Particle Flow for Dual Read-Out Calorimeter

Adelina D'Onofrio[1], Michela Biglietti[1], Biagio Di Micco[2], Roberto Di Nardo[2], Ada Farilla[1] Iacopo Vivarelli[3] , Sofia Vallecorsa[4] , Patrizia Azzi[5]

[1]INFN - Roma Tre

[2]Roma Tre University

[3]University of Sussex

[4]CERN

[5]INFN - Padova

CepC Workshop - Nanjing University

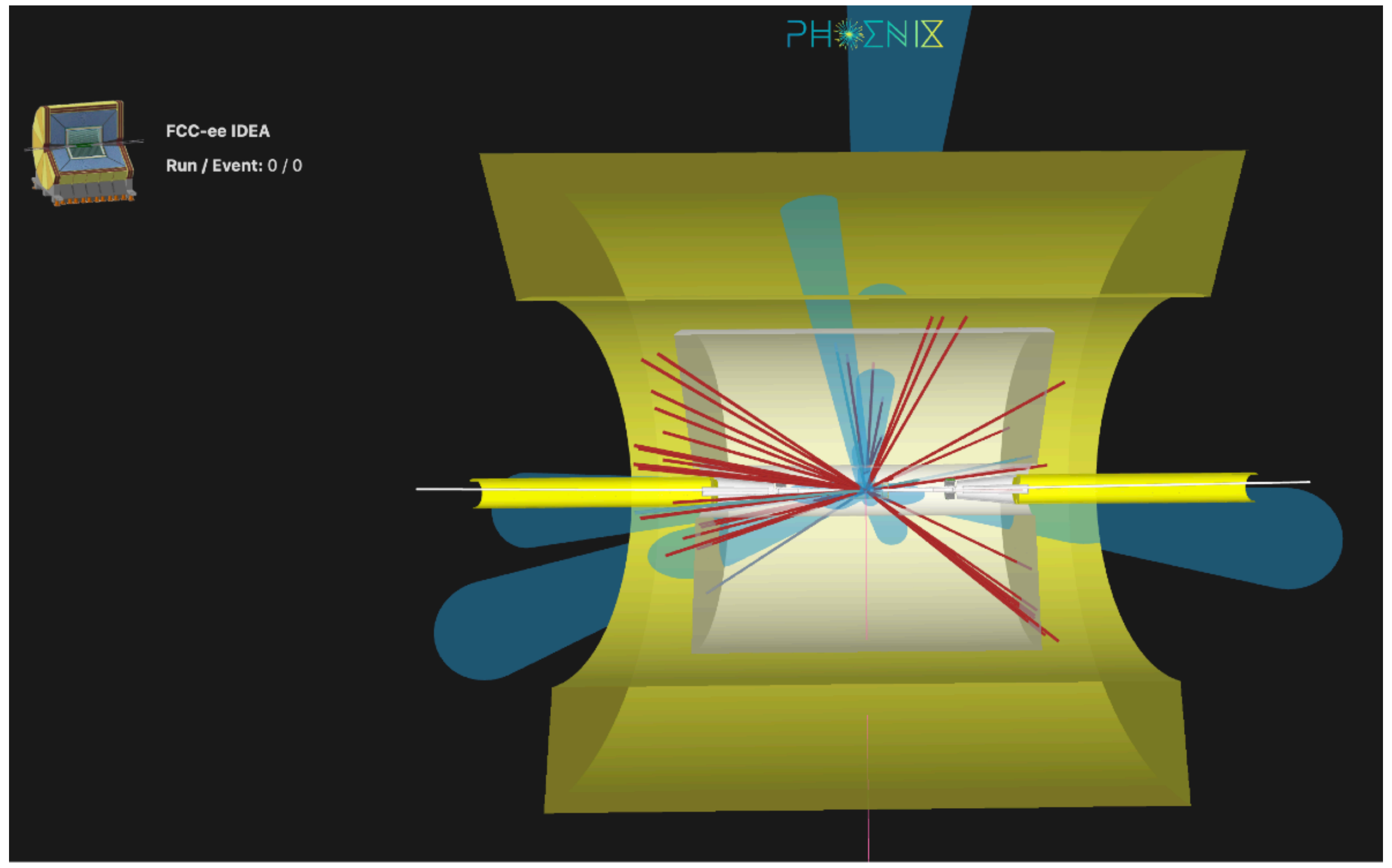
25th October 2022

AIDA innova

INFN Istituto Nazionale di Fisica Nucleare SEZIONE DI ROMA TRE

Outline

- Highlights about dual read-out calorimeter for the IDEA detector (for a detailed overview see the previous talk)
  - Single, dual-readout calorimeter for EM and **HAD** calorimetry
    - But option to have a **crystal, dual-readout** EM section
  - Detector layout
  - Energy measurement
- Particle Flow algorithm for dual-read-out calorimeters
  - Software implementation
  - Neural Network approach used for particle identification and jet energy reconstruction
  - Preliminary results

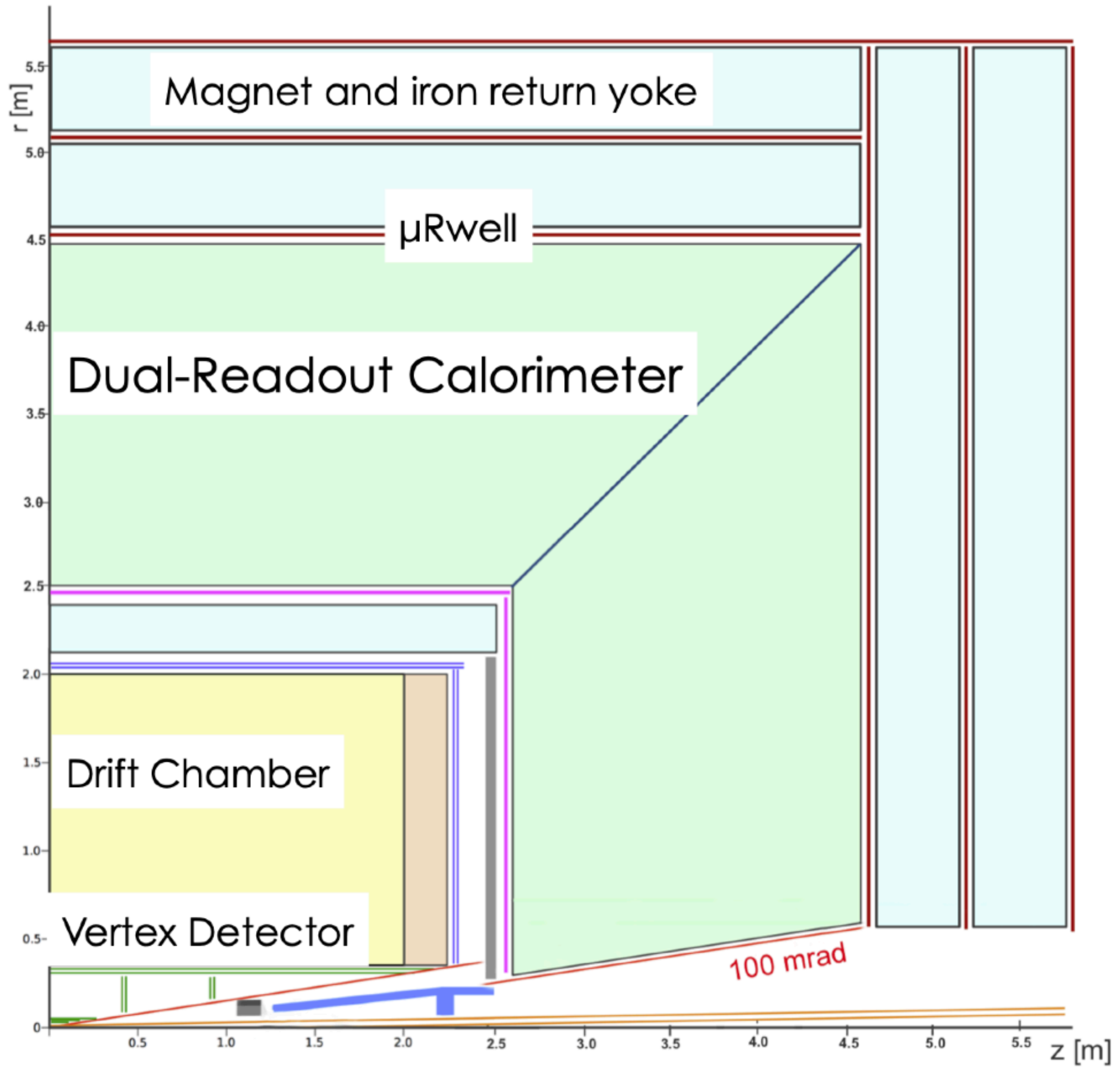


Physics Process	Measured Quantity	Critical Detector	Required Performance
$ZH \to \ell^+\ell^- X$	Higgs mass, cross section	Tracker	$\Delta(1/p_T) \sim 2 \times 10^{-5}$
$H \to \mu^+\mu^-$	$\mathrm{BR}(H \to \mu^+\mu^-)$		$\oplus 1 \times 10^{-3}/(p_T \sin\theta)$
$H \to b\bar{b},\ c\bar{c},\ gg$	$\mathrm{BR}(H \to b\bar{b},\ c\bar{c},\ gg)$	Vertex	$\sigma_{r\phi} \sim 5 \oplus 10/(p \sin^{3/2}\theta)\ \mu\mathrm{m}$
$H \to q\bar{q},\ VV$	$\mathrm{BR}(H \to q\bar{q},\ VV)$	ECAL, HCAL	$\boxed{\sigma_E^{\mathrm{jet}}/E \sim 3-4\%}$
$H \to \gamma\gamma$	$\mathrm{BR}(H \to \gamma\gamma)$	ECAL	$\sigma_E \sim 16\%/\sqrt{E} \oplus 1\%$ (GeV)

