20th International Workshop on Advanced Computing and Analysis Techniques in Physics Research

A novel ML approach for the reconstruction of particle showers with a tracking detector

Elena Graverini and Paul de Bryas

30.11.2021

Contact: paul.debryas@epfl.ch, elena.graverini@epfl.ch

Introduction

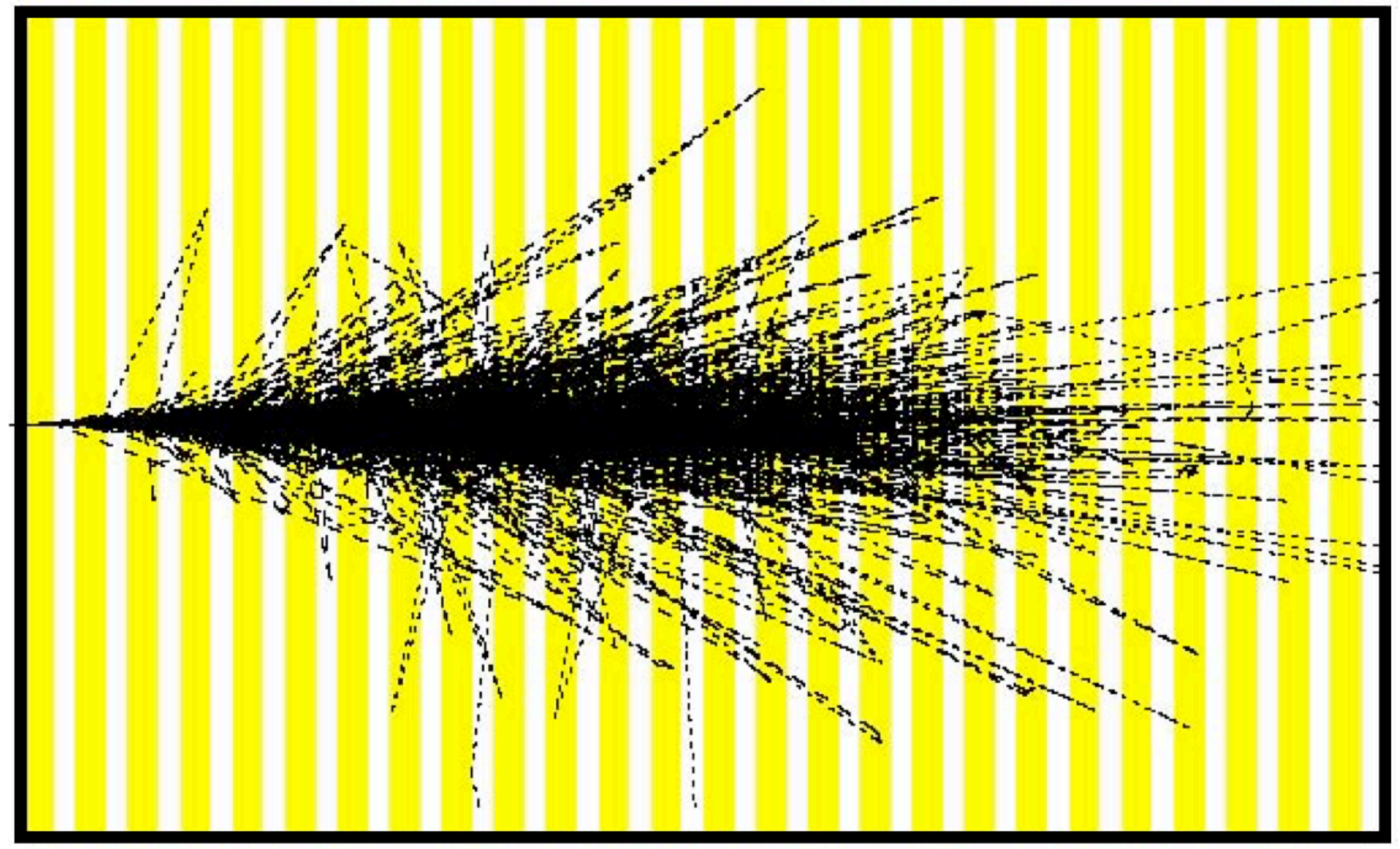
- We will present an approach using Convolutional Neural Networks for the reconstruction of particle showers using informations from a high-granularity **tracking** detector

- ML allows you go beyond calorimetry, also performing tracking

- We are developing this technique to improve the reconstruction of neutrinos at the SND@LHC experiment

- Being able to perform real-time calorimetry adds a lot to the SND@LHC physics case

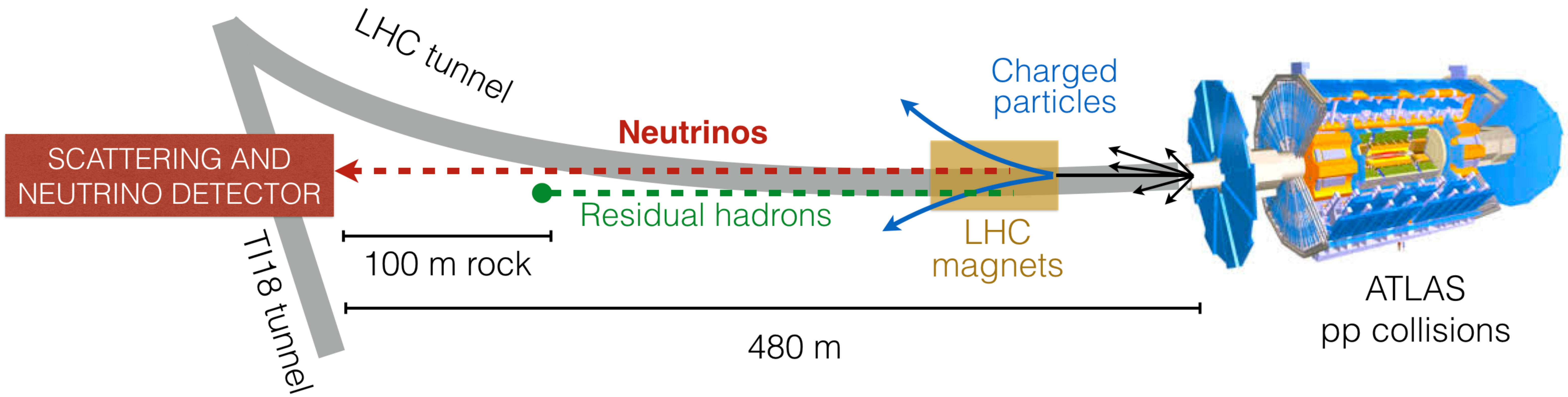
- All of this is an ongoing exploratory work, for the time being.

