Simulation chain and first Neural Network results

Alessandro Villa

Università degli Studi di Pavia (UNIPV)
INFN (Pavia)

02 December 2020

Istituto Nazionale di Fisica Nucleare

DR-calorimeter simulation (Geant4)

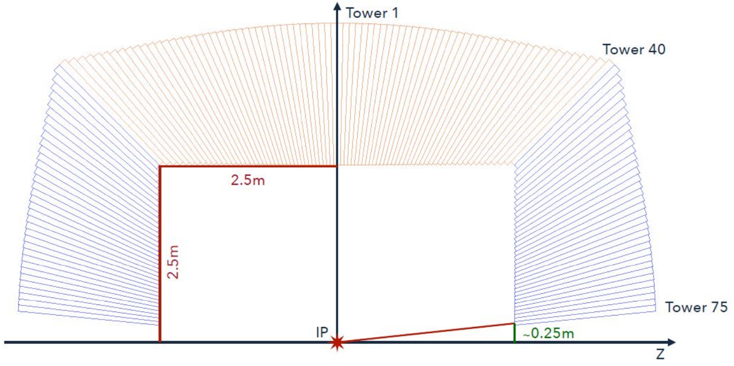
The calorimeter is composed by several towers, covering $\Delta\vartheta$=1.125° and $\Delta\varphi$=10.0° each.

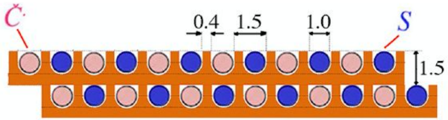
2m long copper based towers.

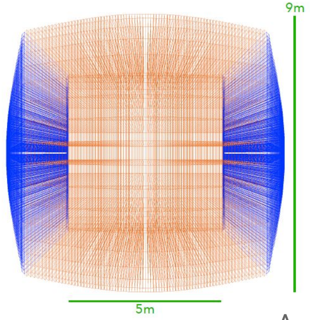
Towers in barrel: 40×2×36 = 2880
Towers per endcap: 36×36 = 1260

Towers are filled with two types of optical fibres for Cherenkov (C) and Scintillating (S) photons.
Each fibre is coupled to a dedicated SiPM.

The results shown later are obtained simulating single 40 GeV e^- or 40 GeV π^- emitted from the IP.







SiPM digitization (pySiPM)

The output from DR-calorimeter simulation and the input required by pySiPM have been modified to be fully compatible.

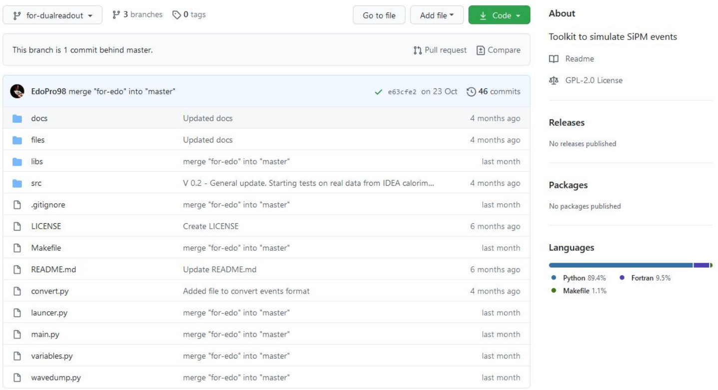
pySiPM can provide useful features such as:

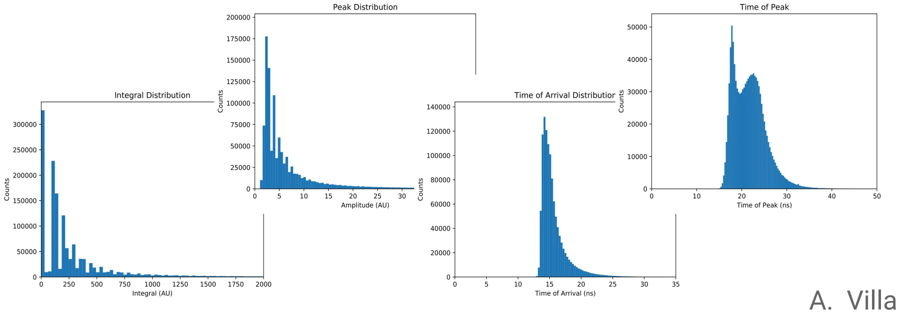
- Peak height
- Charge integral
- Time of Arrival
- Time over Threshold
- Time of Peak

and brings back additional information coming from the DR-calorimeter simulation:

- Fibre ID
- Fibre position (x, y, z)
- Type of fibre (Cherenkov, Scintillating)

It can also provide a digitized waveform as output from each SiPM.