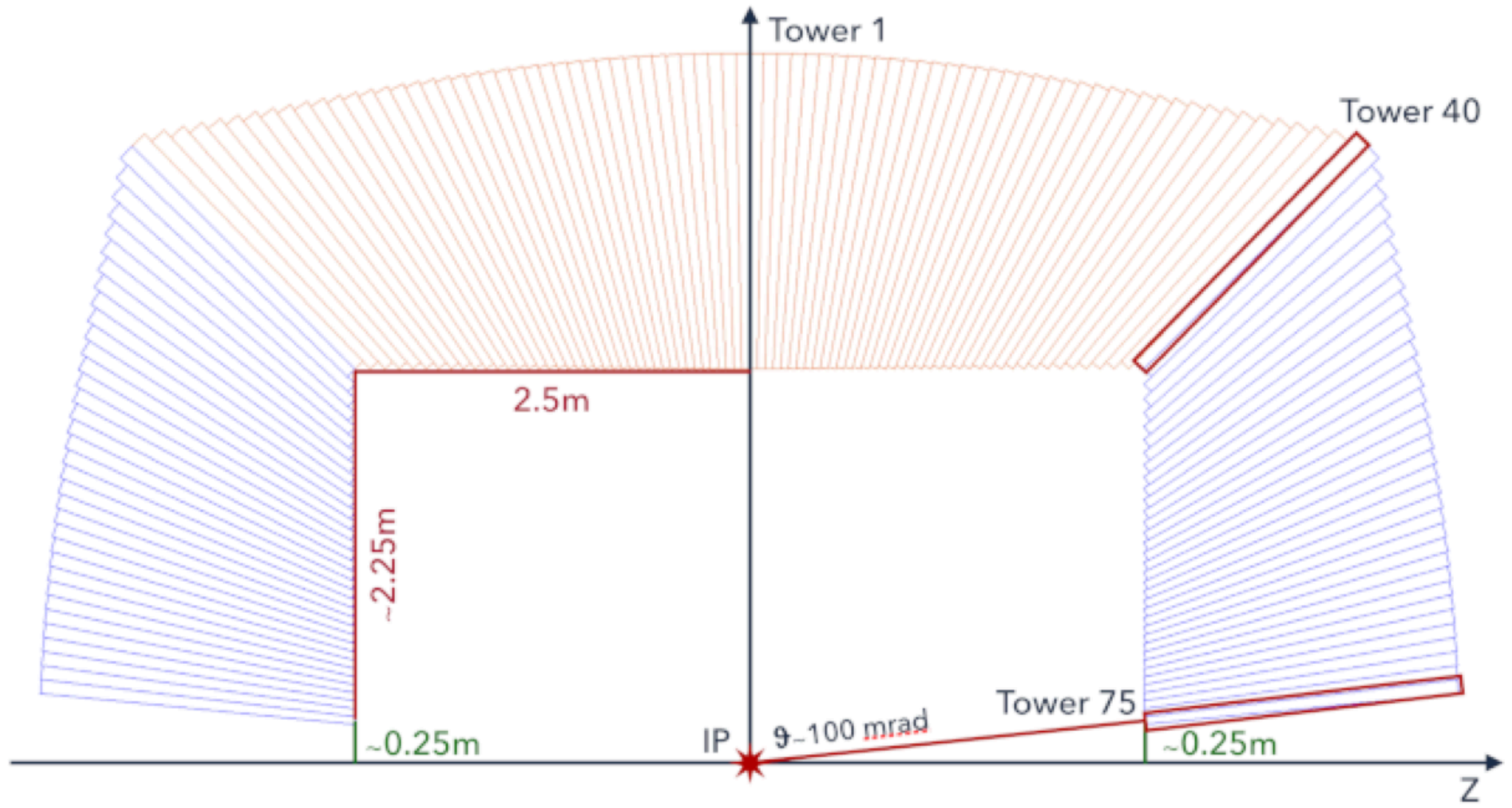
Desired jet energy resolution

Dual Read-out Calorimeter for the IDEA detector

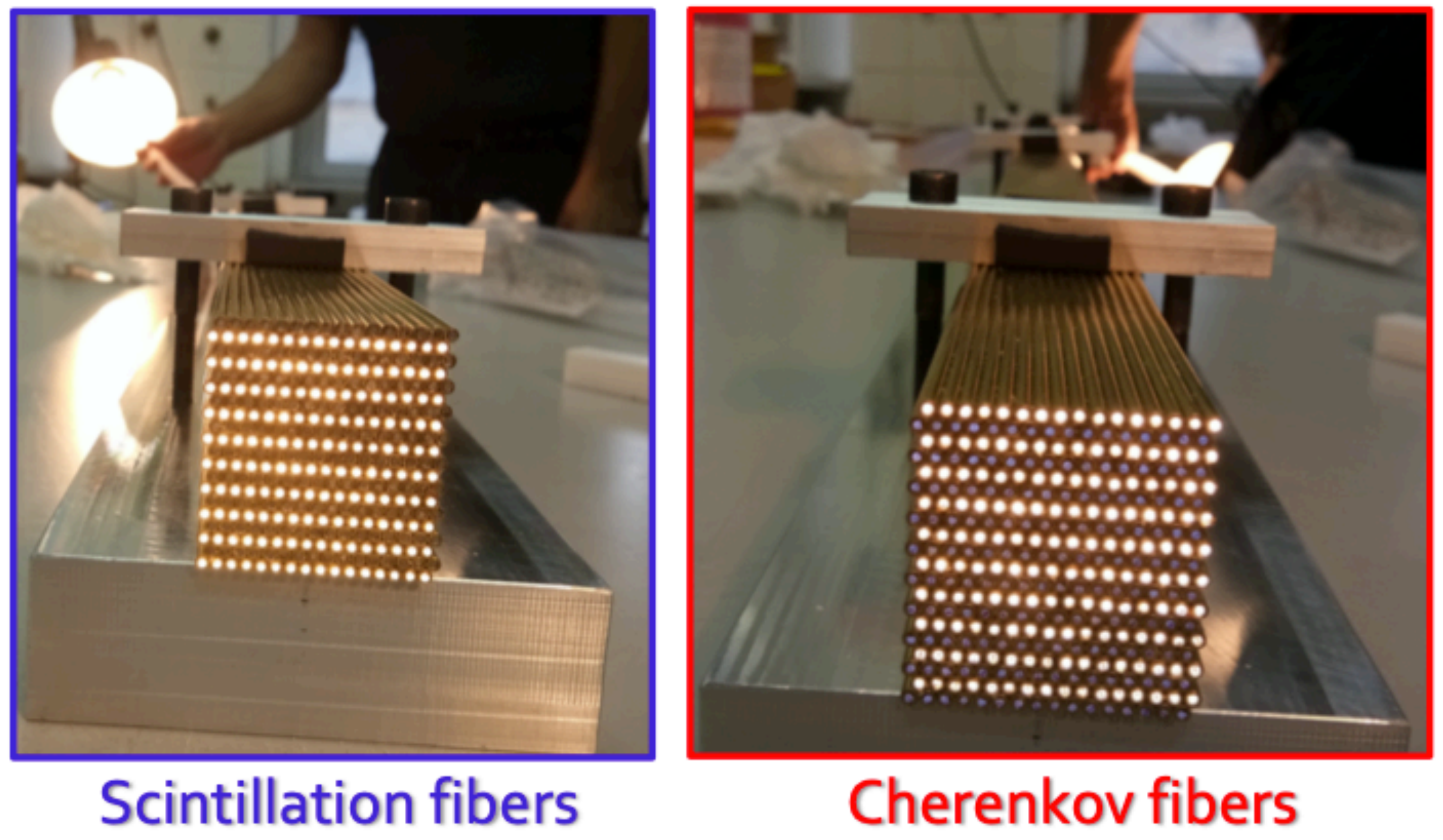
- Dual readout calorimeters aim at improving the energy resolution of hadronic calorimeters
  - Generally driven by the fluctuations between the electromagnetic and the hadronic component of showers
- Measure the hadronic component and the electromagnetic component (dual readout) of the showers separately, to derive proper correction factors to be applied to each component to reconstruct the energy of the impinging hadrons
- Exploit a passive/material - fibre layout where two type of fibres, one sensitive to the usual scintillation process, a second type of fibre producing Cherenkov light when ultra-relativistic particles cross with a speed higher than the speed of light in that fibre



A. IDEA detector transversal view



B. Sketch of a single slice of the IDEA calorimeter

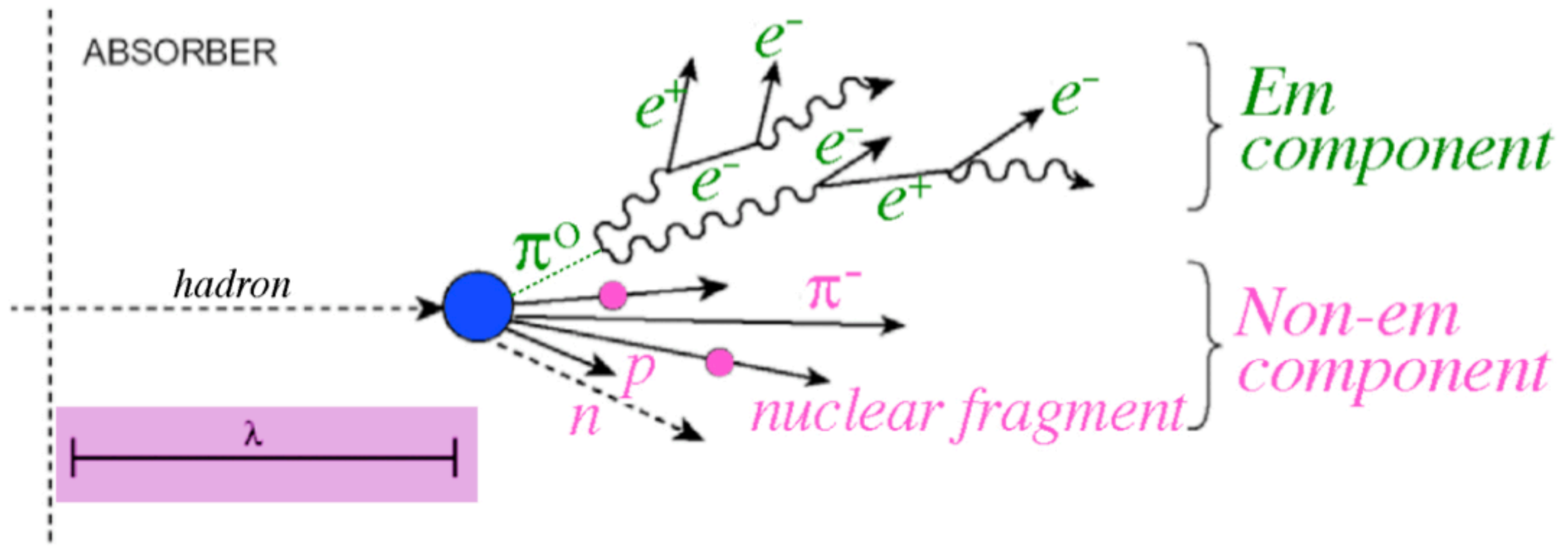


C. Current R&Ds

ParticleFlow Project - Status Report

Energy Measurement

- **Non-compensating calorimeters**: response to EM part different from that to non-EM part.
- The response ratio for electrons and charged hadrons is: $\frac{h}{e} = \eta < 1$
- The EM fraction of the shower, $< f_{em} >$, is energy dependent $\Rightarrow$ Non-linear calorimeter **response** to hadrons
- $< f_{em} >$ fluctuations largely determine energy **resolution** $\Rightarrow$ sampling the hadronic shower with two calorimeters with different *e/h* **boosts energy resolution**