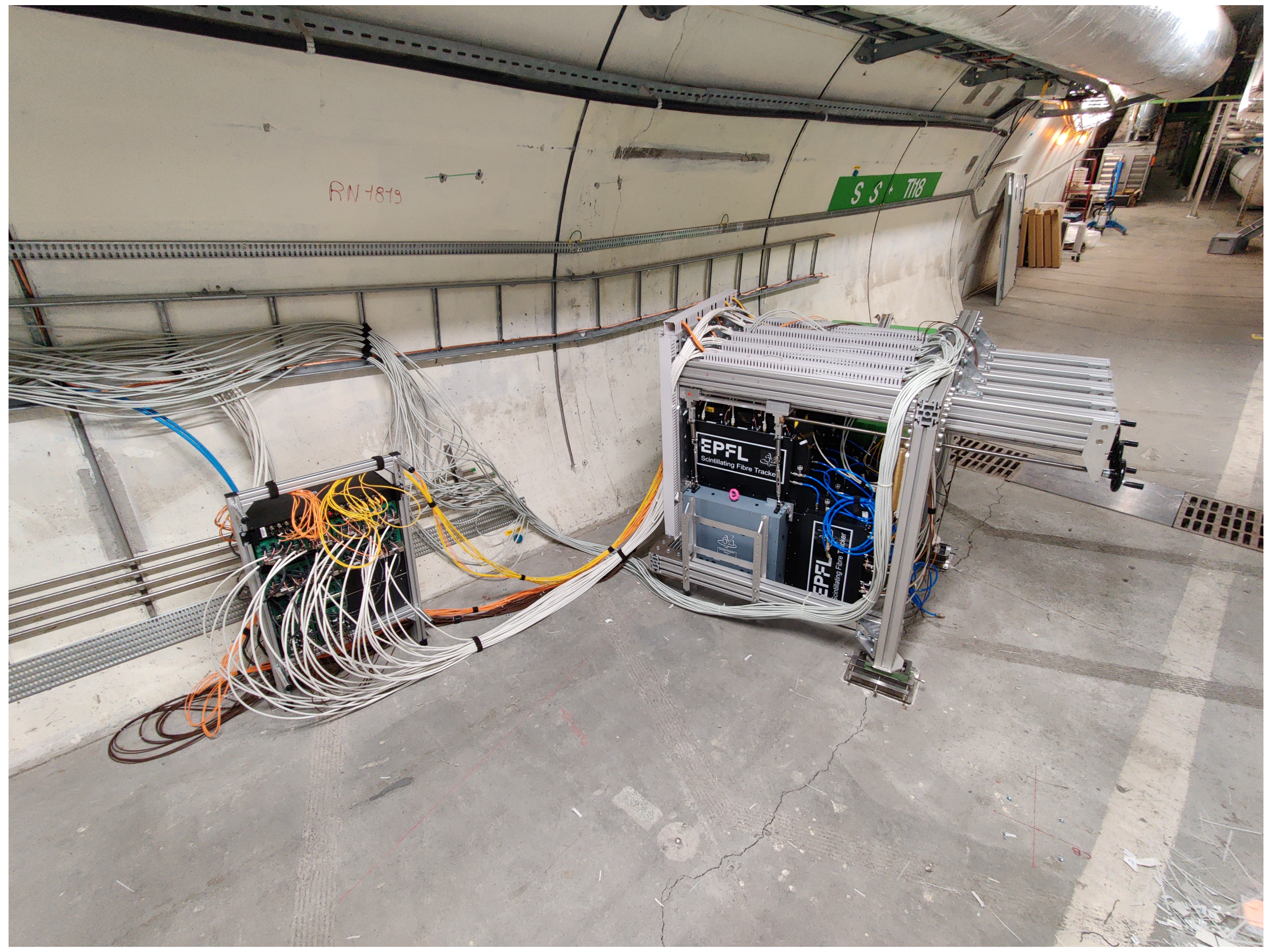


Example of a particle shower

Scattering and Neutrino Detector @ LHC

- SND@LHC is a newly approved experiment at the Large Hadron Collider (LHC)
- Its objective it is to study neutrinos of all flavours produced at the ATLAS interaction point, measuring their cross-sections in the GeV-TeV range for the first time
- Detector can also probe light dark matter scattering signatures
- The tracker of this detector is built at EPFL

