

Università degli Studi di Pavia

---

FACOLTÀ DI SCIENZE MATEMATICHE, FISICHE, NATURALI

Corso di laurea in Scienze Fisiche

**Dual-Readout calorimetry:  
studies of a SiPM-based simulated readout  
and deep learning applications**

Candidato

**Alessandro Villa**

Matricola 462495

Relatore

**Dott. Roberto Ferrari**

Correlatore

**Dott. Lorenzo Pezzotti**



# Contents

<b>Introduction</b>	<b>iii</b>
<b>1 Future colliders</b>	<b>1</b>
1.1 Physics goals . . . . .	1
1.2 Leptonic colliders . . . . .	2
1.3 IDEA Detector concept . . . . .	7
1.3.1 Vertex detector . . . . .	8
1.3.2 Drift chamber . . . . .	9
1.3.3 Magnet system . . . . .	10
1.3.4 Dual-readout calorimeter . . . . .	11
1.3.5 Preshower and muon chambers . . . . .	11
<b>2 Calorimetry and the DR method</b>	<b>15</b>
2.1 Physics of shower development . . . . .	16
2.1.1 Electromagnetic showers . . . . .	16
2.1.2 Hadronic showers . . . . .	20
2.2 The calorimeter response . . . . .	22
2.2.1 Fluctuations . . . . .	26
2.3 Dual-readout compensation . . . . .	27
2.3.1 Working principles . . . . .	27
2.3.2 Dual-readout sampling calorimeters . . . . .	30
<b>3 Silicon Photomultipliers</b>	<b>35</b>
3.1 Working principles . . . . .	35
3.2 SiPM photon detection efficiency . . . . .	39
3.2.1 Linearity and occupancy effect . . . . .	42
3.3 Noise effects . . . . .	42
3.3.1 Dark Count Rate . . . . .	43
3.3.2 Optical Cross-Talk . . . . .	43
3.3.3 Afterpulse . . . . .	44

<b>4 IDEA DR calorimeter simulation</b>	<b>45</b>
4.1 Simulation structure . . . . .	46
4.1.1 Calorimeter simulation . . . . .	46
4.1.2 SiPM response digitization . . . . .	49
4.2 Results from the full simulation chain . . . . .	50
4.2.1 Different configurations . . . . .	50
4.2.2 Time studies . . . . .	50
4.2.3 Occupancy effect . . . . .	56
4.2.4 Energy resolution . . . . .	62
<b>5 Deep neural network applications</b>	<b>73</b>
5.1 Physics benchmark . . . . .	73
5.2 Neural Networks introduction . . . . .	74
5.3 Data preparation . . . . .	83
5.4 Performance . . . . .	84
5.4.1 VGG Network . . . . .	89
5.4.2 Residual Network . . . . .	93
5.5 Energy range extension . . . . .	95
<b>Conclusion</b>	<b>107</b>
<b>Bibliography</b>	<b>109</b>

# Introduction

The discovery of the Higgs boson in 2012 at CERN has opened the door for several new studies in particle physics. The Higgs boson plays a crucial role in the biggest mysteries of modern particle physics and the precise measurement of its properties, together with the properties of the  $W$  and  $Z$  bosons, will provide deep information about the Standard Model (SM) and the physics Beyond it (BSM). At present, several projects are ongoing to design the next high-energy collider and its detectors with the primary task of measuring the Higgs properties.

In the first chapter, a short introduction to the main goals of modern particle physics is presented. The current design of possible future electroweak collider facilities is described, pointing out their capabilities to improve our knowledge of the field. Finally, the IDEA detector concept is outlined, being one of the multi-purpose detector concepts at future electroweak colliders.

The second chapter provides a theoretical background on calorimetry, starting from the interaction between particles and matter and the particle showering mechanism. A description of modern calorimeters is also provided, distinguishing between electromagnetic and hadronic calorimeters. Eventually, the dual-readout compensation technique, cornerstone of the IDEA calorimeter, is introduced.

The third chapter is dedicated to the Silicon PhotoMultipliers. They are the photodetectors envisaged to be coupled to the IDEA dual-readout calorimeter. Their working principle is described, followed by their characteristic features such as gain, photon detection efficiency, linearity and noise sources.

The fourth chapter describes the simulation chain of the IDEA dual-readout calorimeter combining the Monte Carlo simulation of the calorimeter with the readout system simulation, emulating the SiPM response. Several analyses were performed to evaluate the performance and validate the full simulation model. They are described here for the first time.

The last chapter introduces the application of deep learning algorithms to the IDEA calorimeter simulated data. Two different algorithms have been used to perform the very

challenging task of discriminating between neutral pions and photons. Indeed, the high spatial resolution of the IDEA calorimeter allows to reconstruct the two-dimensional pattern of particle showers with great accuracy, thus providing a powerful input for neural-network applications. Encouraging results were found and are extensively reported in the thesis.

# Chapter 1

## Future colliders

The chapter begins with a short description of the most important goals in today’s particle physics, starting from the possibilities opened by the discovery of the Higgs boson. It continues with an overview on the different future projects for high-energy  $e^+e^-$  colliders, both circular and linear (also known as “ElectroWeAK (EWK) factories”). Eventually, the IDEA Detector concept is described as a general-purpose detector proposed for next-generation  $e^+e^-$  EWK factories.

### 1.1 Physics goals

With the discovery of the Higgs boson ( $H$ ) in 2012 by the ATLAS and CMS Collaborations [1, 2], a new era has been opened where the new boson is not only a research object but also a tool for new particle physics studies.

Thanks to the relatively small mass,  $\simeq 125$  GeV, the Higgs boson production is within the reach of high-luminosity future circular colliders. Their main production mechanism is the so-called Higgs-strahlung process, in which the Higgs recoils over a  $Z$ -boson produced at  $E_{cm} \simeq 240$  GeV ( $e^+e^- \rightarrow ZH$ ). In  $ZH$  events, the recoil mass (defined as the invariant mass of the system recoiling against the  $Z$  boson) peaks at the Higgs mass, thus allowing to tag the events independently of the  $H$  decay mode. This process allows to measure the Higgs production cross section in a model-independent way. The other couplings are eventually extracted through branching-ratio and width measurements with a subpercent precision.

The  $H$  couplings to the first Standard Model family particles (i.e. electron, quark up and quark down), because of their small masses and related decay branching ratios, will not be directly measurable at these colliders. However, with beam energies  $\simeq 125.09$  GeV, corresponding to the  $H$  pole mass, leptonic colliders can contribute to set upper limits to the electron Yukawa coupling by taking advantage from the resonant  $H$  production. Also the  $t$ -quark Yukawa coupling and the  $H$  self-coupling will not be directly measurable because their masses are too large for kinematically open production and decay. However, the CERN future electron-positron circular collider (FCC-ee), operating

at  $\sqrt{s} = 350$  GeV, could measure the  $t$  Yukawa coupling with a precision of 10% thanks to higher order corrections of processes at the  $t\bar{t}$  threshold.

## 1.2 Leptonic colliders

Precise measurements of Higgs boson properties, together with those of the  $Z$  and  $W$  bosons, will provide important tests of the SM fundamental physics principles and will be essential for Beyond-the-Standard-Model (BSM) physics studies. At present, the landscape of high energy physics includes four different electroweak-factory proposals:

- the *Future Circular Collider  $e^+e^-$*  (FCC-ee), at CERN;
- the *Circular Electron Positron Collider* (CEPC), in China;
- the *International Linear Collider* (ILC), in Japan;
- the *Compact LInear Collider* (CLIC), at CERN.

Fig 1.1 shows the possible timeline envisaged for the main future high-energy collider projects. In the following sections, a brief description of each one of them will be provided.

### Future Circular Collider $e^+e^-$

A post-LHC circular collider at CERN has been proposed with the name of Future Circular Collider (FCC) project [4]. FCC is staged in a first lepton collider (FCC-ee) [5] phase followed by a hadron collider (FCC-hh) [6] with a final target centre-of-mass energy of 100 TeV. A common tunnel, about 100 km long, is designed to host both of them. With this choice the same facility could potentially house an electron-hadron collider.

FCC-ee is designed to provide the highest possible statistics for the  $Z$ ,  $W$  and  $H$  bosons, and  $t$  quarks. At present, it is supposed to operate at centre-of-mass energies ranging from 88 to 365 GeV in four different ( $\sqrt{s}$ ) operating points:

- $\simeq 91$  GeV, corresponding to the  $Z$  pole;
- $\simeq 160$  GeV, corresponding to the  $W^+W^-$  production threshold;
- $\simeq 240$  GeV, corresponding to the  $ZH$  production threshold;
- $\simeq 340 - 365$  GeV, corresponding to the  $t\bar{t}$  threshold.

The FCC-ee project would fit over the present CERN accelerator complex, where the injector chain would make use of a 6 GeV linac, a damping ring and the CERN SPS as a pre-booster. The baseline layout, sketched in Figure 1.2, is designed with two different interaction points.

The task of increasing our knowledge of many of the electroweak observables, by one or

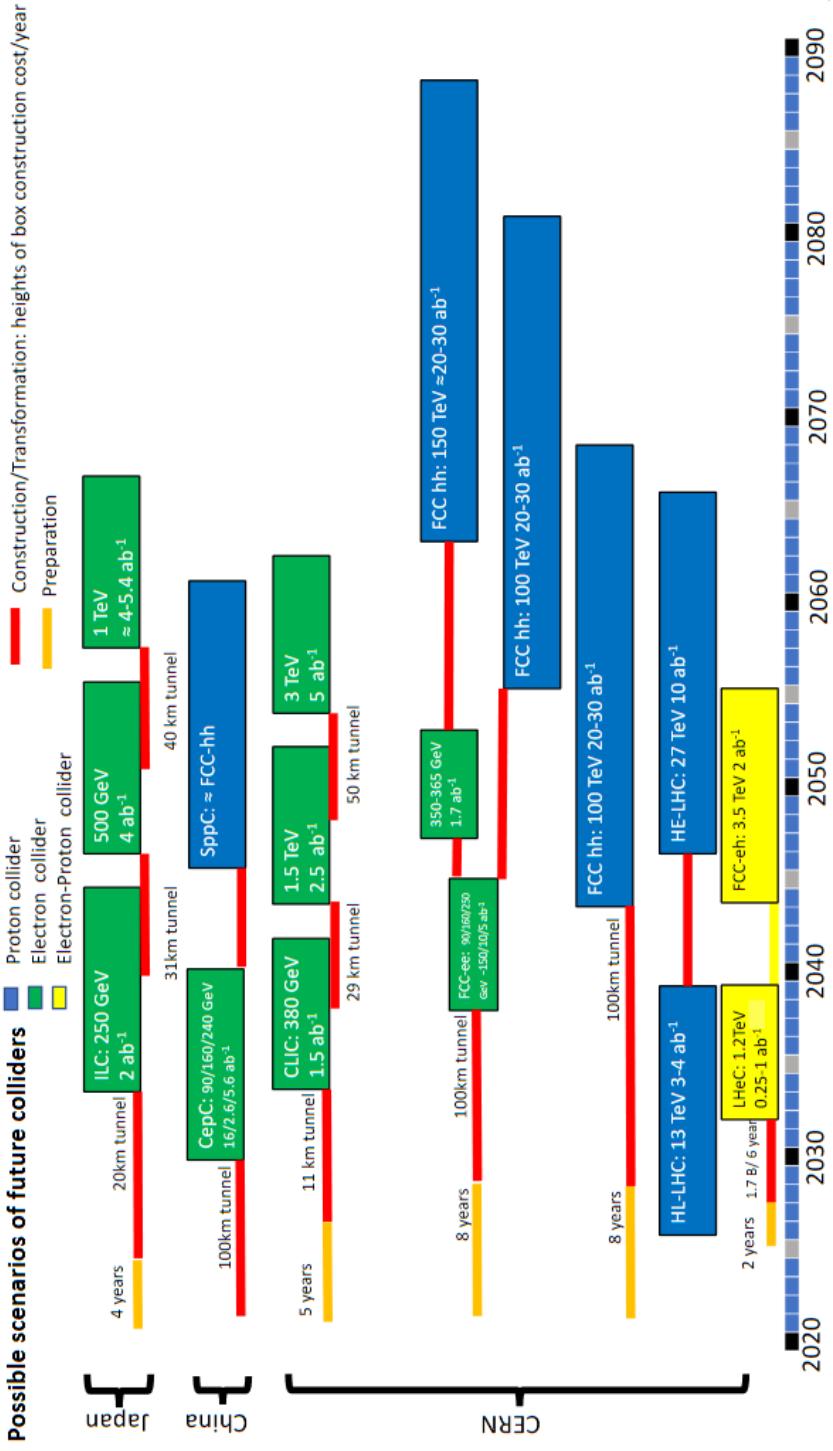


Figure 1.1: Possible timelines of future high-energy colliders. It includes high-energy  $e^+e^-$  (ILC, CLIC, CEPC and FCC-ee),  $pp$  (FCC-hh and HL-LHC) and  $e-p$  (LHeC and FCC-eh) machines. Image from [3].

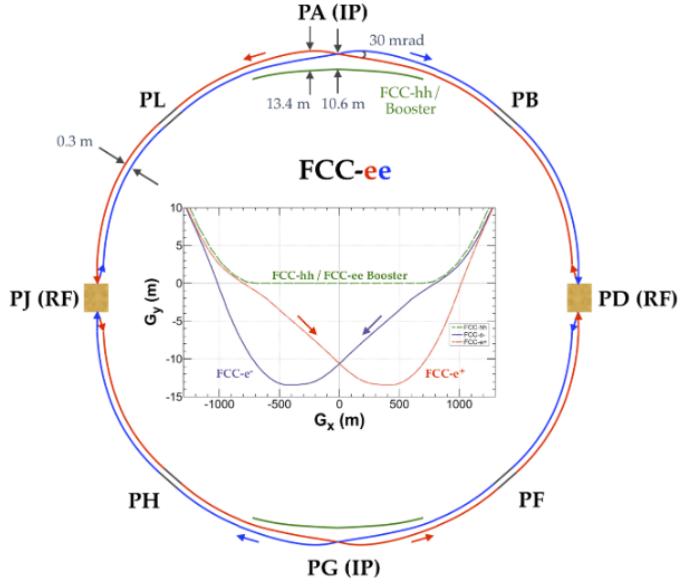


Figure 1.2: Schematic view of the FCC-ee. Image from [4].

two orders of magnitude better than the current one, sets challenging constraints to the integrated luminosities needed. These values range from  $0.2 \text{ ab}^{-1}$ , for the measurement of the top-quark mass and width, to  $100 \text{ ab}^{-1}$ , for the measurement of the effective weak mixing angle. These luminosities can be achieved in a reasonable amount of time only by circular colliders (see Figure 1.3).

### Circular Electron Positron Collider

The Circular Electron Positron Collider (CEPC) is an international project initiated and hosted by China addressing high-luminosity  $e^+e^-$  collisions. It is designed to operate at centre-of-mass energies of 240 GeV (as an  $H$  factory exploiting the  $e^+e^- \rightarrow ZH$  process), 91.2 GeV ( $Z$  pole) and 160 GeV ( $W^+W^-$  threshold).

The present design, similarly to the FCC-ee, foresees a double ring structure, sketched in Figure 1.4 with a circumference of 100 km and two interaction points. The same 100 km-long tunnel could also host a Super Proton-Proton Collider (SPPC) giving, without removing the CEPC ring, the possibility to perform as well electron-proton collisions. As described in the conceptual design report [7, 8], the main accelerator is preceded by a linear accelerator, a damping ring and a booster.

Associated to the three operating  $\sqrt{s}$  values, the instantaneous luminosities are expected to reach  $3 \times 10^{34}$ ,  $32 \times 10^{34}$  and  $10 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$ , respectively. CEPC will produce, over its planned operative time, large samples (more than one million) of Higgs, one trillion of  $Z$  bosons and about 100 million of  $W^+W^-$  events, allowing precise measurements of

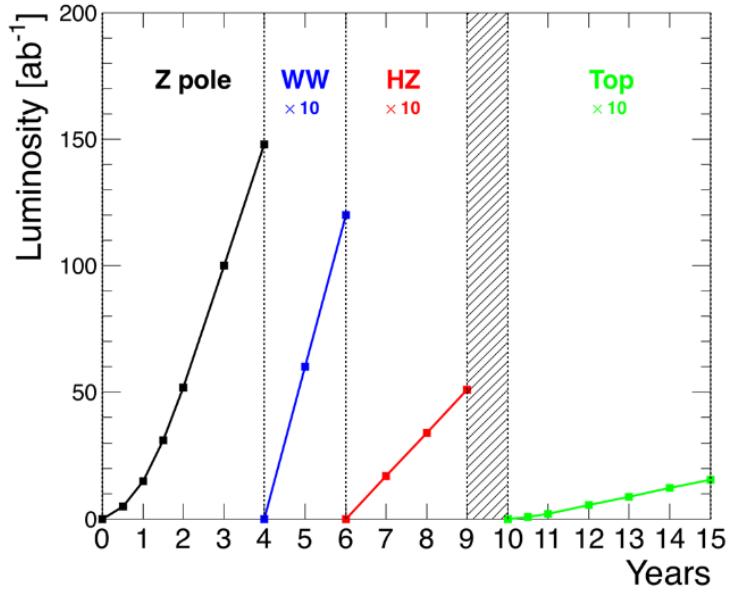


Figure 1.3: Integrated FCC-ee luminosity during 15 years of operation. Image from [5].

many electroweak parameters.

According to [9], the CEPC construction could possibly start within few years in order to be completed by 2030, a very aggressive time schedule compared to the other collider proposals.

### International Linear Collider

The ILC is an  $e^+e^-$  linear collider that, at the current state (ILC250), provides a 250 GeV centre-of-mass energy. The ILC accelerator is based on SuperConducting RadioFrequency (SCRF) cavities already in use at the European X-ray Free Electron Laser facility (E-XFEL), located at DESY/Hamburg.

The design luminosity is  $1.35 - 1.5 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$  at  $E_{cm} = 250 \text{ GeV}$ , with a corresponding integrated luminosity ranging from  $400 \text{ fb}^{-1}$  (in the first years) up to  $2 \text{ ab}^{-1}$  (after future upgrades). The ILC250 is 20.5 km long with two main arms, mostly occupied by the electron and positron linacs, at a 14 mrad crossing angle. The ILC candidate site is in the Kitakami region in northern Japan. A sketch of the ILC250 is shown in Figure 1.5.

In its current state, the SCRF cavities will reach frequency values of 1.3 GHz, providing a gradient of  $31.5 - 35 \text{ MV/m}$  while operating in a cryogenic infrastructure at 2 K.

The first stage of ILC has the task to measure the  $H$  parameters and its model-

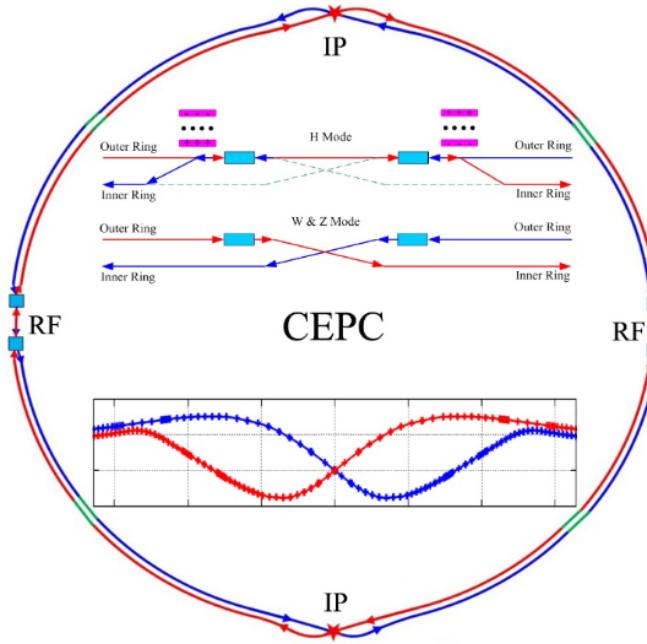


Figure 1.4: Schematic view of the CEPC. Image from [8].

independent determination by studying the  $e^+e^- \rightarrow ZH$  production at  $\sqrt{s} = 250$  GeV. Two energy upgrades are currently foreseen for extending the centre-of-mass energy to 500 GeV and 1 TeV. The main goals for the higher-energy physics runs cover improved precision measurements of the top-quark mass, the top-quark electroweak couplings, the Higgs coupling to the top quark, and the triple-Higgs coupling [10].

### Compact Linear Collider

The Compact Linear Collider (CLIC) is a TeV-scale high-luminosity linear electron-positron collider to be located in the CERN area. The CLIC energy stages, at present, comprises 3 working points, at  $\sqrt{s} = 380$  GeV, 1.5 TeV and 3.0 TeV, with corresponding instantaneous luminosities of 1.5, 3.7 and  $5.9 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$ , respectively. The site length will have to scale, in these stages, from 11 km up to 50 km.

The CLIC project, proposed in 2012 [12, 13, 14] and updated in 2016 [15], will adopt a two-beam acceleration scheme as shown in Figure 1.6, where the electron and positron beams are independently accelerated through the whole chain.

In the first stages, particles are accelerated to 9 GeV using a linac booster. Then they are injected in normal-conducting high-gradient 12 GHz accelerating structures. The two main linacs accelerate beams exploiting normal-conducting  $X$ -band cavities with an

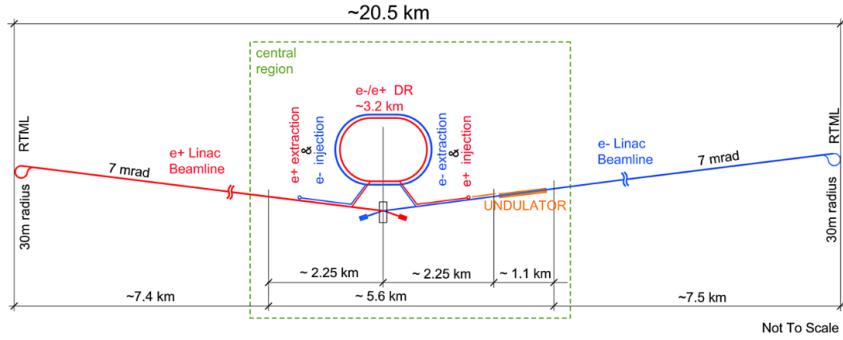


Figure 1.5: Schematic layout of the ILC at 250 GeV staged option [10].

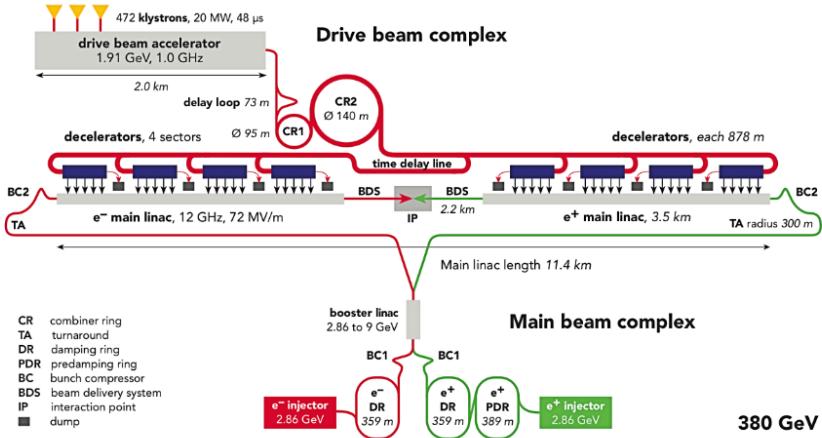


Figure 1.6: Schematic layout of the CLIC at the 380 GeV staged option [16].

accelerating gradient of 100 MV/m, the highest among linear colliders. To reach these extremely high accelerating gradients, a novel drive-beam scheme, using low-frequency klystrons to generate long RF pulses and store their energy in a long, high-current, drive-beam pulse, is designed. This beam pulse is used to generate several short pulses distributed along the main linac.

### 1.3 IDEA Detector concept

IDEA (Innovative Detector for Electron-positron Accelerators) is an innovative multi-purpose detector concept, designed to study electron-positron collisions in a wide energy range as provided by high-luminosity electroweak factories. The detector concept was proposed in 2017 and included in conceptual design reports of both FCC-ee [17] and CEPC [18].

The IDEA Detector concept is sketched in an artistic view in Figure 1.7(a) and in

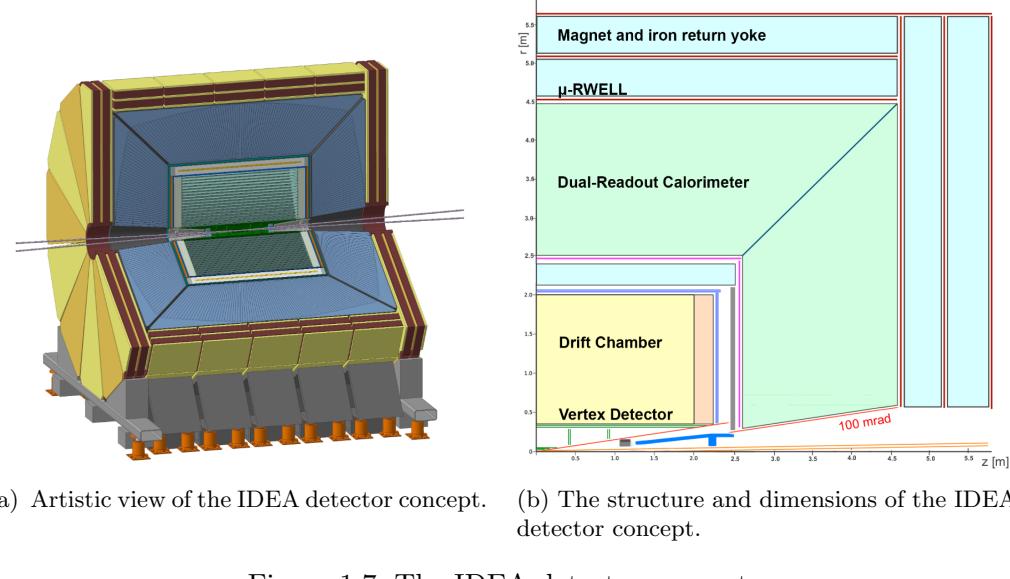


Figure 1.7: The IDEA detector concept.

its structure and dimensions in Figure 1.7(b). The overall detector is composed by a silicon pixel vertex detector, a wire chamber surrounded by a layer of silicon micro-strip detectors, a thin superconducting solenoid coil, a preshower detector, a dual-readout calorimeter and a muon spectrometer. The most innovative elements proposed are:

- an ultra-light drift chamber as main tracker;
- a Dual-Readout (DR) fibre-sampling calorimeter, for both electromagnetic and hadronic shower measurement.

The Drift CHamber (DCH) technology is based on the R&D work done for the upgrade of the MEG experiment (MEG II), designed to search for the charged-lepton flavour violating decay  $\mu \rightarrow e\gamma$ . This work involves the construction of an ultra-light drift chamber providing high momentum resolution and high transparency in terms of radiation length. The IDEA dual-readout calorimeter, on the other hand, stands on the legacy of the DREAM/RD52 Collaboration. The key point is the expected performance of dual-readout fibre-sampling calorimeters in obtaining high-resolution energy measurements for both single hadrons and hadronic jets. All the most important parameters of the IDEA detector components are listed in Table 1.1.

### 1.3.1 Vertex detector

The 1.5 cm beam pipe is surrounded by the IDEA vertex detector composed by active pixel sensors. The structure present a high-resistivity substrate architecture implement-

Vertex technology	Silicon
Vertex inner/outer radius (cm)	1.7/34
Tracker technology	Drift Chamber and Silicon Wrapper
Tracker half length (m)	2.0
Tracker outer radius (m)	2.0
Solenoid field (T)	2.0
Solenoid bore radius / half length (m)	2.1/30.
Preshower absorber	Lead
Preshower $R_{min}/R_{max}$ (m)	2.4/2.5
Calorimeter absorber	Copper
Calorimeter $R_{min}/R_{max}$ (m)	2.5/4.5
Overall height / length (m)	11/13

Table 1.1: Parameters of the different sub-detectors composing IDEA.

ing on-pixel sparsification and data-driven, time-stamped readout. The goal is a thickness of  $0.15 - 0.30\% X_0$  per layer and a power dissipation below  $20 \text{ mW/cm}^2$ .

The tracks of charged particles are measured with very high precision, of the order of  $3 \mu\text{m}$  in the innermost layers. The vertex detector must also be able to precisely reconstruct secondary vertices. The angular coverage extends down to  $\simeq 13^\circ$ , and could be extended with additional silicon disks between the DCH and the calorimeter end caps.

### 1.3.2 Drift chamber

Based on the experience with the KLOE Experiment DCH [19] and the recent DCH for the MEG upgrade [20], a similar detector, with extraordinary transparency to charged particles, has been proposed for the IDEA Experiment.

The chamber is composed by a unique cylindrical volume, coaxial to the beam axis, with an inner radius of 0.35 m and an outer radius of 2 m, for a total length of 4 m. It consists of 112 coaxial layers, at alternating-sign stereo angles, grouped in 24 identical sectors. The total number of drift cells is 56448, with variable size from 12.0 to 14.5 mm. The chamber is operated with a very light gas mixture of 90% He - 10%  $\text{iC}_4\text{H}_{10}$  (isobutane), providing a maximum drift time value of  $\simeq 400 \text{ ns}$ .

In the radial direction the total amount of material is of the order of 1.6% of a radiation length ( $X_0$ ), including the inner and outer cylindrical walls and the contributions due to the gas mixture and the wires. On the other hand, in the forward and backward direction, the total amount of material is equivalent to about 5.0%  $X_0$ , including the inner cylindrical walls and the service end plates, instrumented with front-end electronics, signal and HV cables.

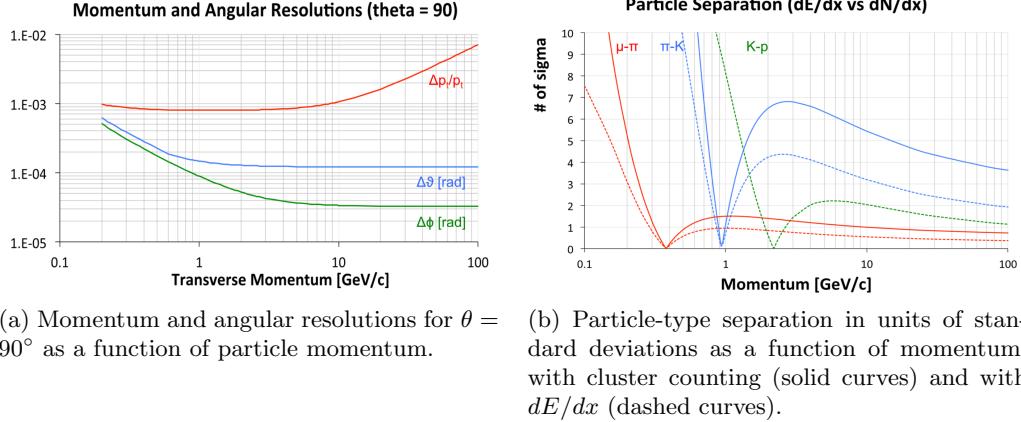


Figure 1.8: IDEA ultra-light drift-chamber performance [17].

In testing MEG-II drift-chamber prototypes [20] with 7 mm cell size, a drift distance resolution of  $100 \mu\text{m}$  has been achieved with both a gas mixture and operating conditions very similar to the ones foreseen for the IDEA DCH. Analytical calculations, in such operating conditions, were used to estimate the momentum and angular resolutions with the results shown in Figure 1.8(a). The drift chamber also offers outstanding particle-identification performance using the cluster counting technology that improves, as well, the spatial resolution.

The ionisation process, by means of which electrons are released, follows a Poisson law, therefore, by counting the total number of ionisation clusters ( $N_{cl}$ ) of a charged track, one can reach a relative resolution on  $N_{cl}$  that scales as  $1/\sqrt{N_{cl}}$ . The expected performance relative to particle separation, in terms of number of standard deviations as a function of the particle momentum, is shown in Figure 1.8(b). In this graph the solid curves refer to the cluster-counting technique, while the dashed ones refer to the expected identification power for the traditional  $dE/dx$  method. As it can be seen, the particle separation by cluster counting performs better over the whole momentum range.