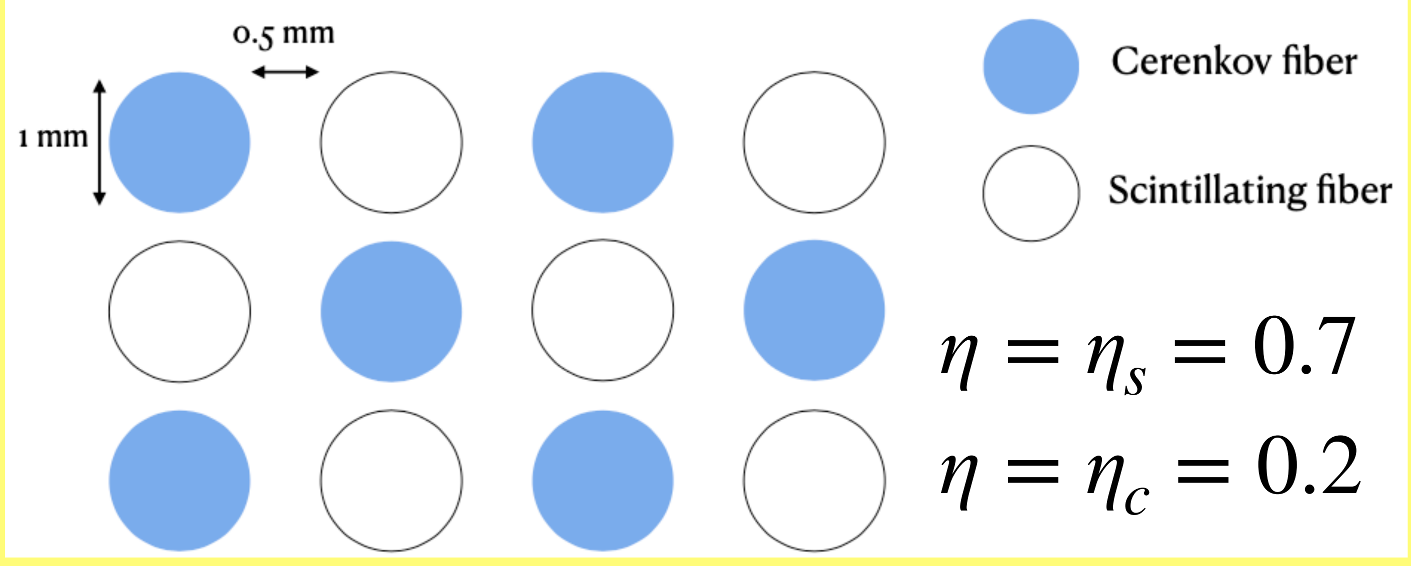


Hadron shower

$$S = E\left[f_{em} + \eta_s \cdot (1 - f_{em})\right]$$

$$C = E\left[f_{em} + \eta_c \cdot (1 - f_{em})\right]$$

$$\Rightarrow \quad \frac{C}{S} = \frac{\left[f_{em} + \eta_c \cdot (1 - f_{em})\right]}{\left[f_{em} + \eta_s \cdot (1 - f_{em})\right]}$$

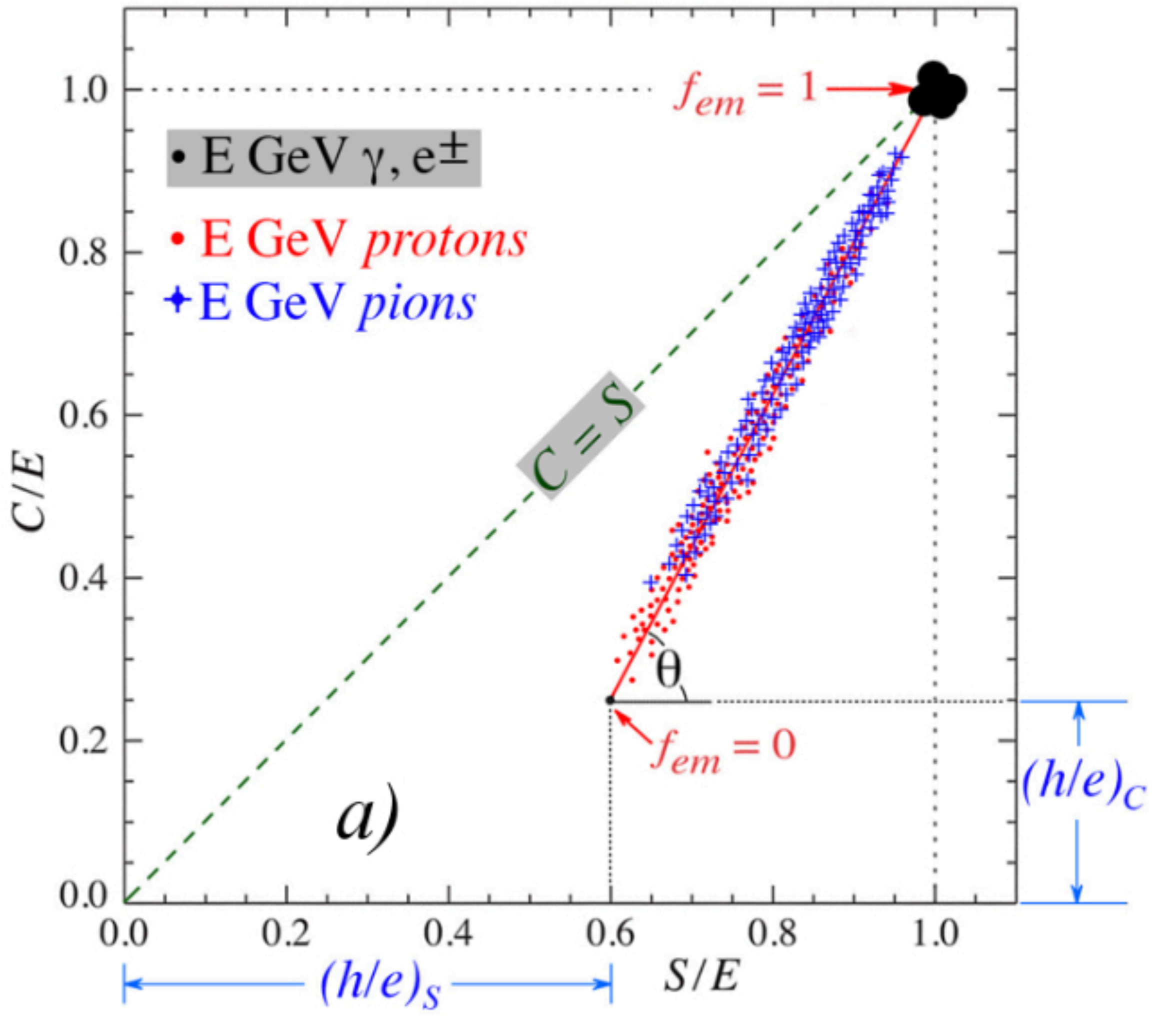


$$\eta = \eta_s = 0.7$$

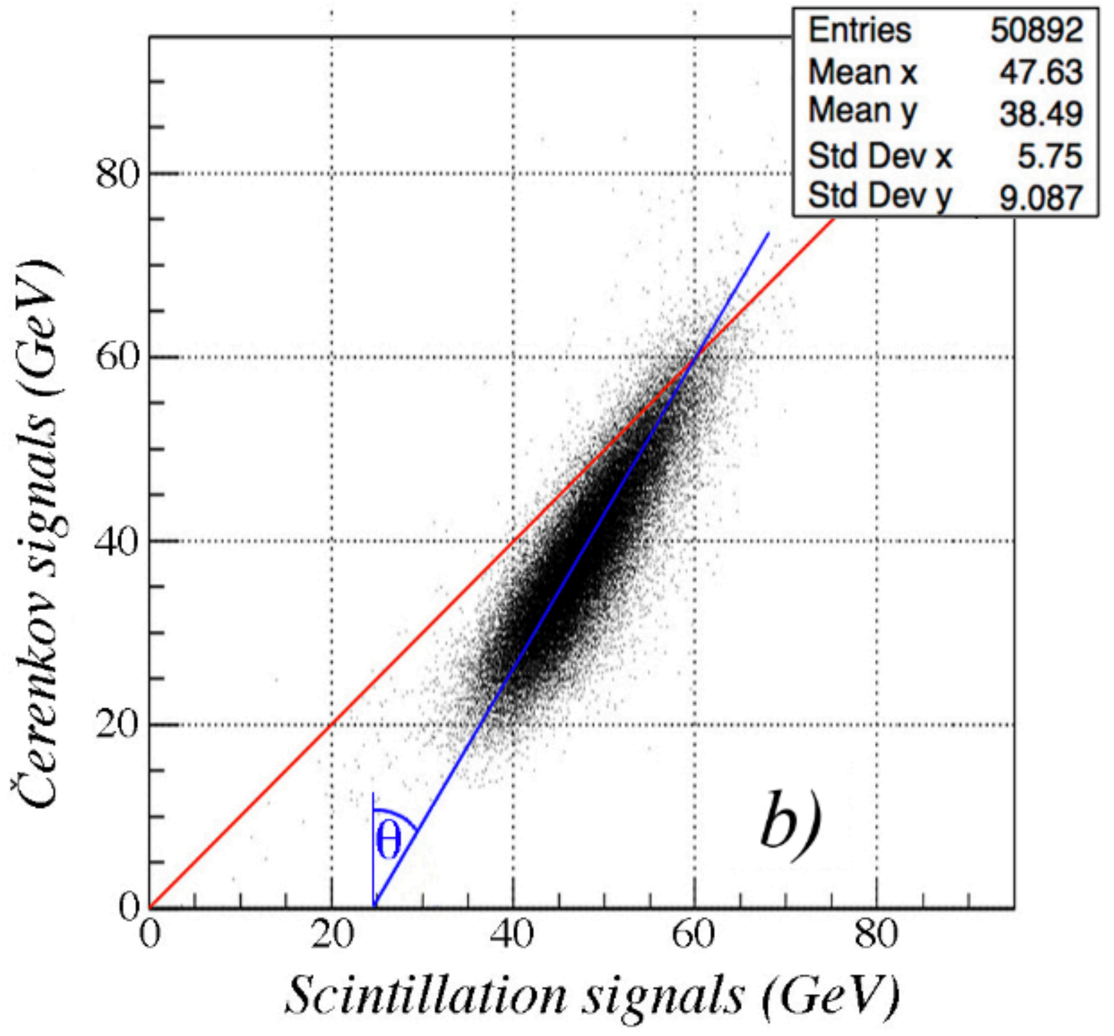
$$\eta = \eta_c = 0.2$$

$$\chi = \frac{1 - \eta_s}{1 - \eta_c} = \cot(\theta)$$

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



a) Scatter plot of C/E versus S/E in a dual-readout calorimeter for p and π



b) Scatter plot of C and S signals for 60 GeV pions in the RD52 lead-fiber calorimeter

The Particle Flow Project

- In general, particle flow algorithms applied to dual read-out calorimeters provide limited performance on the energy resolution of the electromagnetic component of the jets
  - Scenarios with EM calorimeter added in front of the IDEA dual read-out calorimeter —> Cons: calorimeter non-compensation
  - Ongoing R&D for crystal dual read-out calorimeters to fix the compensation issue
- The aim of the project is to build a Neural Network based algorithm that, from a given collection of energy deposits in the calorimeter, is able to completely reconstruct a jet in the detector
- Our goal: maximise the energy resolution of the dual read-out calorimeter exploiting NNs and taking as input all the available kinematic variables
  - NN based particle identification: use as basis a particle flow approach, which aims at identifying each single particle inside a jet
  - NN based jet reconstruction: construct a regression algorithm for particle-jet assignment and jet energy reconstruction

Overview of the Project

1. Geant4 simulation of a jet

2. Extract particle info at the input layer: position, momentum, particle type

   Extract calorimeter info: fiber position, fiber type, collected light by the fiber

3. Clustering

4. NN based particle identification: charged: $e^{\pm}$, $\mu^{\pm}$, $\pi^{\pm}$, $K^{\pm}$, p, neutral: γ, π^{0}, $K_{L,S}$, n, Λ