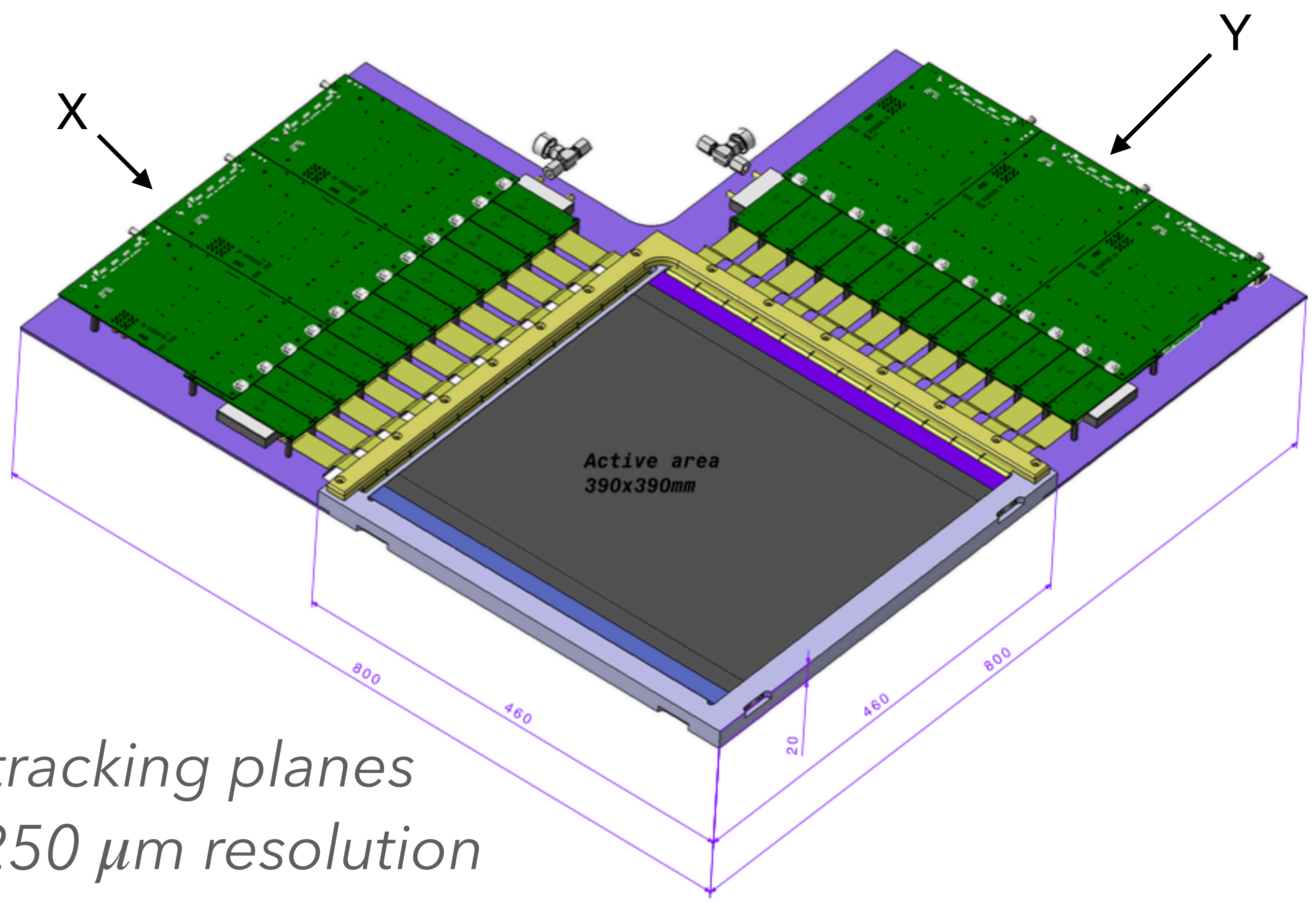




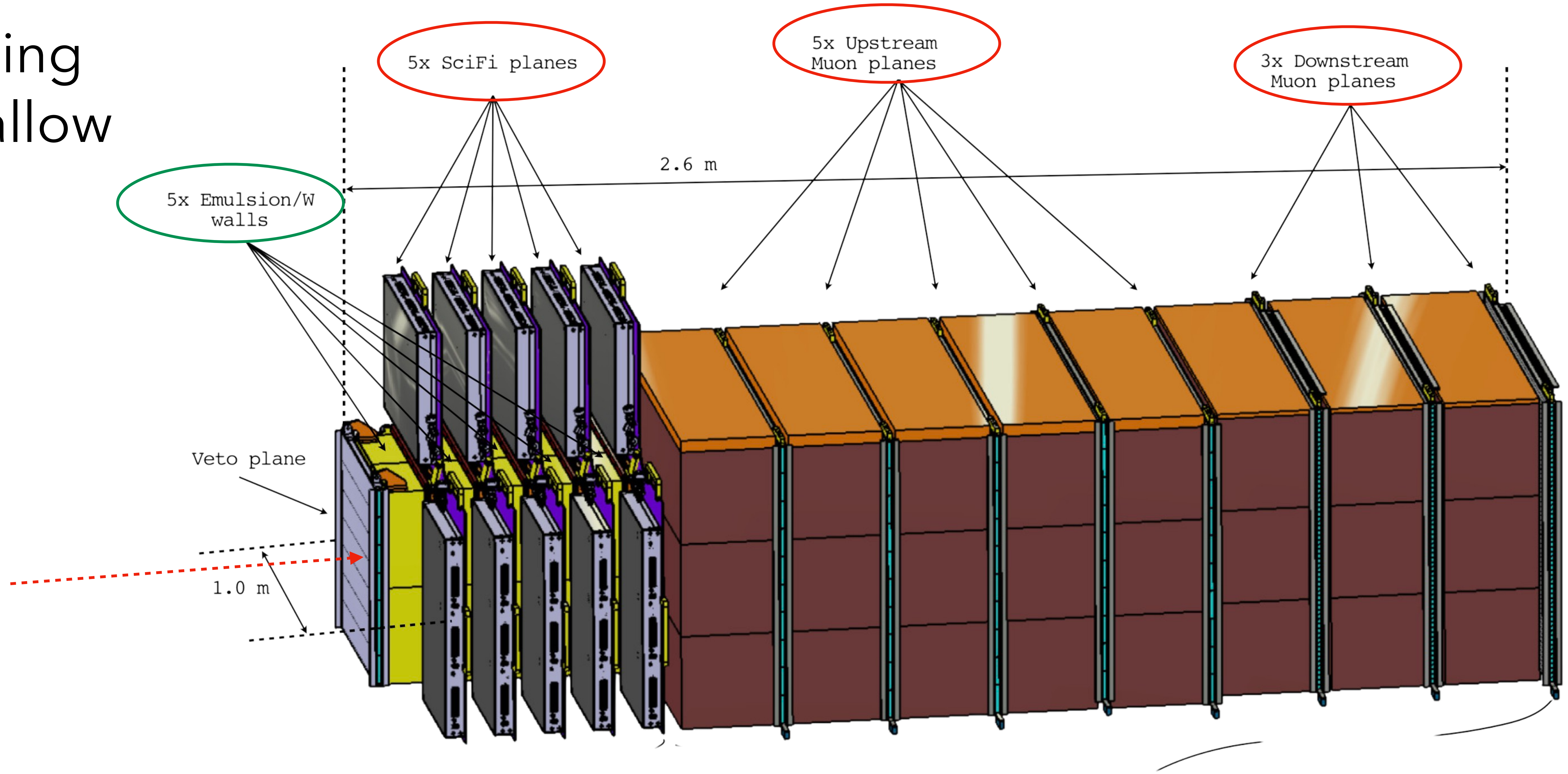
SND@LHC's installation undergoing in TI18 tunnel

- Detector Layout:
  -target region: Emulsion walls (tungsten plates interleaved with nuclear emulsion films) combined with scintillating fibre (SciFi) tracking planes
  -Muon Identification system: iron plates interleaved with scintillating bars.

- Emulsion films are taken out for developing every few months while the SciFi layers allow real-time event analysis.