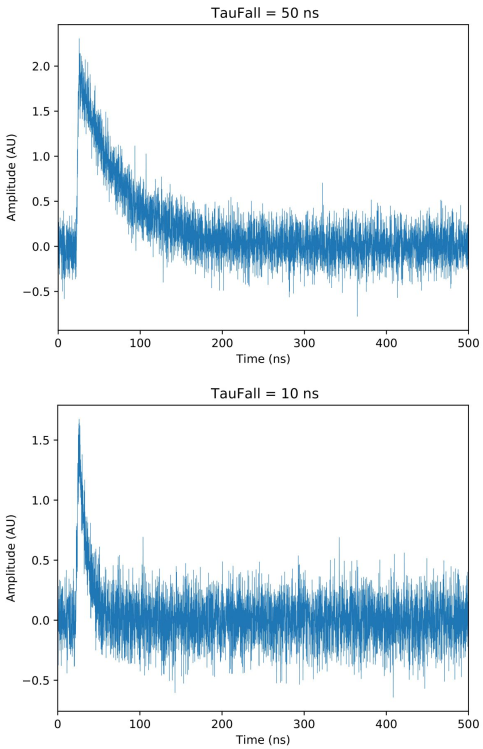
SiPM parameters

Most of the SiPM parameters can be easily modified.
In this way, we can find the best parameter ranges with respect to our goals.

SIGLEN: the length of the signal generated
SAMPLING: the time between two consecutive points

SIZE: the size of the SiPM
CELLSIZE: the size of each single cell
DCR: the Dark Count Rate
XT: the probability of Optical Crosstalk
AP: the probability of Afterpulse
TFALL: the fall time of the signal generated
TRISE: the rise time of the signal generated

INTSTART: the time at which the integration gate starts
INTGATE: the integration time

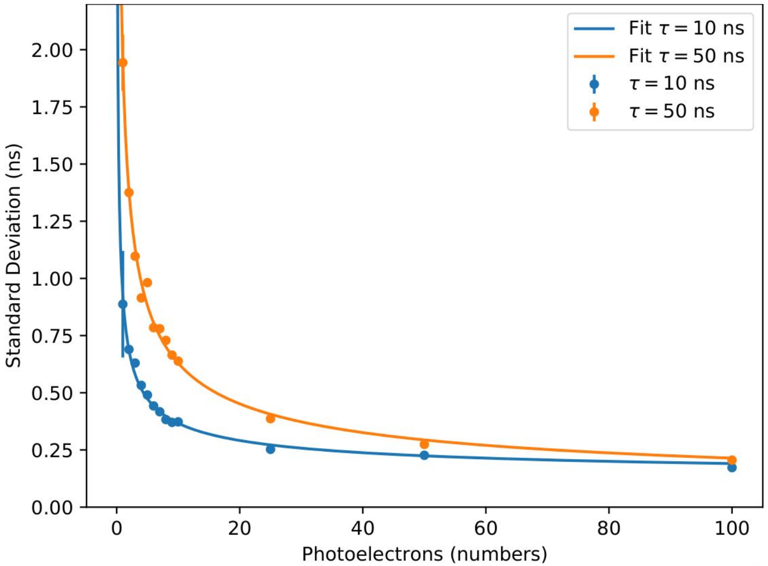


Peak time resolution

With 2 different SiPM fall times (10ns and 50ns) we evaluated the standard deviation of the gaussian fit applied on 10000 waveforms and its behaviour in function of the number of simultaneous photoelectron (n) fired on the SiPM.

This behaviour follows the law:

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



We generate 13 events with an increasing number of photoelectrons (from 1 to 10, then 25, 50 and 100) with the same Geant4 Time of Arrival. In each event we have 10000 activated fibres, and 10000 corresponding SiPMs. We record the Time of Peak from the digitization software and put them in 13 histograms (one for each event). By fitting them with a gaussian function we find Standard Deviation plotted above.