### 1.3.3 Magnet system

The IDEA detector magnet is an ultra-thin and ultra-light (thus “radiation-transparent”) superconducting solenoid. It is 5 m long and has an inner diameter of 4.2 m. The main feature is that the solenoid is positioned between the tracking detectors and the calorimeter, a solution currently employed in ATLAS. This choice requires to maintain the total thickness at the 30 cm level, and below  $1 X_0$  in terms of radiation length, but at the same time the stored energy is reduced by a factor of four and the cost can be halved. In this scenario a relatively low field of 2 T can be produced.

### 1.3.4 Dual-readout calorimeter

The IDEA calorimeter consists of a dual-readout projective fibre-sampling detector. It follows the lessons of the high-resolution fibre calorimeters built by SPACAL [22] and RD52 [23].

The detector is a tower-based, longitudinally unsegmented, fully projective calorimeter sketched in Figure 1.9. The segmentation is chosen to have the shower development confined in a small number of towers and most of the energy deposited inside the volume corresponding to a single tower. This highly simplifies the calibration procedures for which each cell response can be considered individually. Considering that in accelerator-based experiments all the particles come, in good approximation, from the Interaction Point (IP), the projective segmentation can be obtained with a tower-based structure.

DR calorimeters, as described in Section 2.3, are composed by an absorber and two different active media to induce and transport two different signals. A single medium with two different response processes (such as a crystal producing both scintillation and Cherenkov light) may also be used. In the IDEA DR calorimeter, a copper (or brass) matrix is used as absorber, filled with optical fibres as active volumes (the dual-readout method will be described in the next chapter). The possibility to independently read out each fibre with a dedicated Silicon PhotoMultiplier (SiPM), as described in [24], brings a series of advantages especially in terms of spatial and angular resolution, removing the limitations arising from a readout driven by the tower dimension. For example, two particle showering in the same tower, in this way, could still be identified.

The geometry of the calorimeter, from different perspectives, is shown in Figure 1.9. Towers are truncated pyramids pointing to the IP. In such a way, each tower covers a specific region ( $\Delta\theta$ ,  $\Delta\varphi$ ) of the solid angle.

The cylindrical symmetry is achieved through a rotation around the beam axis of a minimal structure called slice. A single slice covers a range of  $10^\circ$  of the  $\varphi$  angle, therefore 36 of these elements cover the full calorimeter volume. In each slice, both barrel and end-cap towers are present, in particular 80 towers for the barrel and 35 for each end cap, with an approximate  $\theta$  coverage of  $1.125^\circ$  per tower.

All the towers are 2 m long, composed by an absorber matrix (copper, in the actual simulation) filled with optical fibres. The active elements are scintillating (polystyrene) and clear-plastic fibres (PolyMethyl MethAcrylate - PMMA). The fibres are  $\sim 1$  mm thick and are disposed in a chess-board like geometry so that each fibre is separated from the closest ones by  $\sim 0.5$  mm of absorber material. This complex geometry has been reproduced within the GEANT4 simulation toolkit [25] and all results in this thesis are obtained with it.

### 1.3.5 Preshower and muon chambers

In the barrel region, just before the calorimeter, the magnet coil, coupled with Micro-Pattern Gaseous Detector (MPGD) chambers, acts as a preshower. In the forward

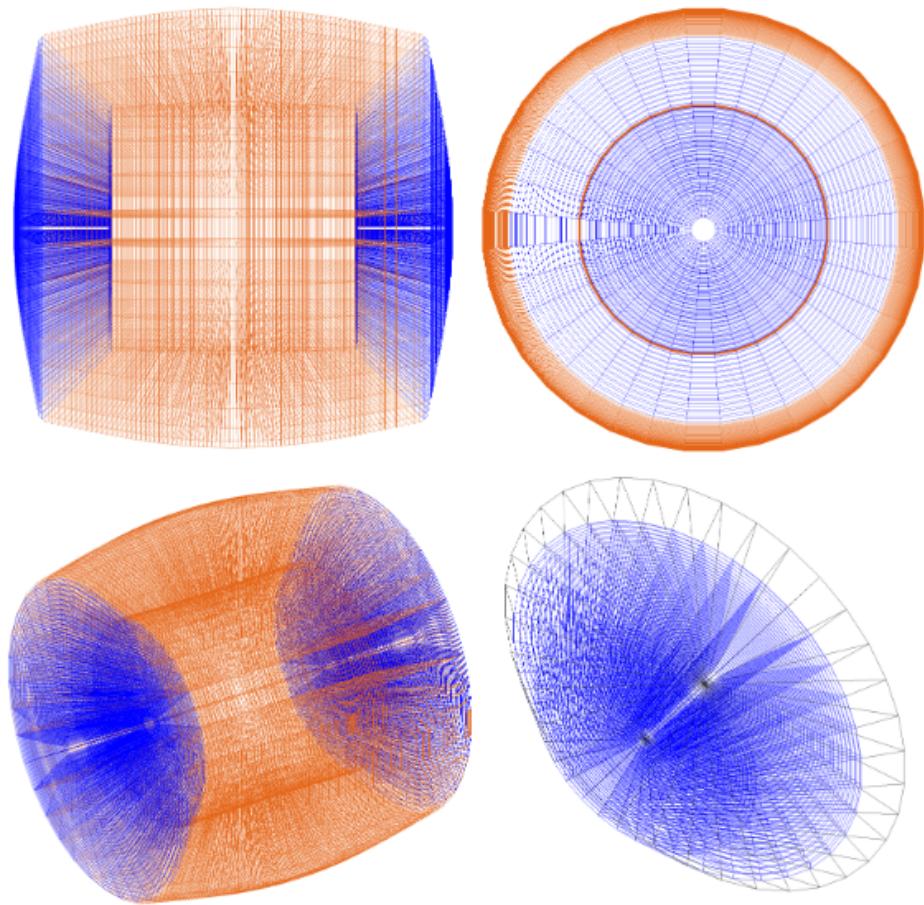


Figure 1.9: IDEA dual-readout calorimetry geometry produced with GEANT4, seen from different perspectives. In orange the towers composing the barrel, in blue the ones composing the end caps.

regions, preshower detectors are located between the drift chamber and the end-cap calorimeters. Externally, the overall detector is surrounded by an iron return joke, to contain the magnetic field, and this volume is also instrumented with MPGD chambers, to allow muon detection and momentum measurement. Both the preshower and the muon chambers are based on the micro-Resistive WELL ( $\mu$ -RWELL) technology. The  $\mu$ -RWELL chambers are compact MPGD detectors, with a single, intrinsically spark protected, amplification stage.

The evaluation of the preshower performance and the single-hit position resolution requirement are still in progress. However, experimental results with small prototypes already show a good acceptance for photons and an efficiency of  $\simeq 30\%$  for tagging  $\pi^0$  from their  $\gamma\gamma$  decay.

Also the requirements for the muon system look to be within reach. Indeed, this technology provides good tracking efficiency, high-voltage stability, a position resolution of  $200-300\ \mu\text{m}$  and a good time resolution, thanks to the fast charge-amplification process.



## Chapter 2

# Calorimetry and the dual-readout method

Originally developed with astrophysical purpose for cosmic-ray studies, the art of calorimetry refers to the detection of particles and the measurement of their properties by total or partial absorption 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 an ideal calorimeter, particles are fully absorbed and their energy transformed into a measurable signal.

The incident particle interacts with the detector (through electromagnetic or nuclear processes) producing a shower of secondary particles, with progressively less and less energy. The energy deposited by the charged particles in the calorimeter active material is used to generate signals, typically in the form of charge or light. Two typical processes exploited are the scintillation light, emitted in response of ionisation, and the Cherenkov light, produced by ultra-relativistic particles.

Calorimeters can be divided into two categories depending on the type of shower they are optimised to detect: electromagnetic calorimeters, used mainly to measure electrons and photons energies through their electromagnetic interactions with the detector material, and hadronic calorimeters, used to measure hadrons energies through their strong and electromagnetic interactions. According to their internal structure, they are also classified as either sampling or homogeneous calorimeters.

Homogeneous calorimeters are built with one type of material that performs two tasks: it produces the shower development and, at the same time, provides the detectable signal. Sampling calorimeters, instead, consist of alternating layers of an absorber, a dense material used to produce the showers, 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 dual-readout calorimetry technique.

A more comprehensive descriptions of the field can be found in [26, 27, 28].

## 2.1 Physics of shower development

The cornerstones for calorimetry are the interaction processes happening between particles and matter. These processes depend on the absorbing medium, the particle type and energy. While absorbing an energetic particle, a cascade of subsequent particles is formed, known as a particle shower. The involved processes and the calorimeter response to the showering particles are the keys for understanding this topic.

### 2.1.1 Electromagnetic showers

Electromagnetic showers are governed by a small number of well understood QED processes. Charged particles (electrons and positrons) lose energy through ionisation and radiative processes, whereas neutral ones (photons) interact with matter by photoelectric effect, Compton scattering and pair production.

Therefore, electrons and positrons ionise the medium under the condition of having enough energy to release atomic electrons from the Coulomb fields generated by the atomic nuclei. The amount of energy released (per unit of path length) is predictable through the Bethe-Block formula restricted to electrons (and positrons) [29]:

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

The stopping power (i.e.  $dE/dx$ ) decreases as the particle energy increases ( $\propto \beta^2$ ). Hence the ionisation process is the greatest source of energy loss for particles at low energy. The radiative energy loss process, known as *bremsstrahlung*, is the dominant source of energy loss for electrons and positrons at energies above 10–100 MeV, depending on the absorber material. Relativistic electrons and positrons radiate photons as a result of the interaction with the atomic electric fields. The energy spectrum of these photons falls off as  $1/E$ , up to the primary particle energy, but in average, a single photon carries a small part of this. The process produces (usually small) changes in electron (or positron) direction, on top of the Coulomb multiple scattering.

At a given energy, the relative importance of ionisation and radiation losses depends on the medium and in particular on its electron density. This density is, in first approximation, proportional to the number of protons in the nuclei ( $Z$ ). 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:

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

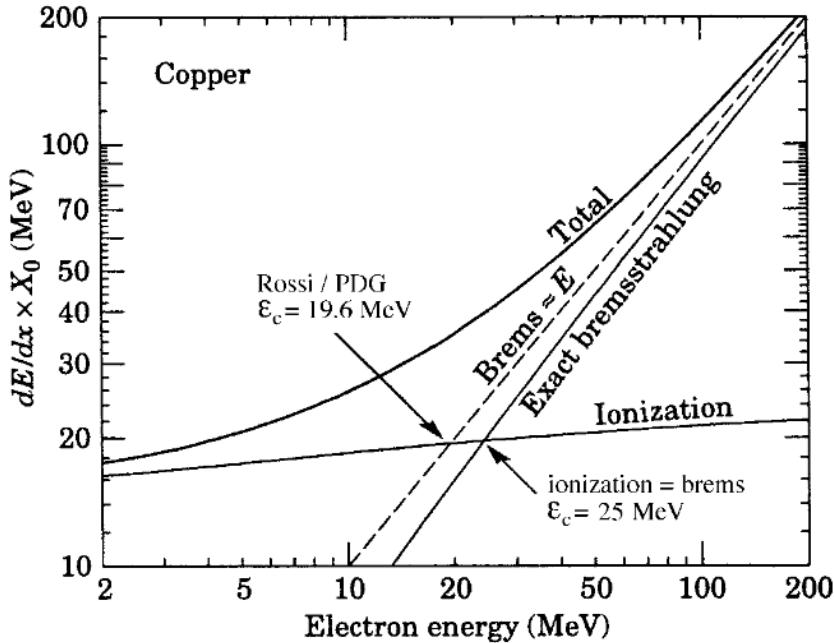


Figure 2.1: Energy losses through ionisation and bremsstrahlung processes by electrons in copper [32].

An example of the energy loss in copper by electron is sketched in Figure 2.1, where the ionisation and radiation contribution are plotted.

On the other hand, the interaction between photons and matter is mainly driven by three different processes: the photoelectric effect, the Compton scattering and the electron–positron pair production.

The photoelectric effect is the process that most likely occurs at low energies. It is characterised by an atom absorbing the photon and releasing an electron. 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 exponent  $n$  ranging between 4 and 5. The photoelectric cross section rapidly decreases with the energy, varying as  $E^{-3}$ , and the process rapidly loses relevance as the energy increases. The Compton process is a scattering process where an impinging photon interacts with an atomic electron transferring enough momentum and energy to have the electron escaping from the atomic Coulomb field. Kinematic variables such as energy transfer and scattering angles can be easily obtained by 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 get the photoelectric effect to overcome.

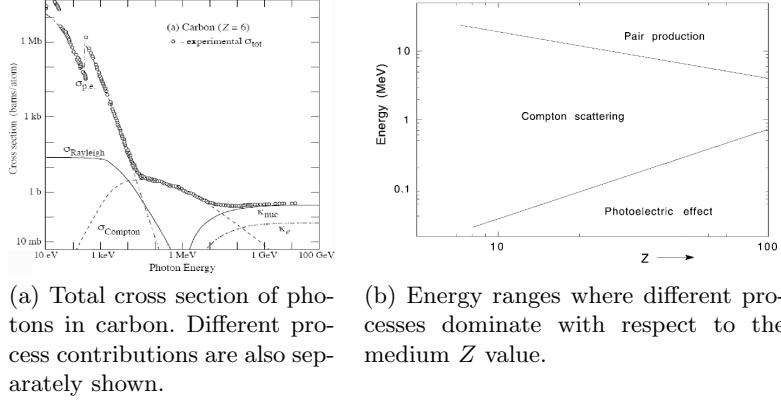


Figure 2.2: Images from [29].

In each step, the photon energy loss is:

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

where  $\xi = E_\gamma/m_e c^2$ . The Compton scattering cross section is much less dependent on the  $Z$  value than the photoelectric one. It is almost proportional to the number of electrons per gram in the absorbing material, which for most elements is approximately the same. Also in this process the cross section decreases as the photon energy increases, but with a linear dependence on  $E$ . Therefore the Compton scattering has more impact than the photoelectric absorption above a certain threshold energy.

The pair production process, differently from the previous ones, has a higher threshold under which the effect can not occur. This threshold is twice the electron rest mass ( $2 \times 511$  MeV) and above that, a photon can produce an electron-positron pair. The cross section for pair production rises with the energy reaching an asymptotic value at energies higher than 1 GeV. For this reason, at high energy, pair production is the most likely process to occur. Besides, the dependence on the medium goes, in first approximation, as  $Z^2$ .

Comparing the cross section of the three processes and their dependence on the photon energy, it is clear that the photoelectric effect dominates at low energy while, for intermediate values, the Compton scattering takes over. At higher energies almost every photon is converted in charged particles through pair production. An example of these contributions is shown in Figure 2.2(a).

Knowing the dependence of the cross sections on the atomic number of the material, ranges of energies where each process dominates can be found and parametrised as a function of the  $Z$  value. A representation is sketched in Figure 2.2(b).

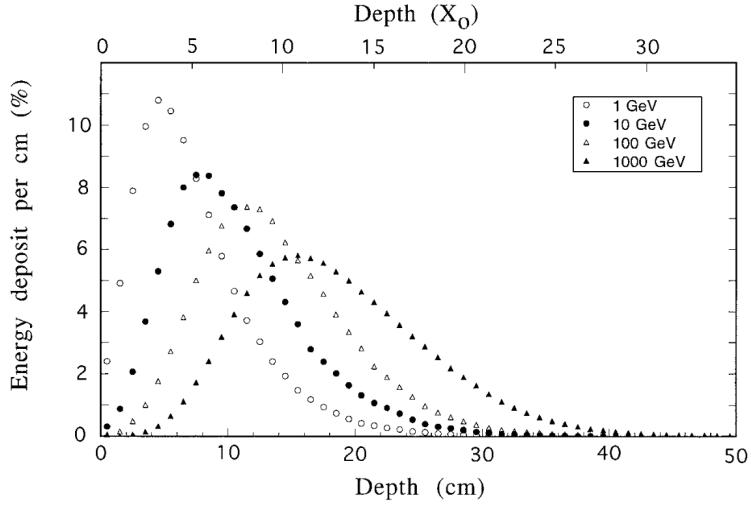


Figure 2.3: The energy deposit as a function of depth, for 1, 10, 100 and 1000 GeV electron showers developing in a block of copper [29].

### Electromagnetic shower principle

At energy values of the order of 1 GeV and higher, electrons, positrons and photons initiate electromagnetic showers in the materials in which they penetrate. At these energies, charged particles lose their energy mostly by bremsstrahlung. The majority of emitted photons are soft and interact with Compton scattering until their absorption through photoelectric effect. Meanwhile, the photons with energy greater than few MeV produce  $e^+ - e^-$  pairs, which eventually radiate more  $\gamma$ 's. The process continues till there are particles with enough energy to give rise to multiplication. The shower maximum is defined as the point at which the number of shower particles produced in this process reaches its maximum. The shower-maximum depth within the absorber increases logarithmically with the energy of the primary particle (see Figure 2.3). The longitudinal shower development is parametrised in term of the radiation length ( $X_0$ ), defined as the distance at which the electron (or positron) loses on average 63% ( $1 - e^{-1}$ ) of its energy by radiation. Describing the shower development as a function of  $X_0$  almost completely hide the dependence on the absorber material.

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:

$$\rho_M = E_s \frac{X_0}{\varepsilon_c} \quad (2.4)$$

where  $E_s$  is defined as  $m_e c^2 \sqrt{4\pi/\alpha} \simeq 21.2$  MeV. This quantity is almost material-independent and, on average, a cylindrical volume with this radius, centred around the shower axis, contains 90% of the shower energy. The lateral spread is mainly caused by two effects: at high energy, electrons and positrons are moved away from the shower axis

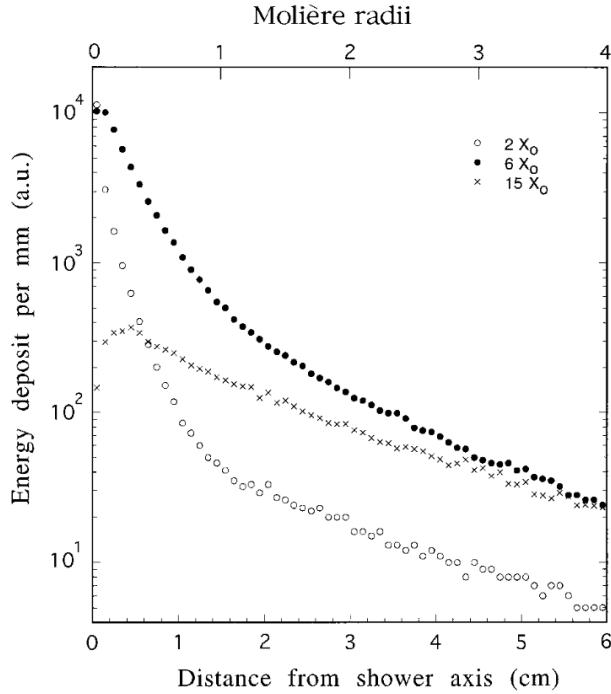


Figure 2.4: The radial distributions of the energy deposited by 10 GeV electron showers in copper, at various depths [29].

because of the deviation occurring in Compton scattering; with lower energies, photons and electrons are (also) produced in isotropic processes and so, lose any correlation with the primary particle direction. Bremsstrahlung photon emission also contributes to the shower lateral spread. Figure 2.4 shows the electromagnetic shower radial profiles at different showering stages.

Showers 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 while 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

With hadronic showers, a new degree of complexity arises since these are also (often mostly) driven by strong interactions. 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.

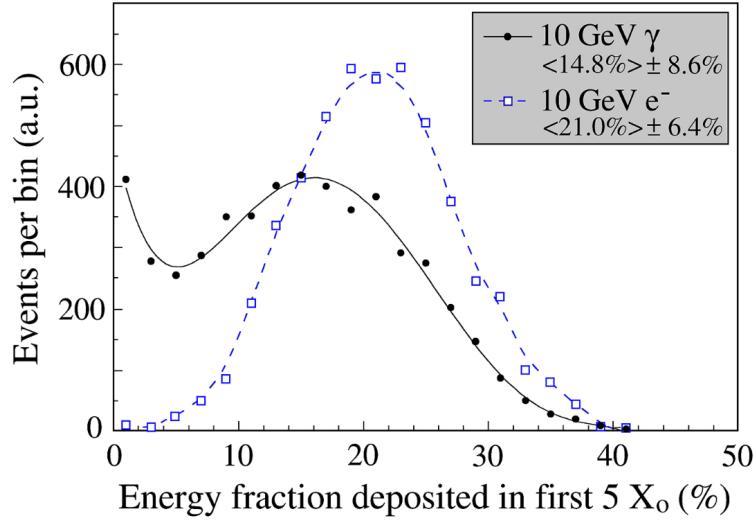


Figure 2.5: Distribution of the energy fraction deposited in the first  $5 X_0$  by 10 GeV electrons and photons showering in lead. Image from [30].

- 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 any signal formation, and therefore is known as invisible energy.

Among the particles in a hadronic shower, neutral pions play an important role. Via their electromagnetic decays, they start electromagnetic showers over-imposed to a hadronic component. The fraction of energy carried by the induced em showers is called electromagnetic fraction ( $f_{em}$ ). On average,  $f_{em}$  increases with the primary particle energy since, in first approximation, neutral  $\pi$ 's are produced with the same percentage at any stage of the hadronic shower, until an energy threshold is reached: the higher the primary particle energy, the more the generations of shower particles, the larger the average em fraction. For charged pion initiated showers, the average electromagnetic fraction has been evaluated to increase with the energy according to the power law:

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

where  $K \simeq 0.8$  and  $E_0$  is a material dependent value of the order of 1 GeV. The behaviour in lead and copper is shown in Figure 2.6.

Another contribution to the complexity is given by the variability of this em fraction event by event and its energy dependence.

Another important difference between em and hadronic showers is the larger spatial profiles of the released energy. Analysing, for example, several charged-pion showers, peaks of signals (with electromagnetic shower dimensions) are produced, every time, at

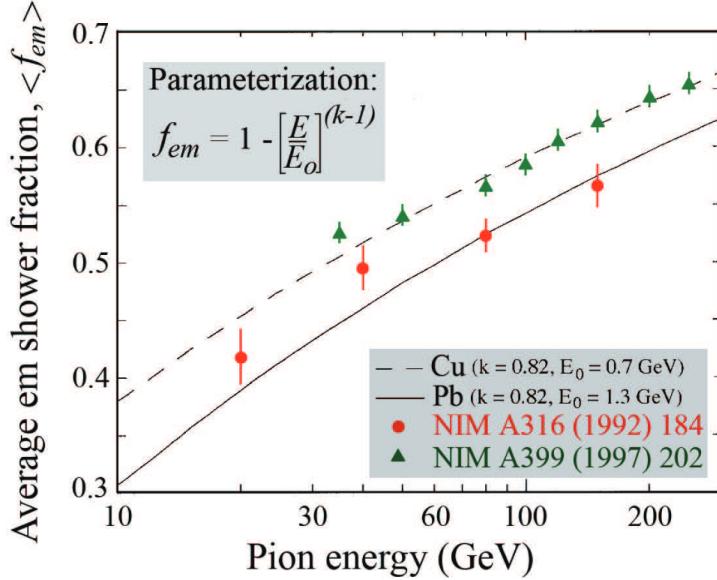


Figure 2.6: Comparison between the experimental results on the em fraction of pion-induced showers in copper-based and lead-based calorimeters. Image from [31].

the place where a  $\pi^0$  is generated. However, as already said, the neutral pion production can occur in the second or third generation of the shower development and such peaks may occur at different depth. An example of 4 different showers from 270 GeV pions is illustrated in Figure 2.7.

The depth of a calorimeter for containing hadronic showers increases logarithmically with energy, as already seen for em showers, but the large longitudinal fluctuations in the showering starting point make the leakage effect non negligible also in configurations that would contain, on average, 99% of the shower.

Laterally, a hadronic shower is better (percentually) contained if the primary particle energy increases. This is due to the fact that the electromagnetic shower fraction increases with energy and the em showers produced tend to develop laterally closer to the shower axis.

## 2.2 The calorimeter response

The calorimeter *response* is defined as the average calorimeter signal per unit of deposited energy. The response is usually expressed in terms of number of photoelectrons or charge (pC) per GeV. A calorimeter with a constant response is said to be *linear*. In the following we compare the response of electromagnetic and hadronic calorimeters.

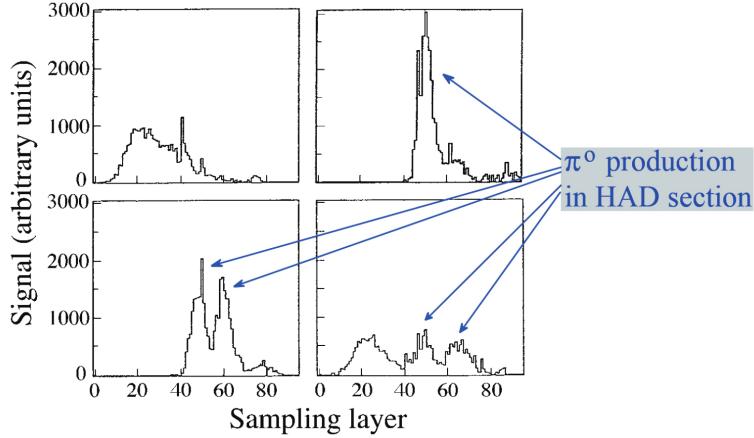


Figure 2.7: Longitudinal profiles for 4 different showers induced by 270 GeV pions in a lead/iron/plastic-scintillator calorimeter [27].

### Electromagnetic calorimeters

Electromagnetic calorimeters are detector optimised for em showers. In these showers, all the energy carried by the primary particle is released by a very large number of particles through processes that may generate signals (excitation or ionisation in the active sensing medium). The number of particles is, on average, proportional to the primary energy. The direct consequence is that em calorimeters are in general linear detectors. A non-linear response is usually an indication of instrumental problems. The most common ones are:

- leakage effects;
- saturation effects of the active components, originated by high localised 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 part of more complex calorimetric systems.

The small dimensions also allow to build em calorimeters operating at colliders both of the homogeneous and sampling type.

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 high-resolution position measurements and high particle-identification efficiencies are required, homogeneous calorimeters are not the best choice because of the larger shower dimensions. 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 a worse energy resolution due to the presence of an absorber material (copper, iron, lead or uranium) that does not contribute to the signal. For electromagnetic sampling calorimeters, typical resolution values are in the range of  $5 - 20\%/\sqrt{E(\text{GeV})}$ . However, they are relatively simple to segment longitudinally and laterally due to their sampling structure often helping to achieve a better spatial resolution and a better particle-identification performance..

### Hadronic calorimeters

Hadronic calorimeters are detector optimised for hadronic showers. At collider experiments, to constrain the detector dimensions, they are always of the sampling type. The response of a hadronic calorimeter is a mixture between the response to the 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 hadrons) is a combination of the two:

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

Since the average  $f_{em}$  value, as already seen, increases with the energy, the response of a non-compensating calorimeter ( $e/h \neq 1$ ) is not constant and non-compensating calorimeters are intrinsically non-linear detectors.

The  $e/h$  value cannot be directly measured. However, it is usually derived from the  $e/\pi$  ratios, measured at various energies. The relationship between  $e/\pi$  and  $e/h$  is:

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

where  $f_{em}$  is the energy-dependent average em fraction. A calorimeter for hadron detection does not present a linear response with respect to the primary particle due to the energy dependence of the em fraction. The relation is represented in Figure 2.8 for compensating and non-compensating calorimeters.

One of the basic reasons, for the different response to the em and hadronic shower components, lies in the fact that in the absorption of the hadronic part, a significant fraction of the energy is *invisible* i.e. is not contributing to the calorimeter signal. The main source of this energy is the energy used to release nucleons from nuclei.

In order to design a linear hadron calorimeter, various compensation methods have been developed. There are two main approaches to obtain  $e/h = 1$ : one is by reducing the

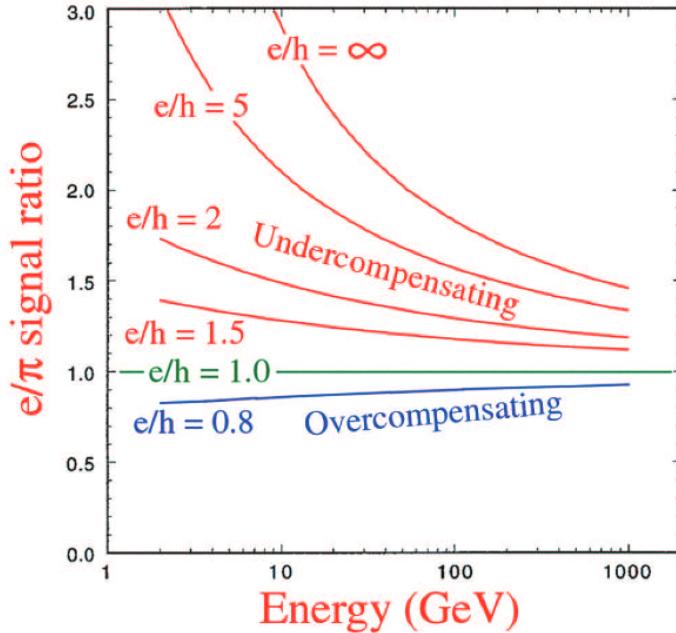


Figure 2.8: Relation between the calorimeter response ratio to em and non-em energy deposits,  $e/h$ , and the measured  $e/\pi$  signal ratios [27].

electromagnetic response and the other one is by increasing the non-em response. The most effective way to reduce the em response of a sampling calorimeter is to use an high- $Z$  material as absorber. The concept is based on the fact that low energy photons release their energy via the photoelectric effect that is highly  $Z$  dependent ( $Z^5$ ). For this reason, when using high- $Z$  absorber materials, causes the fraction of energy deposited in the absorber to increase with respect to the active elements.

The other strategy is known as compensation by neutron-signal boosting or by Signal Amplification through Neutron Detection (SAND). It consists 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 neutrons gives the possibility of an event-by-event correction boosting the signals generated by the neutrons. The most likely process that occurs at low energies for a neutron is the elastic scattering on a (light) 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). Since hydrogen maximises this ratio, the increase of the hydrogen density in the active medium produces an increment of the non-em response. With a proper fine tuning of the sampling fraction, a  $e/h$  ratio of 1 can be reached. 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 different compensation strategy is adopted, the

dual-readout compensation, that will be described in Section 2.3.

### 2.2.1 Fluctuations

Event-by-event fluctuations in the calorimeter response determine the detector energy resolution. The energy resolution is quantified as the reconstructed energy-distribution standard deviation divided by the mean value ( $\sigma/E$ ).

In em calorimeters, four are the main sources of response fluctuation:

- sampling fluctuations;
- signal quantum fluctuations;
- shower leakage fluctuations;
- instrumental-response fluctuations.

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

$$f_{\text{samp}} = \frac{E_{\text{active}}}{E_{\text{passive}} + E_{\text{active}}} \quad (2.8)$$

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

$$\frac{\sigma}{E} = \frac{2.7\% \sqrt{d/f_{\text{samp}}}}{\sqrt{E}} \quad (2.9)$$

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

Fluctuation effects due, for example, to photon statistics in light emitting active materials are grouped under the label of signal quantum fluctuations. Also these fluctuations follow the rules of Poisson statistics, since the contributions are usually uncorrelated. This implies that a second contribution that scales as  $E^{-1/2}$  is added to the relative energy resolution expression.

In leakage fluctuations there are three possibilities: longitudinal, lateral and albedo. These fluctuation are highly dependent on 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 via Monte Carlo simulation of the detector.

Finally, instrumental fluctuations are the ones arising, for example, because of the electronic noise or geometrical inhomogeneities. The electronic noise contribution is usually described with a term that scales as  $1/E$  because it is largely independent of the shower

energy. On the other hand, 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 shown, different contributions with different energy dependence have to be considered. The consequence is that different effects dominate the energy resolution in different energy ranges. These behaviours have to be considered in the design of a calorimetry system in order to optimise the result as a function of the physics cases.

