \qquad \pi^0 \to \gamma + e^- + e^+ \tag{5.1}$$

The two decay mode have very different occurrence probability, in particular the 2γ decay has a branching ratio of $(98.823 \pm 0.034)\%$, instead it is $(1.174 \pm 0.035)\%$ for the gamma-electron-positron decay [8].

The consequence of this behaviour is that a electromagnetic shower is produced from each of the secondary particles. The result is a superposition of two or three (depending on which decay occurs) shower signals similar to the one produced by a single photon. In figure 5.1b the data obtained from a π^0 generation is sketched, it should be noted that, in the event represented, the meson decayed into two photons.

Usually this ID task would be performed applying a number of filters to study the elliptical shape, to find the peaks of charge integral, to measure the distances between the peaks and finally establishing, with a certain probability, the primary particle. The goal is to create a neural network able to accept data associated to an event and make a prediction on the primary particle with a small computational and timing effort.

5.2 Neural Networks introduction

Neural Networks are neural-inspired nonlinear models consisting in a group of artificial *neurons* or *nodes* interconnected through *layers*. This type of architecture is supported by a mathematical structure where each neuron has an activation degree typically ranging from 0 to 1 and each edge is identified by weights and biases.

The simplest neural network structure follows a sequential model where neurons are grouped and linked to each other through *layers* following a
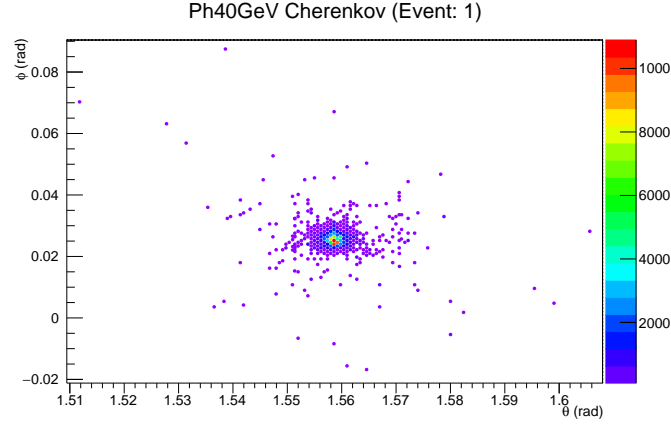
(a) 40 GeV photon.
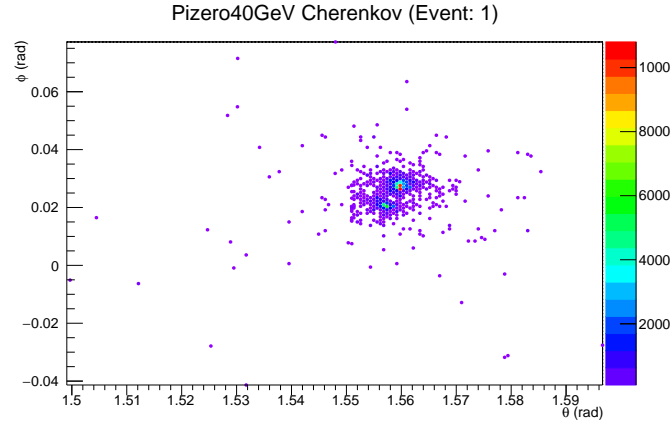
(b) 40 GeV neutral pion.

Figure 5.1: Spatial charge integral distributions from γ and π^0. Each point correspond to an activated fibre and it is represented in cylindrical coordinates, the color indicates the charge integral obtained from the coupled SiPM.

(a) $\pi^0 \to 2\gamma$ (b) $\pi^0 \to \gamma + e^- + e^+$
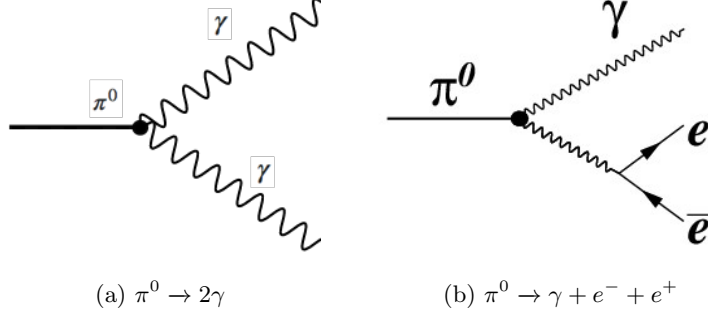
Figure 5.2: π^0 decay modes.

sequential order, but more complex neural networks implement also loops, branching and other different flows between layers. The sequential model present a first input level followed by a series of hidden levels composed by hidden neurons and then connected to the output units. A visual representation is shown in figure 5.3.