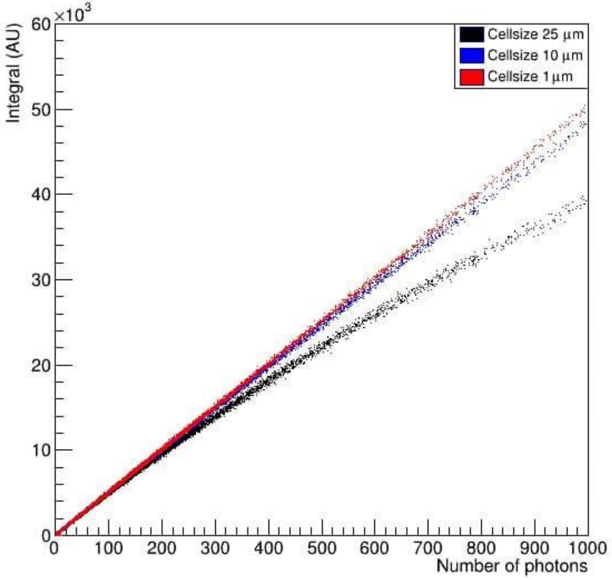
SiPM saturation

By changing the parameter “Cellsize” we can increase the number of cells in each SiPM (fixed size of 1x1 mm^2).

With a 25 μm cell size the saturation effect is clear.

With a 10 μm cell size the saturation effect is negligible.

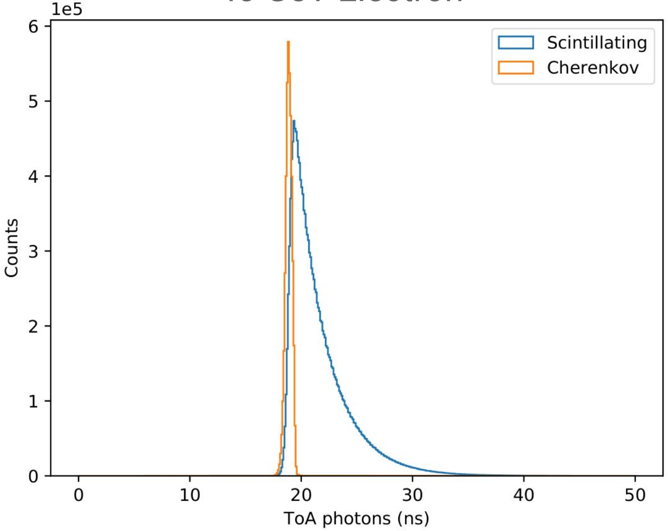
With a 1 μm cell size we have an ideal situation that we include as reference.



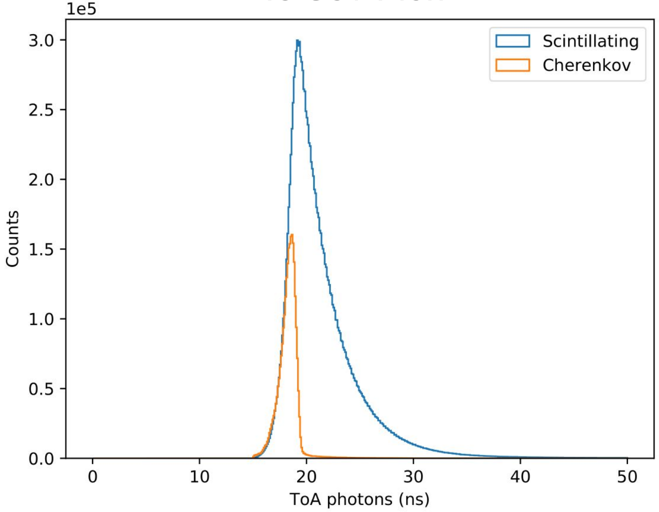
Time distribution of photons

We simulated 1000 events for both particle types.
First we focused on the Geant4 time of arrival of the photons at the end of the fibres.
The long tail for scintillating photons is due to the characteristic emission time of Polystyrene.