All the above fluctuation sources affect hadronic calorimeters as well. For example, sampling fluctuations have a higher impact in hadronic showers due mainly to the lower average number of particles that release energy with respect to em showers. At the same time, in hadronic detectors, resolution is affected also by other fluctuation sources.

In non-compensating calorimeters, event-by-event fluctuations in em component affect the detector response. They can be assumed to contribute to the relative energy resolution through a term roughly scaling as  $cE^{-0.28}$ .

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

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 often described as:

$$\frac{\sigma}{E} = \frac{a_2}{\sqrt{E}} + b. \quad (2.11)$$

A useful approach for removing these fluctuation source is to design a compensating calorimeter.

## 2.3 Dual-readout compensation

All the benefits obtained by compensating calorimeters are achieved by sampling em and non-em components in hadronic showers with the same response ( $e = h$ ). Dual-readout calorimetry is a method that measures the electromagnetic fraction on an event-by-event basis and applies a correction, reaching de facto the  $e/h = 1$  value through the linear combination of two independent energy measurements [33].

### 2.3.1 Working principles

This technique works thanks to the fact that the em component is mostly brought by relativistic electrons and positrons, while the majority of the non-em one is carried by non-relativistic particles. Hence, collecting the Cherenkov signal produced by a hadronic shower is almost equivalent to sampling the em component, thus measuring the electromagnetic fraction. A second simultaneous signal, typically from scintillation light

production, is recorded to provide information on the total ionising component of the shower.

Let's consider scintillation ( $S$ ) and Cherenkov ( $C$ ) signals produced by a hadronic shower developing in a dual-readout calorimeter. 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:

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

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

where  $h/e$  quantifies the two different degree of non-compensation for the two signals. Considering that  $h/e$  are assumed to be constants, the ratio  $C/S$  depends only on  $f_{em}$ :

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

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

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

if an estimate of the  $h/e$  values is provided.

An example of a possible  $C/E$  vs  $S/E$  scatter plot is shown in Figure 2.9 for a generic DR calorimeter, for electrons, pions and protons showers. As expected, data generated from electrons are always around the point  $[1, 1]$  (as  $f_{em} = 1$ ). On the other hand, an hypothetical event with only the non-em component would be at  $[(h/e)_S, (h/e)_C]$ . Signals from pions and protons showers are clustered along a straight line linking these two points. Thanks to the fact that (if the  $h/e$  parameters are constant)  $\theta$  is energy and particle independent, it is possible to define a single parameter:

$$\chi = \frac{1 - (h/e)_S}{1 - (h/e)_C} = \cot \theta, \quad (2.16)$$

that can be estimated with test beam data. Finally, event by event, the energy of each hadron shower, corrected for the  $f_{em}$  value, can be reconstructed using the  $S$  and  $C$  signals as:

$$E = \frac{S - \chi C}{1 - \chi}. \quad (2.17)$$

The ideal DR calorimeter would 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 would have  $(h/e)_C = (h/e)_S$ , in which case the two signals would sample the two components (em and non-em) with the same response giving no information about  $f_{em}$ . Therefore, the best DR calorimeter is the one with the lower  $\chi$  parameter corresponding to the lower  $(h/e)_C$  and the higher  $(h/e)_S$  values possible.

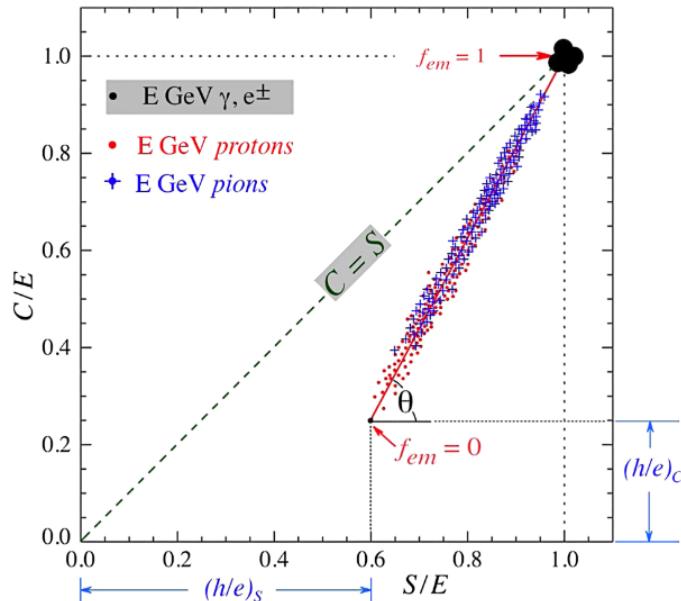


Figure 2.9: The  $(S/E, C/E)$  diagram of the signals from a (generic) dual-readout calorimeter. The hadron events are clustered around the straight (red) line. Data points for protons and pions have different distributions, reflecting differences in the em shower fraction [34].

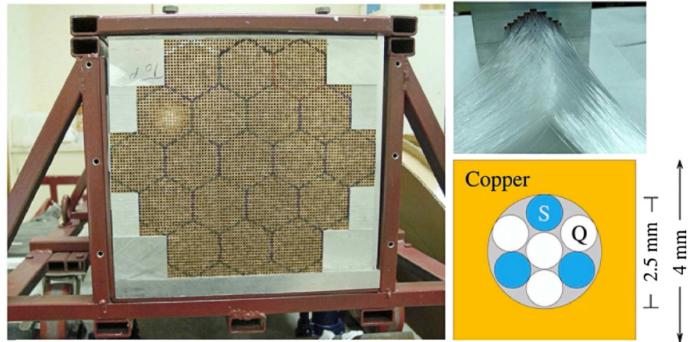


Figure 2.10: The DREAM calorimeter layout [35].

### 2.3.2 Dual-readout sampling calorimeters

At present a few dual-readout calorimeters have been built and tested and more projects are under study.

The first practical demonstration of the dual-readout method was achieved by a R&D study for the Advanced Cosmic Composition Experiment at the Space Station (ACCESS) [36]. 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  $\pi^0$ .

Later, a  $10 \lambda_{int}$  deep calorimeter known as Dual-READout Module (DREAM) was built and tested [37, 38]. The detector was composed by 5580 basic element represented in Figure 2.10. These are 200 cm long, hollow, extruded copper rod of  $4 \times 4 \text{ mm}^2$  cross section in which 3 scintillating and 4 Cherenkov fibres are inserted.

Figure 2.11 shows the energy distributions for  $100 \text{ GeV } \pi^-$  by the two fibre types, after a calibration at the electromagnetic scale. The distributions are peaked at considerably lower values than the ones corresponding to electrons. Populating a scatter plot with Cherenkov and scintillation signals for each event, the correlation between them is evident as shown in Figure 2.12. The two  $h/e$  ratios for the Cherenkov and scintillator DREAM structures were estimated to be 0.21 and 0.77, respectively. The expression of the em fraction 2.14 becomes:

$$f_{em} = \frac{0.21 - 0.77(C/S)}{(C/S)[1 - 0.77] - [1 - 0.21]}. \quad (2.18)$$

Being able to evaluate  $f_{em}$  event-by-event opens the possibility to correct the signals for the effects of non-compensation. Really, you only need to know the  $\chi$  value to eliminate the dependence of the energy estimate on  $f_{em}$  and  $\chi$  is experimentally measurable at a test beam as  $chi = (E-S)/(E-C)$ . In Figure 2.13 the overall Cherenkov signal distribution

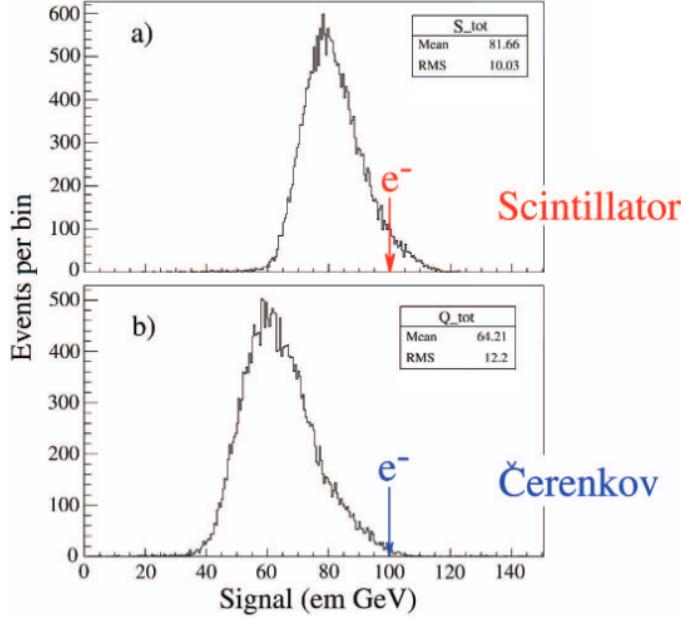


Figure 2.11: Signal distributions for 100 GeV  $\pi^-$  recorded by the scintillating (a) and the Cherenkov (b) fibres. [27].

for 100 GeV  $\pi^-$  (a) is shown, where the asymmetry from  $f_{em}$  fluctuations is evident. In Figure 2.13(b) the events are grouped in subsamples selected on the basis of their  $f_{em}$  value. Each one of these subsamples reproduce a certain region of the overall signal distribution, with the average value that increases with  $f_{em}$  as expected. This representation gives an idea of the effectiveness of the method.

In this process, the energy resolution improved, the signal distribution became much more Gaussian and, most importantly, the hadronic energy was, on average, correctly reproduced, i.e. the linearity greatly improved.

The DREAM results, together with other studies, confirmed the feasibility of the dual-readout compensation method, leading to the IDEA dual-readout calorimetry project.

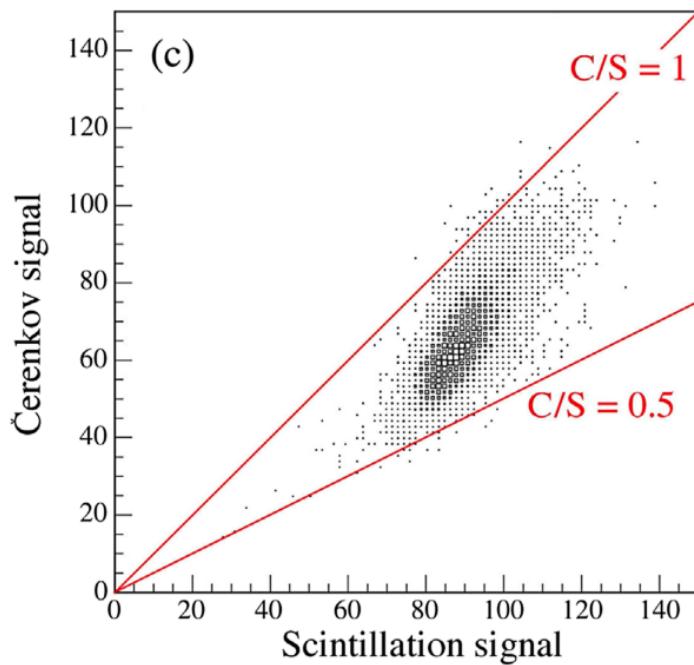


Figure 2.12: The  $S - C$  scatter plot, for  $100 \text{ GeV } \pi^-$ , showing the correlation between the two signals. The signals are expressed in the same units used to calibrate the calorimeter (em GeV) [38].

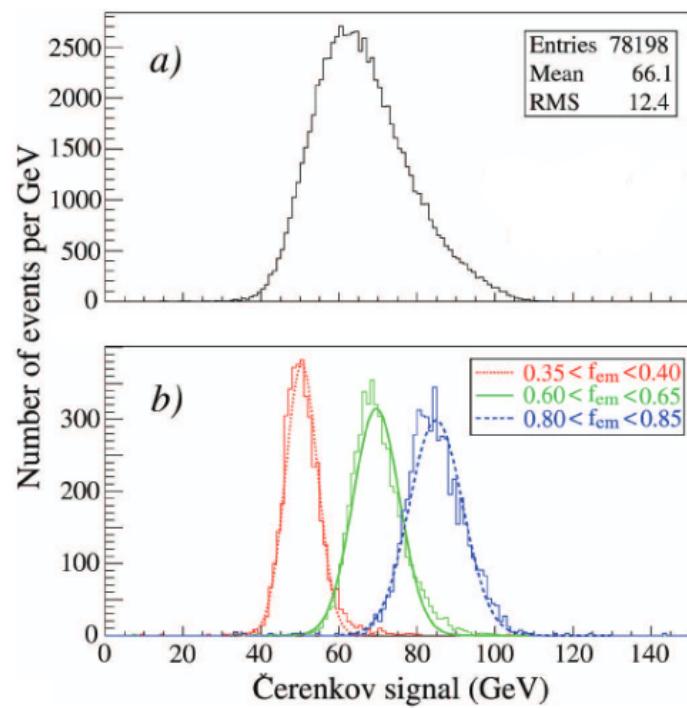


Figure 2.13: The Cherenkov signal distribution for 100 GeV  $\pi^-$  (a) and distributions for subsamples of events selected on the basis of the measured  $f_{em}$  value, using the  $Q/S$  method (b). Image from [27].



# Chapter 3

## Silicon Photomultipliers

Silicon PhotoMultipliers (SiPMs) are solid-state light sensors also referred to as Geiger-mode Avalanche PhotoDiodes (G-APDs) or Multi-Pixel Photon Counters (MPPCs). They represent an interesting alternative to the more standard PhotoMultiplier Tubes (PMTs) with specific advantages as the small size, the insensitive to magnetic fields and the low power consumption.

The chapter will describe the working principles of these detectors (Section 3.1) giving an explanation of their most important features such as single-photon sensitivity, unprecedented photon-number resolving capability, high internal gain, high and tunable Photon Detection Efficiency (*PDE*) (Section 3.2) and excellent time resolution. The last section is dedicated to the most important noise sources.

In the next chapter we will investigate their usage as light readout sensors for the IDEA dual-readout calorimeter.

### 3.1 Working principles

A SiPM is a high-density matrix (up to  $10^4$  per  $\text{mm}^2$ ) of Single-Photon Avalanche Diodes (SPADs) or Avalanche PhotoDiodes (APDs). Each SPAD, also referred to as pixel or cell, is connected with the others in parallel to a single readout output.

The photodiodes are silicon elements based on the presence of an energy gap between the valence band and the conduction band of the semiconductor. In particular, they are characterised by a *p-n* junction operating in the Geiger-Müller regime (a schematic view of the doping structure is shown in Figure 3.1). This condition is obtained by powering the APD a few volts above the breakdown voltage ( $V_{bk}$ ), hence, a high electric field is produced in the depletion region. If a photon, with at least the same energy of the gap, traverses this region, it can be absorbed by producing charge carriers. When operating in the Geiger-Müller mode, the depleted region is immersed in a high electric field (tens of volts) that permits not only the collection of the carriers, but also the production of

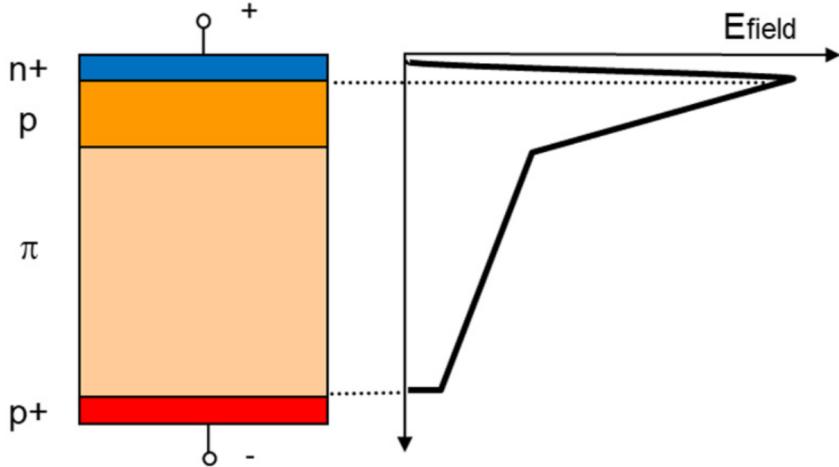


Figure 3.1: A sketch of a SPAD doping structure and electric field in operation mode [40].

avalanche discharges. The peculiarity of the APDs is that the signals generated by one or by more photons are indistinguishable if the photons hit the depleted region at a too short time distance. In other words, these photodetectors work as a binary device that gives a signal, when fired by at least one photon, or nothing, if not (disregarding the possibility of noise effects, see Section 3.3).

To overcome the APD limitations, Silicon PhotoMultipliers are composed by thousands of these small elements connected in parallel as shown in Figure 3.2. In this way, the SiPM can be seen as a collection of binary cells: it can provide quantitative information about the intensity of the incoming light by simply counting the number of fired cells.

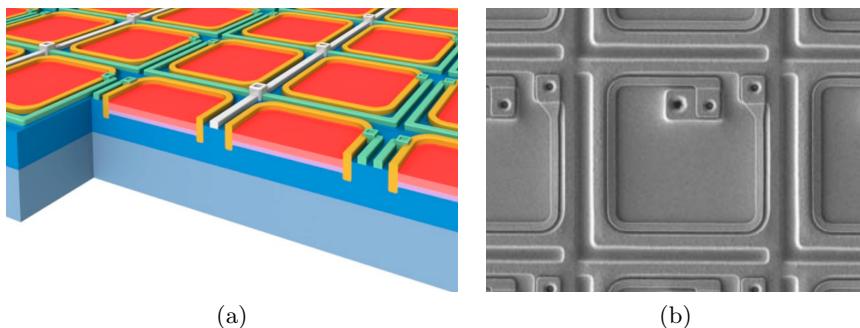


Figure 3.2: An artistic view of a SiPM structure (a). Individual SiPM pixels with metal-composite quenching resistor fabricated around each cell (b). Image from [39].

### Electrical model

The behaviour of a single APD detector is emulated by a commonly accepted electrical model. Developed in the sixties by McIntyre and Haitz, it is sketched in Figure 3.3.

The detector can be interpreted as a depleted region capacitance ( $C_D$ ) connected in parallel to a space-charge resistance ( $R_S$ ) and a switch representing the digital avalanche product phenomenon. The Single Photon Avalanche Detector is directly connected to a quenching resistor ( $R_Q$ ), essential to stop the high current produced by the discharge. The photon-detection process starts with a stationary situation where the capacitance is charged with the voltage  $V_{bias}$  and the switch is set to the OFF position.

When a photon impinges the APD, it can produce an avalanche with a given probability (see *PDE* in Section 3.2). The production of a discharge corresponds to setting the switch to its ON position. The consequence is the discharge of the capacitance through the internal resistance  $R_S$ , to the voltage  $V_{BR}$  (that is lower than the bias voltage:  $V_{bias} = V_{BR} + V_{ov}$ ) with the time constant  $\tau_D \simeq R_S C_D$  [40]. In this phase the external current increases as:

$$I_{EXT}(t) = \frac{V_{ov}}{R_S + R_Q} \left(1 - e^{-t/\tau_D}\right). \quad (3.1)$$

At the end of the avalanche process, the switch comes back to the OFF condition and the capacitor starts to recharge until the voltage  $V_{bias}$  is reached again. The recharge is characterised by a time constant called cell recovery time  $\tau_R \simeq C_D R_Q$ . The external current goes to zero as:

$$I_{EXT}(t) = \frac{V_{ov}}{R_S + R_Q} e^{-t/\tau_R}. \quad (3.2)$$

The two time constants are generally different with  $\tau_R \gg \tau_D$  because  $R_Q$  and  $R_S$  differ up to two orders of magnitude.

At the end of this process the APD is ready for the detection of a new photon.

The conceptual output pulse produced by a single fired APD is sketched in Figure 3.4. Several APDs connected in parallel originate the signal of the whole SiPM. In the next chapter, simulated waveforms from SiPMs will be shown.

### Gain and amplitude

The gain of a single APD is defined as the output charge in unit of the elementary charge. This value can be estimated when the overvoltage and the internal capacitance are known:

$$G = \frac{C_D V_{ov}}{q_e}. \quad (3.3)$$

Considering typical values as  $V_{ov} \simeq 3 - 5$  V and  $C_D \simeq 10$  fF, we obtain typical values for  $G$  of the order of  $10^6$ . Therefore, the SiPMs single-photon sensitivity is well explained by this high single-cell gain.

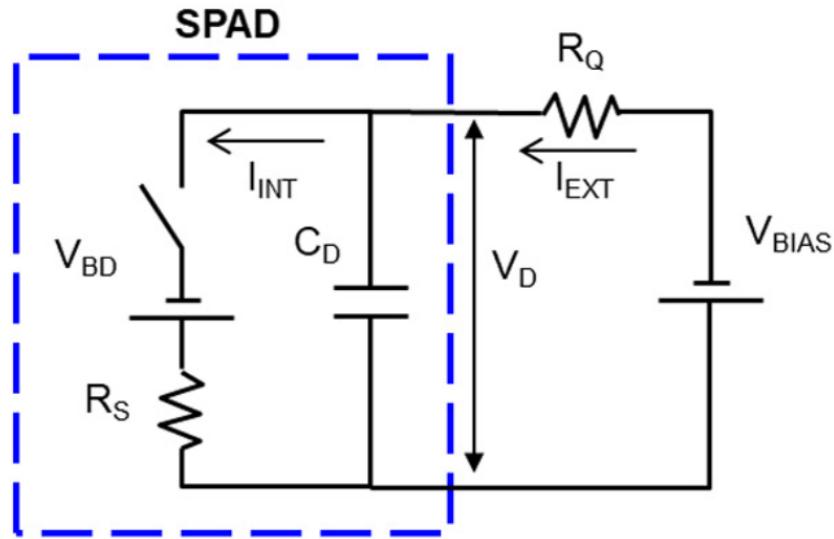


Figure 3.3: Sketch of the equivalent circuit of a single Geiger APD [40].

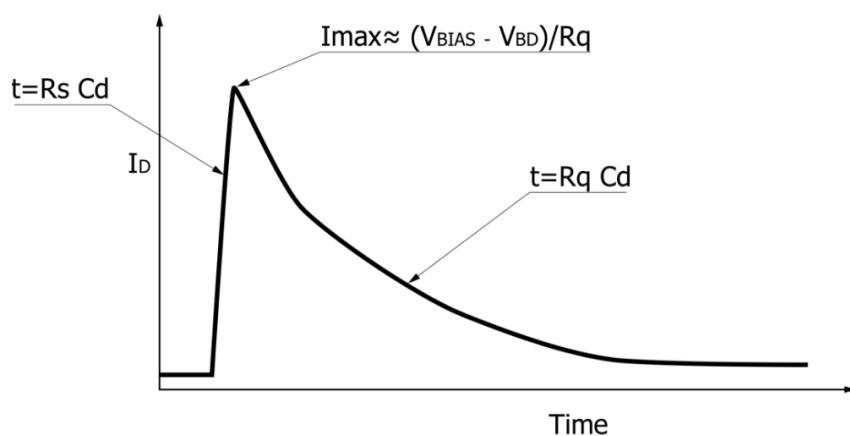


Figure 3.4: Sketch of the conceptual output pulse produced by a single APD [39].

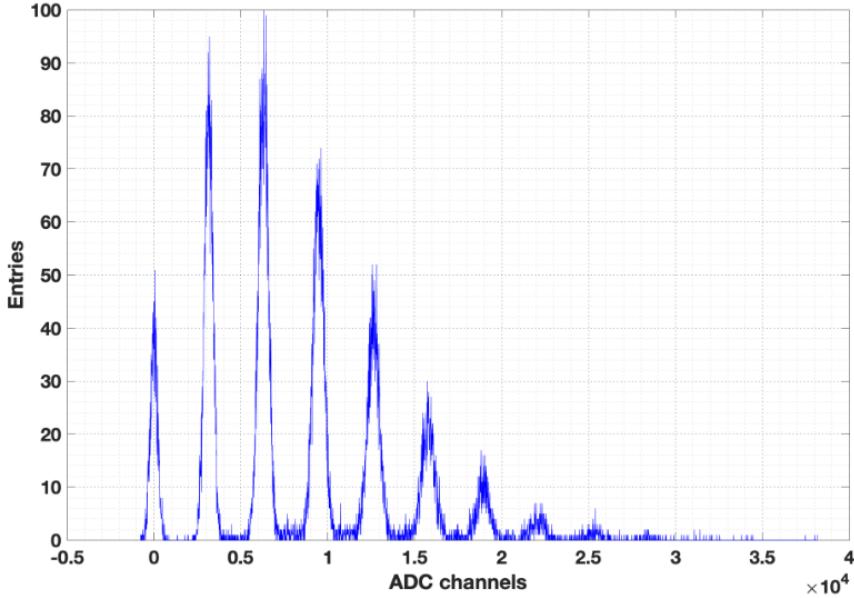


Figure 3.5: Response of Hamamatsu MPPC 13360-1350CS illuminated by a LED: the integrated and digitised signal [42].

A common way to evaluate the gain of a SiPM is to analyse the multi-photon charge spectrum, an example is shown in Figure 3.5. It represents the integrated signal of a SiPM when fired with different number of photons in each event. The histogram is populated with the output charge obtained integrating the signal over a fixed time window. Each peak is associated to events with the same number of fired cells, and, thanks to the high gain, the noise is small compared to the signal intensity. This feature permits to separate different peaks and, hence, correctly identify the number of photons detected.

Moreover, the distance between two consecutive peaks, expressed in unit of charge, can be used to directly estimate the SiPM gain via the relation  $\Delta_{pp} = q_e \cdot G$ .

The number of fired cells ( $N_{FC}$ ) follows the Poisson statistic, giving an increasing spread of the peaks that is proportional to  $\sqrt{N_{FC}}$ . This behaviour sets a maximum number of resolvable peaks, limiting the photon detection capability of these detectors.

## 3.2 SiPM photon detection efficiency

The efficiency of a SiPM in detecting light is one of the most important parameters that characterise the detector. To numerically evaluate this feature, the Photon Detection Efficiency (*PDE*) quantity has to be defined. It represents the probability of producing a signal by the SiPM when it is fired by a single photon and it is quantified as the ratio of the number of fired cells that generate a signal ( $N_{FC}$ ) and the total number of photons hitting the detector  $N_\gamma$  [43].

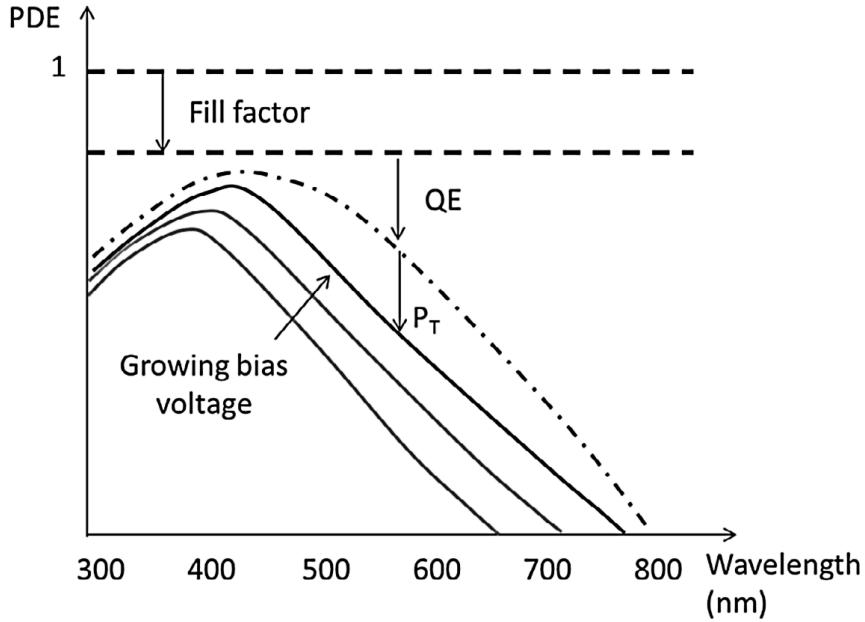


Figure 3.6: Sketch of the PDE on a blue sensitive device [40].

The silicon light-absorption coefficient is highly dependent on the wavelength of the particle and it can vary by several order of magnitude with light from UV to infrared (IR). For this reason, the *PDE* is a function of the wavelength  $\lambda$  of the impinging photon, of the temperature  $T$  and of the overvoltage  $V_{ov}$ . The *PDE* can be written as:

$$PDE = \frac{\langle N_{FC} \rangle}{\langle N_\gamma \rangle} = FF \times QE(\lambda) \times P_T(\lambda, V_{ov}, T) \quad (3.4)$$

where *FF* is the geometrical fill factor, *QE* is the quantum efficiency and  $P_T$  is the avalanche breakdown triggering probability.

In Figure 3.6 a sketch of the effect of each component is shown. The *PDE* is plotted as a function of the wavelength starting from the ideal maximal value of 1, then is reduced by a constant contribution coming from the geometrical fill factor, then the quantum efficiency reduces it in different ways, depending on  $\lambda$ , and finally the trigger probability is involved with his triple dependence on wavelength, temperature and overvoltage.

### Geometrical fill factor

The geometrical fill factor represents the probability that a photon impinging the SiPM traverses the active region. Quantitatively it is defined as the ratio between the sensitive area and the total area of the detector. In practical terms, the composition of SPADs also includes three main inactive regions: the guard ring, which surround each cell to quickly decrease the electric field at the border, the isolation structure between the SPADs in a

SiPM, needed to provide electrical and optical separation, and electrical elements (such as quenching resistors and bias lines). These dead regions contribute in reducing the fill factor.

To minimise the dead region area, a common technique is the trench isolation, consisting in a narrow depleted region placed around the active area, where a dielectric layer is deposited. The trenches and the guard rings are typically designed together to maximise the geometrical fill factor.

With modern technologies the  $FF$  value ranges from 90% for 100  $\mu\text{m}$  down to 30% for 10  $\mu\text{m}$  pixel pitch.

### Quantum efficiency

The quantum efficiency is the probability that an impinging photon crosses the anti-reflective coating layer and generates a pair of carriers. The coating is a stack of different dielectric layers placed on top of the SiPM and there is not a general rule for choosing the optimal layers, since each application has different needs depending on the wavelength to be absorbed.

On the other hand, the probability to produce a  $e-h$  pair by a photon highly depends on the absorption depth in the silicon. In particular, the closer to the junction the photon is absorbed, the higher the probability to produce a pair of carrier is. The absorption depth, in turn, is a function of the wavelength of the photon in such a way that it ranges from  $\simeq 10$  nm for ultraviolet light up to tens of  $\mu\text{m}$  for infrared light.

The wavelength dependence can make it necessary to use a particular coating called *wavelength shifter*, converting photons to other photons with different  $\lambda$ . This process can increase the quantum efficiency and, therefore, the total  $PDE$ .

### Avalanche triggering probability

The avalanche triggering probability represents the probability of a carrier to produce an avalanche while traversing the cell high-field region. At the production of a pair in the depleted region, the two carriers are affected by the electric field and start to drift in opposite direction. The total triggering probability has a contribution from both of them:

$$P_T = P_e + P_h - P_e P_h, \quad (3.5)$$

where  $P_T$  is the total triggering probability,  $P_e$  and  $P_h$  are the triggering probability for single carriers and the last term accounts for the cases when both carriers produce an avalanche. The electron triggering probability, in the same conditions, is more efficient, with respect to the ion triggering one, because of its greater (about twice higher) ionisation rate.

These values are electric-field dependent, hence they vary with the generation position and with the power-up voltage. Moreover, the avalanche triggering probability increases with higher overvoltage and lower temperature and it is also wavelength dependent.

### 3.2.1 Linearity and occupancy effect

In calorimetry measurements, the linearity of the detector with respect to the particle energy is one of the most important features. Considering that the single-charge output is almost constant for each cell of the SiPM, the greatest contribution to the non-linearity of these detectors relies on their digital nature.

The photon detection capability of single APDs is limited to a binary response, hence, two or more impinging photons can not be distinguished if they are temporally separated for less than the cell recovery time.