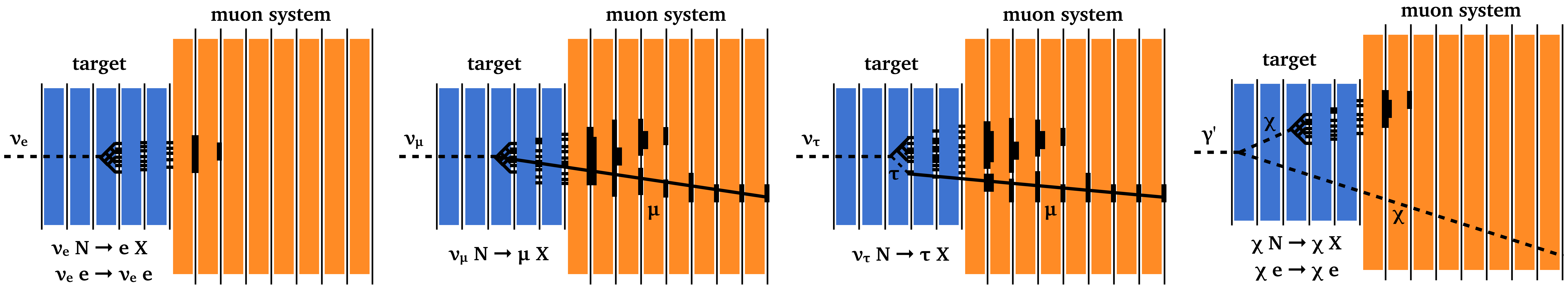


SciFi tracking planes with 250 µm resolution



SND@LHC - Scattering and Neutrino Detector at the LHC

Motivation

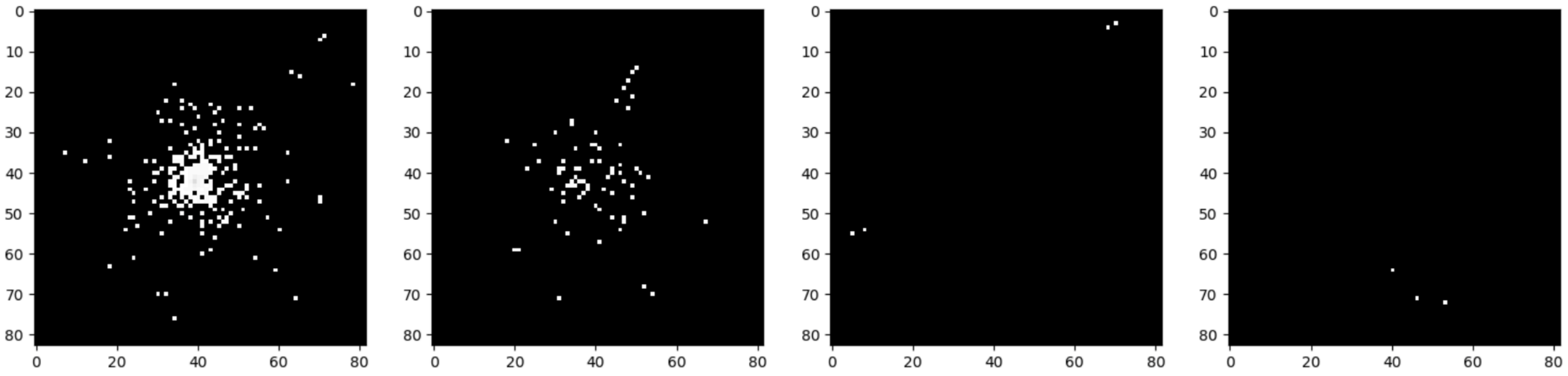


Topology of the events we will be studying

- For real-time event analysis, emulsion walls act as passive materials and SciFi planes behaves as the active layers of a sampling calorimeter
- The goal is to use the information of the SciFi tracker to perform prompt analysis
- Overview of the talk:
  -Energy reconstruction in EM showers (feasibility study)
  -How to deal with ghost hits
  -Energy reconstruction in case of neutrino scattering (EM + HAD showers at SND@LHC)
  -Future Neutrino flavour tagging

Why Machine Learning?

- Shower energy resolution achievable by classical methods (counting hits) is ~22%
- It does not provide flavor tagging, nor use the topological information from the shower, nor from the muon detector
- Feasibility study convinced us to take this direction:
- Objective:
  - Measure the energy of EM shower in the energy range 0-100 GeV.
- Particle gun electron:
  - electrons shot on to the center of the first plane



Example of inputs: here the incoming particle is a 64 GeV electron.