The mathematical representation can be introduced studying a simple two-layer model. Starting from a D-dimension input vector $\boldsymbol{x}$, the first layer apply a linear combination for each hidden neuron in the next level:

$$a_j^{(1)} = \sum_{i=1}^{D} w_{ij}^{(1)} x_i + w_{i0}^{(1)} \tag{5.2}$$

where $j = 1, ..., M$ (M is the layer output dimension) and the matrix $\boldsymbol{W}$ is the weights matrix including biases. The quantities a_j are known as *activations*. On these values a h nonlinear function, called *activation function*, is applied:

$$z_j = h(a_j^{(1)}) \tag{5.3}$$

common activation function are *sigmoid*, *tanh*, *ReLU* (rectified linear activation unit) and *Softmax*. The obtained value are the already introduced hidden units and are used as input of the next layer. The process in the second layer is similar, valuating the following activations:

$$a_k^{(2)} = \sum_{i=1}^{M} w_{jk}^{(2)} z_j + w_{k0}^{(2)} \tag{5.4}$$

where $k = 1, ..., K$ (K is the layer output dimension) and then applying the activation function. In the context of classification neural network, Softmax function (S) is a common choice of activation function in the last layer,
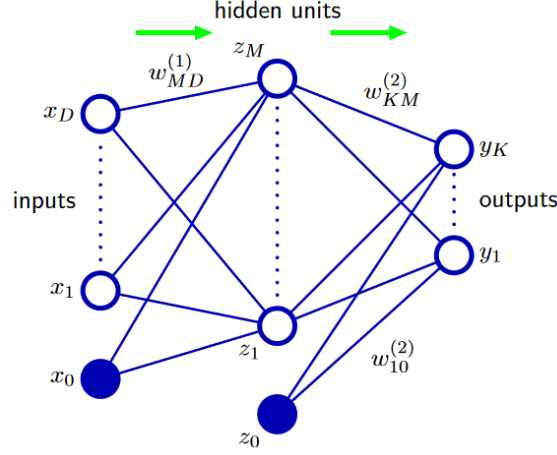
Figure 5.3: Schematic representation of a sequential model. The input, hidden, and output neurons are represented by nodes, and the weight parameters are represented by links between the nodes, for each connection the corresponding bias parameter is denoted by links coming from additional input and hidden variables x_0 and z_0. Green arrows indicate the direction of flow through the network. Figure from [9].

providing the probability associated to each possible label.

$$y_k = S(a_k^{(2)}).\qquad(5.5)$$

The whole process ca be represented in a single equation for each output neuron:

$$y_k = S\left(\sum_{i=1}^{M} w_{jk}^{(2)} h\left(\sum_{i=1}^{D} w_{ij}^{(1)} x_i + w_{i0}^{(1)}\right) + w_{k0}^{(2)}\right).\qquad(5.6)$$

and, to obtain a more compact expression, the values $x_0 = z_0 = 1$ can be introduced:

$$y_k = S\left(\sum_{i=0}^{M} w_{jk}^{(2)} h\left(\sum_{i=0}^{D} w_{ij}^{(1)} x_i\right)\right)\qquad(5.7)$$

making even more clear the two-layer mathematical structure.

Once the neural network is set up a training process has to be performed to make the prediction consistent. In order to do that a large dataset of correctly labelled input has to be provided. During the training, the elements in the weight matrix $\boldsymbol{W}$ are constantly modified to adapt the output to be as closer as possible to the correct label. These are not random corrections but they are conditioned by a error function that quantifies the discrepancy

between the output and the correct result. In classification tasks, a common error function choice is the *cross-entropy* (given by the negative log likelihood):

$$E(\boldsymbol{w}) = -\sum_{n=1}^{N}\sum_{k=1}^{K} t_{nk} \ln y_{nk}(\boldsymbol{x}_n, \boldsymbol{w}) \tag{5.8}$$

where t_n are the target vectors and $y_n = (\boldsymbol{x}_n, \boldsymbol{w})$ are the output vectors, with a dataset dimension of N and K different and mutually exclusive label possibilities.

The training aim is to correct the weights and biases values minimizing this error function.

Once this step is completed the neural network is ready to perform predictions and classify new data.

The study is performed with two different neural network structures that take advantages of few types of layers that will be briefly described.

Dense layer

The dense layer is the simplest and most common layer, but far from being the lightest. It is classified as a fully connected layer, meaning that each output neuron receive an input value from each input node.

Mathematically, this type of layer consist in a matrix application on the input vector ($\boldsymbol{v}_i$) providing an output vector ($\boldsymbol{v}_o$), adding a bias vector ($\boldsymbol{w}_0$):

$$\boldsymbol{v}_o = \boldsymbol{W}\boldsymbol{v}_i + \boldsymbol{w}_0. \tag{5.9}$$

The layer is extremely parameter consuming, let consider a single layer that connects two group of nodes: it is common to use number of neuron that are powers of 2 (32, 64, 128...). So the number of parameters associated to a single dense layer with an input dimension of 64 and an output dimension of 32 is:

$$64 \cdot 32(\text{weights}) + 1 \cdot 32(\text{biases}) = 2080$$

this number can increase very rapidly with the number of neurons.

Being the basic type of layer is often used to graphically represent general neural networks, figure 5.3 shows sequence made by two consecutive dense layers.

Dropout layer

Considering the extremely large number of training parameters, it is common practice to introduce the so called *regularization* techniques to increase the capability of the neural network in generalizing well on new data. The use of a Dropout layer is one of them.
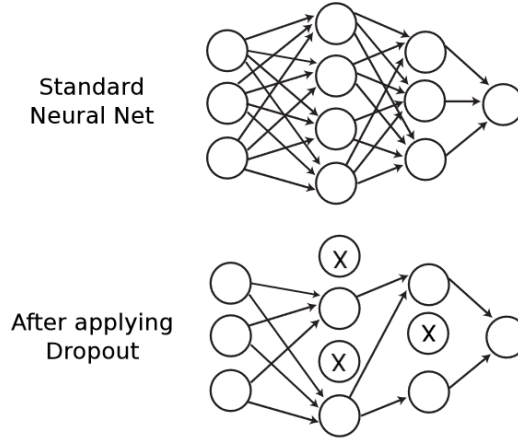
Figure 5.4: Neurons during training are randomly switched off with a probability p. A lighter network is produced reducing correlations between nodes. Figure from [10].