   NN based jet reconstruction: starting from particle lists and their momenta, build a jet

Work in progress

Software Implementation

Input from detector simulation (EDM4HEP) format

↓

Reading using KEY4HEP *code*

↓

Dumping algorithm, input variables for NN training

↓

NN training using *Tensorflow* on CPU/GPU

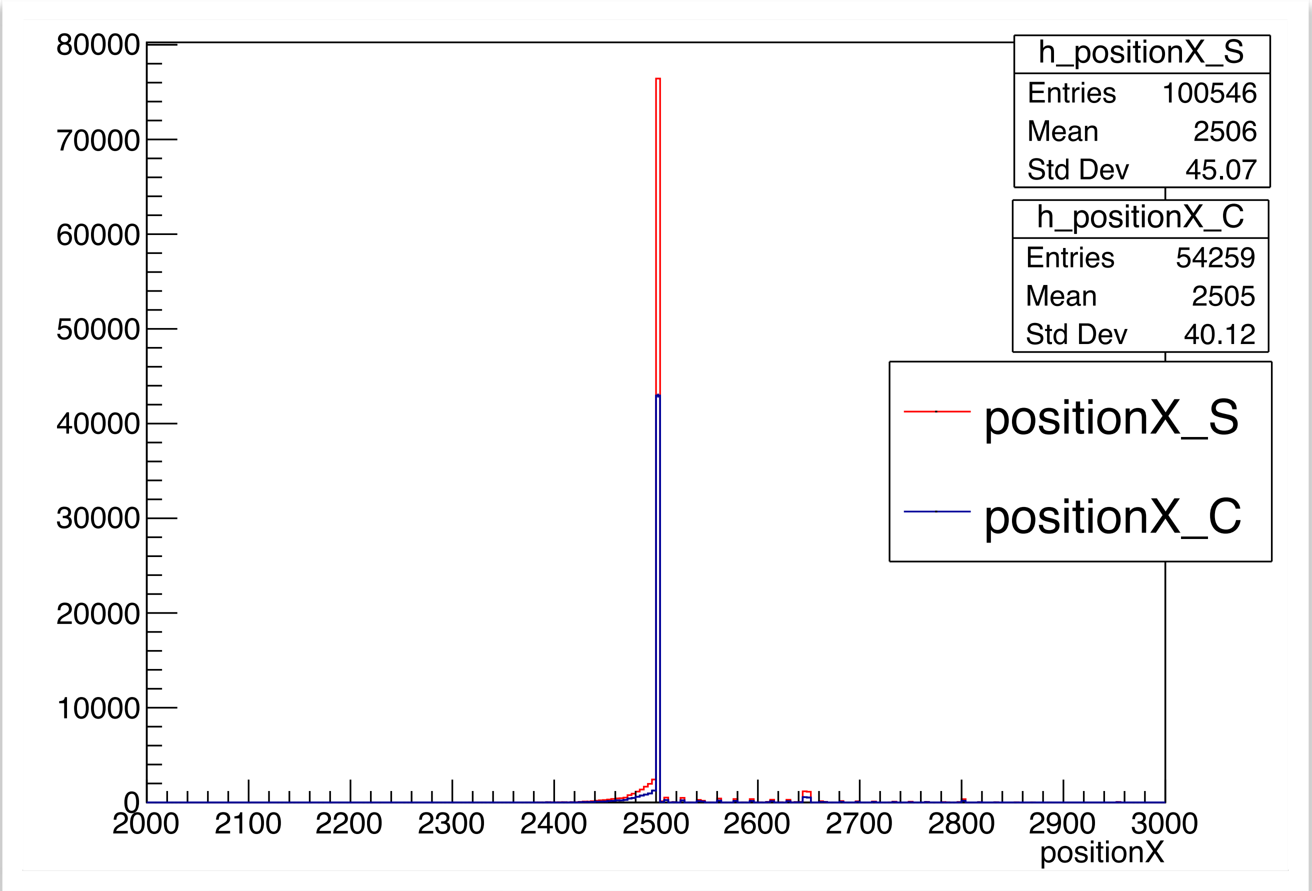
- Geant4-based simulations of the IDEA detector for *e, π, K* with energy and angular uniform distribution (thanks for the inputs!)
- An algorithm was developed that reads KEY4HEP format and produces an output to perform a Neural Network training
  - Interface between KEY4HEP and Pandora
- Target: build a NN able to reconstruct the energy and the position of the impinging particles and identify them
  - Regression and classification (to discriminate *e, π, K*) algorithms implemented in a single NN
- State of the art:
  - NN studies performed on an input sample containing 20 GeV electrons, training performed for energy regression

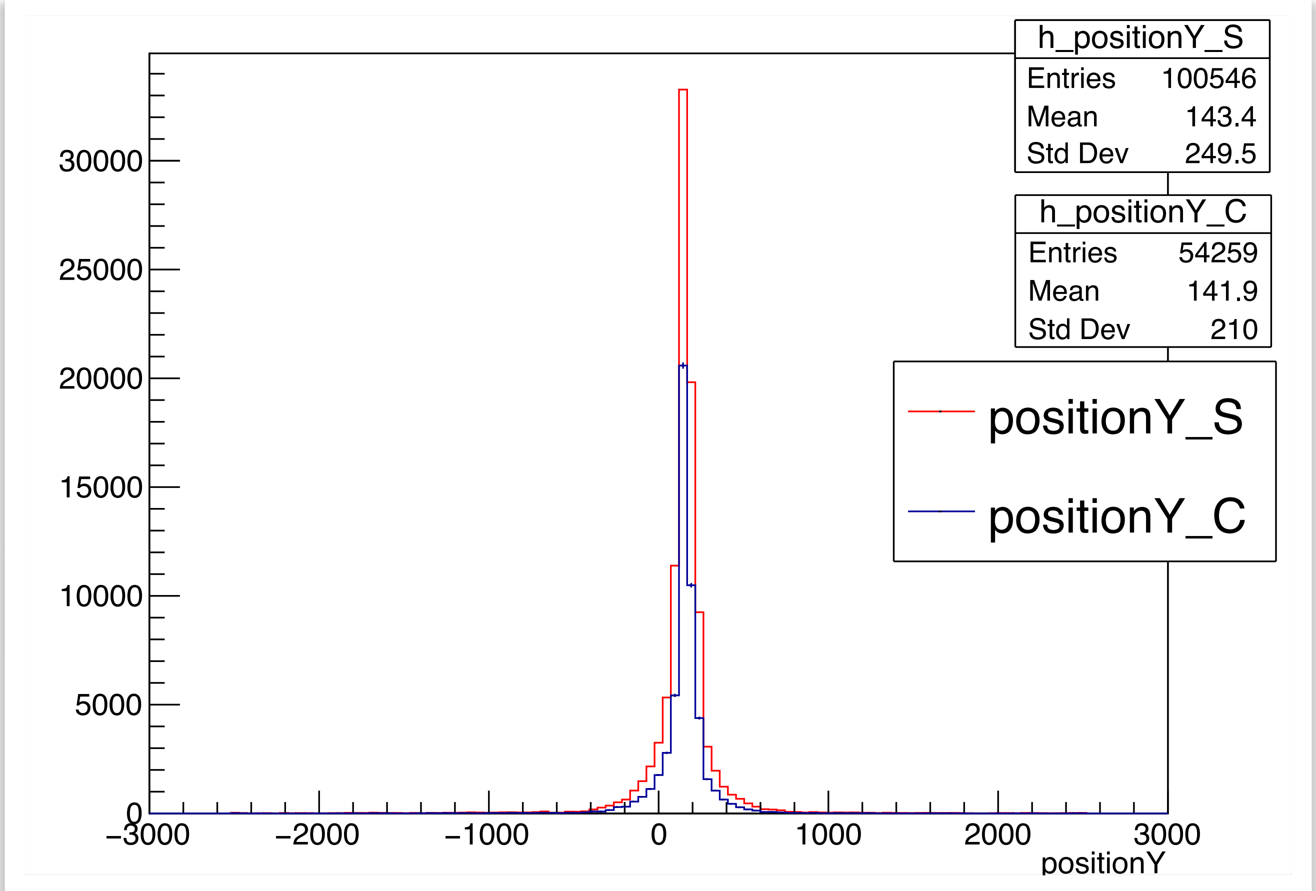
Kinematic distributions - 20 GeV electrons

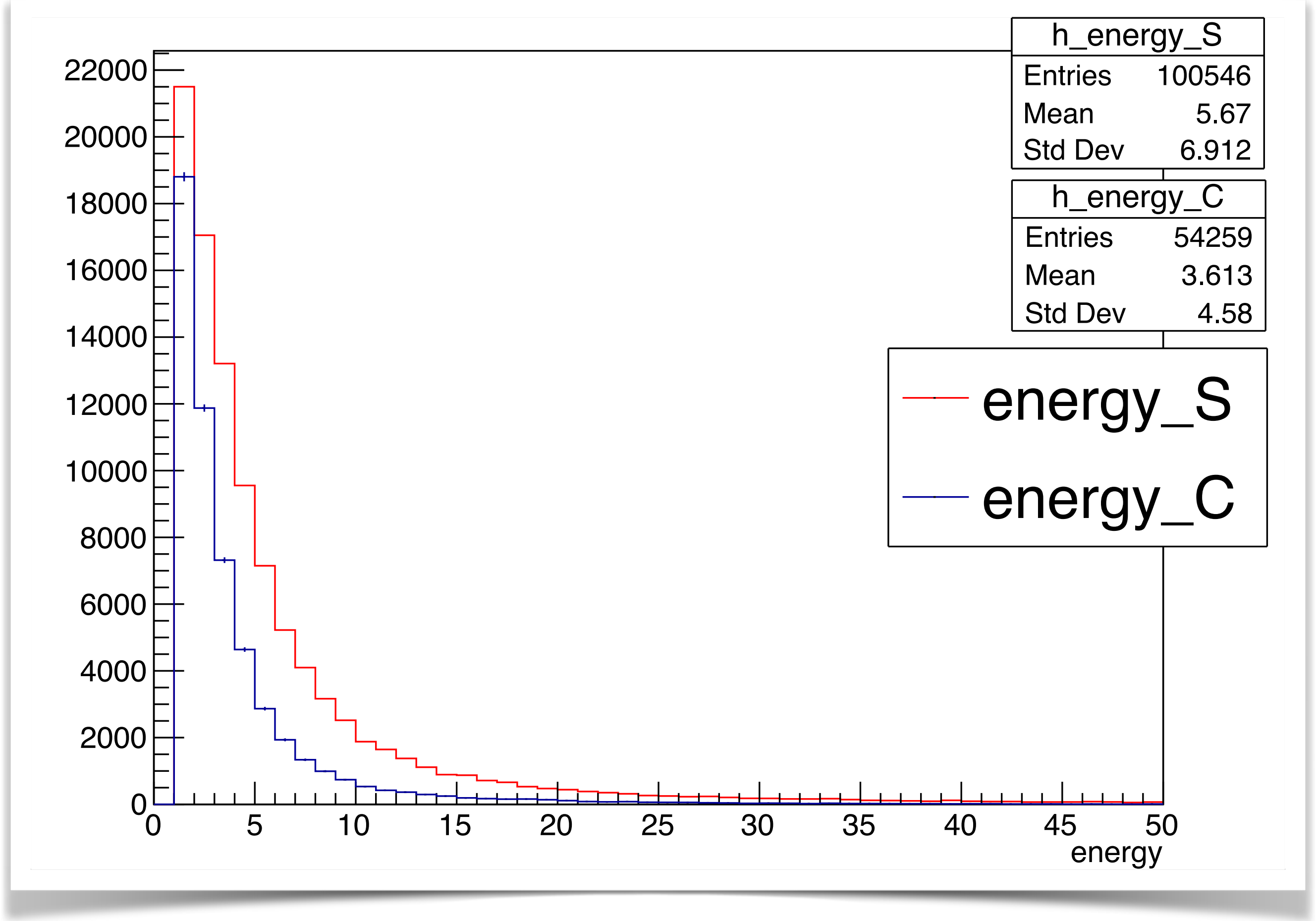
All events used (#102)