40 GeV Electron



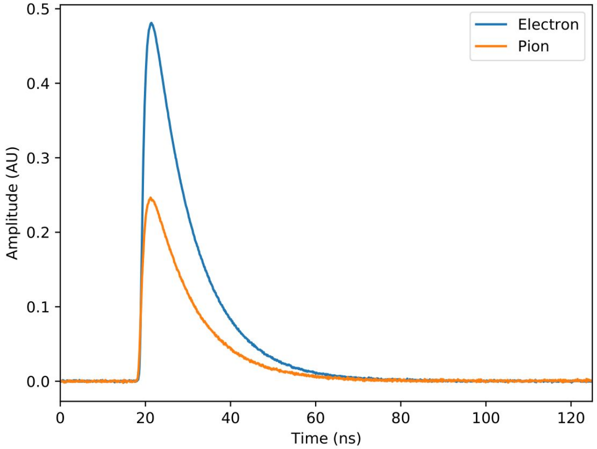
40 GeV Pion



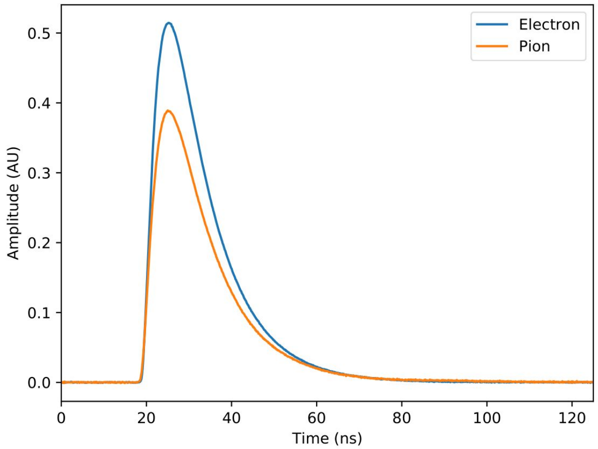
Integrated waveform

Considering one event at a time, we analogically added the signals coming from the same type of fibres. Below there is an example comparing a typical signal from 40 GeV electron and 40 GeV pion.

Cherenkov



Scintillation

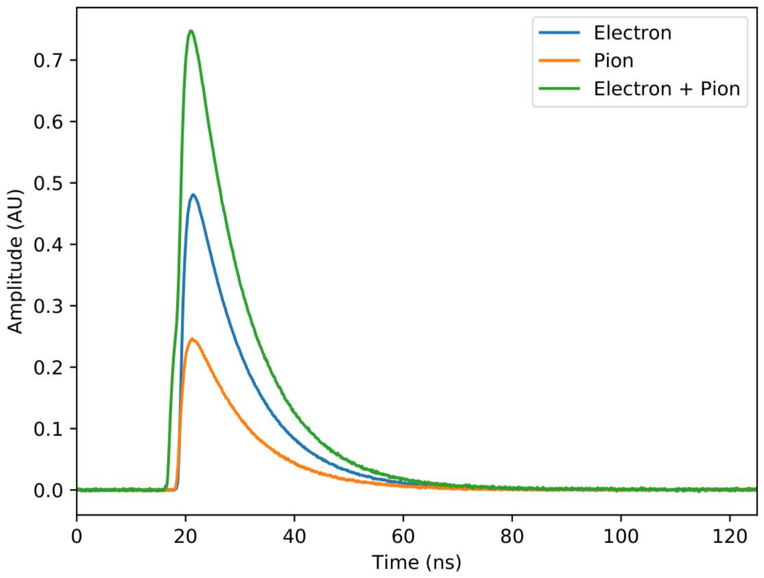


Neural Network Goal – π/e identification

By overlapping the integrated waveform obtained from electron and pion we can see that features such as peak time, rise time, fall time, peak height could be useful in a process of particle ID.

We want to build a Neural Network able to distinguish if the integrated waveform is originated from an electron, a pion or a combination of both.

We set up our data simulating 1000 40 GeV electrons, 1000 40 GeV pions and 1000 events with both 40 GeV electrons and 40 GeV pions. Then we normalized the waveforms to the max value of the sample.



Neural Network optimization

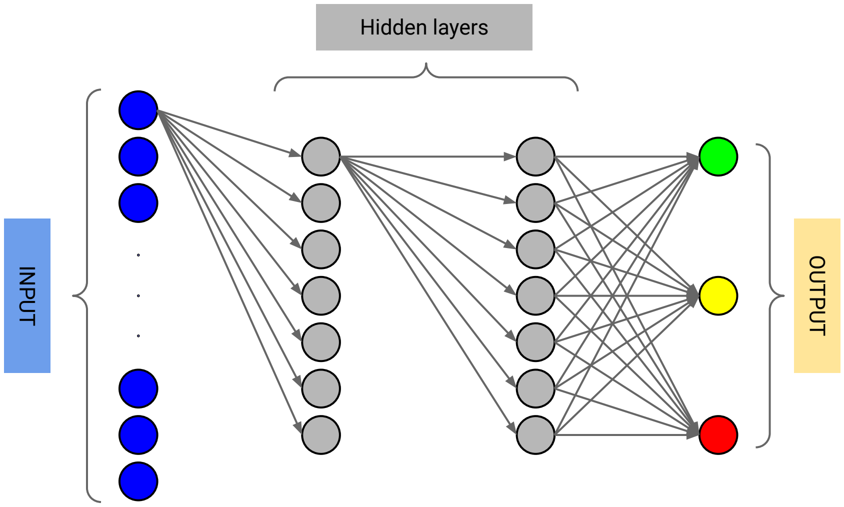
To select an optimal Neural Network structure, we used the Hypertuner.