The basic idea is to reduce the spurious correlations that could occur between neurons in the network, preventing the overfitting. In practical terms, the dropout layer randomly "drops out" neurons (and the corresponding connections) following a probability p. This process is applied in each training step. An intuitive representation is sketched in figure 5.4.

Conv2D layer

A mandatory layer in analyzing images is the convolution layer and it identify a category of NN called Convolutional Neural Networks (CNNs).
A convolution consist in a simple application of a filter to an input that results in an activation. The application of the same filter repeatedly to the same input results in a map of activations called a feature map, this indicate the locations of a detected feature (and its intensity) in an input, such as an image.
Figure 5.5 represents the application of a filter or kernel (K) over a input matrix (I). As can be seen, the feature map is composed considering all the sub-images with the same size of the kernel, each sub-image is associated to an activation value obtained through the formula:

$$a = \sum_{i=1}^{D_K} \sum_{j=1}^{D_K} I_{i,j} \cdot K_{i,j} + b \tag{5.10}$$

where D_K is the dimension of the kernel, typical values are 1, 3, 5, and b is the bias value. An activation is obtained from each sub-image and than corrected by the activation function, these outputs compose the feature map.
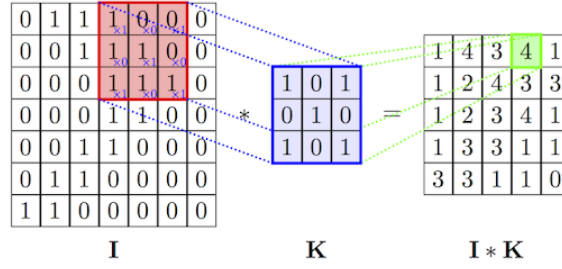
Figure 5.5: Example of convolution applying a single kernel K on the input matrix I. The red square is the sub-image considered and convoluted with the blue matrix giving the activation value in the green box. The procedure is done over all the sub-images producing the feature map on the far right.

The layer output is a group of feature maps obtained from different filters. The number of filters evaluated represents the capability of the layer to identify different patterns.

In terms of trainable parameters, a convolution layer with 32 kernels of 3×3 size is lighter then a dense one:

$$9 \cdot 32(\text{weights}) + 1 \cdot 32(\text{biases}) = 320.$$

MaxPool2D layer

A natural next-step to the convolution layer is represented by the category of pooling layers. This type of layers has the aim of reducing the size of the activation maps.

In particular, MaxPool2D layer consider a sub-matrix with a fixed size, typical 2×2 or 3×3, and record the max value. Performing the process over all the input matrix, the result is a smaller output matrix that keeps the geometrical feature information. A pictorial representation of the layer is shown in figure 5.6.

Figure 5.6: MaxPool2D effect sketched where the same color represent the sub-matrix considered and the corresponding output.

VGGNet structure

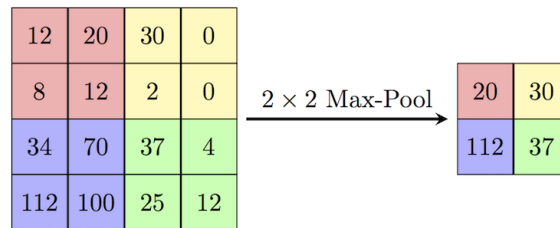The VGGNet structure is a neural network concept introduced in 2015 with the article "Very Deep Convolutional Networks for Large-Scale Image Recognition" [12] (the name VGG is the acronym for Visual Geometry Group, their lab in Oxford).

The proposed, and then accepted as a standard, powerful structure is composed by several small-size filter chain (i.e. kernel size of 1×1 or 3×3), max pooling with size of 2×2 is used after most, but not all, the convolutional layers. The idea is that consecutive small-size filters approximate larger filter effects with an higher number of parameters. Another important characteristic is the large number of filters used: typically deeper layer has greater number of kernels starting from, at least, 32.

In figure 5.7 the structure studied in [12] are listed where groups of two or three convolutional layers are followed by max pooling ones, at the end the last max pooling layer output is flatten and used as input for a group of dense layers till the classification.

A VGGNet like structure has been used to perform the task already introduced and it will be described in detail later (see paragraph 5.4).

ResNet structure

Residual Networks, or ResNet, are another innovative concept of CNN introduced in 2016 with the article "Deep Residual Learning for Image Recognition" by Kaiming He et al. [13].

The structure is composed as a plain convolutional network with small kernel size (same as in VGGNet), this sequence of Conv-layers is splitted in several *residual* blocks. The innovative aspect is that the input of each block is the sum of the input and the output of the previous block. The structure and flow of data are shown in figure 5.8, where two different ResBlock are sketched.

The last block is followed by an average pooling layer (similar to the max pooling layer, but it record the average value and not the max value) and then a structure of consecutive dense layer till the classification one with softmax as activation function.

A ResNet like structure has been used to perform the task already introduced and it will be described in detail later (see paragraph 5.4).