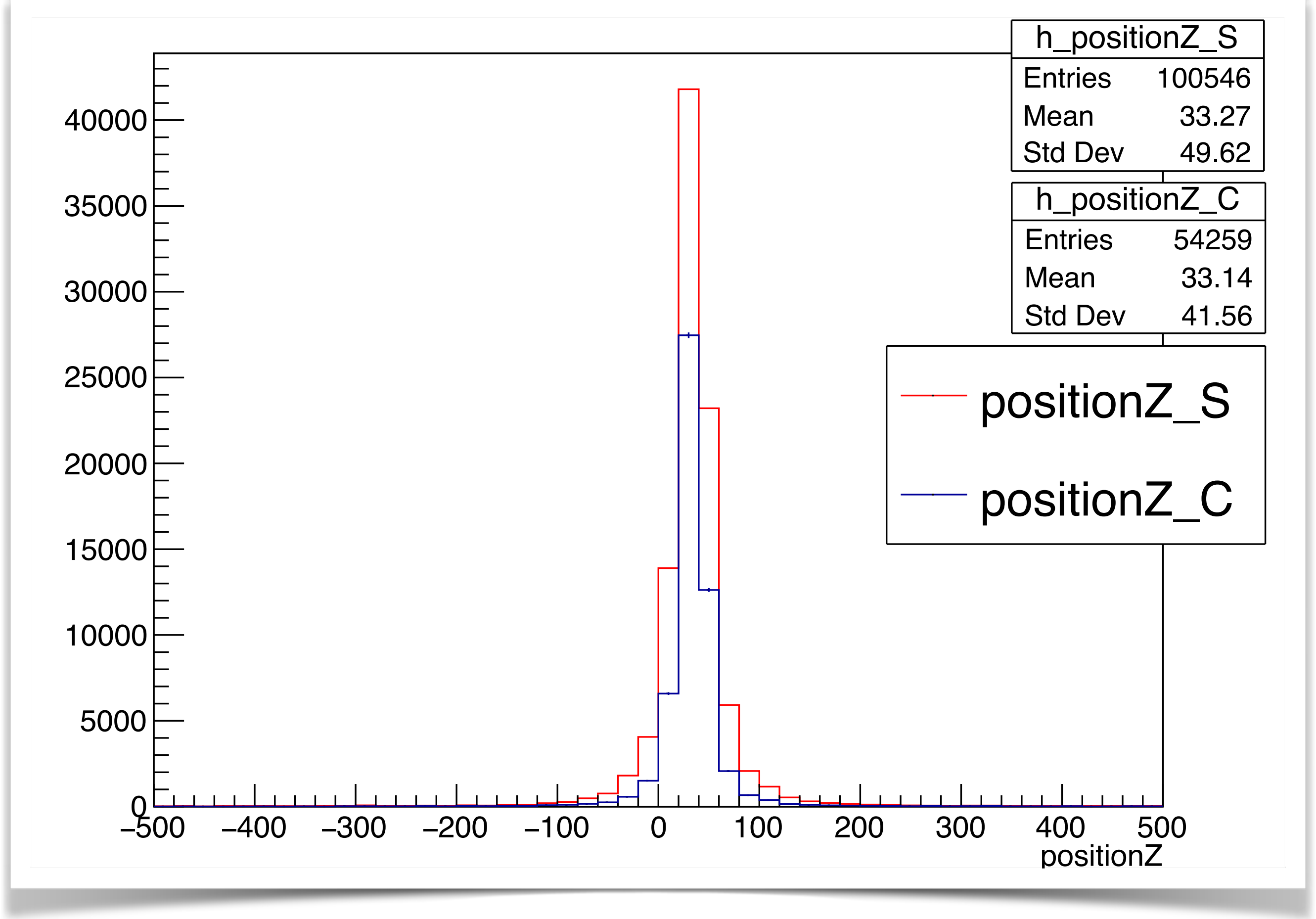
- Position and energy collected in the scintillating (S) **and** Cerenkov (C) fibres in 100 events **simulating** impinging electrons of 20 GeV

Dumper Algorithm output









Energy deposits - 20 GeV electrons

Dumper Algorithm output

Electron deposits

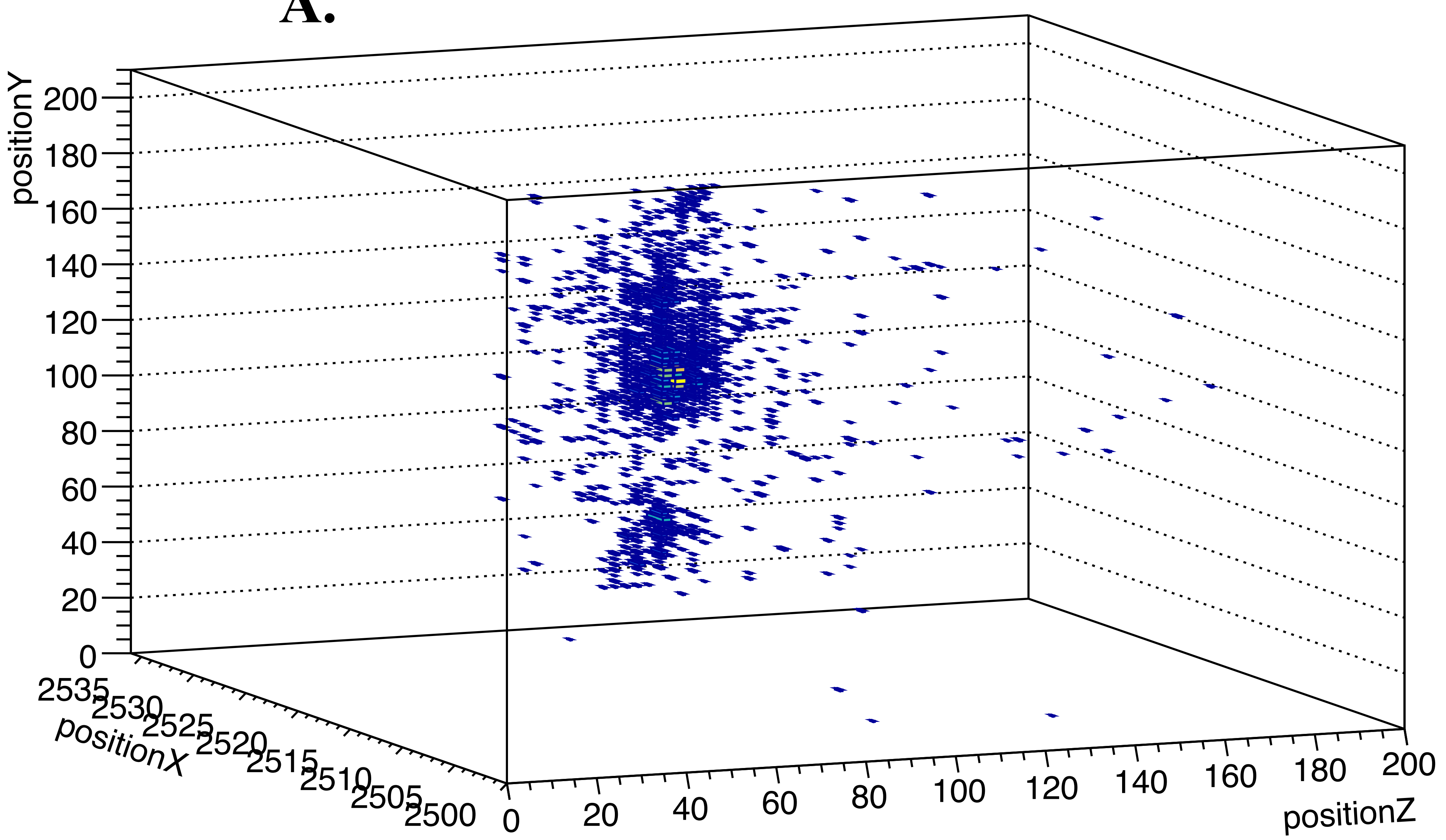
- Energy collected in the scintillating (S) **and** Cherenkov (C) fibres in 100 events **simulating** impinging electrons of 20 GeV

A. Energy deposited in the detector, projected in the *(x,y,z)* space —> combined fibres

B. Energy deposited in the scintillating fibres, polar coordinates

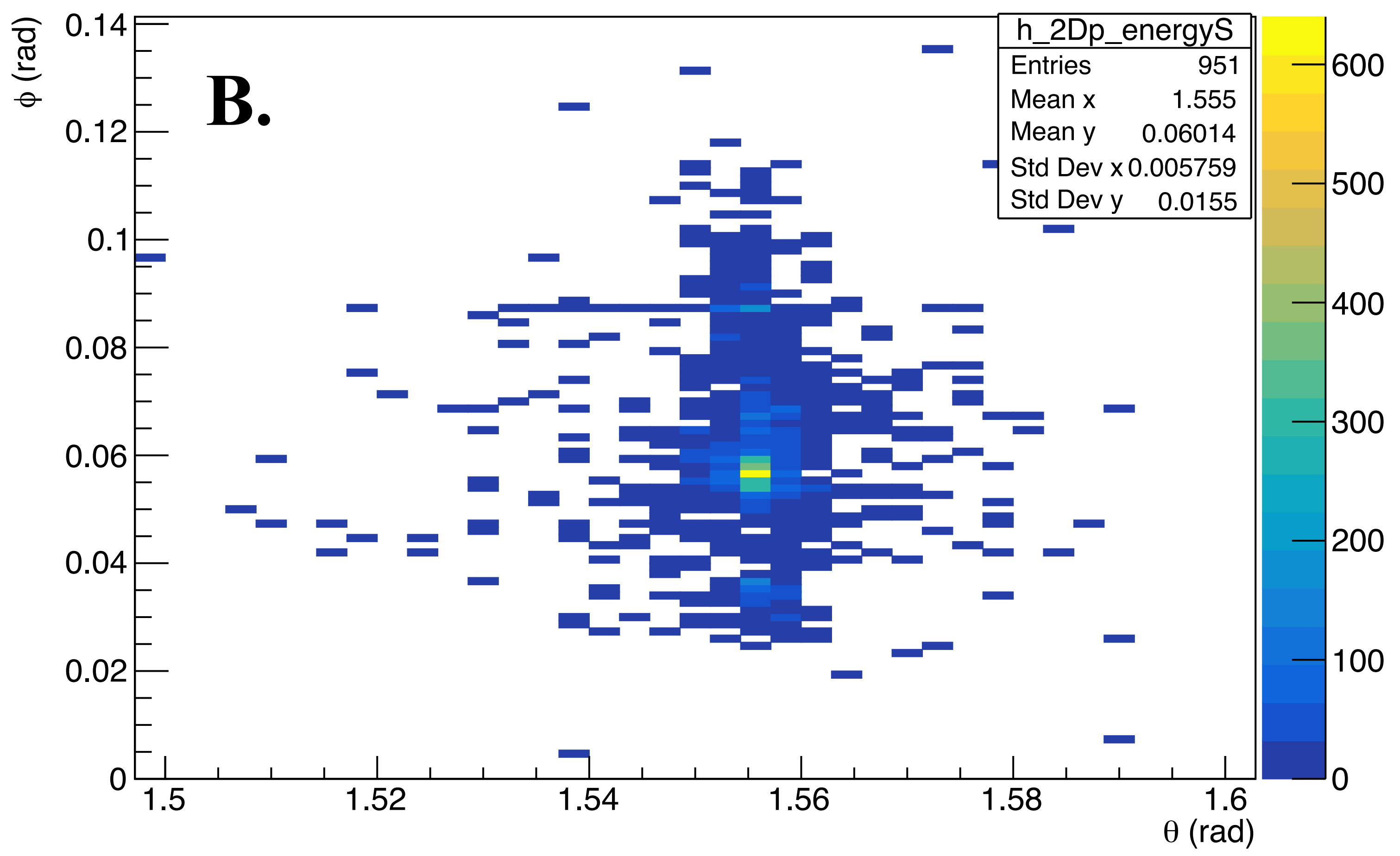
C. Energy deposited in the Cherenkov fibres, polar coordinates