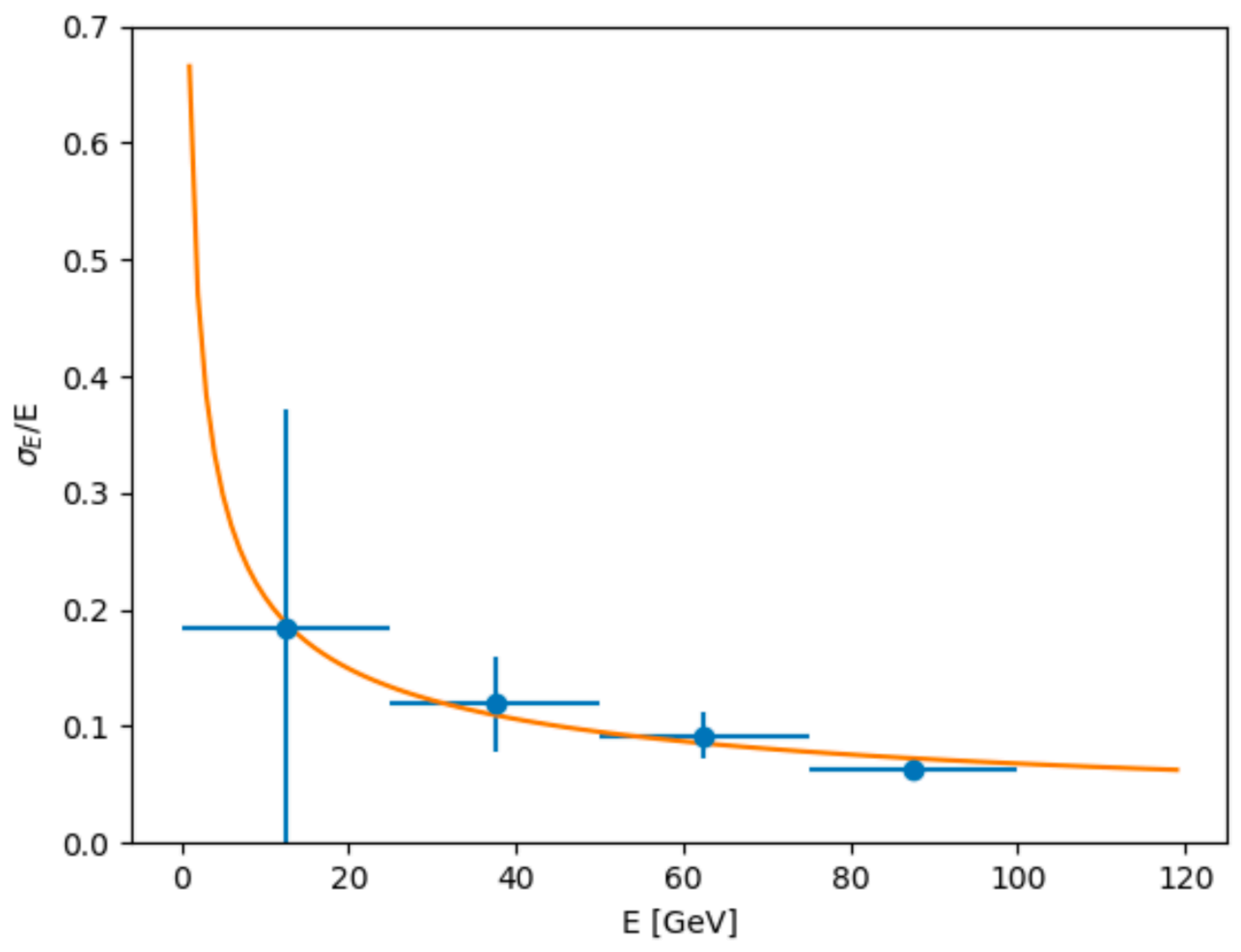
Feasibility study

- Procedure:
  SciFi hits (images) –> convolutional neural network (CNN) –> Energy

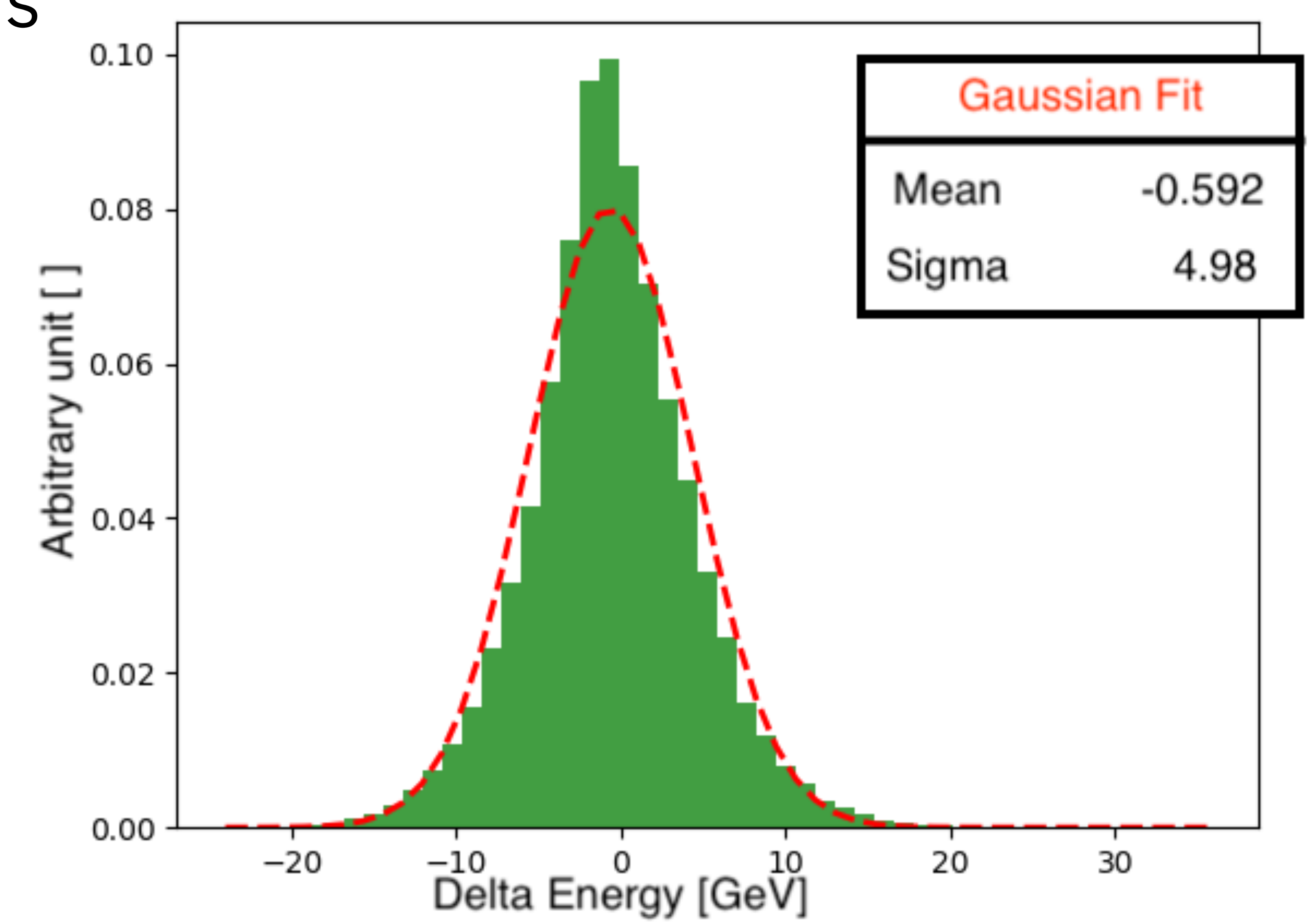
- Analysis of the detector response demonstrates that the target tracker planes behave as a sampling calorimeter.