ConvNet Configuration					
A	A-LRN	B	C	D	E
11 weight layers	11 weight layers	13 weight layers	16 weight layers	16 weight layers	19 weight layers
input (224 × 224 RGB image)					
conv3-64	conv3-64	conv3-64	conv3-64	conv3-64	conv3-64
	LRN	**conv3-64**	conv3-64	conv3-64	conv3-64
maxpool					
conv3-128	conv3-128	conv3-128	conv3-128	conv3-128	conv3-128
		conv3-128	conv3-128	conv3-128	conv3-128
maxpool					
conv3-256	conv3-256	conv3-256	conv3-256	conv3-256	conv3-256
conv3-256	conv3-256	conv3-256	conv3-256	conv3-256	conv3-256
			conv1-256	**conv3-256**	conv3-256
					conv3-256
maxpool					
conv3-512	conv3-512	conv3-512	conv3-512	conv3-512	conv3-512
conv3-512	conv3-512	conv3-512	conv3-512	conv3-512	conv3-512
			conv1-512	**conv3-512**	conv3-512
					conv3-512
maxpool					
conv3-512	conv3-512	conv3-512	conv3-512	conv3-512	conv3-512
conv3-512	conv3-512	conv3-512	conv3-512	conv3-512	conv3-512
			conv1-512	**conv3-512**	conv3-512
					conv3-512
maxpool					
FC-4096					
FC-4096					
FC-1000					
soft-max					

Figure 5.7: Schematic representation of CNN structures studied by Karen Simonyan and Andrew Zisserman. Each column correspond to a CNN with different depth increasing from the left (A) to the right (E), as more layers are added (the added layers are shown in bold. The convolutional layer parameters are denoted as "conv[receptive field size]-[number of channels]".
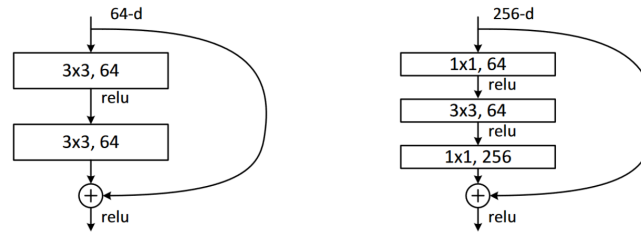

Figure 5.8: Schematic representation of residual blocks with arrows to indicate the data flow. On the left a simple two convolution block, on the right a "bottleneck" building block. Figure from [13].

(a) Cherenkov signal.
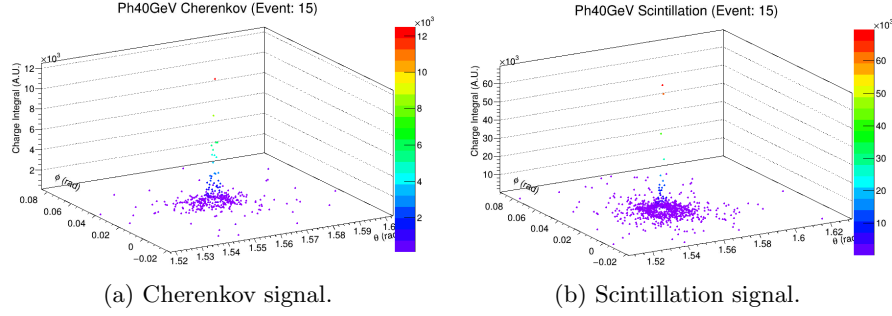
(b) Scintillation signal.

Figure 5.9: 3D-graphs representing Cherenkov and scintillation signals from the same sample event (40 GeV photon as primary particle).

5.3 Data setup

The data are produced through the IDEA DR Calorimeter full simulation, neutral pions and photons are fired from the interaction point to a fixed area of the calorimeter. In the first step particles with the fixed energy of 40 GeV are produced.

Useful data from the simulation are:

- starting fibre spatial coordinates (x, y, z);

- fibre type (Cherenkov or scintillating);

- charge integral from the SiPM digitization software.

The typical coordinate system for a 4π calorimeter is the spherical one so the cartesian coordinates have been transformed in spherical ones. Then, at each point the charge integral has been associated. Finally the are grouped in two subset by filtering with respect to the fibre types. The results obtained can be plotted in a 3D-graph to check the effective correctness of the process, an event taken as example can be seen in figure 5.9.

To train and take advantage of CNNs, the dataset has to be prepared. The input for VGGNet and ResNet has to be a dimension-fixed matrix for each event. Every event will be characterized by two features (Cherenkov charge integral and scintillation charge integral), each one of this represented by a grid reproducing the spatial distribution of the data. The squared area of interest in the $\theta-\phi$ space has been identified as $[(1.51, -0.02), (1.63, 0.10)]$. Different values of grid step in this area has been compared reaching a good compromise between grid shape efficiency and imaging resolution choosing a grid step of 0.0009 rad for both axes. The grid can be represented as 2D-histograms using the sum of all the charge integral values in each bin as bin height (figure 5.10 shows the histograms obtained from the same data in figure 5.9).