Starting from this consideration, the SiPM response can be considered linear when the number of photons  $N_\gamma$  multiplied by the  $PDE$  is small with respect to the number of pixels ( $N_{tot}$ ) composing the SiPM.

On the other hand, while  $N_\gamma \cdot PDE$  increases, the probability of having more than one photon on the same cell becomes not negligible. In this case the occupancy effect of the SiPMs may produce a non-linear response or even saturation. Considering the random nature of the variable  $N_{FC}$  and introducing the number of detected photons as  $N_{pe} = N_\gamma \cdot PDE$ , the occupancy effect can be expressed through the law:

$$N_{FC} = N_{tot} \left[ 1 - \left( 1 - \frac{1}{N_{tot}} \right)^{N_{pe}} \right]. \quad (3.6)$$

The formula relates the number of fired cells  $N_{FC}$  with the total number of cells  $N_{tot}$  and the number of photoelectrons  $N_{pe}$ . In a regime of high incident light ( $N_{pe} \rightarrow \infty$ ), since the value  $N_{tot}$  is usually greater than 100, it can be expressed as:

$$N_{FC} = N_{tot} \left( 1 - e^{-\frac{N_{pe}}{N_{tot}}} \right). \quad (3.7)$$

We will use Equation 3.7 when studying the linearity of a SiPM-based readout for the IDEA dual-readout calorimeter.

It is important to keep in mind that this formula is an approximation: it is based on the assumption of cells uniformly irradiated, no optical cross-talk, no pulse recovery effect. However, considering the complicated construction of a physical model, it represents a good compromise. This statement is also supported by experimental data such as the ones produced by Hamamatsu and shown in Figure 3.7, where the fired cells are plotted as function of  $N_\gamma$ .

## 3.3 Noise effects

When a SiPM is powered up, at least three noise sources are present and must be known [40].

Noises can be distinguished in two categories: primary noise and correlated noise. Primary noises arise when a carrier, generated through thermal generation or tunneling, traverse the depleted region of a SPAD and trigger the Geiger avalanche.

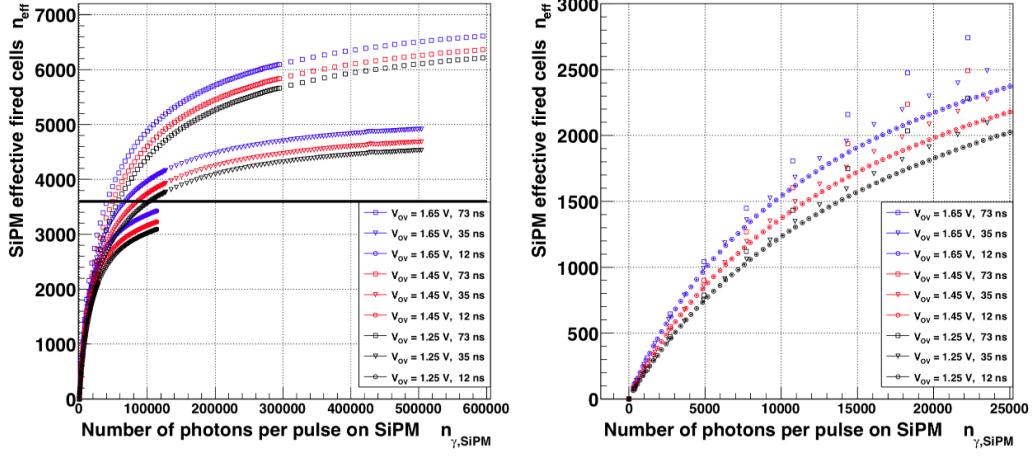


Figure 3.7: Effective fired cells as a function of the number of photons per pulse on the active SiPM surface (over the full and on a reduced interval) for the Hamamatsu S10362-050C type SiPMs, with a 250 ns gate length [44].

On the other hand, a correlated noise is identified as an avalanche discharge triggered by another previous signal in the SiPM. A sketch of the most important noise sources that affect SiPMs is shown in Figure 3.8.

### 3.3.1 Dark Count Rate

In the depleted region, the generation of free carrier is possible also in absence of light. These type of signals are called *dark counts* and are associated to the numerical value of the Dark Count Rate (DCR). With modern technologies the DCR value ranges from tens of kHz to MHz per mm<sup>2</sup>. Since most of the electron-hole pairs are generated thermally or by the electric field, these sources produce pulses not distinguishable from the pulses generated by photoelectrons. Hence, the impact of the dark count rate can be mitigated only through average corrections.

The thermally generated pairs can be reduced by lowering the detector temperature, and the ones produced by the electric field can be reduced by lowering the field itself. However this last has a smaller impact then the thermal one.

### 3.3.2 Optical Cross-Talk

During the avalanche process, there are, on average, 3 emitted photons every 10<sup>5</sup> carrier that cross the junction with at least 1.14 eV (the band gap of silicon). These photons can trigger an avalanche in a neighbour cell, by producing a signal that is classified as correlated noise under the name of *Optical Cross-Talk* (OCT).

The noise production process is roughly instantaneous, making impossible to separate the true signal from the noise. The effect results in a pulse doubled with respect to the

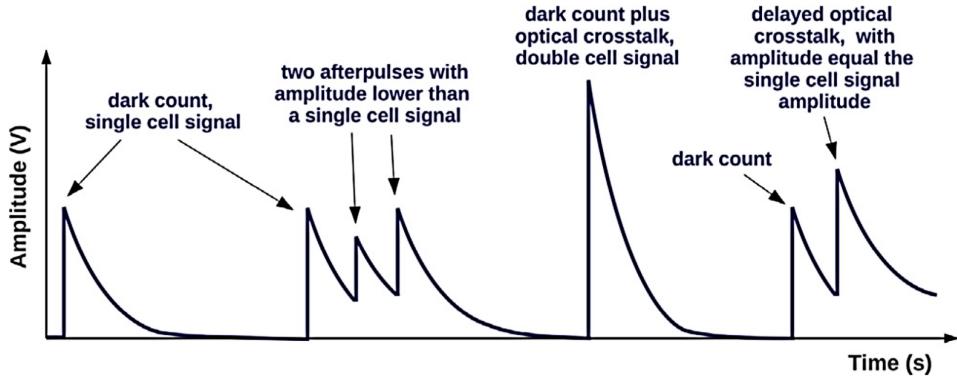


Figure 3.8: Sketch of the different SiPM output noise signals [41].

correct one.

Considering, for example, a SiPM with a gain of  $10^6$ , each avalanche discharge produce on average 30 photons that can reach neighbouring cells and trigger noise signals. At present, to mitigate this effect, SiPM cells are surrounded by elements coated with reflective material reducing the OCT probability to 1 – 2%.

### 3.3.3 Afterpulse

Another source of correlated noise is known as *AfterPulsing* (AP). When an avalanche discharge occurs, carriers can be trapped by impurities in the high-field region. They can be released after a delay time that can range from few ns up to several  $\mu$ s. Once they are released, these carrier produce a signal in the same cell where they were generated. Hence the signal is delayed and its amplitude depends on the delay time: the greater the delay time, the more the cell is recharged, the higher the produced signal amplitude. In first approximation, the charge ratio between the AP and the original pulse follows the law:

$$\frac{Q_{AP}}{Q} = 1 - e^{-\frac{\Delta t}{\tau_r}} \quad (3.8)$$

where  $\Delta t$  is the delay time and  $\tau_r$  is the cell recovery time constant.

The probability of this correlated noise can be lowered by reducing the overvoltage value (with the side effect of reducing the SiPM gain) or by increasing the temperature (producing also an increase of the DCR and OCT probabilities).

Using high-quality silicon lattices, the afterpulse noise probability is reduced to values of 1 – 3%.

## Chapter 4

# IDEA dual-readout calorimeter simulation and SiPM signal digitisation

As already said, a full simulation of the dual-readout calorimeter performance has been implemented. It has been done, in particular, in the context of the R&D work for the preparation of the Conceptual Design Reports (CDRs) and, next, of the Technical Design Reports (TDRs) for both FCC-ee and CEPC. The simulation includes the showering of particles in the detector and the generation and transport of photons till the readout sensors. Finally, the photon time of arrival is used to trigger the simulation of the sensor response which has been extensively studied for this thesis.

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

- the calorimeter simulation, coded in C++ using the GEANT4 toolkit;
- the SiPM response digitisation, coded in a dedicated Python application.

Important aspects of the SiPM response, such as the temporal behaviour, the SiPM non-linear response and the energy resolution will be described in Section 4.2.

The SiPM readout, in which each fibre is coupled to a dedicated SiPM, gives the possibility to sample showers with a unprecedented 2D granularity. In this context the usage of deep neural networks for particle identification studies has been pioneered and will be discussed in the next chapter.

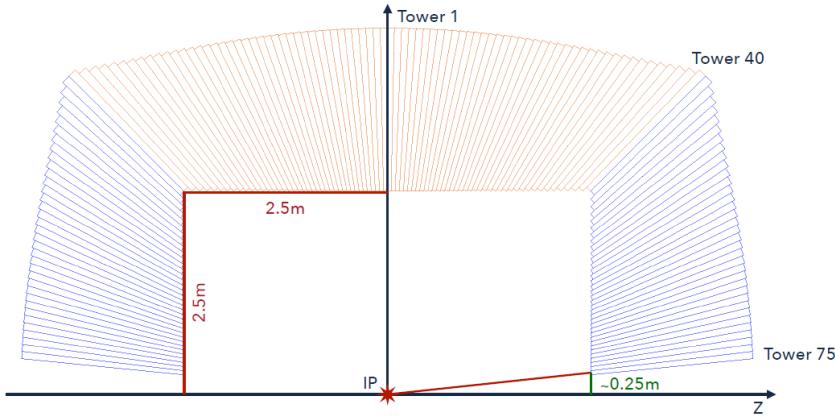


Figure 4.1: A calorimeter slice.

## 4.1 Simulation structure

### 4.1.1 Calorimeter simulation

The simulated calorimeter structure follows the conceptual design shown in Section 1.3. It has a cylindrical symmetry characterised by a barrel and two end-cap regions. This  $4\pi$  structure is obtained through 36 discrete rotations, around the  $z$  axis, of a simpler unit, called a slice. The dimensions of each slice are shown in Figure 4.1. The inner diameter and the inner length are both 5 m, and the overall outer diameter and length are 9 m. Each half slice is composed by 75, 2 m long, towers (40 in the barrel and 35 in the endcap region) and a total of 5400 towers compose the whole calorimeter. To cover a solid angle very close to  $4\pi$ , each tower has a different trapezoidal inner face with dimensions that can vary from  $\sim 5$  cm to  $\sim 8$  cm. The coverage in  $\theta$  is down to  $\sim 0.1$  rad from the beam line.

At present, towers are copper-based with fibres inserted in a chessboard-like geometry. Longitudinally running holes host optical fibres as sensitive elements. The need of a projective geometry makes the detector dimensions increasing with the distance from the IP. New holes (with new fibres) open up at different depths inside the calorimeter, whenever possible, to keep the sampling fraction constant.

As the dual-readout technique needs independent scintillating ( $S$ ) and Cherenkov ( $C$ ) signals, two types of fibres are used (Fig. 4.2). Their characteristics are shown in Tab. 4.1.

The fibre refractive indices determine the light collection capability, as dictated by Snell's law. The signal from the scintillating fibres is parametrised by the deposited energy while the Cherenkov photons are produced accordingly to the Cherenkov emission process.

For each event, the simulation produces as output the following information per each activated fibre:

---

<b>Kuraray SCSF-78 (S)</b>	
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
<b>Mitsubishi SK-40 (C)</b>	
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.

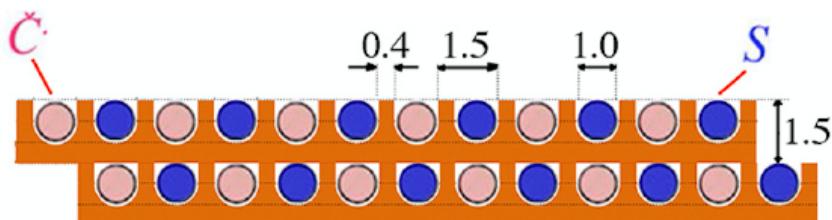


Figure 4.2: Chess-like configuration to dispose fibres in the copper structure.

- Event ID;
- Fibre Type ( $C$  or  $S$ );
- Fibre ID (a unique number per fibre);
- the  $(x, y, z)$  position of the fibre tip closer to the IP;
- the number of photons reaching the SiPM;
- the list of photon times of arrival to the SiPM.

The computation of the light propagation inside fibres is extremely time consuming, so that it has to be fine tuned in order to optimise the full process. In particular, the propagation of  $C$  photons is tracked until the single photon reach the core-cladding boundary. If the emission angle is inside the range of the fibre numerical aperture, the photon is added to the final number of photons. Following a Poissonian probability smearing of being detected, each photon is converted to a photoelectron (p.e.). The time of arrival on the sensor for each photon is estimated as:

$$t_C = t_0 + R \frac{n_C}{c \cdot \cos(\vartheta)} \quad (4.1)$$

where  $n_C = 1.49$  is the fibre refractive index,  $c$  the speed of light,  $R$  the distance from the further end of the fibre and  $t_0$  the time at which the photon reach the core-cladding boundary.

For the  $S$  fibres, instead, the photons are produced by weighting the light yield of the fibres with the actual energy deposited by the interacting particle. The number of photons is smeared with a Poissonian law for the photoelectron conversion and the time of arrival on the sensor is obtained as:

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

where  $n_S = 1.59$  is the refractive index and  $t^*$  a random time that accounts for the scintillator decay time, chosen from an exponential distribution with 2.8 ns as mean value. Considering the internal reflection, the photon path depends on the  $\vartheta$  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  $\alpha = 20.4^\circ$  is the fibre critical angle.

Eventually, the light produced is smeared by two Poissonian distribution, one for the scintillation signal and the other for the Cherenkov one. This procedure correctly reproduces the statistical fluctuations in the scintillation and Cherenkov light production and allows to reproduce within the simulations the desired light yields (p.e./GeV). The simulation is tuned to produce  $\sim 400$  Spe/GeV and  $\sim 100$  Cpe/GeV at the electromagnetic scale. (i.e. for em showers).

### 4.1.2 SiPM response digitization

The output of the G4 simulation is used as input of the second part of the simulation: *pySiPM*, a python-based Monte Carlo simulation able to reproduce the SiPM transfer function and extract features from the waveforms recorded with a digitiser [45].

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

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:** sampling frequency, integration time window, integration starting time.

For each event and fibre, random parameters determine the photon position inside the sensor. Meanwhile the sensor PDE is set at 100% since the actual PDE is accounted within the poissonian smearing applied at the calorimeter simulation level. When a photon impinges on a cell, a control is activated preventing the production of a new signal from the same cell. The control is removed after the recovery time of the cell set at  $10 \cdot \tau_{fall}$ .

The photon times of arrival to the SiPM surface are synchronised to a common time trigger (in our case, the production of the particle at the interaction point). To be consistent with this time-zero (TZ) definition, the ToAs are passed to the digitiser software and used to start the avalanche discharge simulation at the exact time. In this way all the digitised waveforms are synchronised to the same TZ as defined by the GEANT4 trigger.

The generated pulse is a combination of two exponentials characterised by the rise time constant ( $\tau_{rise}$ ) and the decay time constant ( $\tau_{fall}$ ):

$$f(t) = A \cdot \left( e^{-\frac{t+t_0}{\tau_{fall}}} - e^{-\frac{t+t_0}{\tau_{rise}}} \right), \quad (4.3)$$

where  $t_0$  is the photon time of arrival on the SiPM surface.

The output signal for any given SiPM is the sum of all the signals generated by the activated cells.

The information given as output of the SiPM simulation are:

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

- **Computed quantities:** signal integral, peak height, time of arrival, time over threshold, time of peak;
- **Digitised waveform.**

It is also possible to select an analog threshold on the signal height to establish if information has to be recorded. The threshold is defined as a scale factor of the maximum value of the waveform generated by a single photoelectron (neglecting the electrical noise). In the results shown later a 1.5pe-suppression has been applied using a threshold of 1.5 of the signal height from a single photon. Such a filter is useful to neglect signals from DCR.

## 4.2 Results from the full simulation chain

### 4.2.1 Different configurations

The results shown in this chapter are obtained considering different SiPM parameter configurations.

They have been chosen in a common parameter space identified by checking the portfolio of SiPMs produced by Hamamatsu [46]. Two are the parameters that have been changed in our studies:

- the decay time constant of the signal, the selected values being 10 ns and 50 ns;
- the pixel size, the selected values being 10  $\mu\text{m}$ , 15  $\mu\text{m}$  and 25  $\mu\text{m}$ .

The other parameters have not been modified for this work. Their values are listed in the Table 4.2.

An example of generated waveform, in response to 3 simultaneous photoelectrons with a time of arrival of 20 ns, is plotted in Figure 4.3. The impact of the two different decay-time constants is evident.

Other waveform examples are shown in Figures 4.4 and 4.5, where signals generated by DCR and after-pulse noise, respectively, are included.

### 4.2.2 Time studies

An important aspect that has to be studied is the time evolution of the signals. For that, data from 1000 events have been generated. In each event a 20 GeV electron, produced at the interaction point, showered in tower 1 of the GEANT4 DR calorimeter simulation. We firstly analysed the distribution of the time of arrival of the photons converted at the SiPMs surface (i.e. the time recorded in the GEANT4 simulation output).

The distributions obtained for  $C$  and  $S$  photons are plotted in Figure 4.6, where each entry corresponds to a single-photon time of arrival. As expected, the distribution of the  $C$  photon time of arrival is extremely narrow due to the prompt production of photons

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

Table 4.2: SiPM parameters kept constant in the simulation configuration file. For the values of the parameters that have been varied see the text.

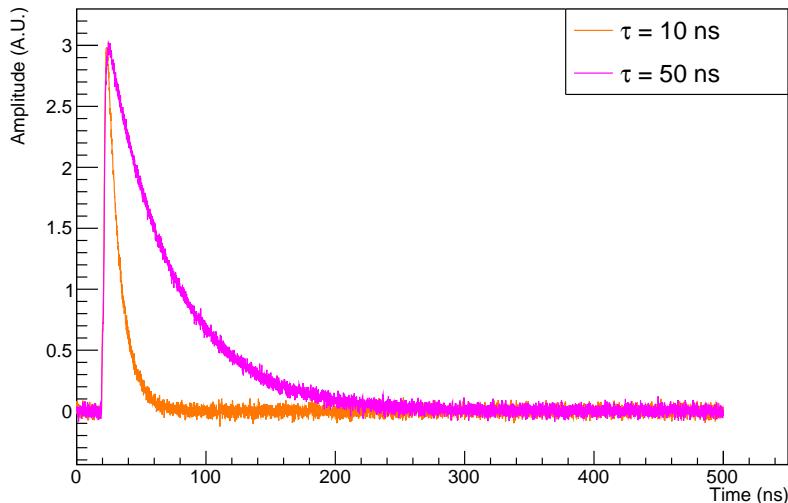


Figure 4.3: Single waveforms generated in response to 3 simultaneous photoelectrons with time of arrival of 20 ns while considering two identical conditions except for the decay-time constant.

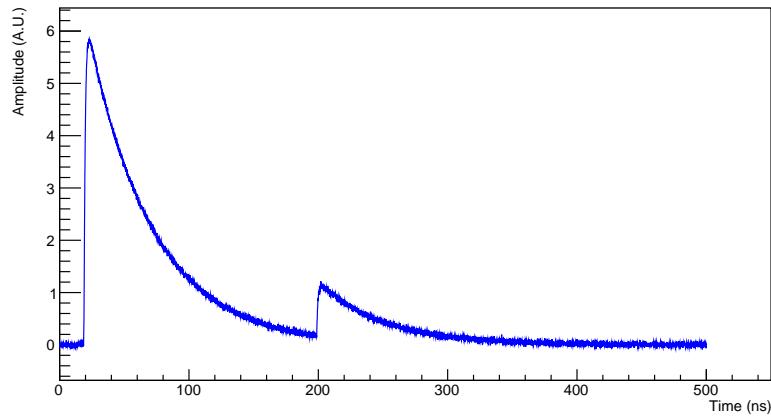


Figure 4.4: Single waveforms generated in response to 6 simultaneous photoelectrons with time of arrival of 20 ns. A peak generated by DCR noise is identified at  $\simeq 200$  ns.

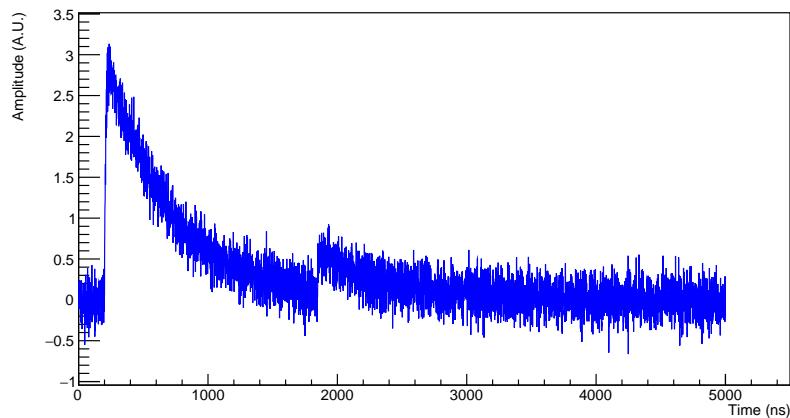


Figure 4.5: Single waveforms generated in response to 3 simultaneous photoelectrons with time of arrival of 20 ns. A peak generated by AP noise is identified at  $\simeq 180$  ns.

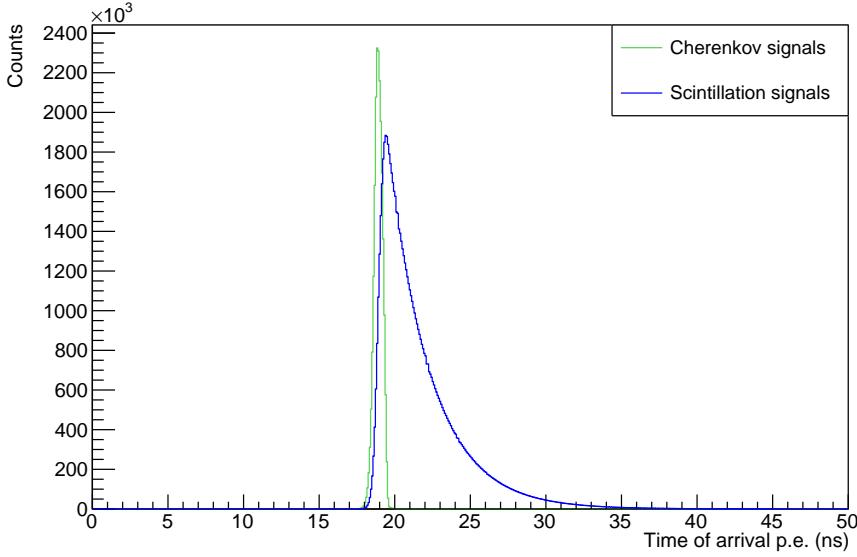


Figure 4.6: The distributions of the time of arrivals of photons at the SiPMs surface, separately for Cherenkov and scintillation fibres.

at the passage of relativistic charged particle in the fibres. Conversely, the  $S$  photon time distribution shows an exponential tail due to the emission time constant of the polystyrene-based scintillators ( $\tau = 2.8$  ns).

These data are then used as input for *pySiPM*. The SiPM parameters are chosen as described in Paragraph 4.2.1. In this context the most interesting parameter is the SiPM decay time constant.

Figures 4.7(a) and 4.7(b) show the peaking-time distributions for the two different SiPM configurations.

These data can be compared looking for differences as a function of the SiPM configuration. Figure 4.8(a) and (b) show the peaking-time distributions for Cherenkov and scintillation signals, respectively, in the two SiPM configuration. As expected, a broadening occurs with a larger effect for the configuration with  $\tau_{fall} = 50$  ns. In this case, the wider response function, on top of the electrical noise, causes a lost in time-of-peak resolution quantified by the broadening shown in the plots.

The impact of noise on the time-of-peak resolution is also dependent on the total number of photons impinging on the same SiPM. In particular, the time-of-peak resolution improves 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 signal time of peak has been recorded, plotted in an histogram and fitted with a Gaussian function.

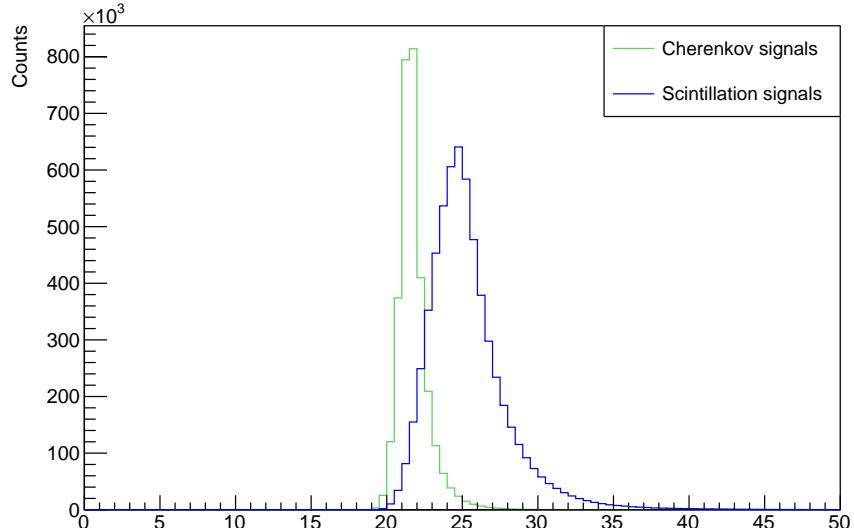
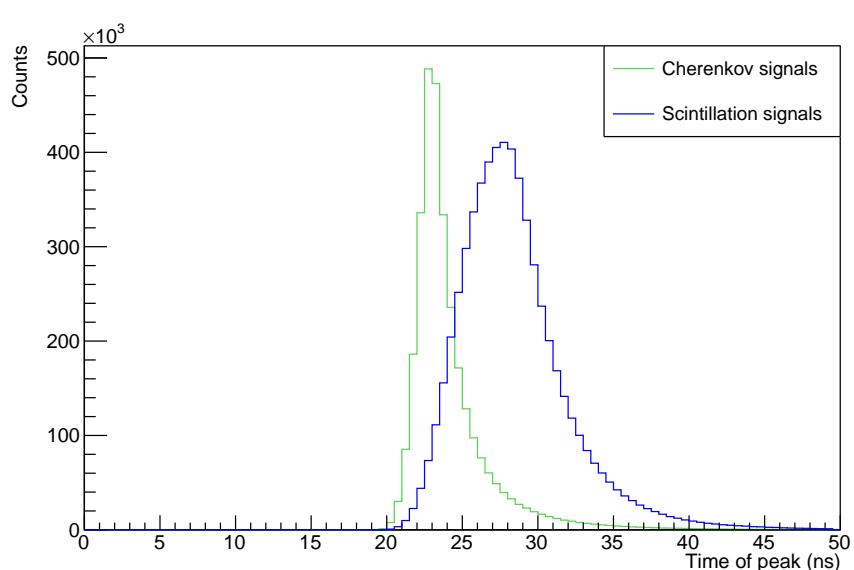
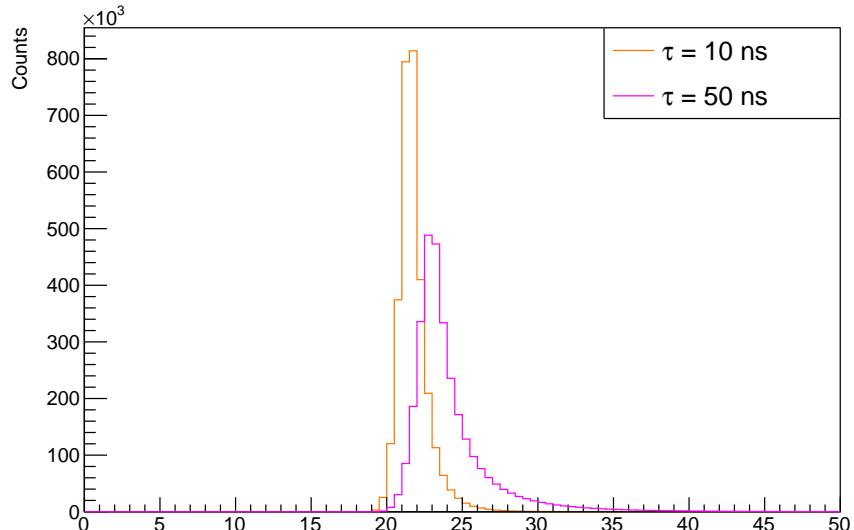
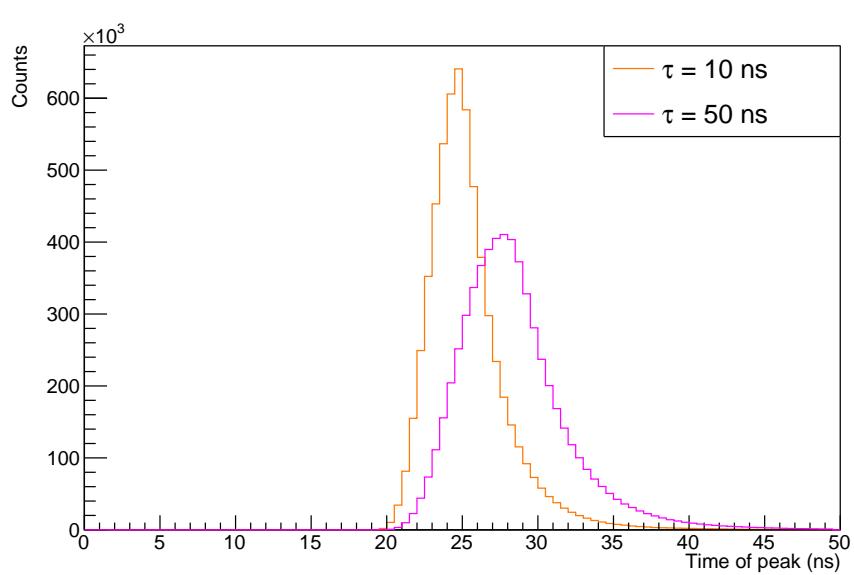
(a)  $\tau_{fall} = 10$  ns.(b)  $\tau_{fall} = 50$  ns.

Figure 4.7: Time-of-peak distributions comparing in each histogram the signals for Cherenkov and for scintillation fibres. Different figures correspond to different decay time constants. Results for 20 GeV electrons.



(a) Cherenkov fibres.



(b) Scintillation fibres.

Figure 4.8: Time-of-peak distributions comparing in each histogram, the signals generated with two different SiPM decay time constants, (a) for the Cherenkov, (b) for the scintillation fibres. Results for 20 GeV electrons.

Number of photons	$\sigma$ with $\tau_{fall} = 10$ ns (ns)	$\sigma$ 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 a Gaussian fit of the time-of-peak distributions under different conditions.

The standard deviation of these Gaussian fits is used as estimator for the resolution and has been recorded and reported in Table 4.3.

It is interesting to plot these data and study the behaviour of the standard deviation as a function of the number of simultaneous photons ( $n$ ), as shown in Figure 4.9. The distribution is well interpolated with a function of the form:

$$\sigma = \frac{A}{\sqrt{n}} + B. \quad (4.4)$$

The best estimate for the parameters is:  $A = 0.8712$  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, described in Paragraph 3.1, is an important characteristic that has to be deeply studied since it may impact on the calorimeter signal linearity.

Figure 4.10, the charge integral as a function of the true number of p.e., shows a quantitative example: while for small photon numbers a linear relation is found, a non-linearity effect is evident for events with many photons.

The simplest way to tackle this effect is to (possibly) identify a calibration analytical function that reproduces the charge integral with respect to the number of impinging p.e.