In this way we could create several NN with different hyperparameters in predisposed ranges.

The hyperparameters we studied are:
- number of hidden layers (2 - 10 layers)
- number of neurons in each layer (8 - 64 neurons)
- learning rate (10^{-4}, 10^{-5}, 10^{-6})

We produced 1000 NN and to each one the Hypertuner associated a score.

At the end of this process we obtained and saved the best model structure.



Cherenkov and Scintillation NNs

We performed this process with both the types of data and we obtained the following neural network structure.

Cherenkov

Number of layers: 8

Layer neurons:
[56, 24, 32, 8, 32, 32, 56, 3]

Learning rate: 10-5

Score: Val_loss = 0.023

Scintillation

Number of layers: 5

Layer neurons:
[40, 64, 32, 40, 3]

Learning rate: 10-5

Score: Val_loss = 0.026

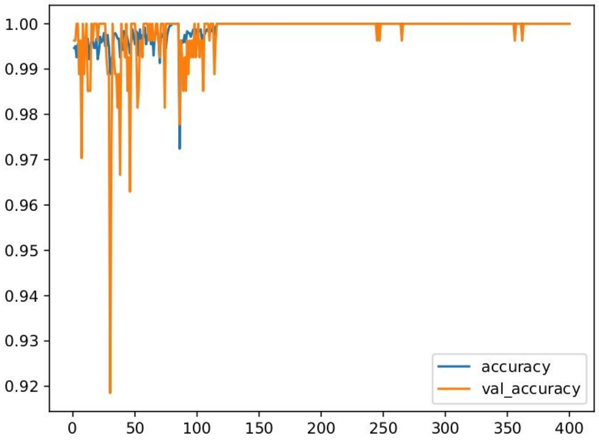
Cherenkov vs Scintillation performances - 1/3

Accuracy: the classic probability associated to the correct label prediction.
The evolution through 400 epoches is shown.

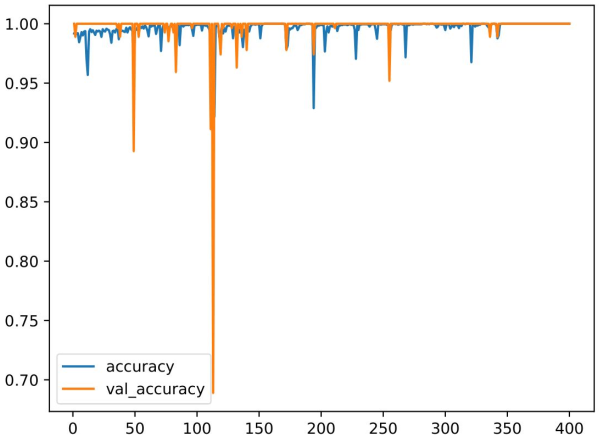
$$acc = \frac{n_c}{N}$$

n_c is the number of correct predictions
N is the total number of predictions

Cherenkov



Scintillation



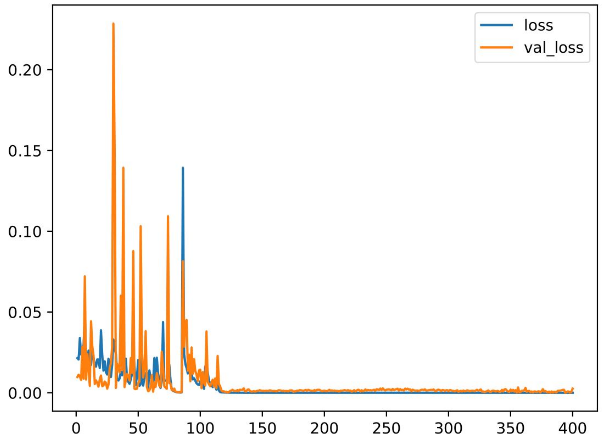
Cherenkov vs Scintillation performances - 2/3

Loss function: the Neural Network attempt to minimize this function modifying weights.