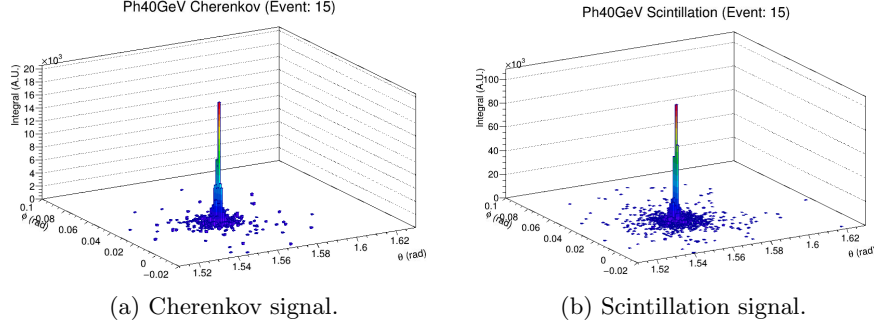(a) Cherenkov signal.   (b) Scintillation signal.

Figure 5.10: 2D-histograms representing Cherenkov and scintillation signals from the same sample event (40 GeV photon as primary particle).

Hence the input matrix for each event has a shape of 133 (height) $\times$ 133 (width) $\times$ 2 (features). A 2D visualization can be seen in figure 5.11.

The complete process of data preparation is applied on 10000 events of photons and 10000 events of neutral pions. All the events have been labelled (γ and π^0), normalized to the max value of 1 (same normalization constant for all the dataset), shuffled and splitted in two sub-set dedicated to training (80%) and to validation (20%).

5.4 Performances

Analyses on the prepared data have been performed using the two different convolutional neural network structures already introduced, i.e. VGG Network and Residual Network. Some of the *hyperparameters* of the CNNs, namely all the parameters that identify the structure of the networks such as the number of layers, the number of filters and the kernel size in convolutional layers, the number of neurons in dense layers.
This section shows the comparison between CNNs with different hyperparameters and details of the best VGGNet and ResNet obtained.

5.4.1 VGG Network

The VGG neural networks used for the particle ID task follow the structure described in section 5.2.
The input and output are well defined by the data preparation and the classification task. The elaborated data set the input as ma 3D matrix with dimensions of $133 \times 133 \times 2$, meanwhile the aim of distinguishing between two types of primary particle set the last layer as a dense layer with two neurons with softmax as the activation function.

(a) Cherenkov signal pre data setup.

(b) Scintillation signal pre data setup.

(c) Cherenkov signal post data setup.
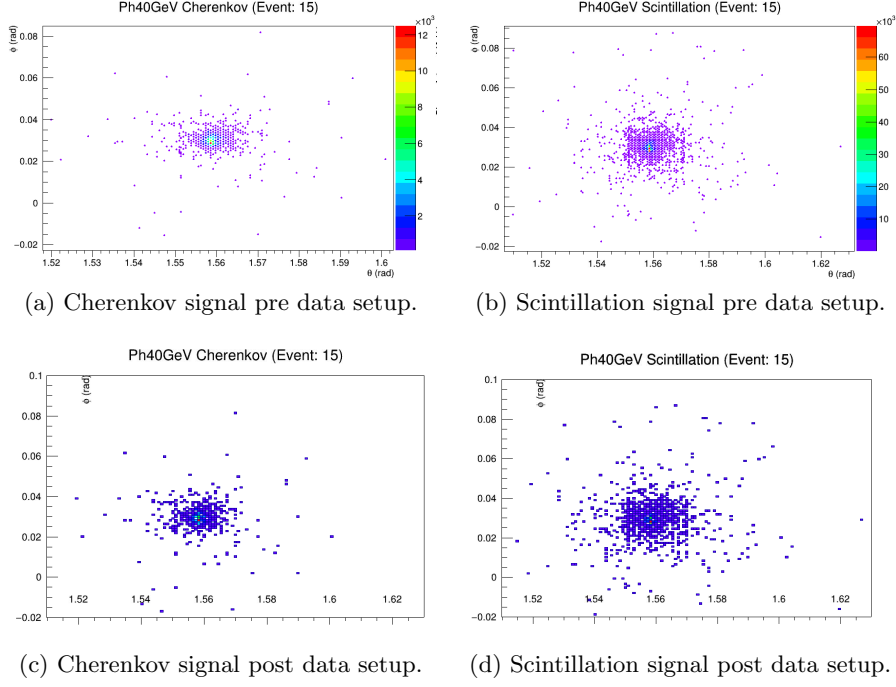

(d) Scintillation signal post data setup.

Figure 5.11: 2D-vision of data before and after data setup.

The optimization process has to take into account a very large number of hyperparameters combinations, to make it simple and tidy the layers can be divided in two groups separated by the flatten one: the "convolutional half" and the "dense half". The halves have been analyzed one at a time keeping fixed the other one and ranging the hyperparameters in common values.
In table 5.12 the structure of the six different VGGNet tested are shown with the accuracy value and the time of training to evaluate the performances.
The neural networks labelled as VGGNet A, VGGNet B and VGGNet C have the aim to optimize the dense half modifying the Dropout probability. Once the best dense half has been fixed the VGGNet D, E and F have been tested to select the best hyperparameters in the convolution half. The conclusion from this study set the best structure as the VGGNet D with an accuracy of 98.875 and a training time of 212 s/epoch.