Combined fibres

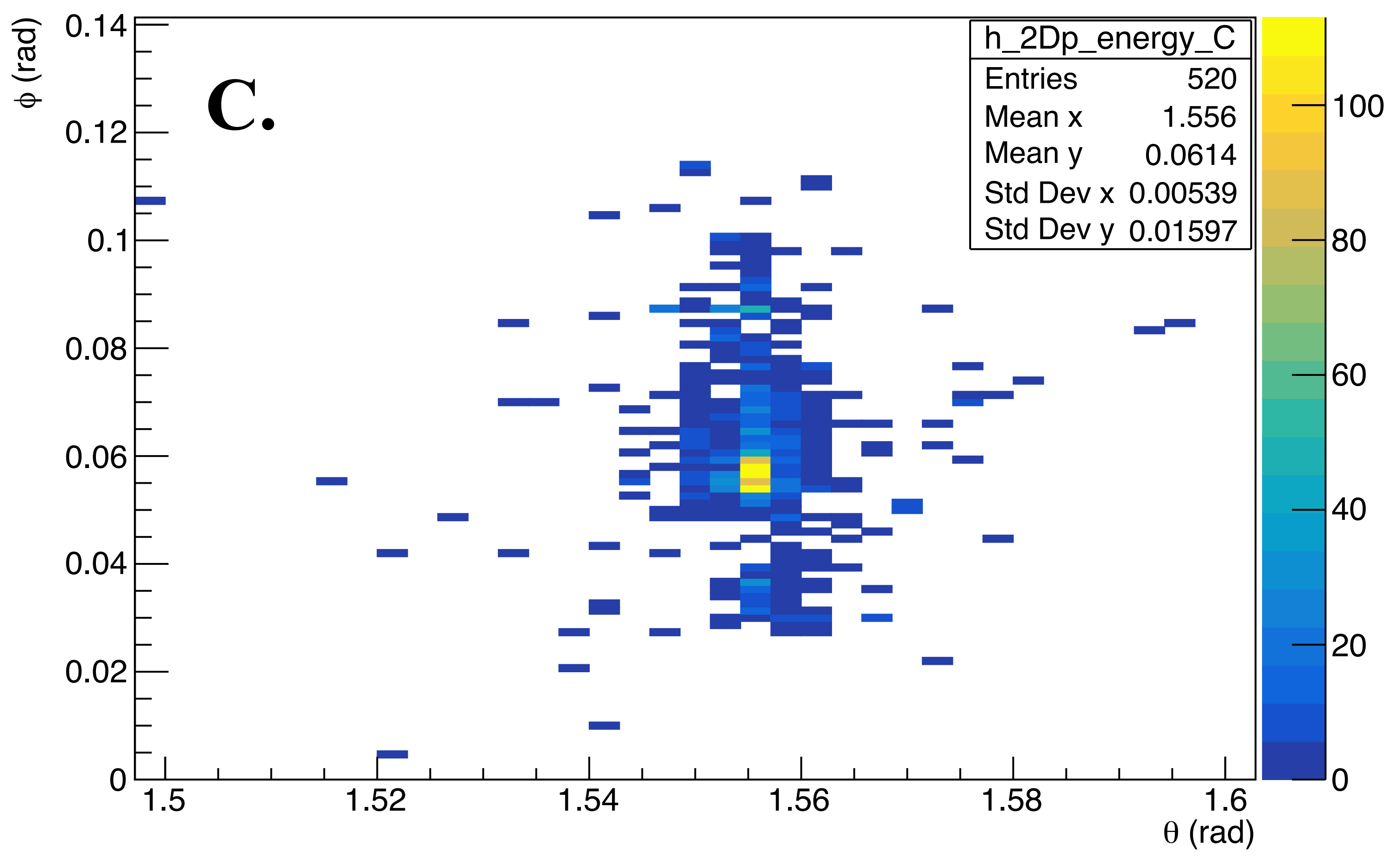


3D plot, cartesian coordinates

Scintillating fibres



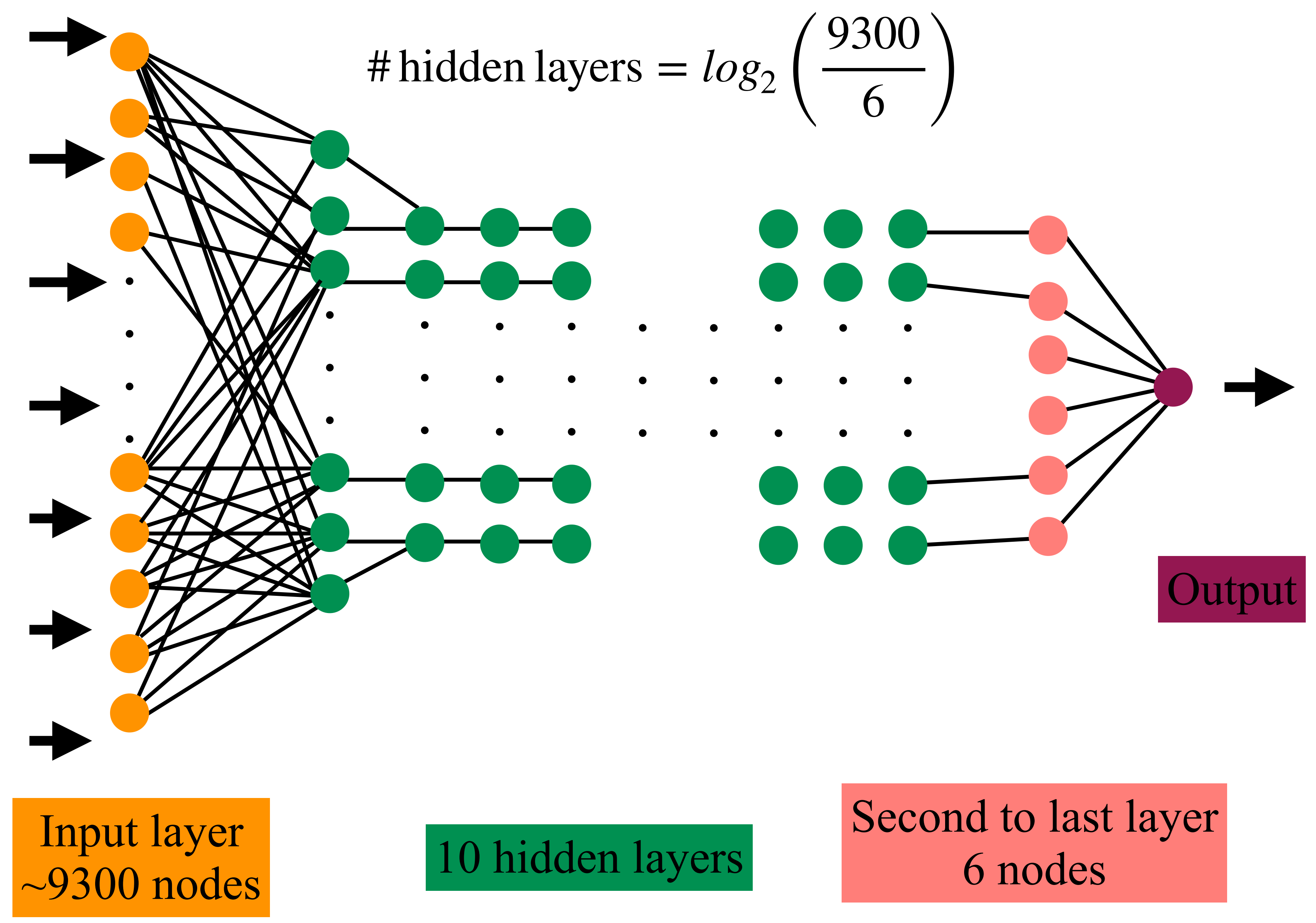
Cherenkov fibres



2D plots, polar coordinates

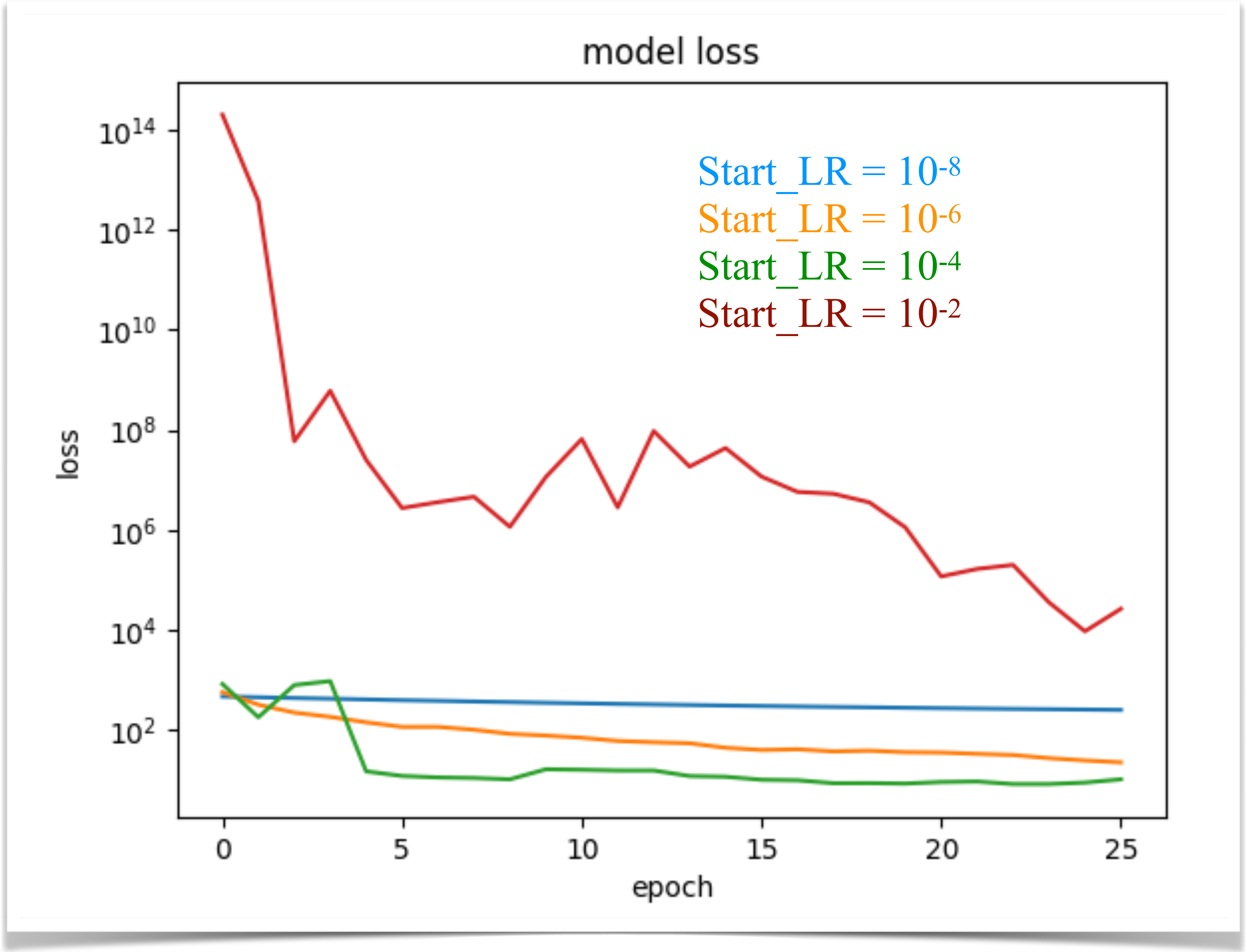
NN training using Tensorflow on GPUs

- Tensorflow, interfaced with Keras, is used to build and train a NN on GPUs
- Inputs: energy and position of each hit in the shower generated by the impinging electron and recorded in both S&C fibres—> NN input: 6 kinematic variables (*E, x, y, z, t, flag*) times hit multiplicity (~ 9300 info per event, 100 simulated events used)
  - Maximum hit multiplicity: ~1500 per event
  - Zero padding approach: if the number of hits in the event is less than the max hit multiplicity, set to zero the remaining positions in the array
- # initial nodes = # input info
  - Exploit the average hit multiplicity * 6 kinematic variables as #initial nodes to reduce the complexity of the problem
- # hidden layers = 10
- At each layer the number of nodes halves



NN training using Tensorflow - Model