- The CNN exhibits a resolution of 5% at E = 100 GeV and it is almost unbiased

- Show the feasibility of using such ML algorithm for real time analysis

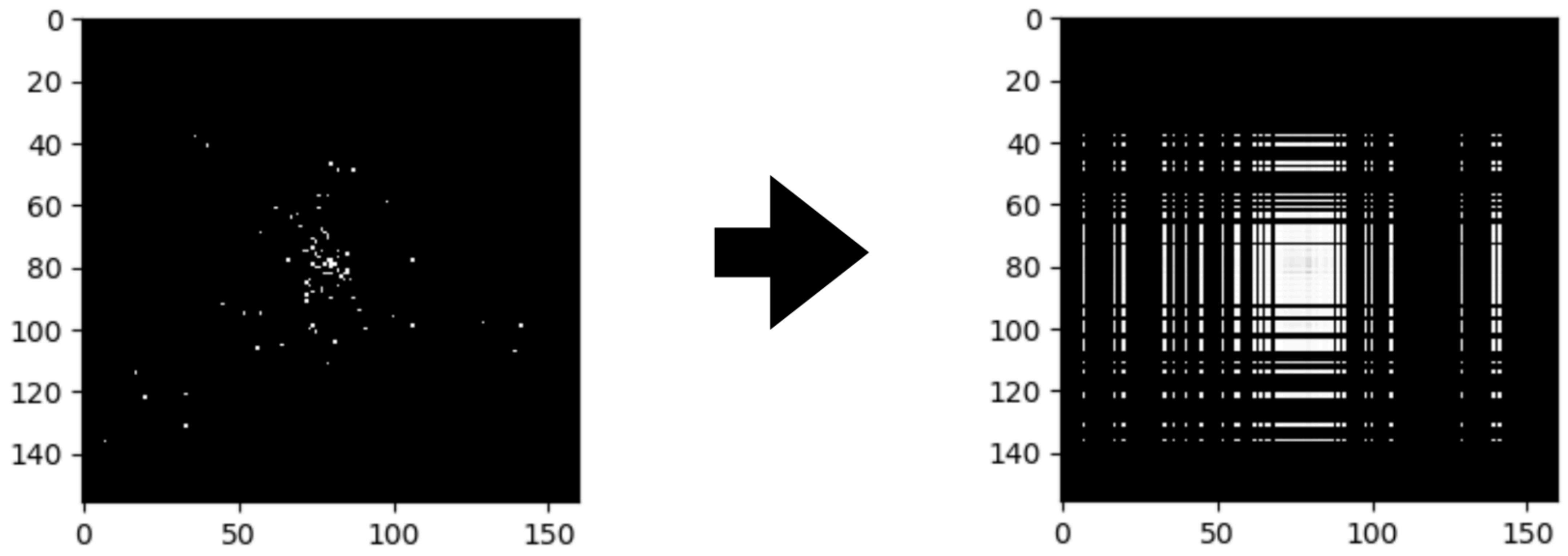


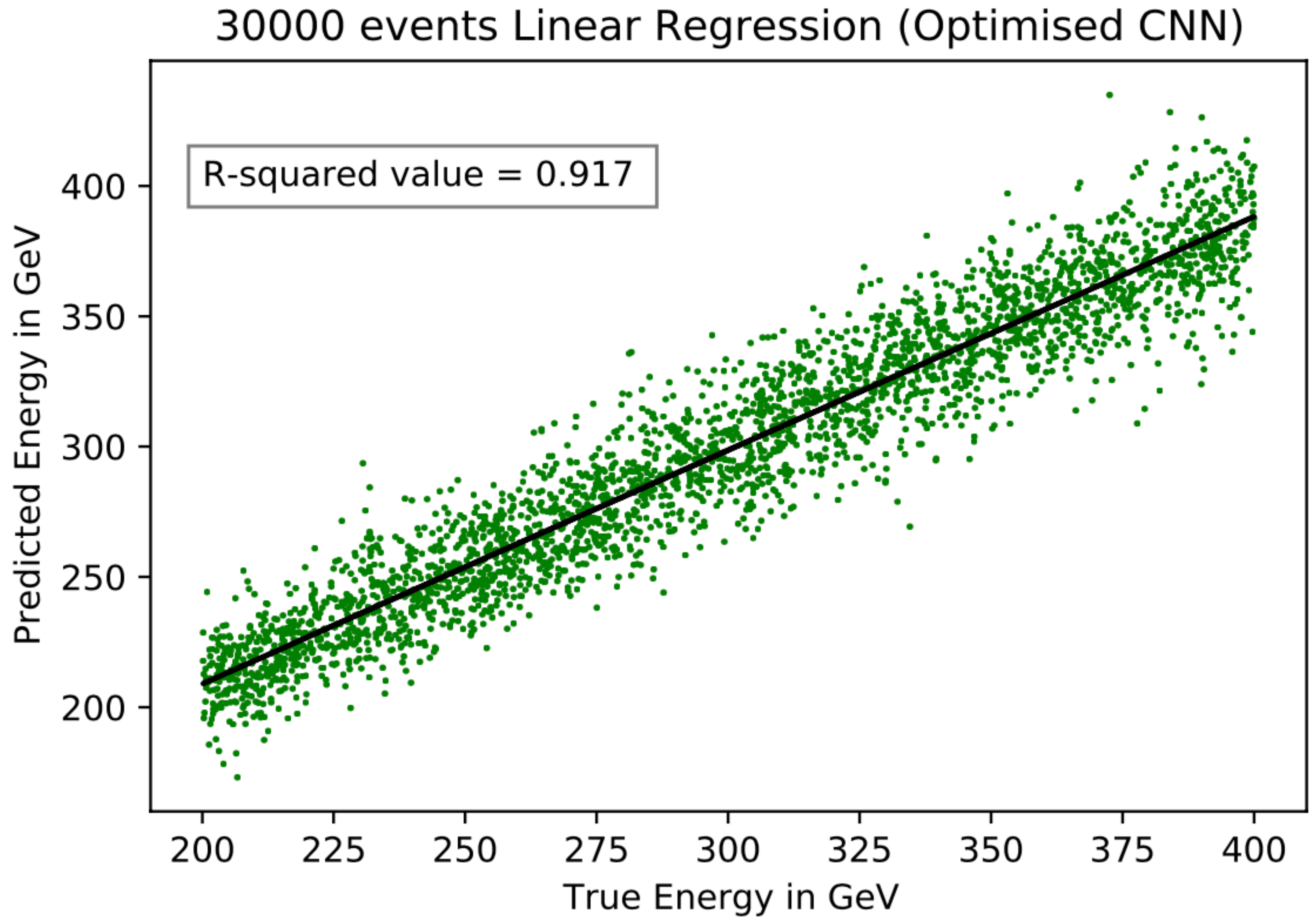
$$\frac{\sigma_E}{E} = \sqrt{\left(\frac{a}{\sqrt{(E)}}\right)^2 + \left(\frac{b}{E}\right)^2 + c^2}$$