Once identified the VGGNet D as the best structure for the task deeper study have been performed.
The training is applied on a training dataset and simultaneously validated on a validation dataset, this can give a double view during the process. The accuracy is probably the most important aspect for a neural network, it is evaluated as the number of correct predictions over the total number of predictions (n_c/N). Figure 5.13 shows the behaviour of the accuracy during

	VGGNet A	VGGNet B	VGGNet C	VGGNet D	VGGNet E	VGGNet F
	Input (133x133x2)	Input (133x133x2)	Input (133x133x2)	Input (133x133x2)	Input (133x133x2)	Input (133x133x2)
	Conv2D (32,3x3,ReLU)	Conv2D (32,3x3,ReLU)	Conv2D (32,3x3,ReLU)	Conv2D (32,3x3,ReLU)	Conv2D (32,3x3,ReLU)	Conv2D (32,3x3,ReLU)
	Conv2D (32,3x3,ReLU)	Conv2D (32,3x3,ReLU)	Conv2D (32,3x3,ReLU)	Conv2D (32,3x3,ReLU)	Conv2D (32,3x3,ReLU)	Conv2D (32,3x3,ReLU)
	MaxPool	MaxPool	MaxPool	MaxPool	MaxPool	MaxPool
	Conv2D (64,3x3,ReLU)	Conv2D (64,3x3,ReLU)	Conv2D (64,3x3,ReLU)	Conv2D (64,3x3,ReLU)	Conv2D (32,3x3,ReLU)	Conv2D (64,3x3,ReLU)
	Conv2D (64,3x3,ReLU)	Conv2D (64,3x3,ReLU)	Conv2D (64,3x3,ReLU)	Conv2D (64,3x3,ReLU)	Conv2D (32,3x3,ReLU)	Conv2D (64,3x3,ReLU)
	MaxPool	MaxPool	MaxPool	MaxPool	MaxPool	MaxPool
	Conv2D (128,3x3,ReLU)	Conv2D (128,3x3,ReLU)	Conv2D (128,3x3,ReLU)	Conv2D (64,3x3,ReLU)	Conv2D (64,3x3,ReLU)	Conv2D (128,3x3,ReLU)
	Conv2D (128,3x3,ReLU)	Conv2D (128,3x3,ReLU)	Conv2D (128,3x3,ReLU)	Conv2D (64,3x3,ReLU)	Conv2D (64,3x3,ReLU)	Conv2D (128,3x3,ReLU)
	MaxPool	MaxPool	MaxPool	MaxPool	MaxPool	MaxPool
	Conv2D (128,3x3,ReLU)	Conv2D (128,3x3,ReLU)	Conv2D (128,3x3,ReLU)	Conv2D (128,3x3,ReLU)	Conv2D (64,3x3,ReLU)	Conv2D (256,3x3,ReLU)
	Conv2D (128,3x3,ReLU)	Conv2D (128,3x3,ReLU)	Conv2D (128,3x3,ReLU)	Conv2D (128,3x3,ReLU)	Conv2D (64,3x3,ReLU)	Conv2D (256,3x3,ReLU)
	MaxPool	MaxPool	MaxPool	MaxPool	MaxPool	MaxPool
	Conv2D (128,3x3,ReLU)	Conv2D (128,3x3,ReLU)	Conv2D (128,3x3,ReLU)	Conv2D (128,3x3,ReLU)	Conv2D (64,3x3,ReLU)	Conv2D (256,3x3,ReLU)
	Conv2D (128,3x3,ReLU)	Conv2D (128,3x3,ReLU)	Conv2D (128,3x3,ReLU)	Conv2D (128,3x3,ReLU)	Conv2D (64,3x3,ReLU)	Conv2D (256,3x3,ReLU)
	MaxPool	MaxPool	MaxPool	MaxPool	MaxPool	MaxPool
	Flatten	Flatten	Flatten			
	Dense(256,ReLU)	Dense(256,ReLU)	Dense(256,ReLU)			
	Dense(256,ReLU)	Dropout(0.3)	Dropout(0.5)			
	Output(2,SoftMax)	Dense(256,ReLU)	Dense(256,ReLU)			
		Dropout(0.3)	Dropout(0.5)			
		Output(2,SoftMax)	Output(2,SoftMax)			
Accuracy (%)	97,275					
Time per Epoch (s)	230	230				

Figure 5.12: Table of different VGGNet structure studied. Performances are valuated looking at accuracy and training time values.

the training process by overlapping the results from training and validation datasets.

The loss function is another important feature that can be studied. It has been already introduced as error function 5.8, in the specific case it get simpler noting that there are only two labels ($K = 2$). A loss value can be evaluated at the end of each training epoch to monitor the improvements in each training step. Figure 5.14 shows the behaviour just described.

The smoothness of both accuracy and loss plots is an indication of the good dataset size, if a too small dataset would be used the graphs would present high spikes.

Results on test data are often represented in confusion matrices, matrices showing the percentage probability of a neural network in identifying categories emphasizing true positive, true negative, false positive and false negative. The confusion matrix obtained with the VGGNet selected is shown in figure 5.15.