- Model loss: MeanSquaredError(), $\frac{1}{n}\sum_{i=1}^{n}\left(y_{\text{true}} - y_{\text{pred}}\right)^2$, optimised with respect to the simulated energy of the incoming electrons
- Adam, a stochastic optimiser, is used as optimiser to minimise the loss [Refence]
- Testing different START learning rate



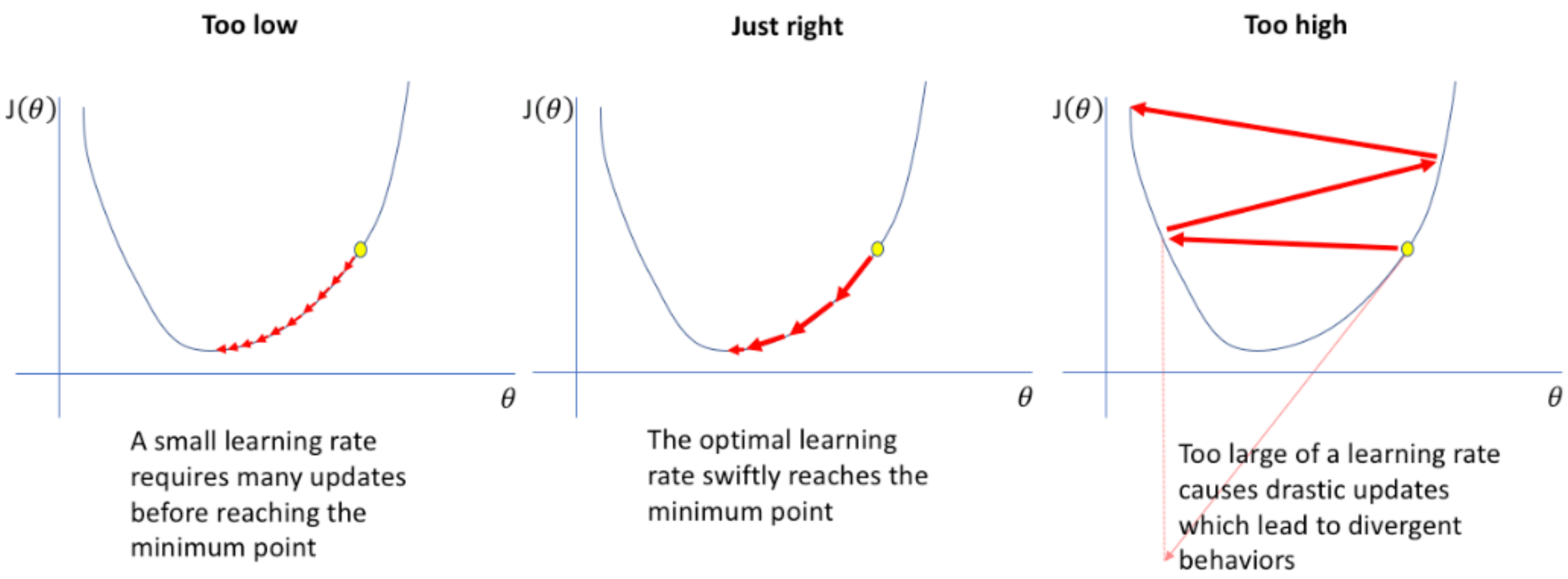
Conclusions and Next steps

- To do: Update NN plugging in angular variables
- To do: Train the NN on a newly simulated sample containing electrons with uniform energy (up to 125 GeV) and angular distributions, and info about the initial spatial coordinates of the impinging electrons at truth level —> **Simulation done**
  - Perform the hyper parameter optimisation (*i.e.*: #layers, #epochs)—> **Work in progress**
- Determine the energy and position resolution from NN, for electrons
- Repeat the above procedure also for π, K, μ, γ
- Plan to move to Pytorch for better optimisation with Pandora
- Long term goal: NN-based particle identification and jets reconstruction

Thanks a lot for listening!

Back-Up Slides

Learning rate



Simulated events:
e^+e^- —>$Z(\nu\nu)H(\gamma\gamma)$

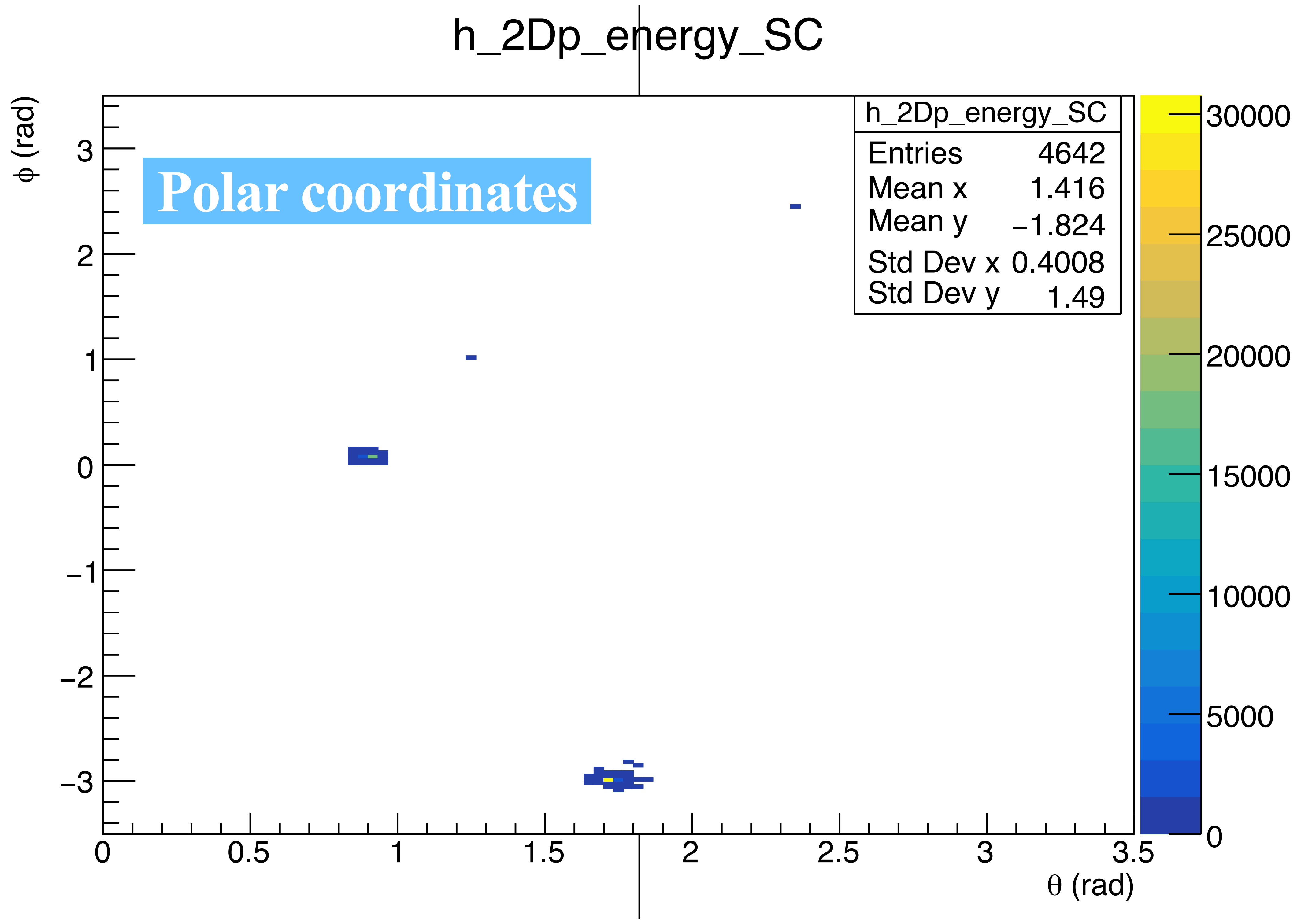
Energy deposits - e^+e^- —>Z(νν)H(γγ)