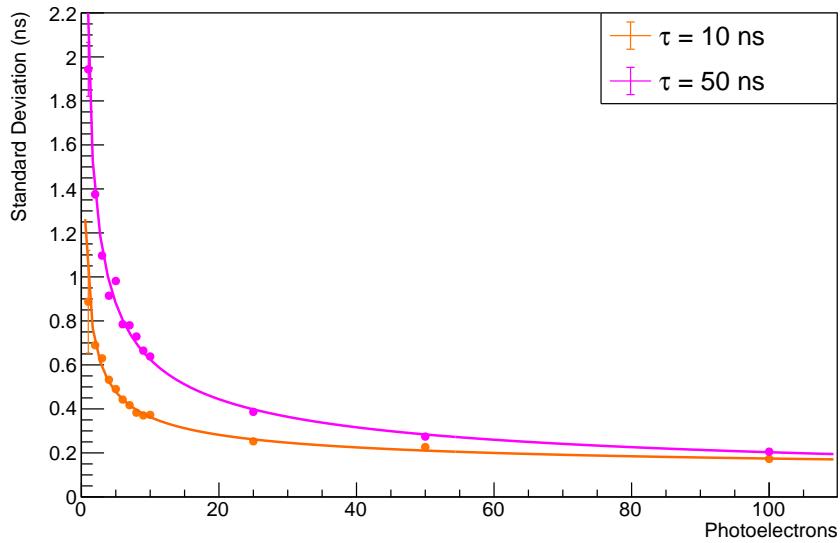


Figure 4.9: Standard deviation of the peaking time distribution for events with a fixed number of photoelectron, as a function of the number of simultaneous photoelectrons. Study performed with two different SiPM decay-time constants.

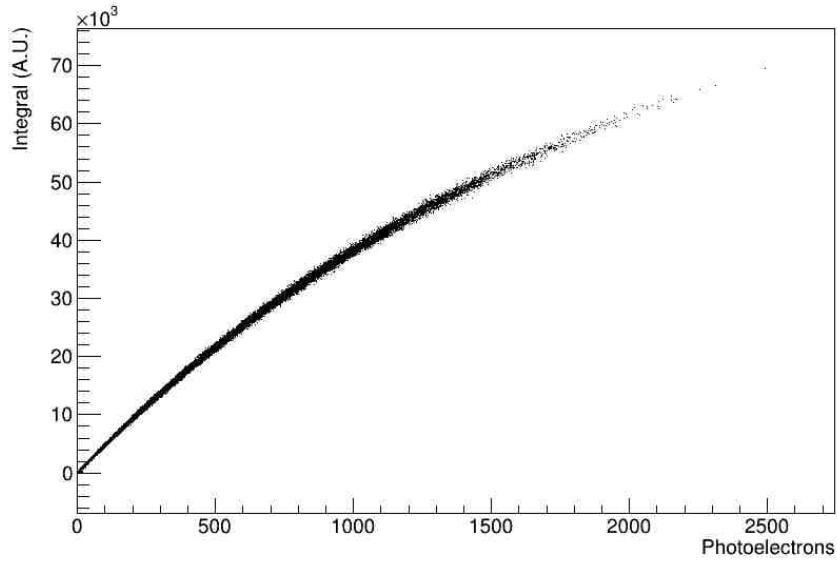


Figure 4.10: Example of a (simulated) non-linearity behaviour for a SiPM with a cell size of  $25\text{ }\mu\text{m}$ .

Starting by assuming that, in the range from 2 to 10 photoelectrons, in our SiPM configurations (i.e. 10000, 4356 or 1600 cells in each SiPM), the response is linear, 10000 SiPMs have been fired 9 times with an increasing number of simultaneous p.e.s (from 2 up to 10).

As a sanity check of the above assumption, the charge integral distribution from the three different configurations has been compared and no bias was found, as shown in Figure 4.11. The mean and the RMS values have been recorded and their behaviour with the number of photons, sketched in Figure 4.12, was fitted with a straight line allowing to obtain the initial calibration law:

$$I(n) = A \cdot n + B \quad (4.5)$$

with  $A = 49.43 \pm 0.03852$  and  $B = 2.491 \pm 0.2516$  (Fig. 4.12).

The parameter  $A$  represents the contribution to the charge integral associated to a single photoelectron. Instead,  $B$  is the pedestal originated mostly by the DCR events inside the integration time window. This contribution can be compared to analytical calculation. In the simulation, we defined a DCR of 200 kH and an integration window of 300 ns. The pedestal contribution due to DCR can be estimated as follows:

$$B_{DCR} = 2 \cdot 10^5 \cdot 3 \cdot 10^{-7} p.e. = 6 \cdot 10^{-2} p.e. = 6 \cdot 10^{-2} \cdot 49.43 = 2.96 \text{ (in good agreement with the fit result).}$$

In the following, the pedestal (fitted) value is always subtracted to the charge-integral value of each SiPM.

The occupancy effect has been studied for electromagnetic showers developing in the IDEA calorimeter simulation produced by single electrons with energies of 20, 40, 60 and 80 GeV.

The non-linearity effects have been quantified using the relation  $I = A \cdot n$  (after pedestal subtraction) as reference. The results obtained with 10000 events of single 40 GeV  $e^-$ , for different SiPM configurations, are shown in Figure 4.13, for the Cherenkov and scintillation signals independently. For each event, there is one entry for each fibre (i.e. SiPM) that got a signal. Each point correspond to a single SiPM. Clearly, the smaller the number of cells, the greater the non-linearity effect.

Moreover, since, as expected, the scintillation fibres have on average  $\sim 4$  times more p.e.s than the Cherenkov ones, the  $S$  signal is more affected than the  $C$  one.

This process can be extended by independently integrating, the number of photoelectrons and the charge, over the full event. The result is represented in Figure 4.14 where each point correspond to a single event.

As can be seen, non-linearity effects show up even for 10  $\mu\text{m}$  cell SiPMs. To mitigate this problem, an analytical correction can be introduced via the formula:

$$N_{fired} = N_{cells} \cdot \left[ 1 - \exp \left( -\frac{N_{p.e.}}{N_{cells}} \right) \right], \quad (4.6)$$

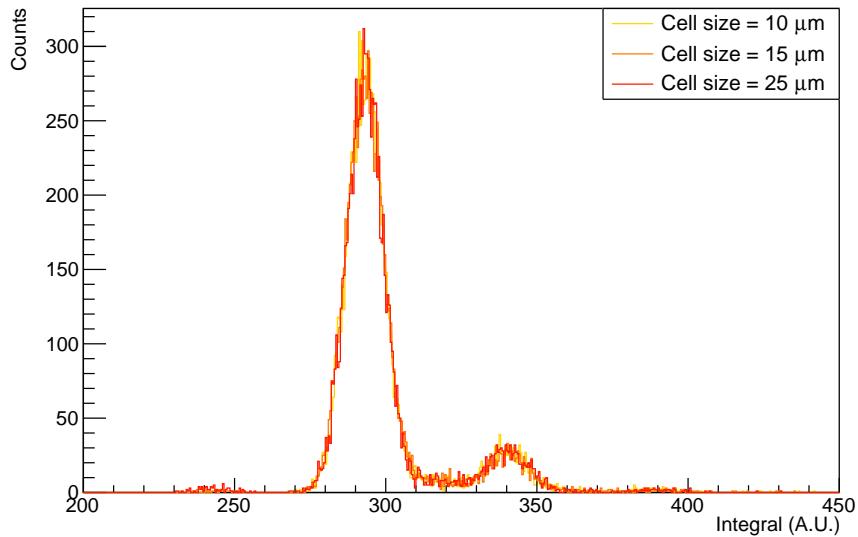


Figure 4.11: Charge integral distributions with different SiPM cell sizes considering 6 simultaneous photoelectrons. The distribution shape (in the full range considered, from 2 up to 10 simultaneous p.e.) is independent of the number of cells.

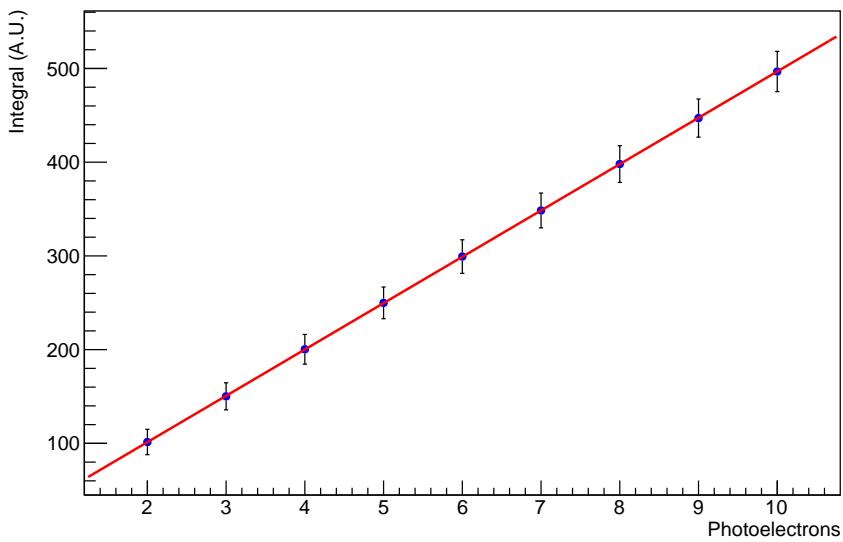
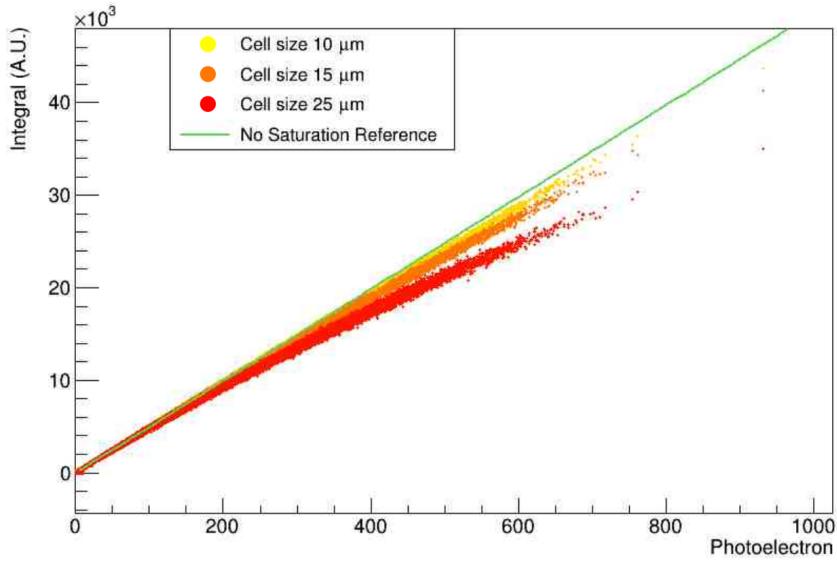
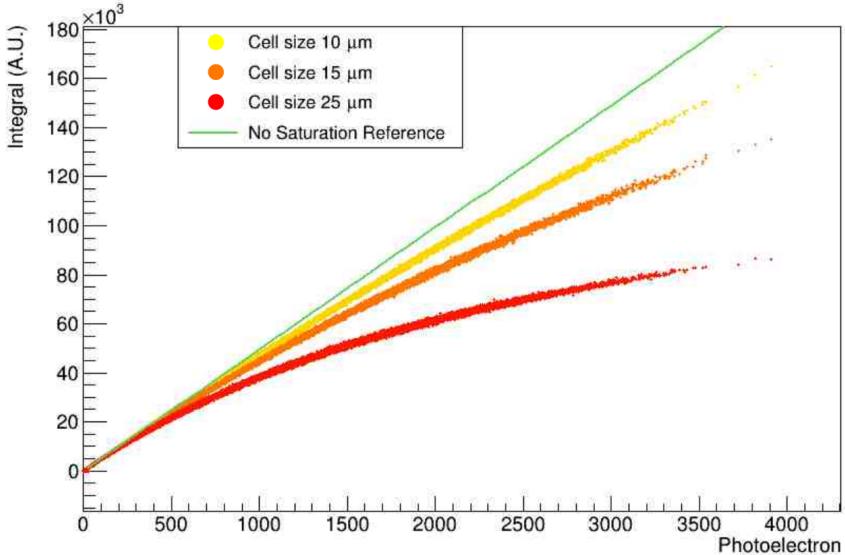


Figure 4.12: Charge integral as a function of the number of photoelectrons, in the range from 2 to 10 p.e.s. A linear fit has been performed to find the calibration law in the linear regime.

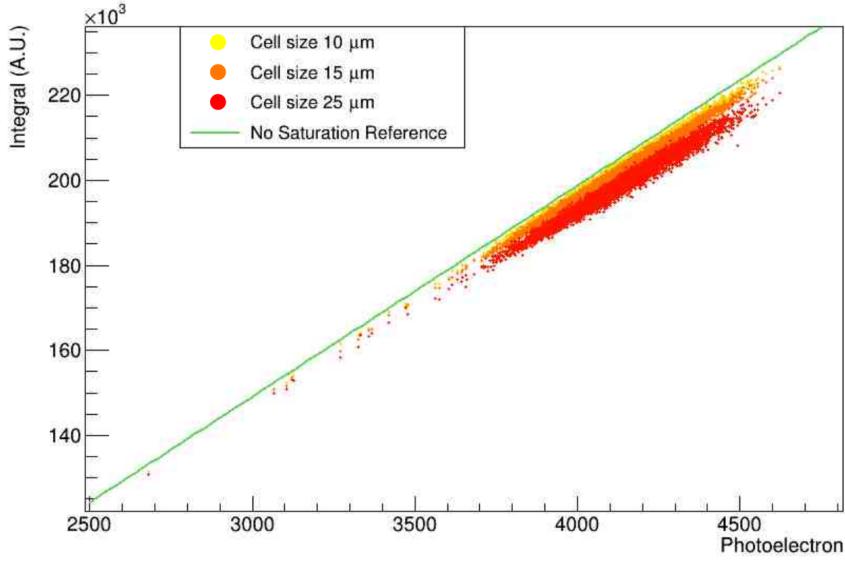


(a) Cherenkov signals.

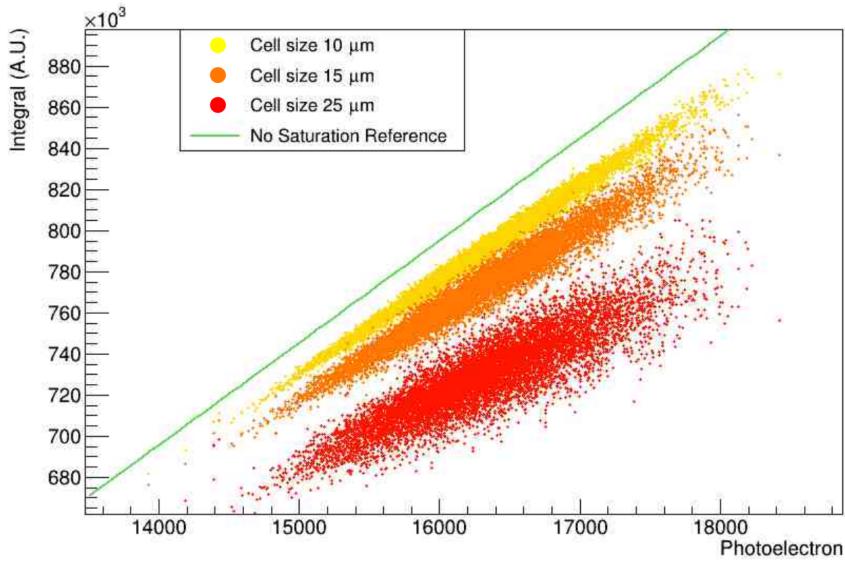


(b) Scintillation signals.

Figure 4.13: The plots, (a) for the Cherenkov and (b) for the scintillation signal, show the impact of non-linearity effects for single 40 GeV electrons, with different SiPM cell sizes. For each event, there is one entry for each fibre (i.e. SiPM) that got a signal. Each point corresponds to a SiPM with different values of cell size. The linear behaviour, as extrapolated from the fit to the plot in Figure 4.12, is included as reference.



(a) Cherenkov signals.



(b) Scintillation signals.

Figure 4.14: The plots, (a) for the Cherenkov, (b) for the scintillation signal, show the non-linearity effects with single 40 GeV electrons. Each of the 10000 points correspond to a single event where the charge and the number of photoelectrons was integrated over all the fired fibres. Different values of the SiPM cell size are represented with different colours. The linear behaviour, as extrapolated from the fit to the plot in Figure 4.12, is included for reference.

The above correction has been applied by modifying the integral values as follows:

$$Q_{corr} = -AN_{cells} \left[ \ln \left( 1 - \frac{Q}{AN_{cells}} \right) \right], \quad (4.7)$$

where the linear relation to convert  $N_{cells}$  ( $N_{fired}$ ) to the corresponding integral value  $Q_{corr}$  ( $Q$ ), through the constant  $A$  (as fitted in the linear regime region), has been applied.

The results obtained can be found in Figures 4.15 and 4.16 where the data, for SiPMs with a cell size of  $10 \mu\text{m}$ , are compared with and without the analytical correction.

The discrepancy from the linearly extrapolated reference quantifies the effect of the occupancy when performing the energy reconstruction task. The percentage difference has been evaluated through the formula  $\frac{E_{linear} - E}{E_{linear}}$ , and the values obtained fill the histograms in Figure 4.17.

After applying the analytical correction a clear improvement in terms of readout linearity is found as shown in Figure 4.18.

This whole process has been performed simulating electrons with energies of 20, 40, 60 and 80 GeV. Mean and standard deviation of the gaussian fit of the percentage discrepancy have been recorded and their values are plotted in Figure 4.19 where a non-linearity well below the 1% is obtained for both Cherenkov and scintillation signals, after applying the analytical corrections for a  $10 \mu\text{m}$  cell size.

#### 4.2.4 Energy resolution

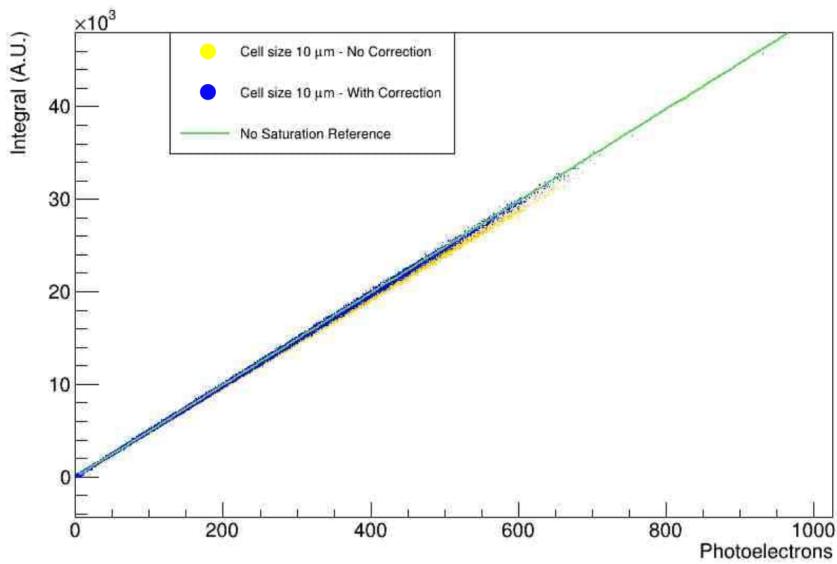
Dual-readout calorimeters are typically calibrated at the electromagnetic scale. This is especially useful in a leptonic collider because, having easily access to reference “candles” such as electrons and positrons from  $Z$  decays, the calorimeter can be precisely calibrated during all the life of the experiment.

The first step to study the energy resolution is to calibrate the calorimeter.

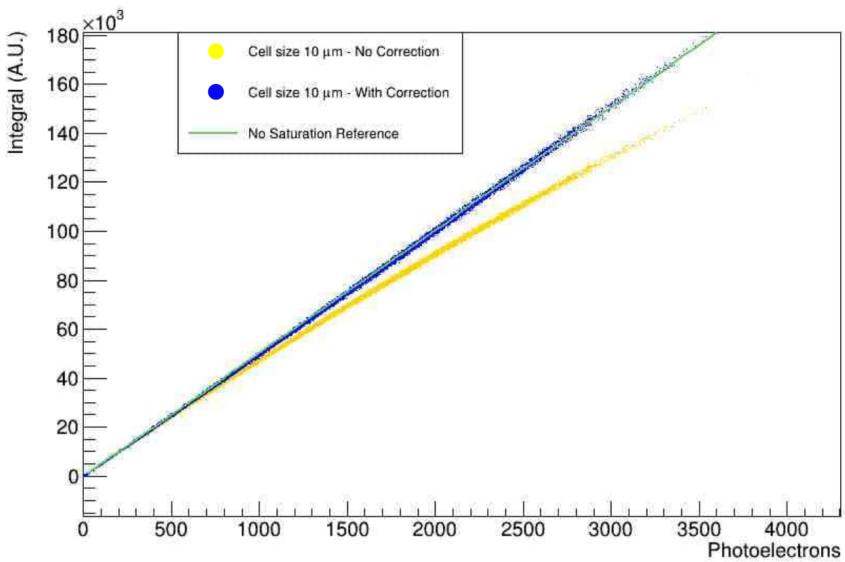
The calibration has been performed with 40 GeV electrons applying the 1.5pe-suppression already introduced in Paragraph 4.1.2. 10000 events with single 40 GeV electrons have been fired from the interaction point obtaining the charge integral distributions for scintillation and Cherenkov signals. The calibration constants obtained to transform these data in energy distributions centred around 40 GeV are:  $k_S = 4.998 \times 10^{-5} \text{ GeV/A.U.}$  and  $k_C = 2.023 \times 10^{-4} \text{ GeV/A.U.}$

Two analogue calibration constants have been obtained starting from the number of photoelectron distribution with values of:  $k_{pe,S} = 2.48 \times 10^{-3} \text{ GeV/p.e.}$  and  $k_{pe,C} = 1.00 \times 10^{-2} \text{ GeV/p.e.}$

Starting from these calibration constants, the energy distributions can be obtained either from the number of p.e.s (pre SiPM digitisation simulation) or from the charge integral

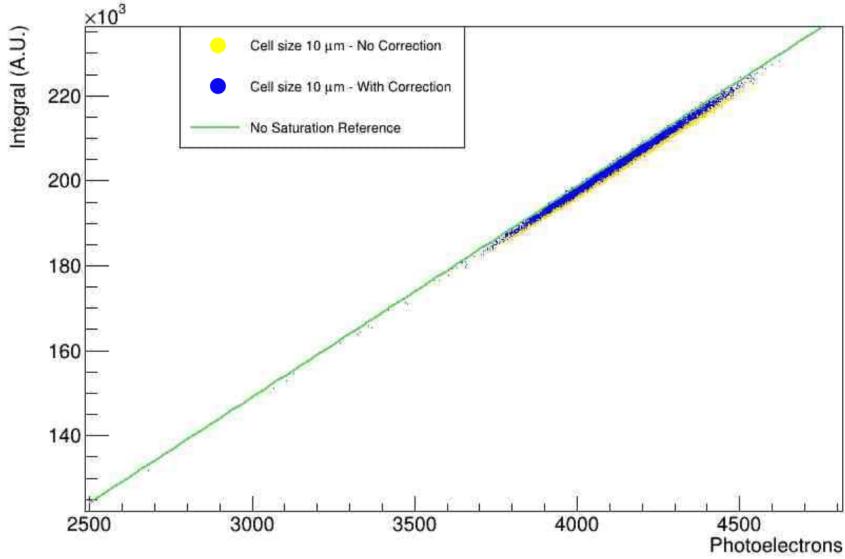


(a) Cherenkov signals. Points are single SiPMs.

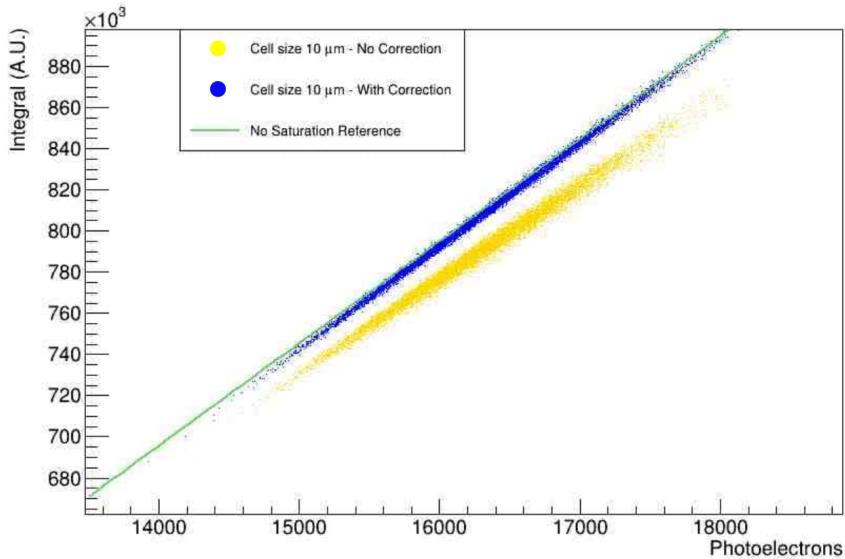


(b) Scintillation signals. Points are single SiPMs.

Figure 4.15: Similar plots to Figures 4.13. Each point correspond to a  $10 \mu\text{m}$  cell-size SiPM, with or without applying the analytical correction. Two series of data are presented to show the effectiveness of the analytical correction. The linear behaviour, as extrapolated from the fit to the plot in Figure 4.12, is included as reference. Results for 40 GeV electrons.

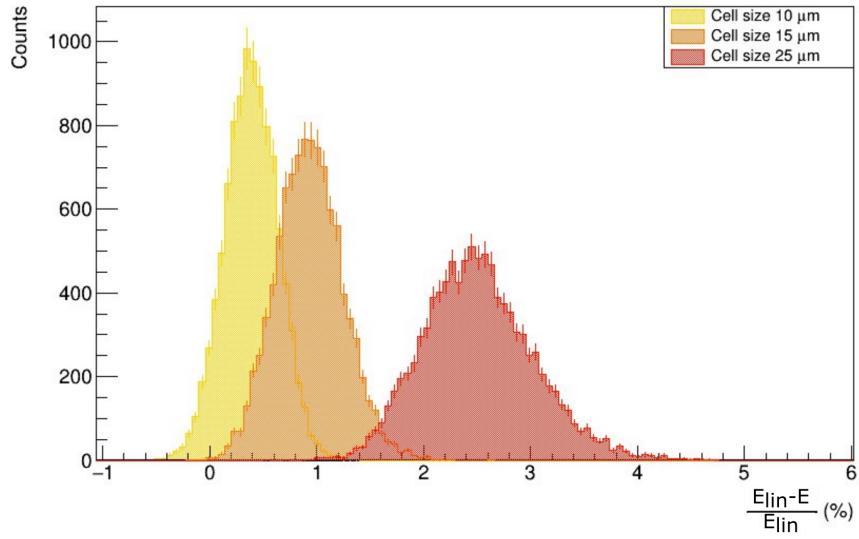


(a) Cherenkov signals. Points are single events.

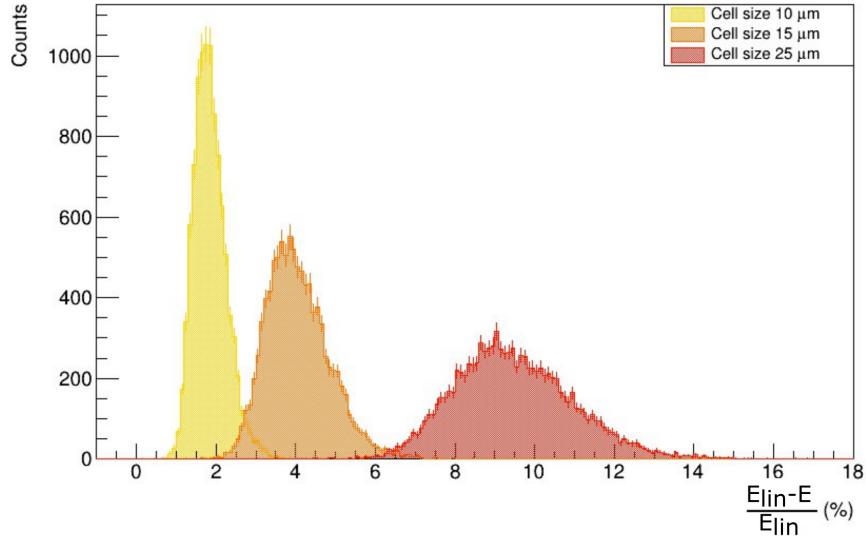


(b) Scintillation signals. Points are single events.

Figure 4.16: Similar plots to Figures 4.14. Each of the 10000 points correspond to a single event where integral and number of photoelectrons have been added over the fired SiPMs. Two series of data (with or without applying the analytical correction) are presented to show the effectiveness of the analytical correction. The no saturation line shown in Figure 4.12 has been added as reference. Results for 40 GeV electrons.

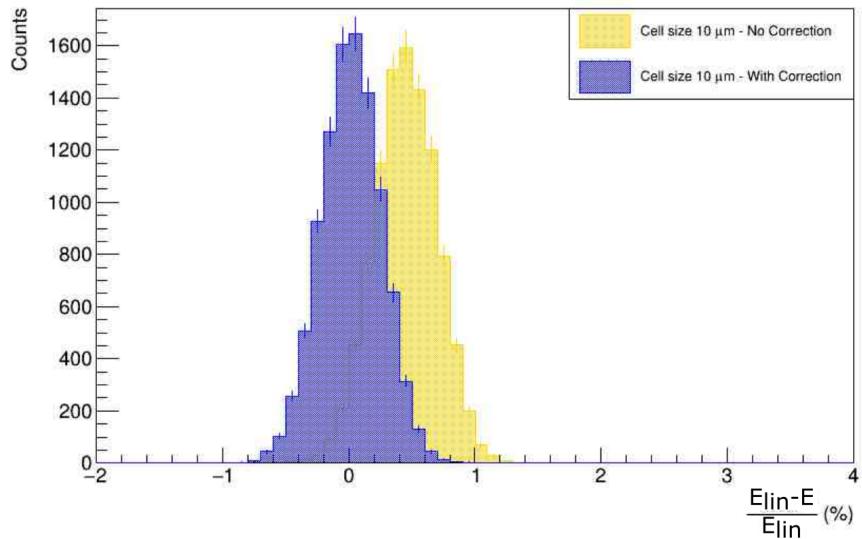


(a) Cherenkov signals.

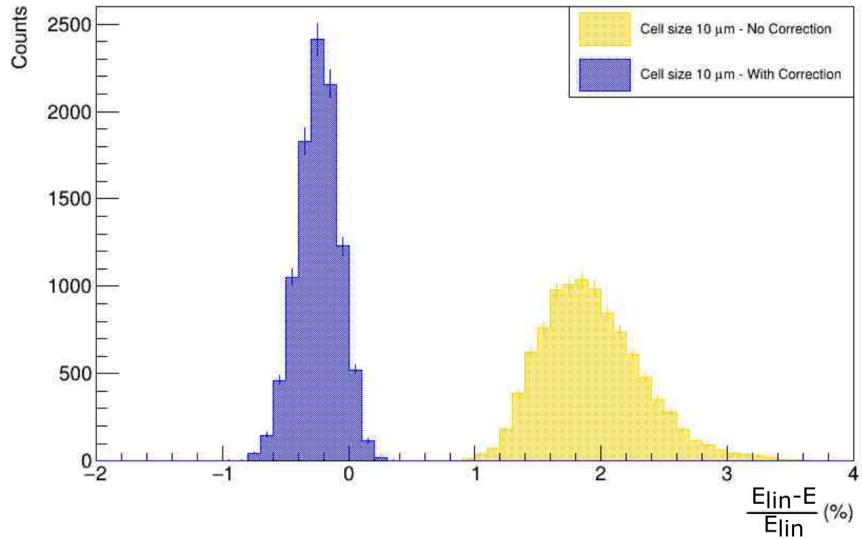


(b) Scintillation signals.

Figure 4.17: Percentage discrepancy distribution considering events with single 40 GeV electrons. Different colours correspond to different cell-size values.