5.4.2 Residual Network

The residual neural networks used for the particle ID task follow the structure described in section 5.2.
The input and output of the ResNets has to coincide with the ones for VG-GNets, due to the same data preparation process and the same classification task. The input data will have a 3D matrix shape ($133 \times 133 \times 2$) and the last layer is set to be the a two-neurons dense layer with the softmax activation
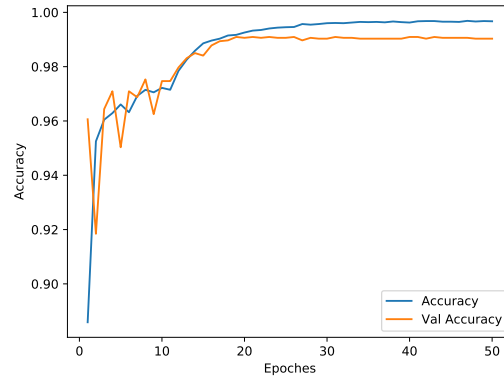
Figure 5.13: Accuracy behaviour over 50 training epoches using the selected VGGNet. Train and validation accuracy are shown with different colors.
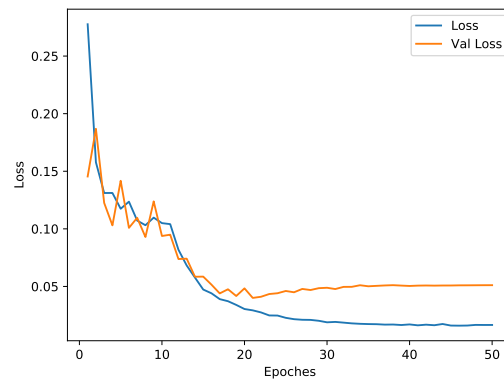

Figure 5.14: Loss behaviour over 50 training epoches using the selected VGGNet. Train and validation loss are shown with different colors.
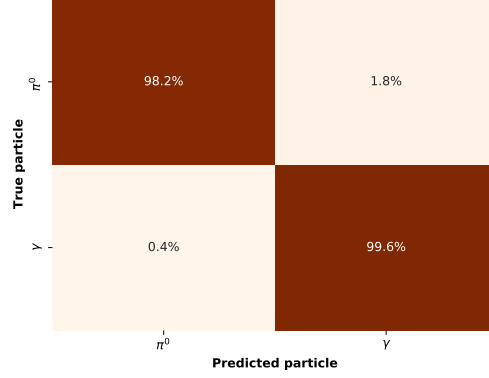
Figure 5.15: Confusion matrix obtained on test date using the selected VG-GNet. Values are normalized on rows (i.e. on the true label).

	ResNet B	ResNet A	ResNet C	ResNet D	ResNet E	ResNet F
	Input (133x133x2)	Input (133x133x2)	Input (133x133x2)	Input (133x133x2)	Input (133x133x2)	Input (133x133x2)
	Conv2D (32,3x3,ReLU)	Conv2D (32,3x3,ReLU)	Conv2D (32,3x3,ReLU)	Conv2D (32,3x3,ReLU)	Conv2D (32,3x3,ReLU)	Conv2D (32,3x3,ReLU)
	Conv2D (32,3x3,ReLU)	Conv2D (32,3x3,ReLU)	Conv2D (32,3x3,ReLU)	Conv2D (32,3x3,ReLU)	Conv2D (32,3x3,ReLU)	Conv2D (32,3x3,ReLU)
					Conv2D (32,3x3,ReLU)	Conv2D (32,3x3,ReLU)
	Conv2D (64,3x3,ReLU)	Conv2D (64,3x3,ReLU)	Conv2D (64,3x3,ReLU)	Conv2D (64,3x3,ReLU)	Conv2D (32,3x3,ReLU)	Conv2D (64,3x3,ReLU)
	Conv2D (64,3x3,ReLU)	Conv2D (64,3x3,ReLU)	Conv2D (64,3x3,ReLU)	Conv2D (64,3x3,ReLU)	Conv2D (32,3x3,ReLU)	Conv2D (64,3x3,ReLU)
					Conv2D (32,3x3,ReLU)	Conv2D (64,3x3,ReLU)
	Conv2D (64,3x3,ReLU)	Conv2D (64,3x3,ReLU)	Conv2D (64,3x3,ReLU)	Conv2D (64,3x3,ReLU)	Conv2D (32,3x3,ReLU)	Conv2D (64,3x3,ReLU)
	Conv2D (64,3x3,ReLU)	Conv2D (64,3x3,ReLU)	Conv2D (64,3x3,ReLU)	Conv2D (64,3x3,ReLU)	Conv2D (32,3x3,ReLU)	Conv2D (64,3x3,ReLU)
					Conv2D (32,3x3,ReLU)	Conv2D (64,3x3,ReLU)
	Conv2D (64,3x3,ReLU)	Conv2D (64,3x3,ReLU)	Conv2D (64,3x3,ReLU)	Conv2D (64,3x3,ReLU)	Conv2D (32,3x3,ReLU)	Conv2D (64,3x3,ReLU)
	Conv2D (64,3x3,ReLU)	Conv2D (64,3x3,ReLU)	Conv2D (64,3x3,ReLU)	Conv2D (64,3x3,ReLU)	Conv2D (32,3x3,ReLU)	Conv2D (64,3x3,ReLU)
					Conv2D (32,3x3,ReLU)	
	GlobalAveragePooling	GlobalAveragePooling	GlobalAveragePooling	GlobalAveragePooling		
	Dense(256,ReLU)	Dense(128,ReLU)	Dense(256,ReLU)	Dense(256,ReLU)		
	Dense(256,ReLU)	Dense(128,ReLU)	Dropout(0.3)	Dropout(0.5)		
	Output(2,SoftMax)	Output(2,SoftMax)	Dense(256,ReLU)	Dense(256,ReLU)		
			Dropout(0.3)	Dropout(0.5)		
			Output(2,SoftMax)	Output(2,SoftMax)		
Accuracy (%)	97,3					
Time per Epoch (s)						

Figure 5.16: Table of different ResNet structure studied. Performances are valuated looking at accuracy and training time values. Note that the thicker lines divide the different Recurrent blocks as described in section 5.2.

function.