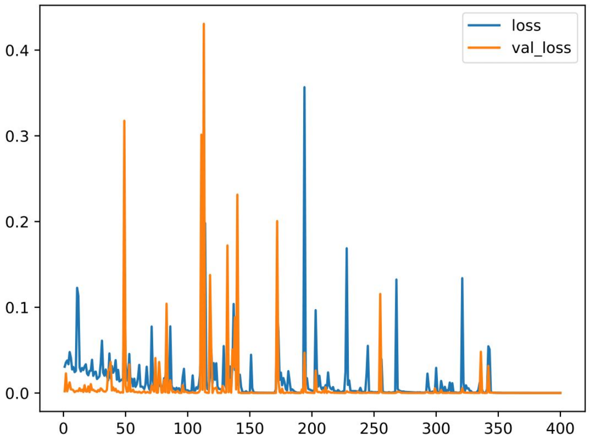
We choose sparse_categorical_crossentropy: $-\sum_i p_i \log q_i$

p_i is true probability for each label
q_i is the probability associated to a label from the NN

Cherenkov



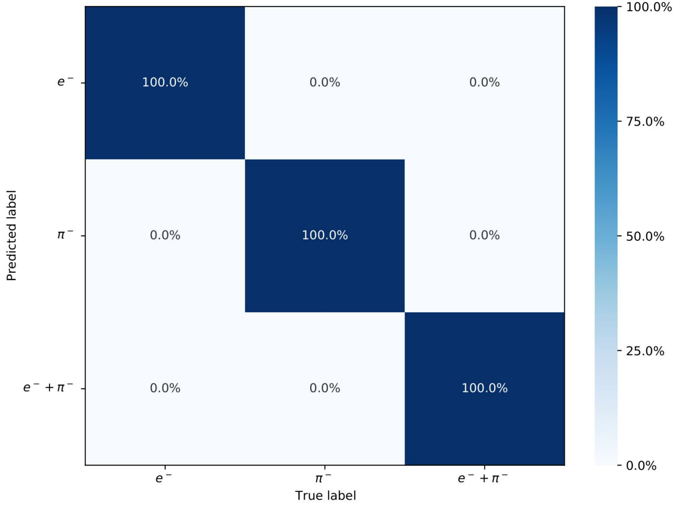
Scintillation



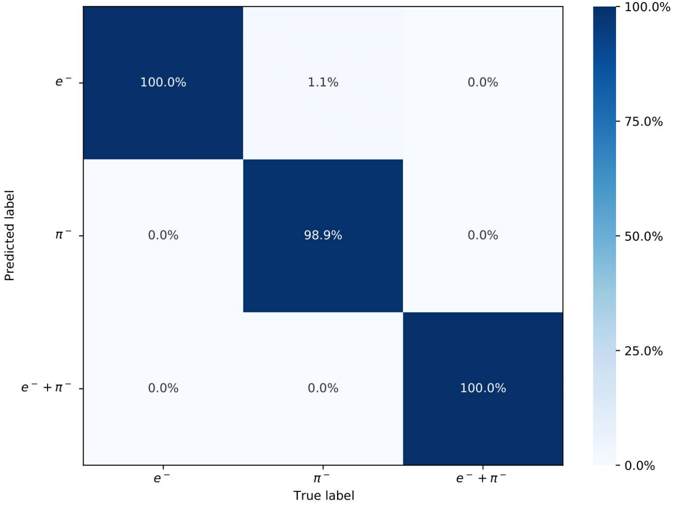
Cherenkov vs Scintillation performances - 3/3

Confusion Matrix applied on a sample of 300 events.
Values are normalized dividing by the total event number for each true label.

Cherenkov



Scintillation



Conclusion

The Geant4 + Digitization chain is well oiled and can give many useful informations.

The exercise of particle ID at 40 GeV is giving good results and confirmed that Cherenkov signals are more efficient in this process.

The process will be extended in energy range (e.g. 1-60 GeV) and in type of particles fired (e.g. photons).
It will be interesting to see the accuracy loss with these extensions. In particular in respect to the energy range, because the peak height will be a weaker factor of discrimination.

The introduction of Recurrent layers could be an important upgrade, because they are commonly used and optimized in time-evolving problems.

Any comment and suggestions will be appreciated.

Backup

Unbiased Neural Network

We wanted to prove that our Neural Networks are unbiased, so that they don't learn information from nowhere.
Confusion Matrix applied on a sample of 3000 events randomly generated and randomly associated to labels.