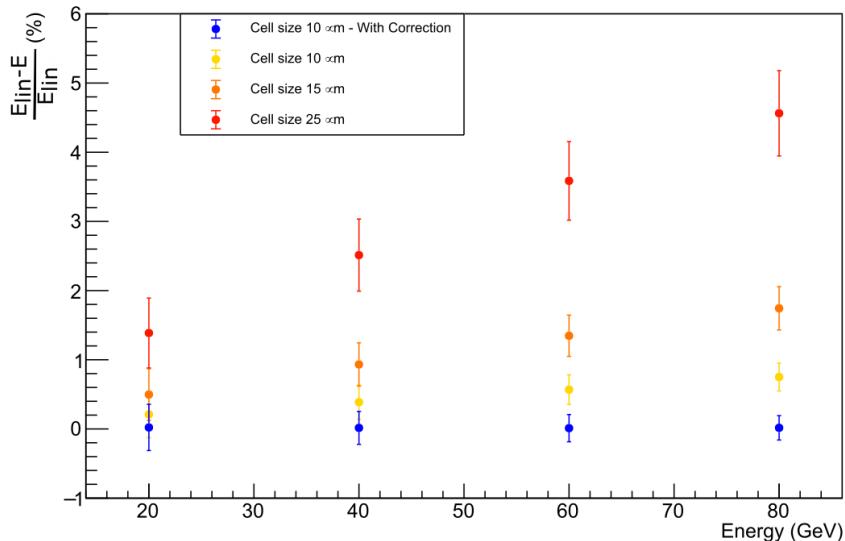


(a) Cherenkov signals.

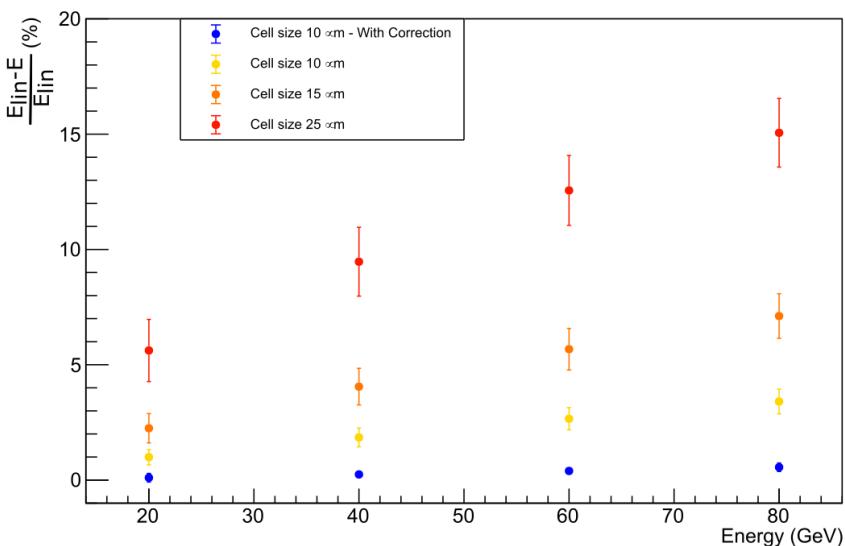


(b) Scintillation signals.

Figure 4.18: Percentage discrepancy distribution considering events with single 40 GeV electrons to evaluate the effect of the analytical correction applied on data obtained with a cell size of 10  $\mu\text{m}$ .



(a) Cherenkov signals.



(b) Scintillation signals.

Figure 4.19: Percentage discrepancy behaviour with respect to the primary particle energy. The results for the different cell sizes are shown, also compared with the results obtained performing the analytical correction to the  $10 \mu\text{m}$  cell-size data.

(post SiPM digitisation simulation). The two types of distribution are compared in Figure 4.20.

The energy distributions obtained from the charge integrals are slightly wider, as expected due to the introduction of the electronic noise. However, the effect is really minimal considering the fact that, in each event, hundreds of SiPMs are active.

The energy distributions have been fitted with a Gaussian function to obtain mean, standard deviation and the respective errors. By doing this with different primary electron energies, the energy resolution can be studied. Mean and standard deviation are listed in Table 4.4.

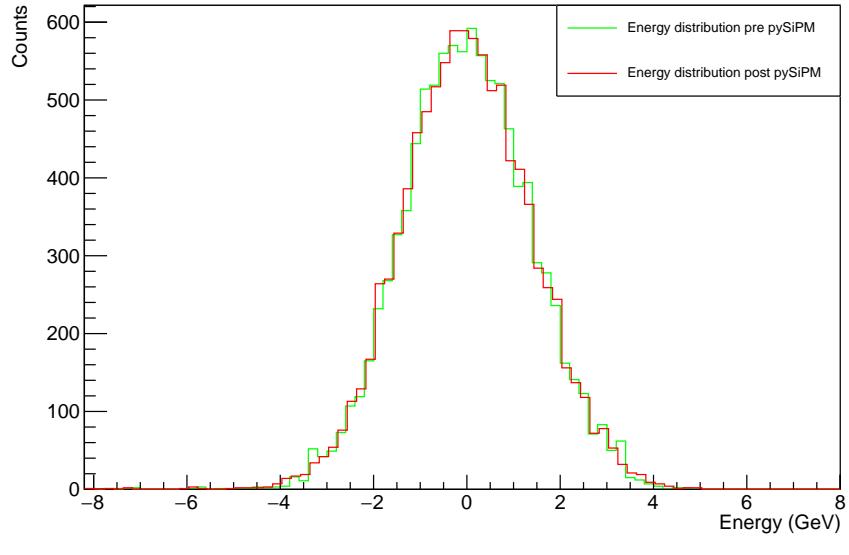
Then, fitting the  $\sigma/E$  values obtained from photoelectron number in the energy range 5 – 80 GeV, it can be seen that the distribution is well described by the function:

$$\frac{\sigma}{E} = \frac{A}{\sqrt{E}} \oplus B, \quad (4.8)$$

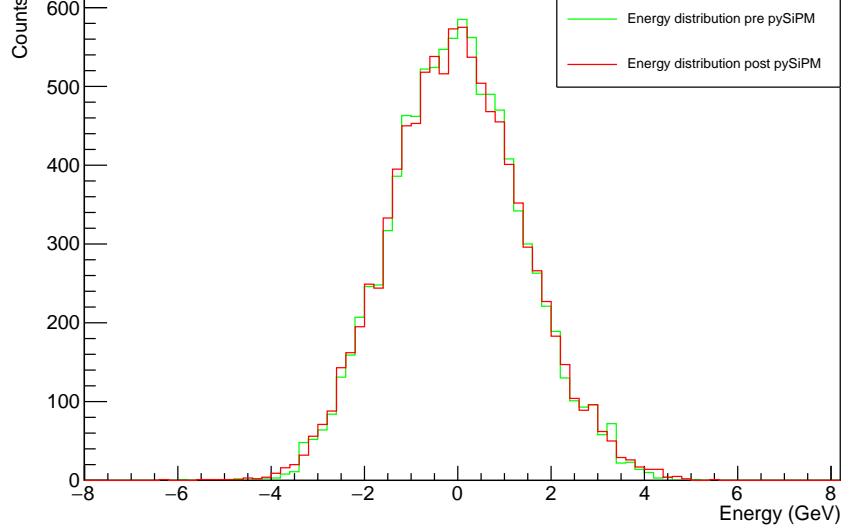
with parameter values of  $A = 18.97\%$  and  $B = 1.677\%$ , for scintillation signals, and  $A = 21.01\%$  and  $B = 0.386\%$ , for Cherenkov signals.

On the other hand, adding the SiPM digitization, an uncorrelated electronics noise contribution scaling as  $C/E$  has to be added in quadrature.

The plots in Figure 4.21 show that the energy resolution points after the SiPM digitization are almost identical with respect to the one obtained from the calorimeter-only simulation data. Therefore the impact on the resolution, due the introduction of the SiPM simulation software, is negligible.



(a) Cherenkov signals.



(b) Scintillation signals.

Figure 4.20: Charge integral distributions generated with data from GEANT4 (*pre pySiPM*) or from the full simulation (*post pySiPM*). The mean values (that, however, have very little differences) have been subtracted.

Pre pySiPM			
	True E (GeV)	Mean E (GeV)	Standard Deviation (GeV)
<b>Scintillation</b>	5	4.89	0.41
	20	19.82	0.92
	40	40.00	1.39
	60	60.28	1.78
	80	80.58	2.16
<b>Cherenkov</b>	5	4.77	0.48
	20	19.76	0.92
	40	40.00	1.36
	60	60.40	1.66
	80	81.86	1.91
Post pySiPM			
	True E (GeV)	Mean E (GeV)	Standard Deviation (GeV)
<b>Scintillation</b>	5	4.91	0.42
	20	19.90	0.92
	40	40.00	1.42
	60	60.65	1.83
	80	81.18	2.22
<b>Cherenkov</b>	5	4.80	0.45
	20	19.88	0.91
	40	40.00	1.36
	60	60.63	1.66
	80	81.12	1.92

Table 4.4: 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 photoelectron number in the top rows and with results from the full simulation in the bottom rows.

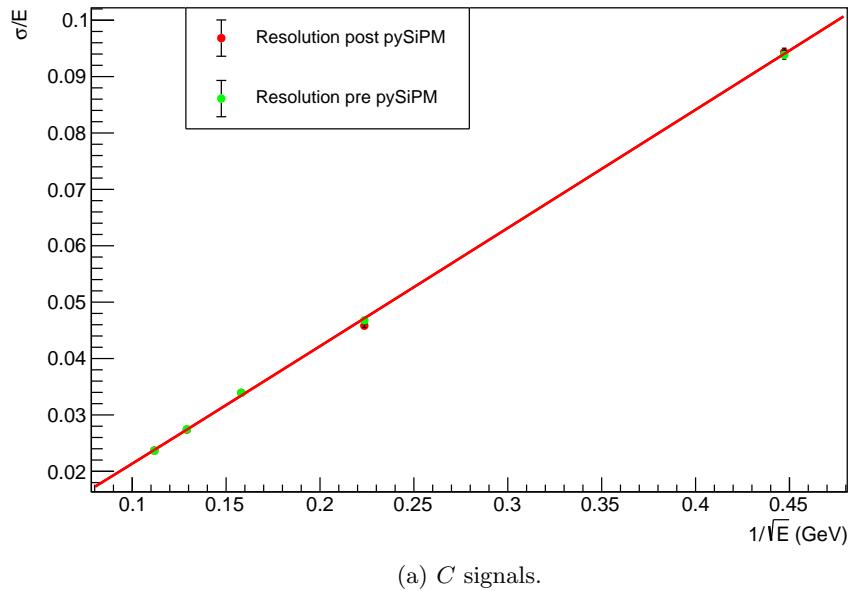
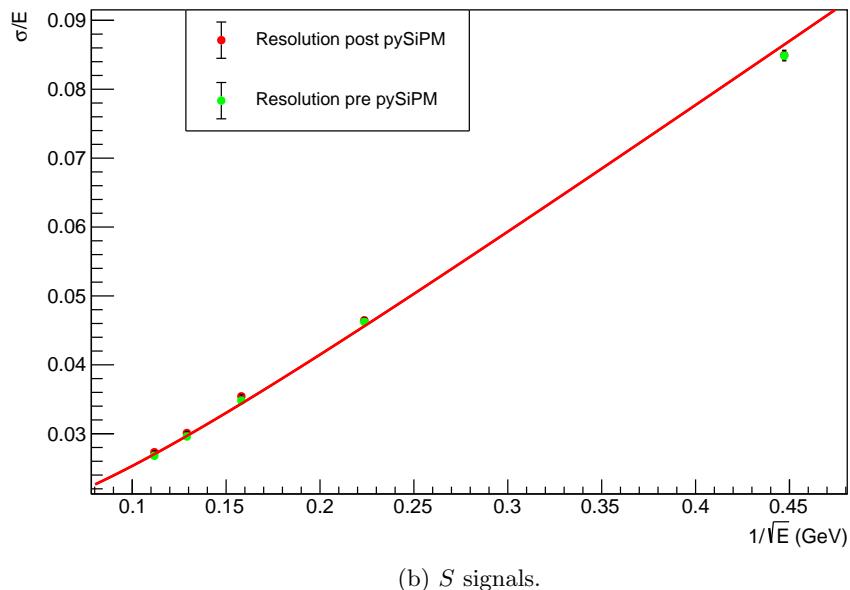
(a)  $C$  signals.(b)  $S$  signals.

Figure 4.21: Calorimeter resolution plot where the data have been fitted with the Formula 4.8.



# Chapter 5

## Deep neural network applications on image processing and particle identification

Machine learning algorithms are powerful and versatile tools that can improve the process of data analysis through a learning process from known training datasets of a neural network that is then used for unknown data categorisation. 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 inside which properly labelled large datasets can be provided, for training, via test beams and simulations.

This chapter describes a possible application of deep neural networks on data generated through the IDEA DR Calorimeter full simulation with the aim to identify particles by analysing the spatial distribution of the deposited energy.

An introduction to the computational techniques is provided, briefly describing the role of each component and the common structures that are used in our study.

Section 5.3 describes how data from the IDEA dual-readout calorimeter can be shaped in order to be passed as inputs in image processing neural networks. After that, the Neural Network (NN) structures are presented in details showing the performance obtained in the training and testing phases, on particle ID tasks.

### 5.1 Physics benchmark

The chosen task is to recognise and distinguish neutral pions and photons by analysing the spatial energy distribution in a fixed area on the calorimeter inner surface.

As already introduced in Paragraph 2.1.1, high-energy 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 close to the shower axis and the remnant energy will be absorbed by the surrounding fibres (an example is shown in Figure 5.1(a)). On the other hand, the  $\pi^0$  meson has a different behaviour. It decays in two main modes:

$$\pi^0 \rightarrow 2\gamma, \quad \pi^0 \rightarrow \gamma e^- e^+ \quad (5.1)$$

that have a very different occurrence probability. Indeed, the branching ratio amounts to  $(98.823 \pm 0.034)\%$ , for the  $2\gamma$  decay, and  $(1.174 \pm 0.035)\%$  for the  $\gamma e^- e^+$  decay [47]. An electromagnetic shower is produced from each of the final-state particles. The result is a superposition of two or three (depending on which decay occurs) em shower almost completely overlapped. In Figure 5.1(b) the data obtained from a  $\pi^0$  event is shown. It should be noted that, in the event represented, the meson decayed into two photons. The two peaks correspond to the core of the two electromagnetic showers.

Usually this ID task would be performed by applying a number of filters to study the shower shape, to find the peaks of charge integral, to measure the distances between the peaks and finally establishing, with a certain probability, the final-state category and the primary-particle type. The goal is to set up a neural network able to accept data associated to such an event and make a prediction on the primary particle in a computationally effective way.

It is important to underline that this task is almost impossible with calorimeters presently operating at colliders due to their much coarser granularity with respect to the IDEA calorimeter.

## 5.2 Neural Networks introduction

Neural Networks are neural-inspired nonlinear models consisting in a group of artificial *neurons* or *nodes* interconnected to each other and grouped in *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 as described in the following.

The simplest neural network structure follows a sequential model where neurons are grouped in *layers* and linked following a sequential order, but more complex neural networks implement also loops, branching and other different flow models between layers. The sequential model presents 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 5.2.

The mathematical representation can be introduced studying a simple two-layer model. Starting from a D-dimension input vector  $\mathbf{x}$ , at the first hidden layer, for each hidden

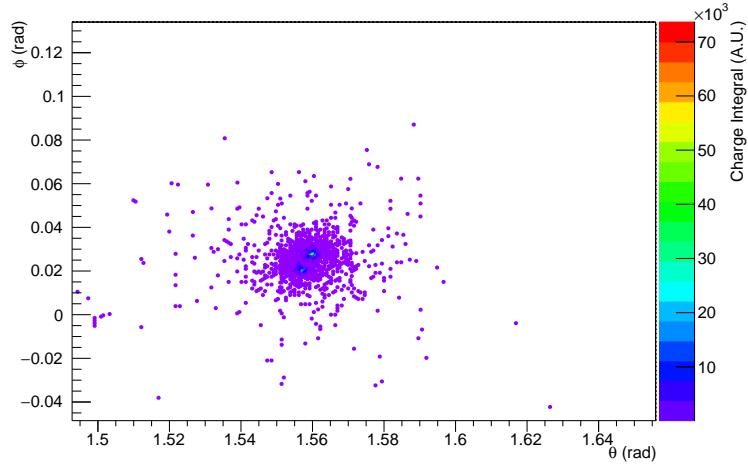
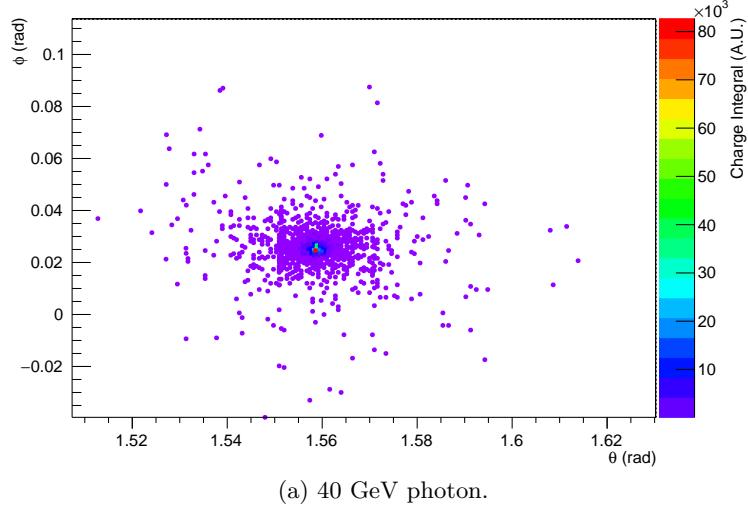


Figure 5.1: Spatial charge integral distributions from  $\gamma$  and  $\pi^0$ . Each point corresponds to an activated scintillating fibre and it is represented in the  $(\theta, \phi)$  spherical coordinates. The colour indicates the charge integral obtained from the coupled SiPM. Results form the IDEA dual-readout full simulation.

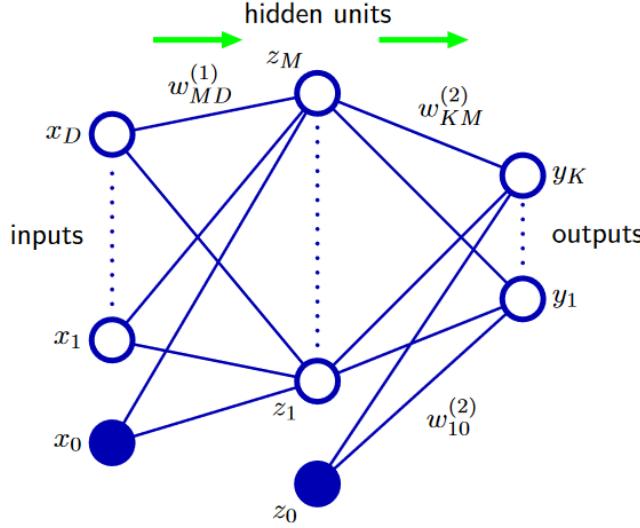


Figure 5.2: Schematic representation of a sequential model. The input, hidden, and output neurons are represented by nodes, and the weight parameters are represented by links between the nodes, for each connection the corresponding bias parameter is denoted by links coming from additional input and hidden variables  $x_0$  and  $z_0$ . Green arrows indicate the direction of flow through the network. Figure from [48].

neuron, a linear combination of the input vector is calculated:

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

where  $j = 1, \dots, M$  ( $M$  is the layer output dimension) and the matrix  $\mathbf{W}$  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:

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

common activation function are *sigmoid*, *tanh*, *ReLU* (rectified linear activation unit) and *Softmax* (for a description see [53]). The obtained value are the hidden-unit neuron activation values and are used as input of the next layer. The process in the second layer is similar, evaluating, for each neuron, the following linear combinations:

$$a_k^{(2)} = \sum_{j=1}^M w_{jk}^{(2)} z_j + w_{k0}^{(2)}, \quad (5.4)$$

where  $k = 1, \dots, K$  ( $K$  is the layer output dimension), and then applying the activation function. In the context of the classification neural networks, the Softmax function ( $S$ )

is a common choice. In fact this activation function normalises the sum of all the output neurons to 1, therefore the values can be associated to the probability of classifying an event with the corresponding label:

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

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

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

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

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

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 effective. In order to do that, a large dataset of correctly labelled input data has to be provided. It is important to divide the dataset in at least two subsets dedicated one to the training process and one to the validation process. This separation is essential to obtain validation results that are not affected by the training process. During the training, the elements in the weight matrix  $\mathbf{W}$  are constantly modified to adapt the output to be as closer as possible to the correct label. These corrections are quantified by an error function that indicates 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):

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

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

The ideal neural network corresponds to the one that satisfy the condition  $t_{nk} = y_{nk}$  for each  $n$  and  $k$ . Being their value bounded to be either 0 or 1 by the target vectors, the minimum cross-entropy value is 0. With a non-ideal NN, the cross-entropy value increases with the worsening of the performance.

The training aim is to correct weights and biases values in order to minimise the 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 advantage of different layer types described in the following.

## 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 ( $\mathbf{v}_i$ ) providing an output vector ( $\mathbf{v}_o$ ) after the inclusion of a bias vector ( $\mathbf{w}_0$ ):

$$\mathbf{v}_o = \mathbf{W}\mathbf{v}_i + \mathbf{w}_0. \quad (5.9)$$

The layer is extremely parameter consuming, let consider a single layer that connects two group of nodes: it is common to use numbers 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, it is often used to graphically represent general neural networks, for example, Figure 5.2 shows a 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 *regularisation* techniques to increase the capability of neural networks in generalising well with new data. The use of a Dropout layer is one of them. The basic idea is to reduce the spurious correlations that could occur between neurons in the network, preventing the *overfitting*, a condition in which the model fits very close the training dataset but may fail in fitting unknown new data. 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 5.3.

## Conv2D layer

A useful layer in image processing is the convolution layer. It identifies a category of NN called Convolutional Neural Networks (CNNs).

A convolution consists 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. It indicates the locations of a detected feature (and its intensity) in an input image.

Figure 5.4 represents the application of a filter or kernel ( $K$ ) over an input matrix ( $I$ ). As it can be seen, the feature map is composed considering all the sub-images with the

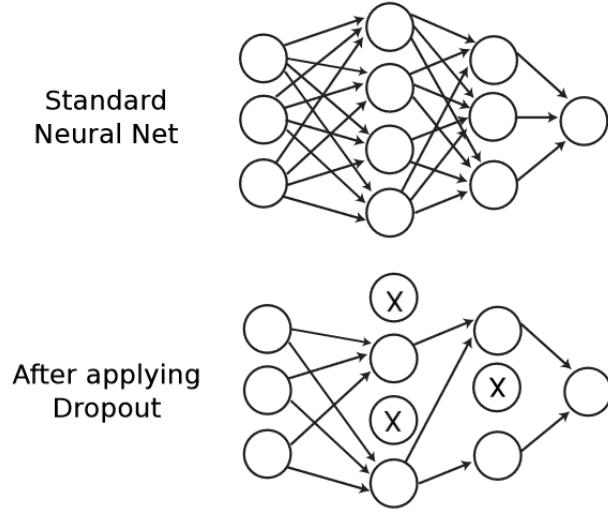


Figure 5.3: Neurons during training are randomly switched off with a probability  $p$ . A lighter network is produced reducing correlations between nodes. Figure from [49].

same size of the kernel; each sub-image is associated to an activation value obtained through the formula:

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

where  $D_K$  is the dimension of the kernel, typical values are  $1 \times 1$ ,  $3 \times 3$ ,  $5 \times 5$ , and  $b$  is the bias value. An activation is obtained from each sub-image and then corrected by the activation function. These outputs compose the feature map.

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  $3 \times 3$  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, a MaxPool2D layer considers a sub-matrix with a fixed size, typically  $2 \times 2$  or  $3 \times 3$ , and records the max value. By performing the process over all the input sub-matrices, the result is a smaller output matrix that keeps the geometrical feature informations. A pictorial representation of the layer is shown in Figure 5.5.

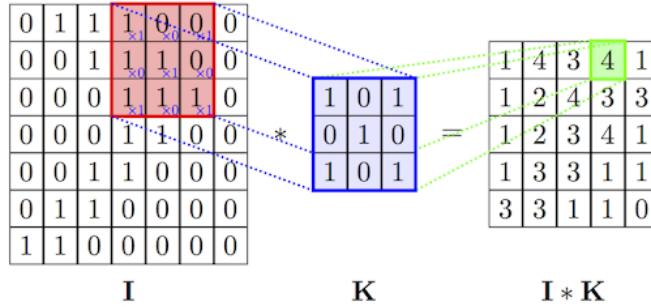


Figure 5.4: Example of convolution applying a single kernel  $K$  on the input matrix  $I$ . The red square is the sub-image considered and convoluted with the blue matrix giving the activation value in the green box. The procedure is done over all the sub-images producing the feature map on the far right.

### VGGNet structure

The VGGNet structure is a neural network concept introduced in 2015 in the article “Very Deep Convolutional Networks for Large-Scale Image Recognition” [51] (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 chains (i.e. kernel size of  $1 \times 1$  or  $3 \times 3$ ) and max pooling with size of  $2 \times 2$  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 a deeper layer has a greater number of kernels starting from, at least, 32.

In Figure 5.6, the structures studied in [51] are listed where groups of two or three convolutional layers are followed by max pooling ones and, at the end, the last max pooling layer output is flattened and used as input for a group of dense layers till the classification.

A VGGNet like structure has been set up and used to perform the task of  $\pi^0/\gamma$  discrim-

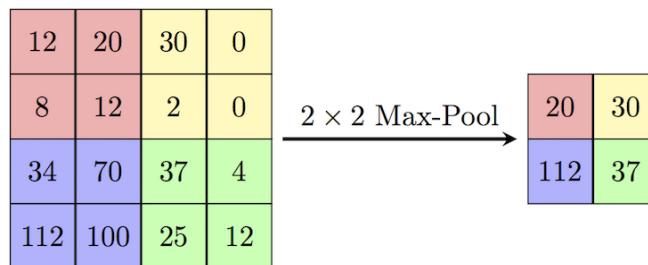


Figure 5.5: MaxPool2D effect sketched where the same colour represents the sub-matrix considered and the corresponding output.

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 <b>LRN</b>	conv3-64 <b>conv3-64</b>	conv3-64 conv3-64	conv3-64 conv3-64	conv3-64 conv3-64
maxpool					
conv3-128	conv3-128	conv3-128 <b>conv3-128</b>	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 <b>conv1-256</b>	conv3-256 conv3-256 <b>conv3-256</b>	conv3-256 conv3-256 <b>conv3-256</b>
maxpool					
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 <b>conv1-512</b>	conv3-512 conv3-512 <b>conv3-512</b>	conv3-512 conv3-512 <b>conv3-512</b>
maxpool					
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 <b>conv1-512</b>	conv3-512 conv3-512 <b>conv3-512</b>	conv3-512 conv3-512 <b>conv3-512</b>
maxpool					
FC-4096					
FC-4096					
FC-1000					
soft-max					

Figure 5.6: 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]”. Image from [51].

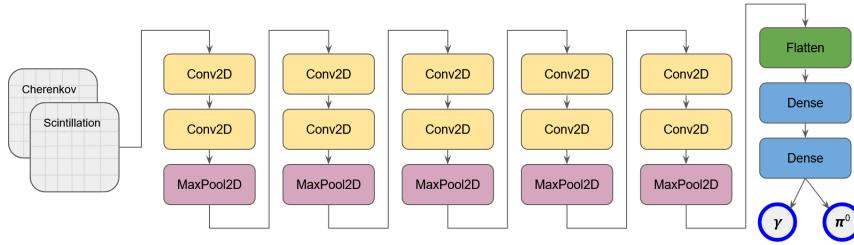


Figure 5.7: Schematic representation of the VGG Network used in our task.

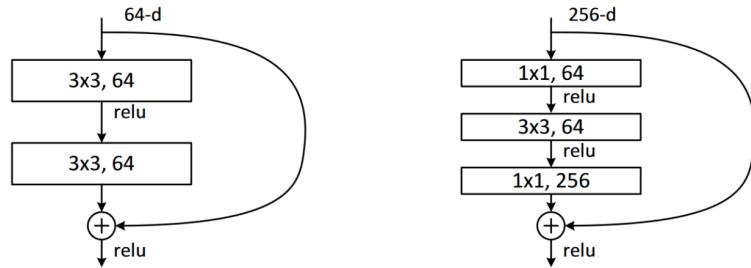


Figure 5.8: Schematic representation of residual blocks with arrows to indicate the data flow. On the left, a simple two convolutional blocks, on the right, a “bottleneck” building block. Figure from [52].

ination and will be described in detail in Paragraph 5.4. A schematic representation of the structure in our context is shown in Figure 5.7.

### ResNet structure

Residual Networks, or ResNet, are another innovative concept of CNN introduced in 2016 in the article “Deep Residual Learning for Image Recognition” by Kaiming He et al. [52].

The structure is composed as a plain convolutional network with small kernel size (same as in VGGNet) and the sequence of convolutional 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; in this way the connection between layers is no more sequential due to the fact that some of the connections skip hidden layers. 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 records the average value and not the max value) and then a structure of consecutive dense layers till the classification one, with Softmax as activation function.

A ResNet like structure has been set up and used to perform the task of  $\pi^0/\gamma$  discrimina-

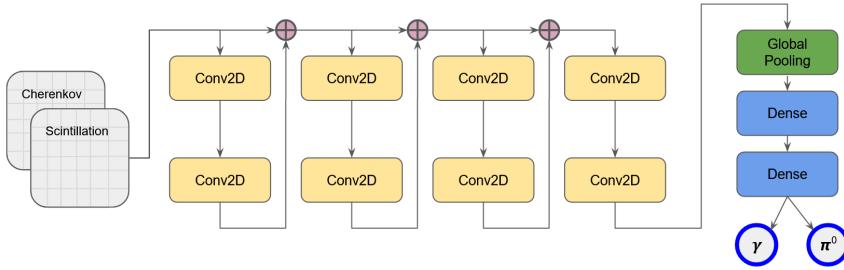


Figure 5.9: Schematic representation of the Residual Network used in our task.

tion and to compare the result with VGGNet. It will be described in detail in Paragraph 5.4. A schematic representation of this structure in our context is shown in Figure 5.9.

### 5.3 Data preparation

The data are produced with the IDEA DR Calorimeter full simulation, neutral pions and photons are fired from the interaction point to tower 1 of the calorimeter. In the first step, photons and neutral pions with fixed energy of 40 GeV are simulated.

The simulation input data used for this study are:

- spatial coordinates of the fibre inner tip ( $x, y, z$ );
- fibre type (Cherenkov or scintillating);
- charge integral from the SiPM digitization software.

The typical coordinate system for a  $4\pi$  calorimeter is the spherical one so the cartesian coordinates have been transformed in spherical ones. Then, at each  $(x, y, z)$  point, the charge integral has been associated to a point in the  $(\theta, \phi)$  space. Finally, the input data are grouped in two subsets by filtering with respect to the fibre types. The results obtained can be plotted in a 3D-graph to visually check the effective correctness of the process. An event display can be seen in Figure 5.10.

To feed the CNNs, the dataset has to be properly prepared. The input for the VGGNet and the ResNet has to be a dimension-fixed matrix. Every event will be characterized by two features (Cherenkov integral and scintillation integral), each one represented by a grid reproducing the spatial distribution of the data. The squared area of interest in the  $(\theta, \phi)$  space has been selected (in radians) as  $(1.51, 1.63)$ , along  $\theta$ , and  $(-0.02, 0.10)$ , along  $\phi$ . Different values of the grid step within this area have been compared by searching a compromise between grid shape efficiency and imaging resolution. A grid step of 0.0009 rad for both axes, corresponding to a single fibre per bin, has been selected. The grid can be represented as a 2D-histogram where the total charge integral in each bin is the bin height. (Figure 5.11 shows the histograms obtained from the same data as in Figure 5.10).