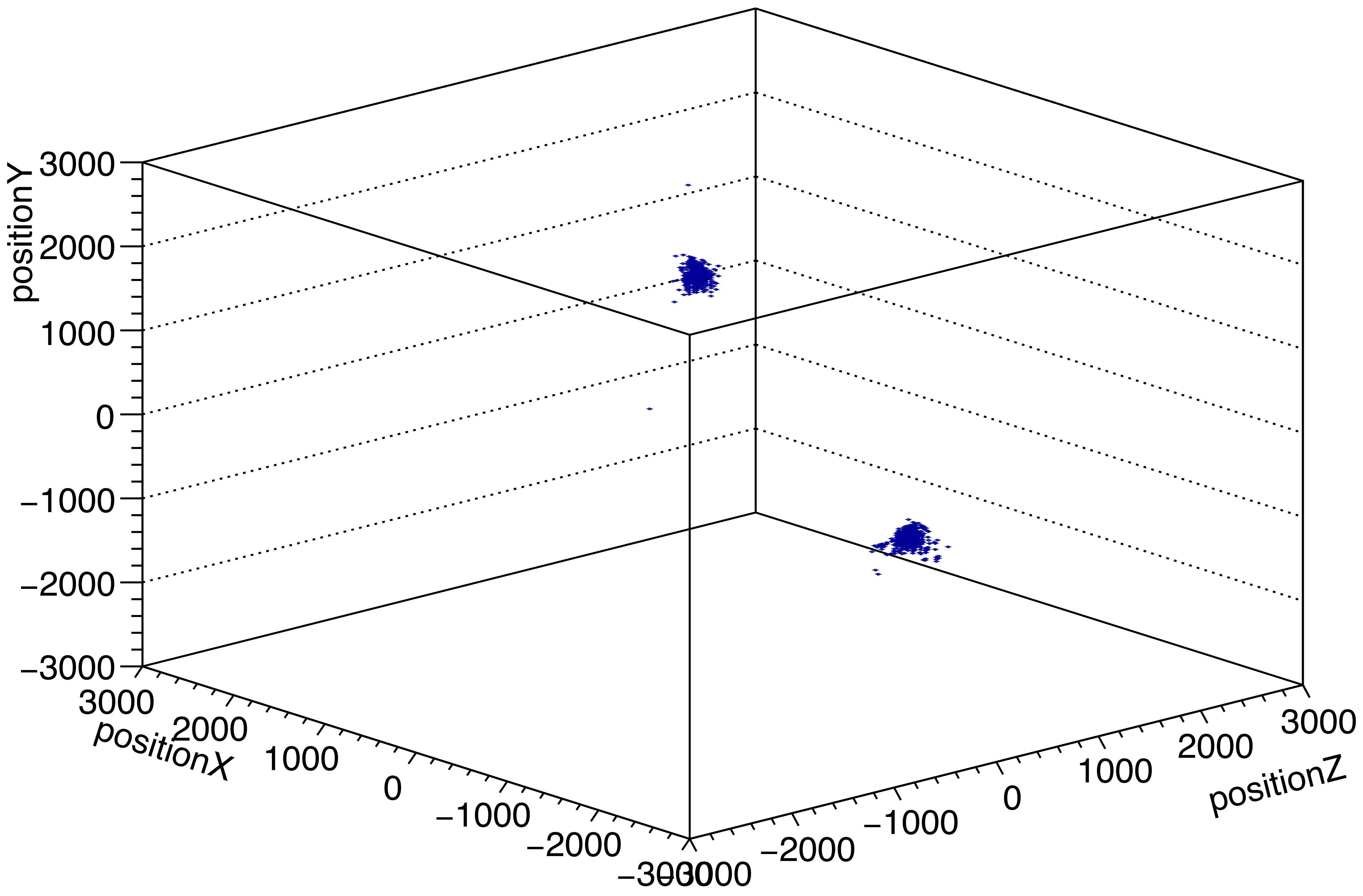
Dumper Algorithm output

Combined fibres

Photon deposits

Input file: EDMOutput_Higgs.root



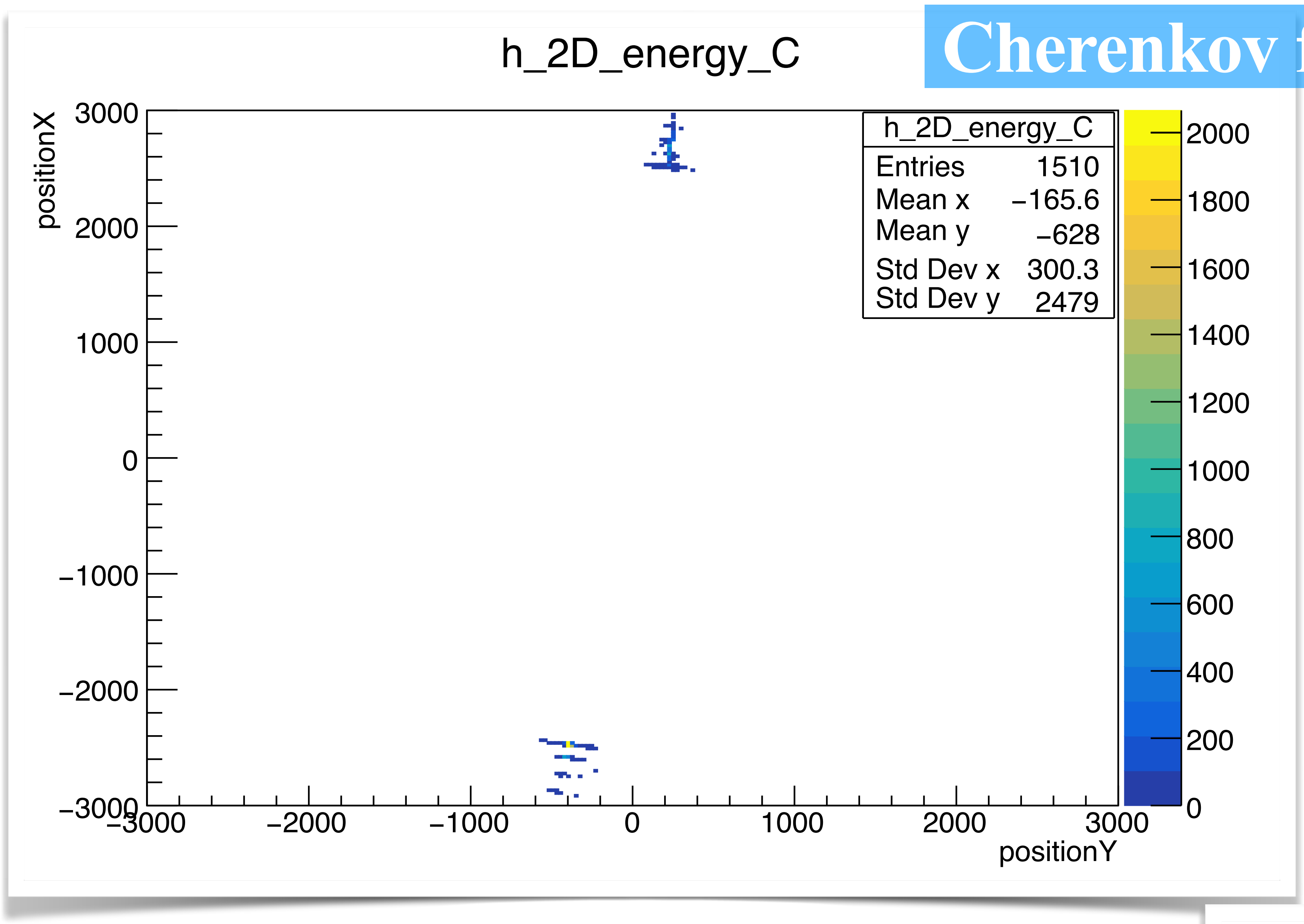


3D plot, cartesian coordinates

Energy deposits - $e^+e^- \longrightarrow Z(\nu\nu)H(\gamma\gamma)$

Dumper Algorithm output

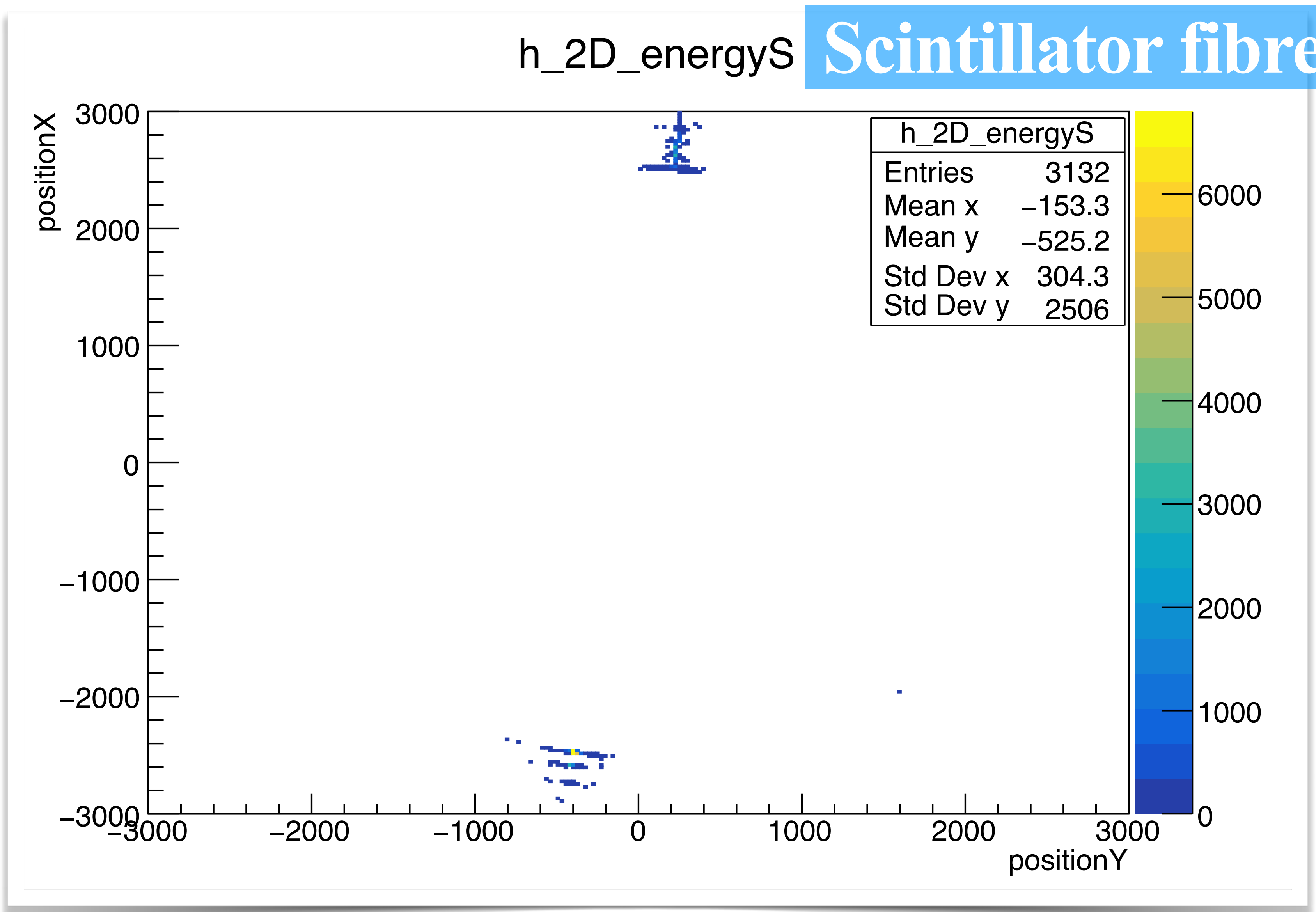
Cherenkov fibres



Photon deposits

Input file: EDMOutput_Higgs.root

Scintillator fibres



Combined fibres