The articulation of the middle layers follows the typical ResNet structure presented in section 5.2.

As done for the VGGNet optimization, also this structure has been divided in two parts identified by the *GlobalAveragePooling* layer ("colvolutional half" and "dense half"). In table 5.16 the six different ResNet with different hyperparameters are listed. ResNet A, B, C and D are dedicated to the optimization of the dense half. Once the best dense half structure has been found, the ResNet E and F have been trained to find the best ResNet for the task. As schematically shown in the table, ResNet D has the best performances with an accuracy of 97.275% and a training time of 192 s/epoch.
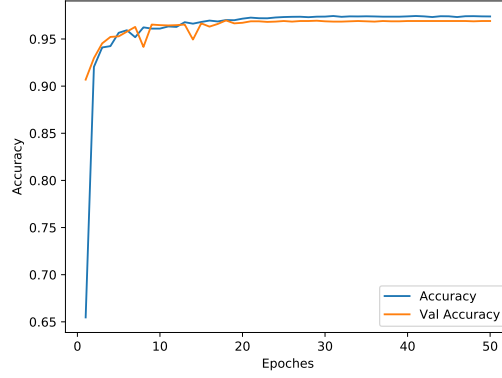
Figure 5.17: Accuracy behaviour over 50 training epoches using the selected ResNet. Train and validation accuracy are shown with different colors.

Residual Network and VGG Network performances have been studied in same conditions such as batch size (128) and number of epoches (50) and also with the same accuracy and loss evaluation.
Both accuracy and loss values have been recorded during the training process and their behaviour with respect to the training epoches is shown in figures 5.17 and 5.18.
Also in this case a confusion matrix has been produced on validation data reaching the results shown in figure 5.19.
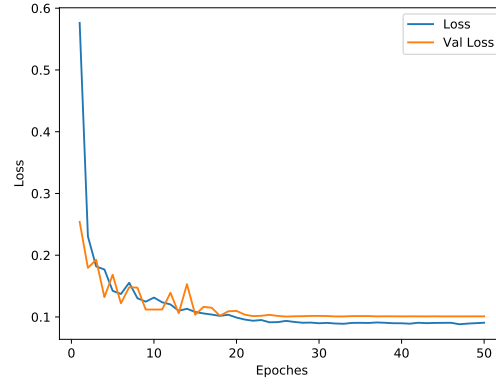
5.5 Energy range extension

aaa

Figure 5.18: Loss behaviour over 50 training epoches using the selected ResNet. Train and validation loss are shown with different colors.

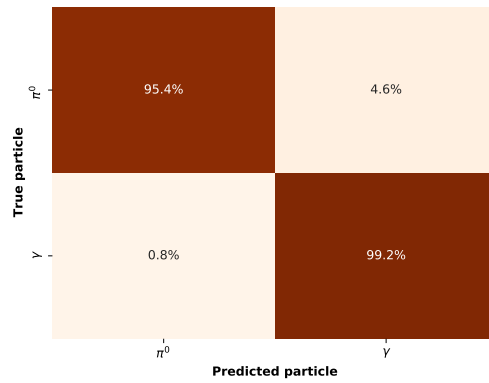
Figure 5.19: Confusion matrix obtained on test date using the selected ResNet. Values are normalized on rows (i.e. on the true label).

Chapter 6

Conclusion

aaa

Thanks

aaa

Bibliography

[1] R. Wigmans, «Calorimetry - Energy Measurement In Particle Physics», second ed., in:International Series of Monographs on Physics, vol. 168, Oxford University Press, 2017.

[2] C. W. Fabjan, F. Gianotti, «Calorimetry for particle physics», Rev. Mod. Phys. 75(2003) 1243-1286.

[3] W. R. Leo, «Techniques for nuclear and particle physics experiments», second ed., Springer-Verlag, 1994.

[4] R. Wigmans and M. Zeyrek, «On the differences between calorimetric detection of electrons and photons», Nucl. Instr. and Meth. A485 (2002) 385

[5] L.Pezzotti. «Dual readout calorimetry development for future collider experiments». Master thesis. Università degli Studi di Pavia, 2016.

[6] R.Santoro. «International Conference on New Photo-Detectors (PD15) (July, 2015)». In:Moscow, Troitsk(2015) (cit. on p. 132).

[7] Hamamatsu SiPMs lineup https://www.hamamatsu.com/eu/en/product/optical-sensors/mppc/mppc_mppc-array/all_products/index.html

[8] P.A. Zyla et al. «Particle Data Group», Prog. Theor. Exp. Phys. 2020, 083C01 (2020).

[9] C.M. Bishop «Pattern recognition and machine learning». Springer, 2006.

[10] C.M. Bishop «A high-bias, low-variance introduction to Machine Learning for physicists». arXiv:1803.08823v3, 2019.

[11] Jason Brownlee «Deep Learning for Computer Vision: Image Classification, Object Detection, and Face Recognition in Python». Machine Learning Mastery, 2019.

[12] Karen Simonyan and Andrew Zisserman ≪Very Deep Convolutional Networks for Large-Scale Image Recognition≫. arXiv:1409.1556v6, 2015.

[13] Kaiming He et al. ≪Deep Residual Learning for Image Recognition≫. arXiv:1512.03385v1, 2015.