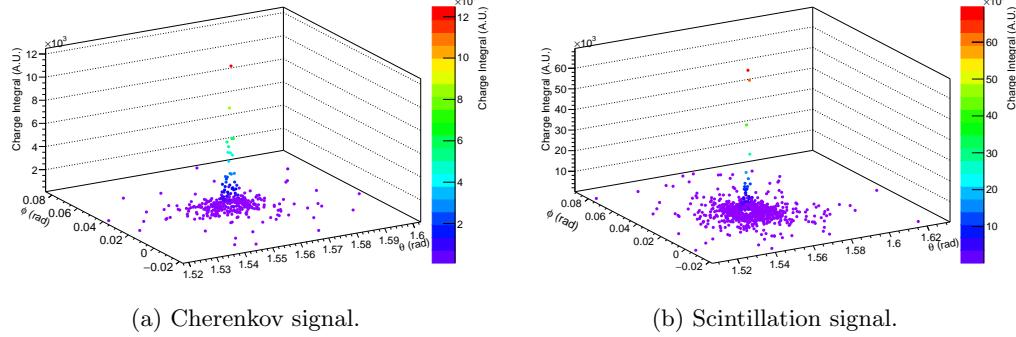


Figure 5.10: 3D-graphs representing Cherenkov and scintillation signals from the same simulated event (with a 40 GeV photon as primary particle).

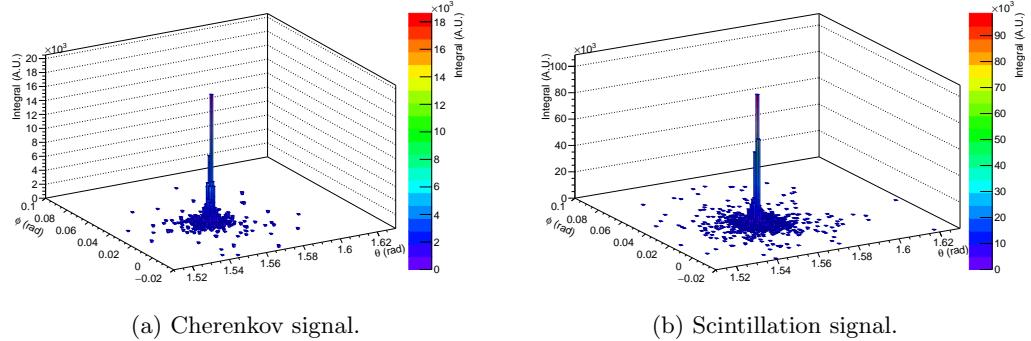


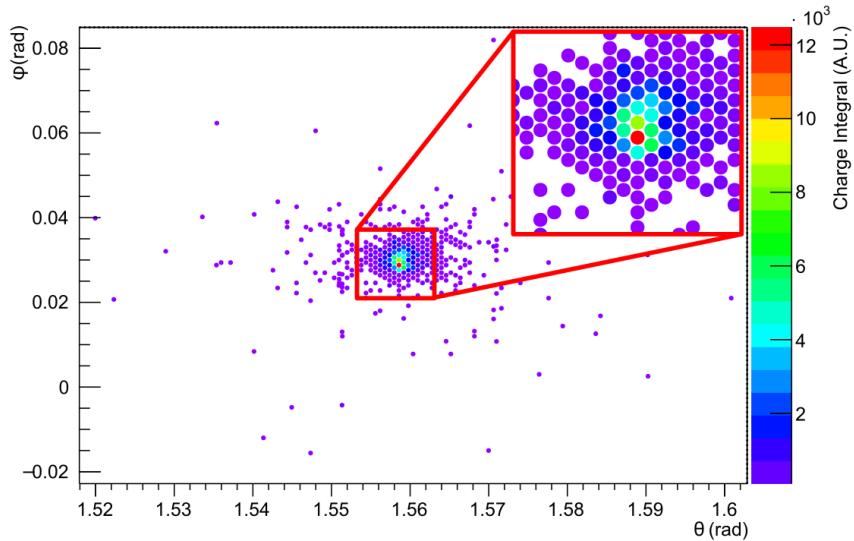
Figure 5.11: 2D-histograms representing Cherenkov and scintillation signals from the same sample event (40 GeV photon as primary particle).

Hence the input matrix for each event has a dimension of 133 (height)  $\times$  133 (width)  $\times$  2 (features). A 2D visualisation can be seen in Figures 5.12, 5.13, 5.14 and 5.15.

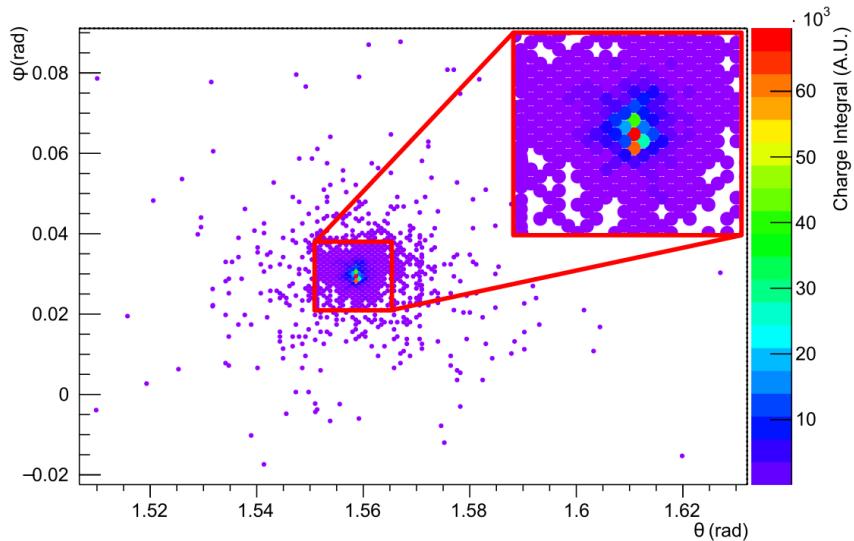
The dataset is made of 10000 events of photons and 10000 events of neutral pions. All the events have been labelled ( $\gamma$  or  $\pi^0$ ), normalized to the max value of 1, shuffled and splitted in two sub-set dedicated to training (80% of the whole statistics) and to validation (20% of the whole statistics).

## 5.4 Performance

Analyses of the performance of the two different convolutional neural network structures have been performed. 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

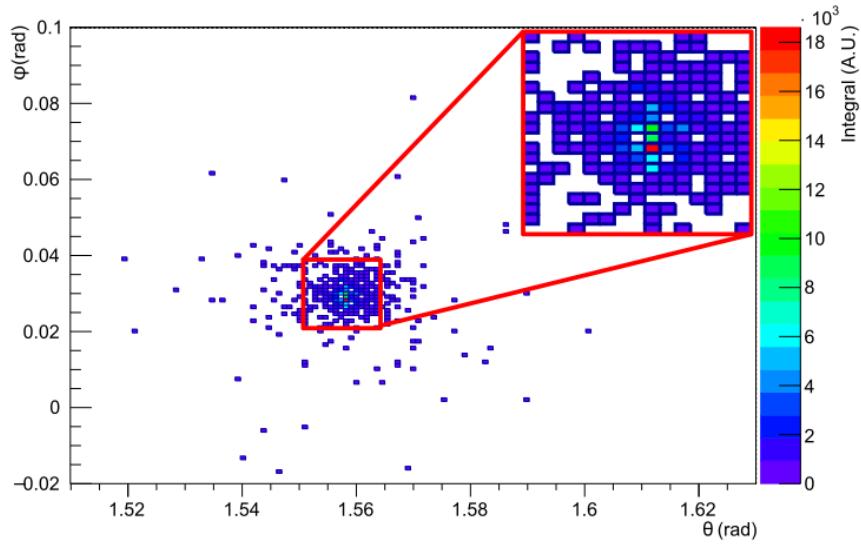


(a) Cherenkov signal pre data preparation.

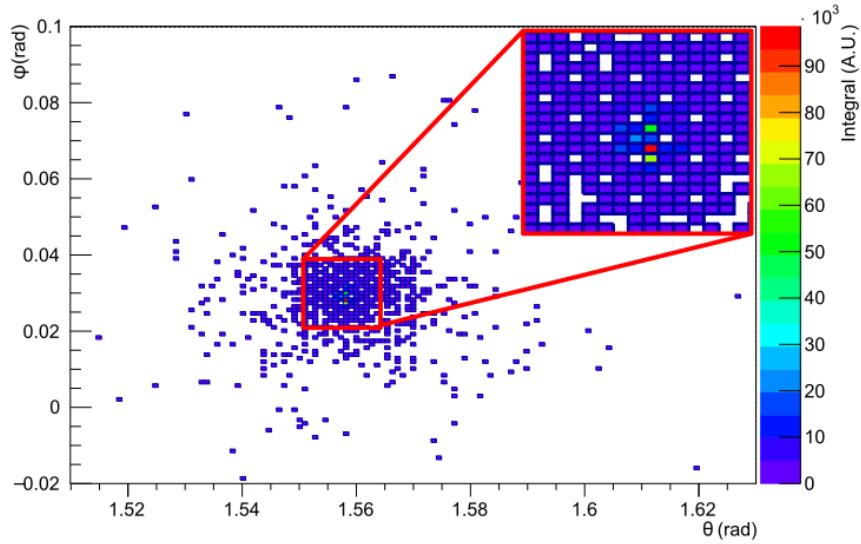


(b) Scintillation signal pre data preparation.

Figure 5.12: 2D-vision of data before data preparation. Data from a 40 GeV photon.



(a) Cherenkov signal post data preparation.



(b) Scintillation signal post data preparation.

Figure 5.13: 2D-vision of data after data preparation. Data from a 40 GeV photon.

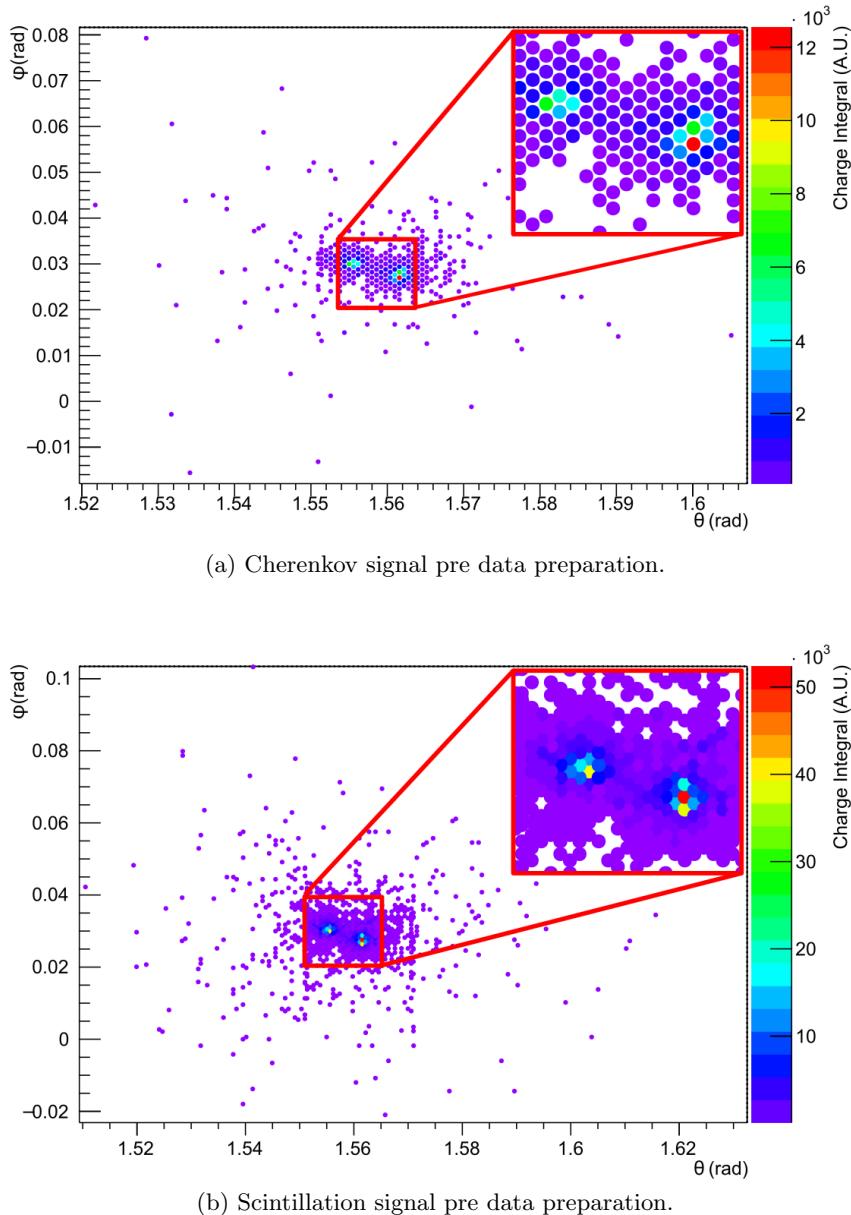
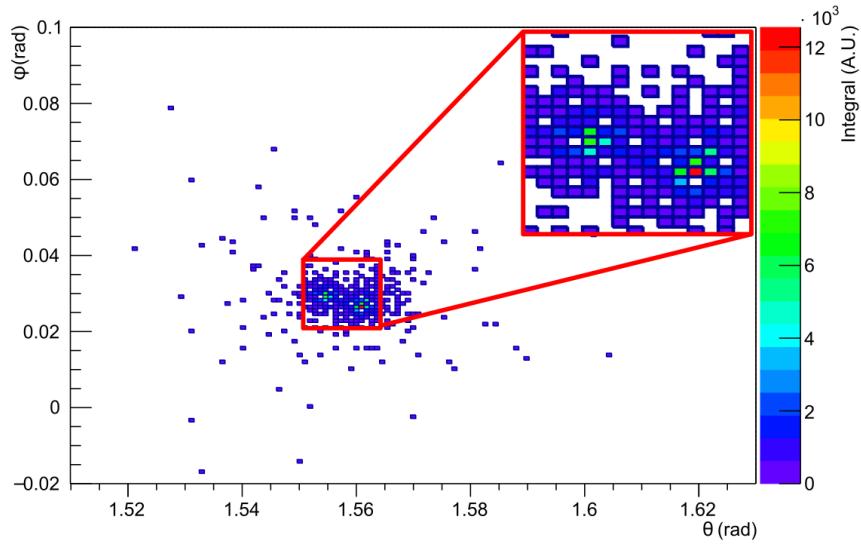
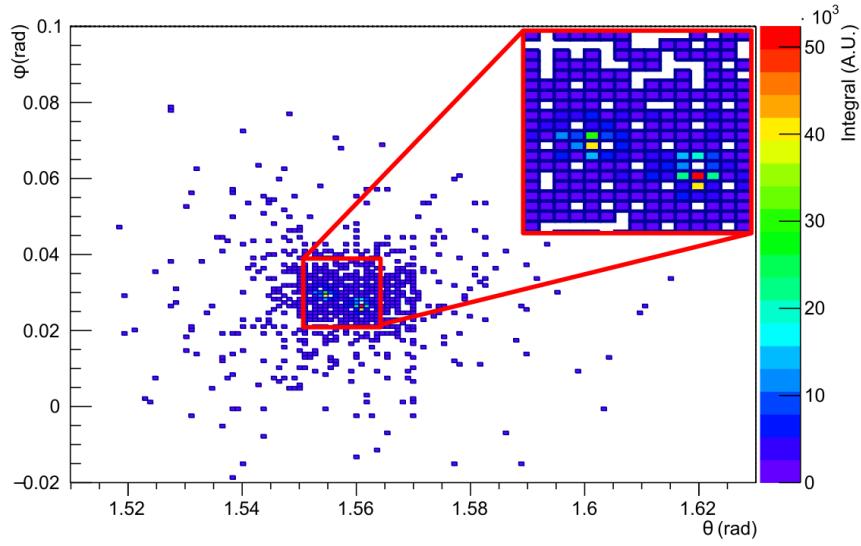


Figure 5.14: 2D-vision of data before data preparation. Data from a 40 GeV pion.



(a) Cherenkov signal post data preparation.



(b) Scintillation signal post data preparation.

Figure 5.15: 2D-vision of data after data preparation. Data from a 40 GeV pion.

number of filters and the kernel size in convolutional layers, the number of neurons in dense layers, have been modified to identify the optimal set of values.

This section shows the comparison between CNNs with different hyperparameters and details of the best performance obtained with the VGGNet and the ResNet.

#### 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 input is a 3D matrix with dimensions of  $133 \times 133 \times 2$ . The aim of distinguishing between two types of primary particle sets the output layer as a dense layer with two neurons, with Softmax as the activation function.

The optimisation 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 analysed one at a time by keeping fixed the other one and changing the hyperparameters within common value ranges.

In Table 5.1, the structures of the six different VGGNet tested are shown together with the accuracy value and the training time needed to evaluate the performance. The training process have been performed with 10 threads at a clock speed of 5 GHz.

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 VGGNet D, E and F have been tested to select the best hyperparameters in the convolution half. The conclusion from this study identifies VGGNet D as the best structure, with an accuracy of 98.875% and a training time of 212 s/epoch.

Further studies on the performance have been done using VGGNet D. The training was done with a training dataset and validated at each epoch on a validation dataset. The accuracy was evaluated as the number of correct predictions over the total number of predictions ( $n_c/N$ ). Figure 5.16 shows the behaviour of the accuracy during the training process and compares the results for the training and the validation datasets.

The loss function is another important feature that can be studied. It has been already introduced as the error function 5.8. A loss value can be evaluated at the end of each training epoch to monitor the improvements in each training step. Figure 5.17 shows its behaviour as a function of the training epoches.

The smoothness of both accuracy and loss plots is an indication of the good dataset size; the graphs would typically show spikes, if a too small dataset were used.

Results on classification problems 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

VGGNet A	VGGNet B	VGGNet C	VGGNet D	VGGNet E	VGGNet F
Input (133 × 133 × 2)					
Conv2D (32, 3 × 3,ReLU)					
Conv2D (32, 3 × 3,ReLU)					
MaxPool (2x2)					
Conv2D (64, 3 × 3,ReLU)					
Conv2D (64, 3 × 3,ReLU)					
MaxPool (2x2)					
Conv2D (128, 3 × 3,ReLU)					
Conv2D (128, 3 × 3,ReLU)					
MaxPool (2x2)					
Conv2D (128, 3 × 3,ReLU)					
Conv2D (128, 3 × 3,ReLU)					
MaxPool (2x2)					
Conv2D (128, 3 × 3,ReLU)					
Conv2D (128, 3 × 3,ReLU)					
MaxPool (2x2)					
Flatten	Flatten	Flatten	Flatten	Flatten	Flatten
Dense(256,ReLU)	Dense(256,ReLU)	Dense(256,ReLU)	Dense(256,ReLU)	Dense(256,ReLU)	Dense(256,ReLU)
Dropout(0.3)	Dropout(0.5)	Dropout(0.5)	Dropout(0.5)	Dropout(0.5)	Dropout(0.5)
Dense(256,ReLU)	Dense(256,ReLU)	Dense(256,ReLU)	Dense(256,ReLU)	Dense(256,ReLU)	Dense(256,ReLU)
Dropout(0.5)	Dropout(0.5)	Dropout(0.5)	Dropout(0.5)	Dropout(0.5)	Dropout(0.5)
Output(2,SoftMax)	Output(2,SoftMax)	Output(2,SoftMax)	Output(2,SoftMax)	Output(2,SoftMax)	Output(2,SoftMax)
Acc.	97.275%	97.650%	97.660%	98.875%	97.299%
Time	230 s/epoch	229 s/epoch	228 s/epoch	170 s/epoch	240 s/epoch

Table 5.1: Table listing the different VGGNet structures studied. Performance is evaluated by looking at the accuracy and the training time.

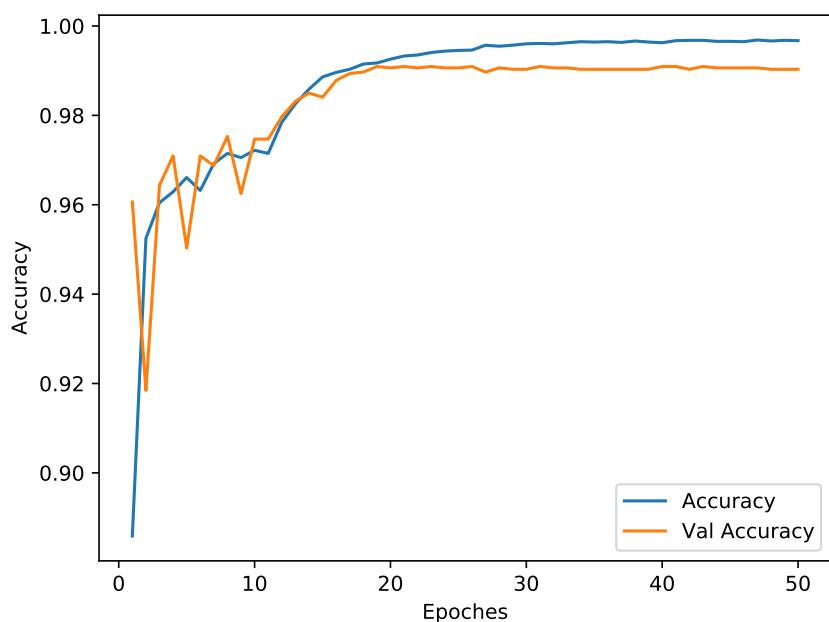


Figure 5.16: Accuracy behaviour over 50 training epoches using the selected VGGNet. Train and validation accuracy are shown with different colours. Results for a 40 GeV  $\pi^0/\gamma$  events simulated with the IDEA calorimeter simulation chain.

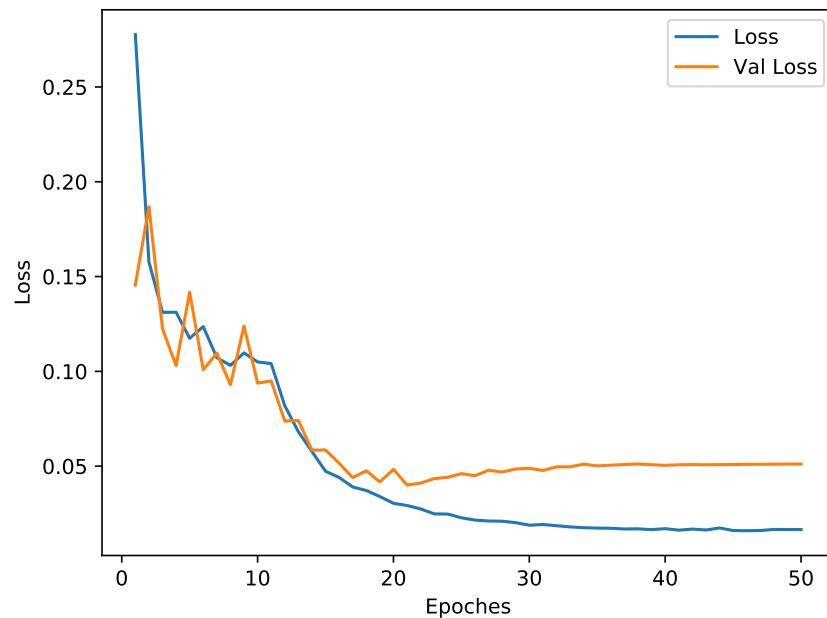


Figure 5.17: Loss behaviour over 50 training epoches using the selected VGGNet. Train and validation losses are shown with different colors. Results for a 40 GeV  $\pi^0/\gamma$  events simulated with the IDEA calorimeter simulation chain.

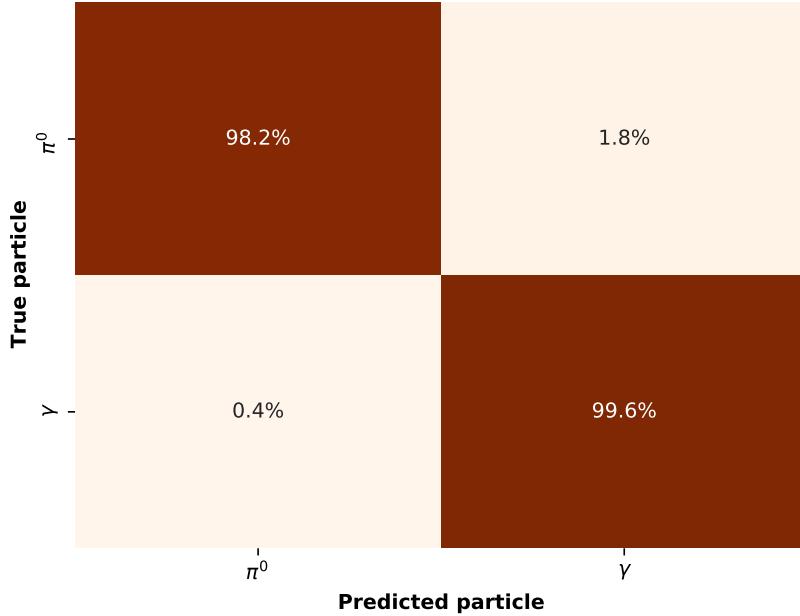


Figure 5.18: Confusion matrix obtained with test data using the selected VGGNet. Values are normalized on rows (i.e. on the true label). Results for a 40 GeV  $\pi^0/\gamma$  events simulated with the IDEA calorimeter simulation chain.

obtained with the selected VGGNet (trained for 50 epoches) is shown in Figure 5.18.

#### 5.4.2 Residual Network

The residual neural networks used for the particle ID task follows the structure described in Section 5.2.

The input and output of the ResNets has to coincide with the ones for VGGNets, 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 a two-neurons dense layer with the Softmax activation function, corresponding to the probability of identifying a photon or a neutral pion.

As done for the VGGNet optimization, also this structure has been divided in two parts identified by the *GlobalAveragePooling* layer (“convolutional half” and “dense half”). In Table 5.2 the six different ResNet with different hyperparameters are listed. ResNet A, B, C and D are dedicated to the optimisation 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 convolutional 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.

Table 5.2: Table of different ResNet structure studied. The performance is evaluated by looking at the accuracy and training time values. Note that the thicker lines divide the different Recurrent blocks as described in Section 5.2.

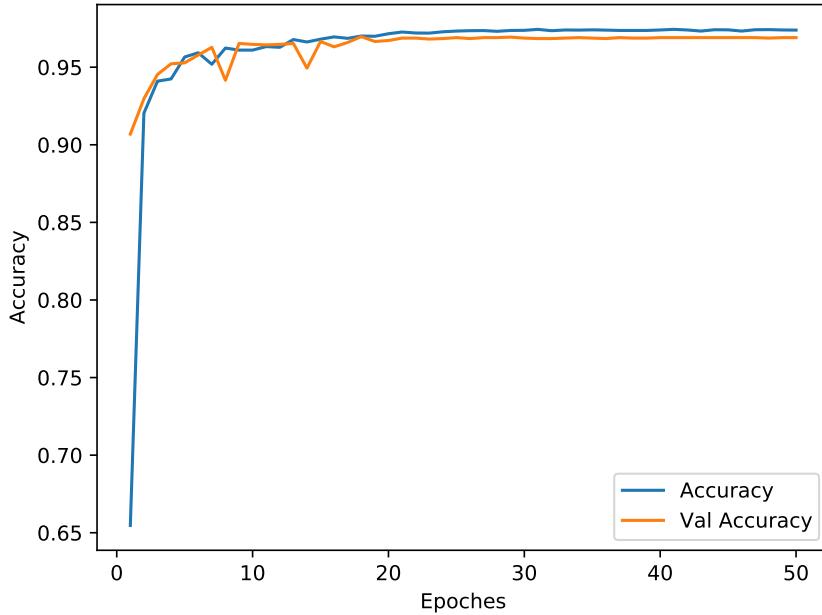


Figure 5.19: Accuracy behaviour over 50 training epoches using the selected ResNet. Train and validation accuracy are shown with different colours. Results for a 40 GeV  $\pi^0/\gamma$  events simulated with the IDEA calorimeter simulation chain.

The Residual Network performance, as the VGG Network one, has been studied under the same conditions such as batch size (128 events) and number of training epoches (50), using also the same accuracy and loss evaluation functions.

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 5.19 and 5.20. Also in this case a confusion matrix has been produced on validation data reaching the results shown in Figure 5.21.

## 5.5 Energy range extension

Once the NN performances have been studied in the simplified case of photons and pions of 40 GeV energy, the best structures (i.e. VGGNet D and ResNet D) have also been used to discriminate  $\pi^0$  and  $\gamma$  with energies ranging from 1 to 80 GeV. This range will likely cover the one at future circular electroweak colliders where a particle, with unknown energy, interacts with the calorimeter and the neural network has to distinguish if the particle is a  $\gamma$  or a superposition of  $2\gamma$ 's from a  $\pi^0$  decay. This task is only possible thanks to the high-granularity of the IDEA calorimeter. Standard today's calorimeters do not have the granularity required to perform this type of discrimination.

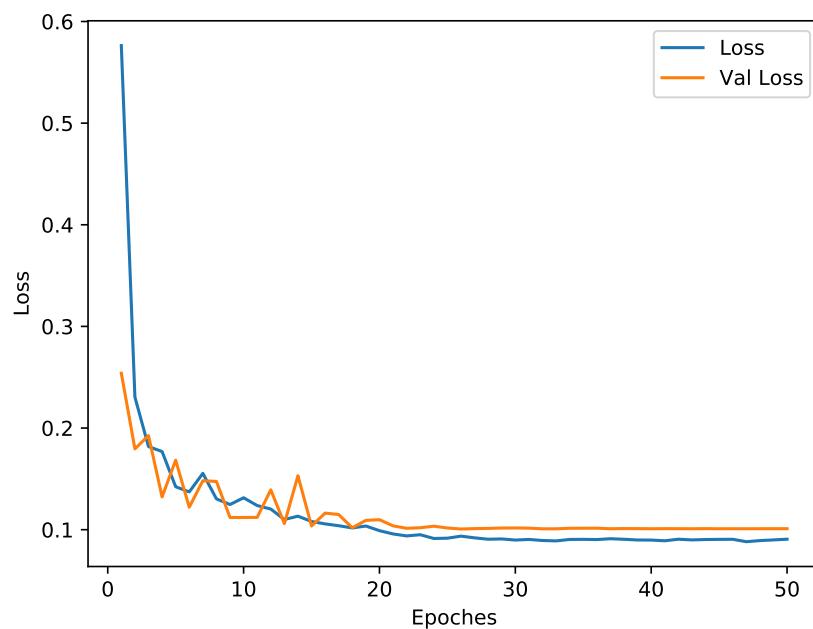


Figure 5.20: Loss behaviour over 50 training epoches using the selected ResNet. Train and validation loss are shown with different colours. Results for a 40 GeV  $\pi^0/\gamma$  events simulated with the IDEA calorimeter simulation chain.

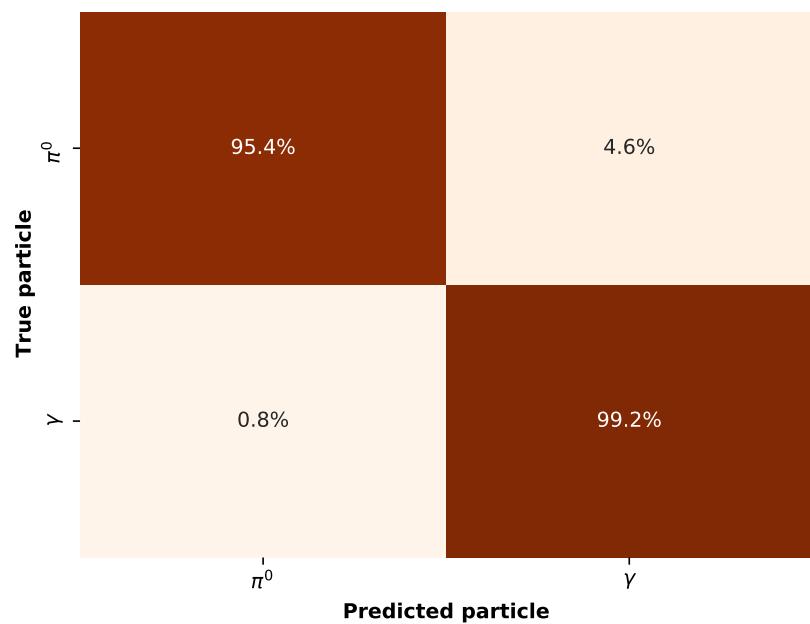


Figure 5.21: Confusion matrix obtained with test data using the selected ResNet. Values are normalized on rows (i.e. on the true label). Results for 40 GeV  $\pi^0/\gamma$  events simulated with the IDEA calorimeter simulation chain.