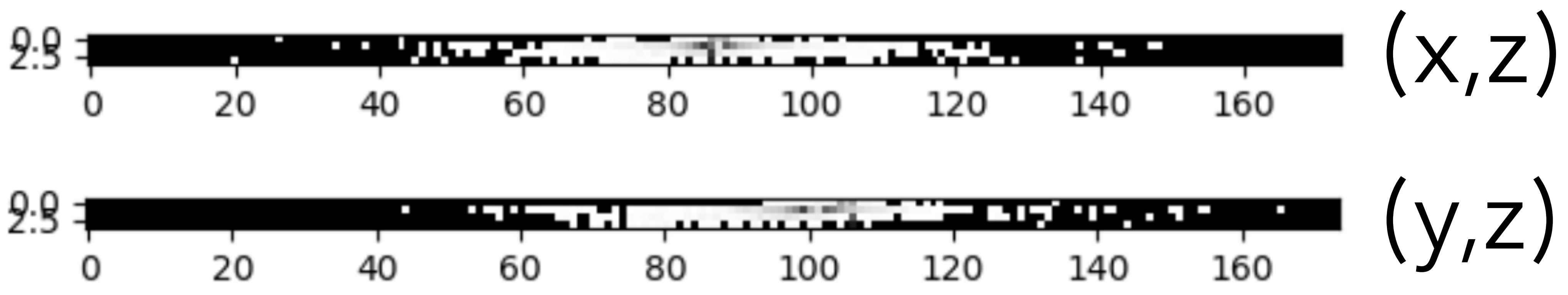


Ghost hit problem

- Explain the pb!
- The architecture of the CNN was changed in order to use only the (x,z) and (y,z) projections of the simulated hits on the target tracker.
- This new architecture exhibits an average fractional energy resolution of 5.7% with PG data.
- keep the same resolution with less information!