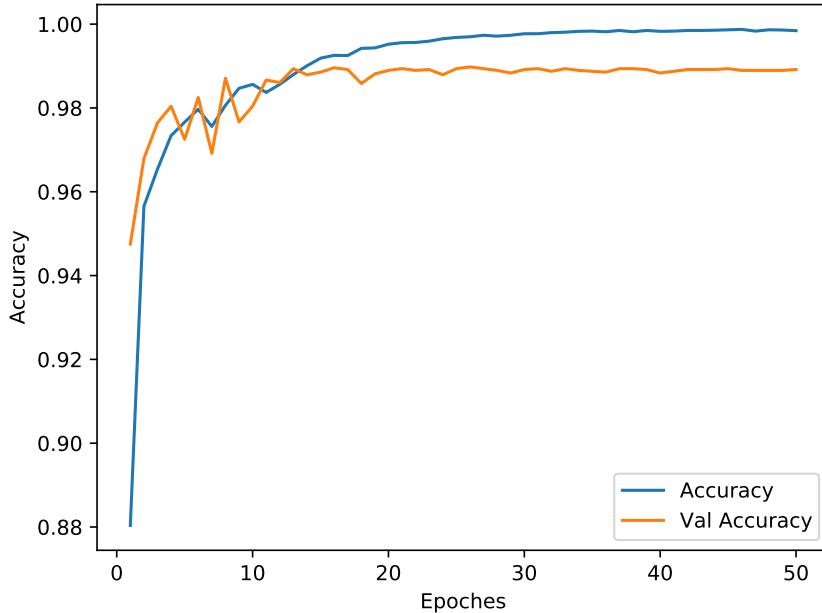


Figure 5.22: Accuracy behaviour over 50 training epoches using the selected VGGNet on events with variable energies. Train and validation accuracy are shown with different colours. Results for  $1 - 80$  GeV  $\pi^0/\gamma$  events simulated with the IDEA calorimeter simulation chain.

To perform this study, a set of 15000 events for each particle type has been produced, in which the energy is uniformly distributed in the considered range. In this study we neglected the  $\pi^0 \rightarrow e^+ e^- \gamma$  process. The dataset production and preparation have been done following the same procedure previously described.

Once again the two neural networks have been trained over the 80% of the data, reaching similar results in accuracy and loss performance (see Figures 5.22, 5.23, 5.24 and 5.25). The validation process on the remaining 20% of the dataset has provided the confusion matrices shown in Figure 5.26.

## ROC and AUC

A way to evaluate in more detail the correctness of a neural network in a classification task is the production of the Receiver Operating Characteristic curve (ROC curve) and the Area Under the ROC Curve (AUC).

Starting from the idea that the NN goal is to identify a neutral pion signal in a set of data where photon signals are also present, the focus has to be placed on the output

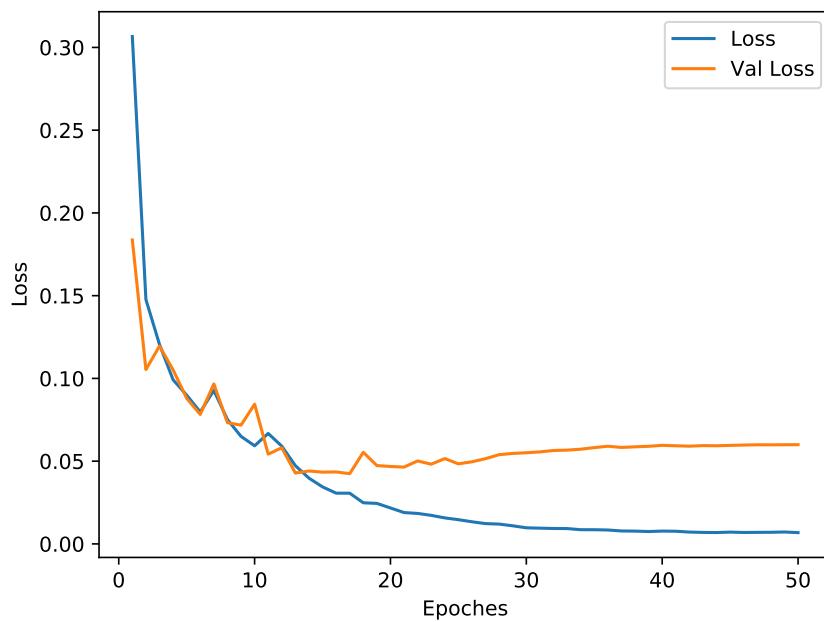


Figure 5.23: Loss behaviour over 50 training epoches using the selected VGGNet on events with variable energies. Train and validation accuracy are shown with different colours. Results for  $1 - 80$  GeV  $\pi^0/\gamma$  events simulated with the IDEA calorimeter simulation chain.

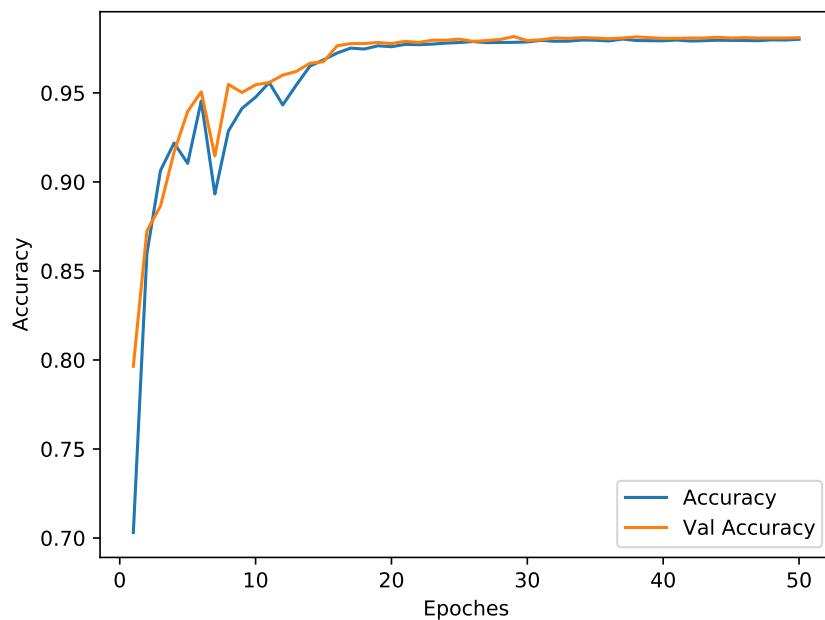


Figure 5.24: Accuracy behaviour over 50 training epoches using the selected ResNet on events with variable energies. Train and validation accuracy are shown with different colours. Results for  $1 - 80$  GeV  $\pi^0/\gamma$  events simulated with the IDEA calorimeter simulation chain.

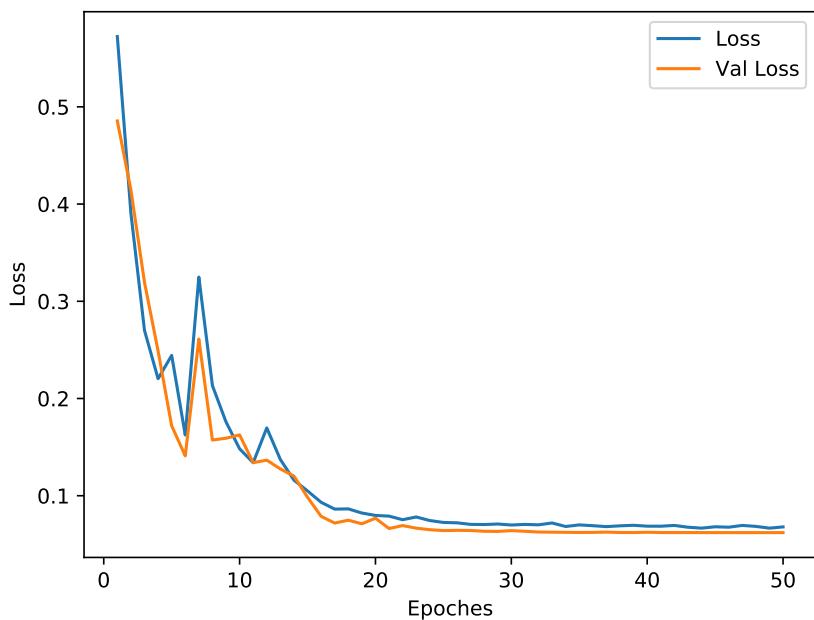
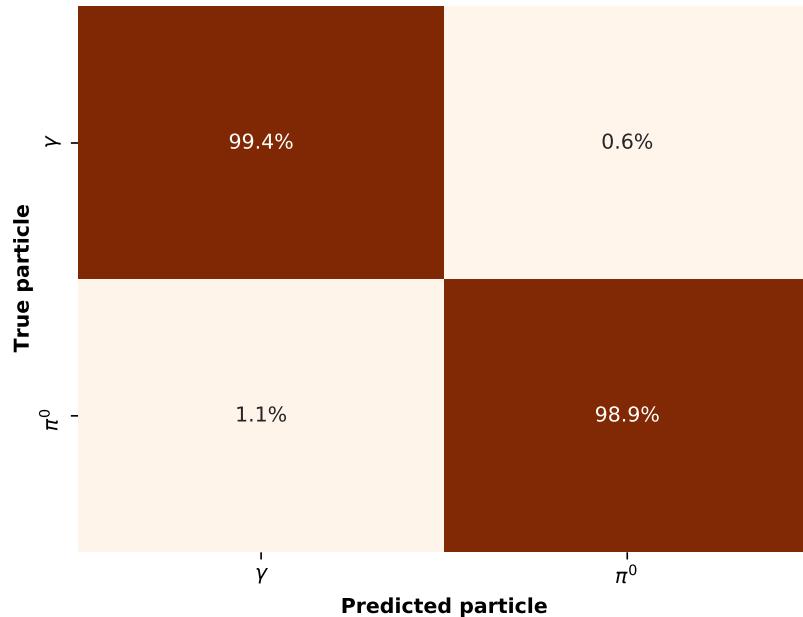
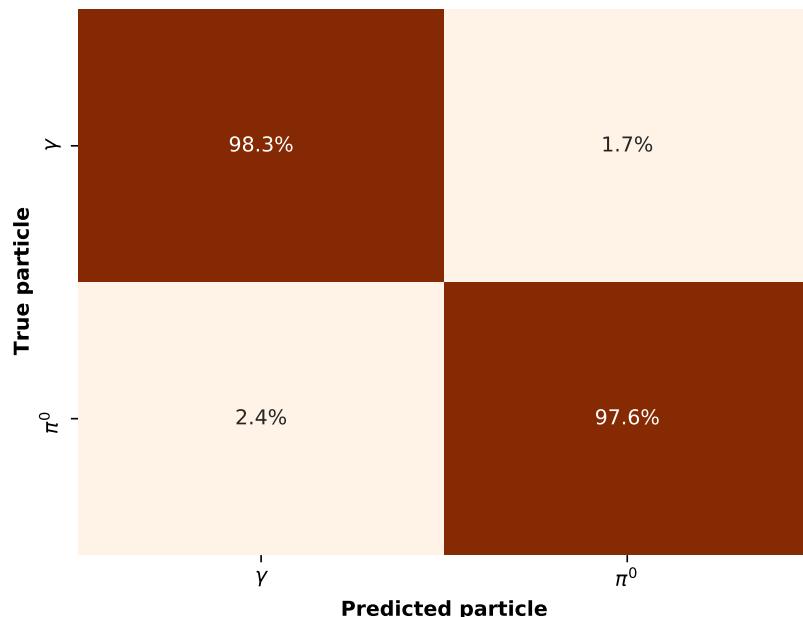


Figure 5.25: Loss behaviour over 50 training epoches using the selected ResNet on events with variable energies. Train and validation accuracy are shown with different colours. Results for  $1 - 80$  GeV  $\pi^0/\gamma$  events simulated with the IDEA calorimeter simulation chain.



(a) VGGNet structure.



(b) ResNet structure.

Figure 5.26: The confusion matrices obtained in the validation process for data produced by  $\gamma$  and  $\pi^0$  with energies in the range 1 – 80 GeV.

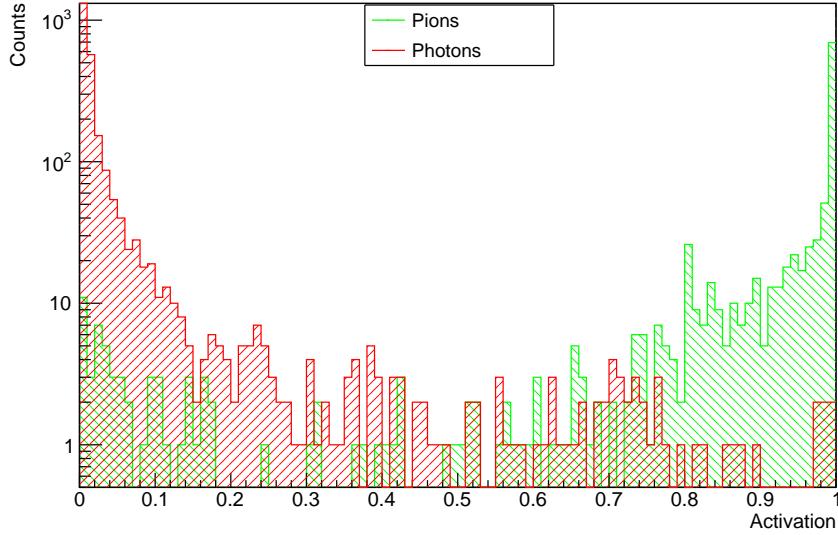


Figure 5.27: Distributions of the  $\pi^0$  neuron activation value for the VGGNet, divided in pion and photon events. Results for  $1 - 80$  GeV  $\pi^0/\gamma$  events simulated with the IDEA calorimeter simulation chain.

neuron associated to the  $\pi^0$ .

A relevant plot, as the one shown in Figure 5.27 for the VGGNet, is the distribution of activation values associated to the neuron of interest, dividing them in two groups depending on the true particle producing the data. In case of data from  $\gamma$ s, the closer to 0 is the pion activation value, the more correct is the NN prediction. On the other hand, in case of data from  $\pi^0$ s, the closer to 1 is the pion activation value, the more correct is the NN prediction.

A neural network shows perfect classification capabilities when the two distribution are perfectly separated, one all below the 0.5 activation value and the other one all above.

In 2-classes classification tasks, the threshold to consider a neuron activated is set at 0.5 as default. Changing this value, different confusion matrices are produced, and the ROC curve is a practical way to summarise these matrices by changing the threshold. The ROC curve is obtained by plotting a point for each threshold value assigning as  $x$  coordinate the efficiency in detecting a pion (i.e. the fraction of neutral pions correctly recognised) and as  $y$  coordinate the rejection of photon events (i.e. the fraction of photons that are recognised as photons).

The ideal ROC curve is the one passing on the point  $(1, 1)$  that corresponds to the perfect efficiency and rejection conditions.

The area under this curve is called AUC and represents a numerical value that easily allows to compare different NN or conditions (the ideal AUC value is 1).

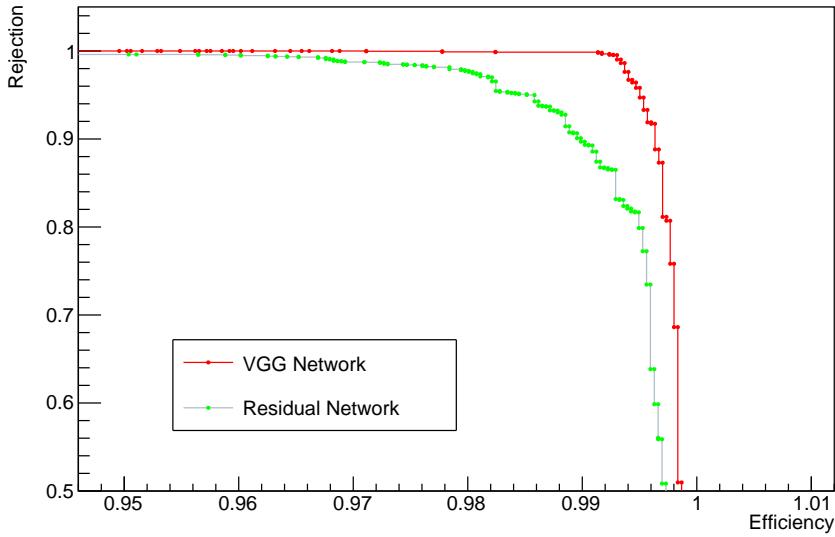


Figure 5.28: ROC curve comparison for the two neural networks. Results for  $1 - 80$  GeV  $\pi^0/\gamma$  events simulated with the IDEA calorimeter simulation chain.

The ROC curves obtained from the selected VGGNet and ResNet are shown in Figure 5.28 and the AUC values are 0.9982 and 0.9957 respectively.

ROC curves are also produced by dividing the validation dataset in sub-samples depending on the energies. As the graphs 5.29 and 5.30 show, in the range  $1 - 80$  GeV the performance of the two neural networks are almost energy independent.

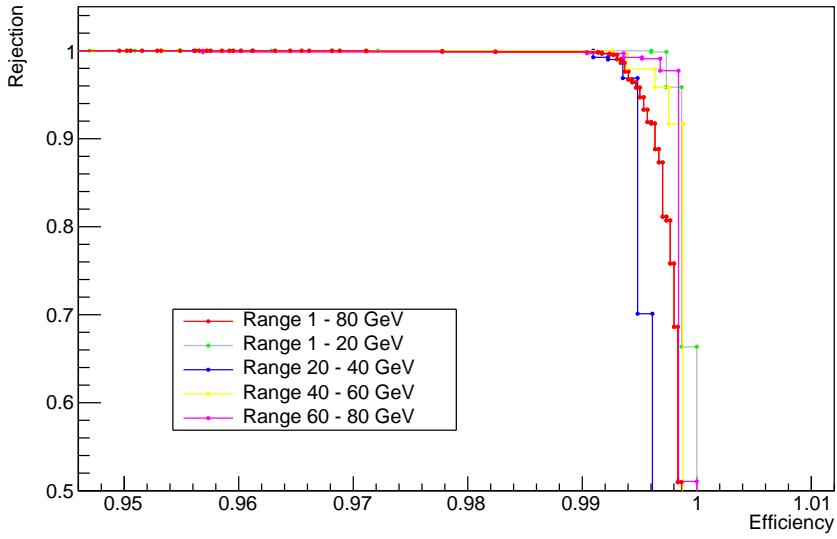


Figure 5.29: ROC curves for the selected VGG Network over subsamples of events with different energies, as described in the legend. Results for  $1 - 80$  GeV  $\pi^0/\gamma$  events simulated with the IDEA calorimeter simulation chain.

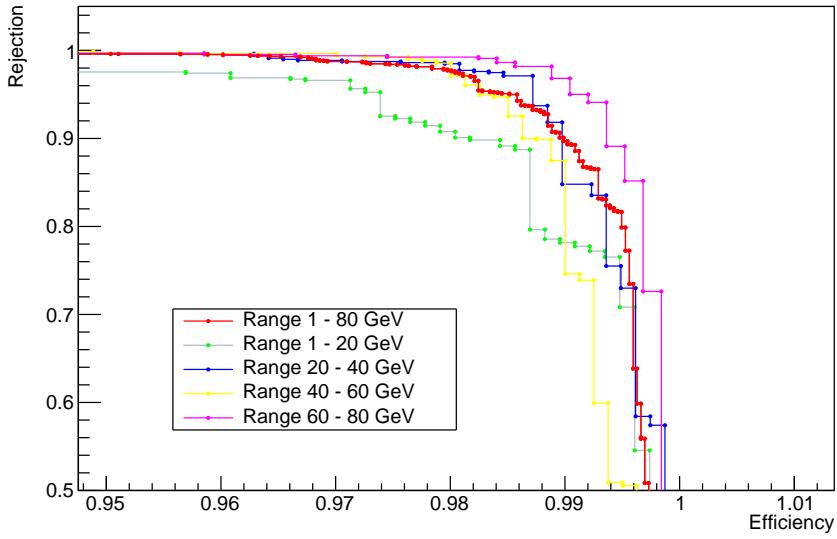


Figure 5.30: ROC curves for the selected Residual Network over subsamples of events with different energies, as described in the legend. Results for  $1 - 80$  GeV  $\pi^0/\gamma$  events simulated with the IDEA calorimeter simulation chain.



# Conclusion

The thesis describes my personal contribution in the context of the IDEA experiment for future  $e^+e^-$  electroweak factories and its dual-readout calorimeter simulation.

The main goals in modern particle physics have been described focusing on the studies on the Higgs boson. A picture of the main future colliders has been provided, pointing out their capabilities and limitations. The IDEA detector concept has been outlined in all its parts, being one of the multi-purpose detector proposed for future  $e^+e^-$  electroweak factories.

The second chapter has detailed the theoretical background on calorimetry, from the basic concepts to more modern techniques. The focus has been placed on the dual-readout compensation technique, cornerstone of the IDEA calorimeter, explaining its strengths and weaknesses.

The third chapter has been dedicated to Silicon PhotoMultipliers (SiPMs), photo-detectors that represent today's alternative to PhotoMultiplier Tubes for the IDEA dual-readout calorimeter. The working principle of these devices has been described along with their characterisation features as gain, photon detection efficiency, linearity and noise effects.

The last two chapters has described my contributions to the IDEA dual-readout calorimeter software. It has begun with the study of the simulation chain (modelling both the calorimeter response and the SiPM transfer function) to evaluate the performance and validate the full simulation.

In this context, I have shown that the impact of the SiPM digitisation on the temporal distributions behaves as expected. The readout linearity has been studied, showing that a 1% linear detector for electromagnetic shower detection, considering both the calorimeter and the readout system, can be achieved, at least in the energy range studied (20 – 80 GeV).

Eventually, the development of deep learning algorithms has been shown, together with the achievement of excellent results in performing particle identification tasks, like distinguishing between neutral pions and photons on the basis of their spatial shower distribution. Two neural network (NN) structures were considered, a VGG-like NN and a

ResNet-like NN. I have shown that they can achieve excellent performance in identifying  $\pi^0$ 's with accuracy values of 98.9% (VGGNet) and 97.6% (ResNet) while rejecting  $\gamma$ 's with probabilities of 99.4% (VGGNet) and 98.3% (ResNet). These results demonstrate the great potential of deep neural network methods applied on data from the IDEA dual-readout calorimeter. The high granularity of the detector and the capability of deep neural network to effectively handle huge amount of data may provide the key for a significant step forward for future calorimetry at high-energy  $e^+e^-$  colliders.

# Bibliography

- [1] ATLAS Collaboration. «Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC». In: Physics Letters B7161-29(2012). DOI:10.1016/j.physletb.2012.08.020
- [2] CMS Collaboration. «Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC». In: Physics Letters B71630-61(2012). DOI:10.1016/j.physletb.2012.08.021
- [3] European Strategy Group, «Supporting note for briefing book 2020», CERN/ESG/052019
- [4] Future Circular Collider Study. «Volume 1: Physics Opportunities. Conceptual Design Report», preprint edited by M. Mangano et al. CERN accelerator reports, CERN-ACC-2018-0056, Geneva, December 2018. Published in Eur. Phys. J. C
- [5] Future Circular Collider Study. «Volume 2: The Lepton Collider (FCC-ee) Conceptual Design Report», preprint edited by M. Benedikt et al. CERN accelerator reports, CERN-ACC-2018-0057, Geneva, December 2018. Published in Eur. Phys. J. ST
- [6] Future Circular Collider Study. «Volume 3: The Hadron Collider (FCC-hh) Conceptual Design Report», preprint edited by M. Benedikt et al. CERN accelerator reports, CERN-ACC-2018-0058, Geneva, December 2018. Published in Eur. Phys. J. ST
- [7] The CEPC Study Group, «CEPC Conceptual Design Report: Volume 1 - Accelerator», arXiv:1809.00285 [physics.acc-ph]
- [8] The CEPC Study Group, «CEPC Conceptual Design Report: Volume 2 - Physics & Detector», arXiv:1811.10545 [hep-ex]
- [9] The CEPC Accelerator Study Group, «The CEPC input for the European Strategy for Particle Physics - Accelerator», arXiv:1901.03169 [physics.acc-ph]
- [10] J. Brau and et al., «The international linear collider: A global project», 2018
- [11] H. Bauer et al. «The International Linear Collider Technical Design Report - Volume 2: Physics». In: ILC-REPORT-2013-040(2013)

- [12] M. Aicheler et al., «A Multi-TeV Linear Collider Based on CLIC Technology: CLIC Conceptual Design Report», CERN-2012-007, 2012, DOI: 10.5170/CERN-2012-007
- [13] L. Linssen et al., «CLIC Conceptual Design Report: Physics and Detectors at CLIC», CERN-2012-003, 2012, DOI: 10.5170/CERN-2012-003
- [14] P. Lebrun et al., «CLIC Conceptual Design Report: The CLIC Programme: Towards a Staged e+e-Linear Collider Exploring the Terascale», CERN-2012-005, 2012, DOI:10.5170/CERN-2012-005
- [15] CLIC and CLICdp Collaborations. «Updated Baseline for a staged Compact Linear Col-lider». In:CERN-2016-004(2016). DOI:10.5170/CERN-2016-004
- [16] A. Robson et al., «The Compact Linear e+e-Collider (CLIC): Accelerator and Detector», arXiv:1812.07987 [physics.acc-ph]
- [17] «Future Circular Collider Study. Volume 2: The Lepton Collider (FCC-ee) Conceptual Design Report», preprint edited by M. Benedikt et al. CERN accelerator reports, CERN-ACC-2018-0057, Geneva, December 2018. Published in Eur. Phys. J. ST
- [18] The CEPC Study Group, «CEPC Conceptual Design Report: Volume 2 - Physics & Detector», arXiv:1811.10545 [hep-ex]
- [19] M. Adinolfi et al., «The tracking detector of the KLOE Experiment», Nucl. Instr. andMeth. A 488 (2002) 51[33]
- [20] A. M. Baldini et al., «MEG Upgrade Proposal», arXiv:1301.7225 [physics.insdet]
- [21] ALICE Collaboration, «Expression of Interest for an ALICE ITS Upgrade in LS3», ALICE-PUBLIC-2018-013, 2018
- [22] M. Livan, V. Vercesi, R. Wigmans, «Scintillating-fibre Calorimetry», CERN Yellow Report, CERN 95-02
- [23] S. Lee, M. Livan, R. Wigmans, «Dual-readout calorimetry», Rev. Mod. Phys. 90 (2018)025002
- [24] M. Antonello et al., «Test of a dual-readout fiber calorimeter with SiPM light sensors», Nucl. Instr. and Meth. A899 (2018) 52
- [25] S. Agostinelli et al., «GEANT4 - A simulation toolkit», Nucl. Instr. and Meth. A506(2003) 250
- [26] R. Wigmans, «Calorimetry - Energy Measurement In Particle Physics», second ed., in:International Series of Monographs on Physics, vol. 168, Oxford University Press, 2017

- [27] R. Wigmans et al. «The art of calorimetry». in:Proceedings of the INternational School of Physics "Enrico Fermi", DOI 10.3254/978-1-60750-630-0-31
- [28] C. W. Fabjan, F. Gianotti, «Calorimetry for particle physics», Rev. Mod. Phys. 75(2003) 1243-1286
- [29] W. R. Leo, «Techniques for nuclear and particle physics experiments», second ed., Springer-Verlag, 1994
- [30] R. Wigmans and M. Zeyrek, «On the differences between calorimetric detection of electrons and photons», Nucl. Instr. and Meth. A485 (2002) 385
- [31] D. Acosta et al., «Lateral shower profiles in a lead/scintillating fiber calorimeter», Nucl. Instr. and Meth. A316 (1992) 184
- [32] Caso, C. et al. «Particle Data Group», Eur. Phys. J.C3,1 (1998)
- [33] S. Lee, M. Livan, R. Wigmans, «Dual-readout calorimetry», Rev. Mod. Phys. 90 (2018)025002
- [34] S. Lee et al. «On the limits of the hadronic energy resolution of calorimeters». In:Nuclear Inst. and Methods in Physics Research, A882148-157(2018). DOI:10.1016/j.nima.2017.10.087
- [35] N. Akchurin et al. «Hadron and jet detection with a dual-readout calorimeter». In:Nuclear Inst. and Methods in Physics Research, A537537-561(2005). DOI:10.1016/j.nima.2004.07.285
- [36] A. Sill V. Nagaslaev and R. Wigmans. «Beam tests of a thin dual-readout calorimeter for detecting cosmic rays outside the Earth's atmosphere». In:Nuclear Inst. and Methods in Physics Research, A462411-425(2001). DOI:10.1016/S0168-9002(01)00185-1
- [37] N. Akchurin et al. «Muon detection with a dual-readout calorimeter». In:Nuclear Inst. and Methods in Physics Research, A533305-321(2004). DOI:10.1016/j.nima.2004.05.134
- [38] N. Akchurin et al. «Hadron and jet detection with a dual-readout calorimeter». In:Nuclear Inst. and Methods in Physics Research, A537537-561(2005). DOI:10.1016/j.nima.2004.07.285
- [39] A. Ghassemi, K. Sato, K. Kobayash, Technical Note:MPPC,  
[https://www.hamamatsu.com/resources/pdf/ssd/mppc\\_kapd9005e.pdf](https://www.hamamatsu.com/resources/pdf/ssd/mppc_kapd9005e.pdf)
- [40] C. Piemonte and A. Gola. «Overview on the main parameters and technology of modern Silicon Photomultipliers». In:Nuclear Inst. and Methods in Physics Research, A9262-15(2019). DOI:10.1016/j.nima.2018.11.119

- [41] F.Acerbi and S.Gundacker. «Understanding and simulating SiPMs». In:Nuclear Inst. and Methods in Physics Research, A92616-35(2019). DOI:10.1016/j.nima.2018.11.118
- [42] V. Arosio et al. «An Educational Kit Based on a Modular Silicon Photomultiplier System». In:arXiv(2014).  
<https://arxiv.org/abs/1308.3622v2>
- [43] P. Eckert et al. «Characterisation Studies of Silicon Photomultipliers». In:Nuclear Inst. and Methods in Physics Research, A620217-226(2010). DOI:10.1016/j.nima.2010.03.169
- [44] T. Bretz et al. «Dynamic range measurement and calibration of SiPMs». In:JINST,11P03009(2016). DOI:10.1088/1748-0221/11/03/P03009
- [45] R.Santoro. «International Conference on New Photo-Detectors (PD15) (July, 2015)». In:Moscow, Troitsk(2015) (cit. on p. 132)
- [46] Hamamatsu SiPMs lineup  
[https://www.hamamatsu.com/eu/en/product/optical-sensors/mppc/mppc-mppc-array/all\\_products/index.html](https://www.hamamatsu.com/eu/en/product/optical-sensors/mppc/mppc-mppc-array/all_products/index.html)
- [47] P.A. Zyla et al. «Particle Data Group», Prog. Theor. Exp. Phys. 2020, 083C01 (2020)
- [48] C.M. Bishop «Pattern recognition and machine learning». Springer, 2006
- [49] C.M. Bishop «A high-bias, low-variance introduction to Machine Learning for physicists». arXiv:1803.08823v3, 2019
- [50] Jason Brownlee «Deep Learning for Computer Vision: Image Classification, Object Detection, and Face Recognition in Python». Machine Learning Mastery, 2019
- [51] Karen Simonyan and Andrew Zisserman «Very Deep Convolutional Networks for Large-Scale Image Recognition». arXiv:1409.1556v6, 2015
- [52] Kaiming He et al. «Deep Residual Learning for Image Recognition». arXiv:1512.03385v1, 2015
- [53] «Layer activation functions» from Keras API reference,  
<https://keras.io/api/layers/activations/>