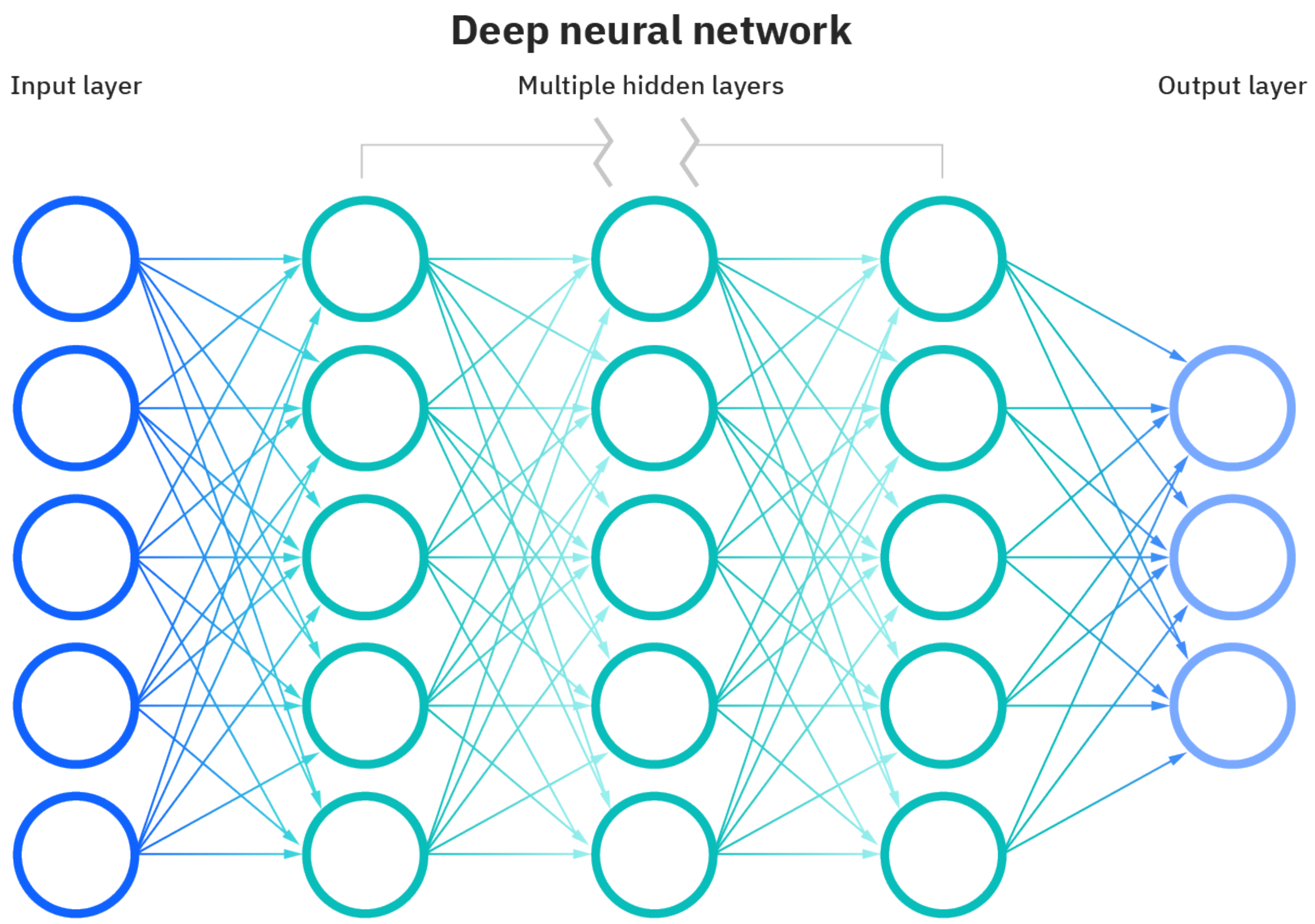


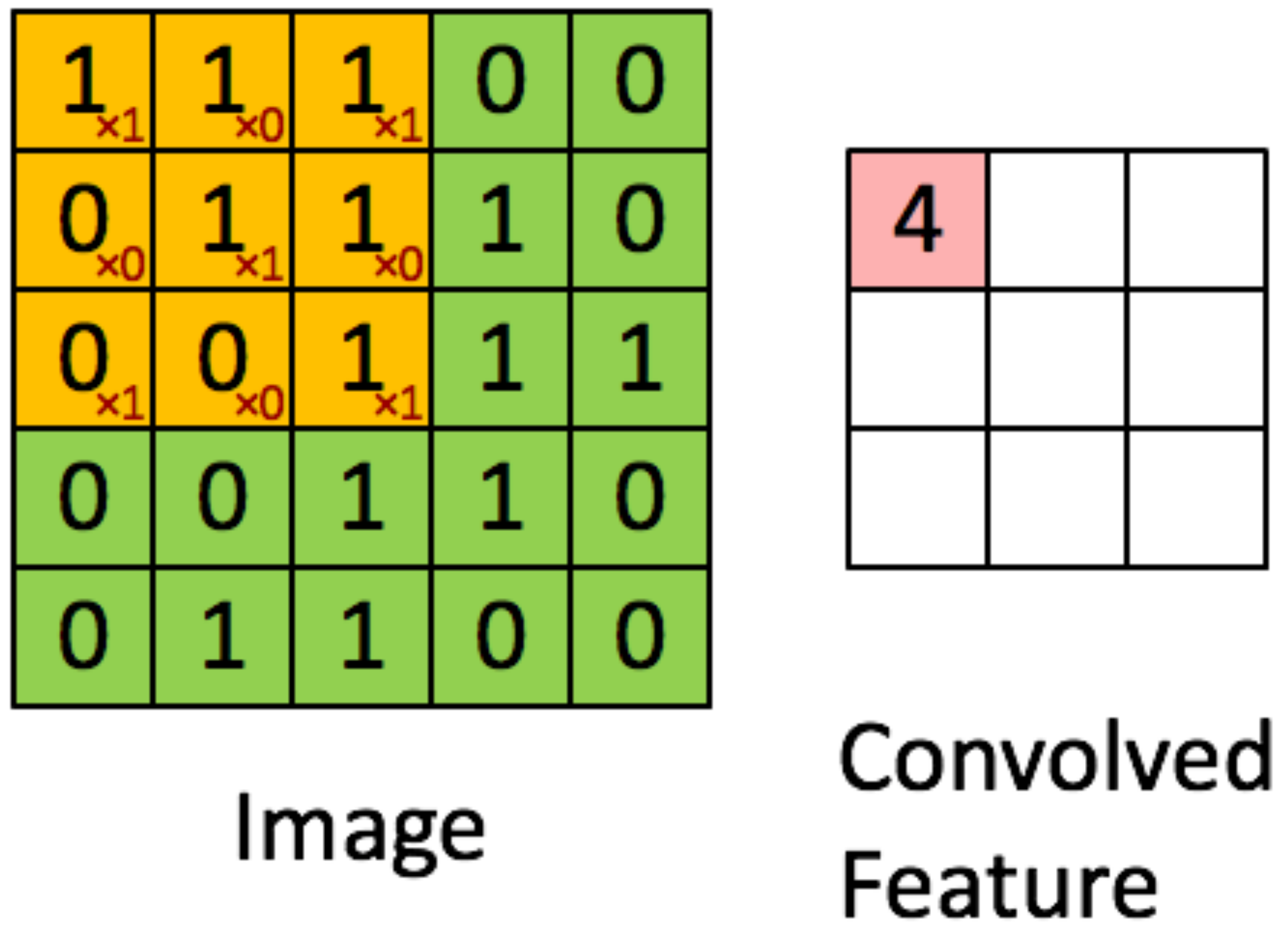


Inputs of the new CNN

Artificial Neural Network

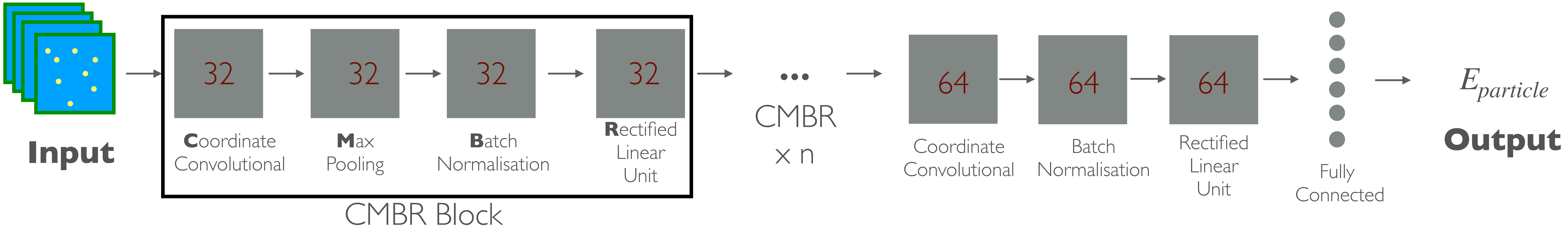
- Artificial Neural Networks algorithms are subset of ML
- They are comprised of one input layer, hidden layers, and an output layer.
- Each node of a layer is connected to all the nodes of the previous layer.
- Weights (–>) are the parameters of the NN
- CNN are a type of Artificial Neural Networks originally designed for large number of inputs (pixels)
- The convolution operation allows to reduce drastically the number of NN parameters and to keep track of the spatial information between adjacent nodes.





Convolutional Neural Network

CNN Architecture



Loss functions:

smooth_l1_loss

$$l(x,y) = \sum_n z_n$$

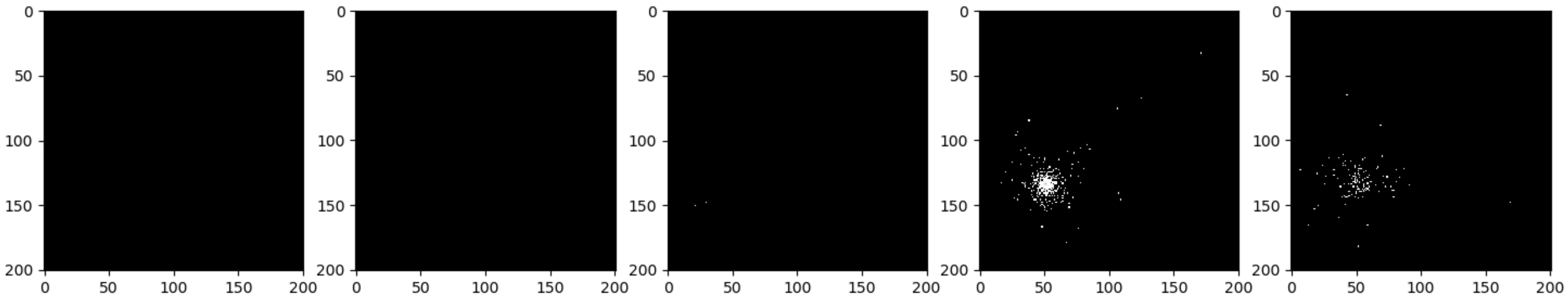
$$l(x,y) = \begin{cases} \frac{0.5\,(x_n - y_n)^2}{\beta} & if \;\; |x_n - y_n| < \beta \\ |x_n - y_n| - 0.5\beta & otherwise \end{cases}$$

- Our architecture is composed of n CMBR blocks where n should be proportional to the number of inputs.
- A CMBR block is a succession of 4 different operations:
  -CoordConv= Convolution operation but invariant under translation
  -MaxPool = take maximum value to reduce dimentionality
  -BatchNorm = normalisation to avoid very large value
  -ReLu = Removing negative values to increases the nonlinear properties of the loss function

Neutrino energy reconstruction

- Objective:
  - Measure the energy of an EM shower induced by a neutrino in the energy range 100-5000 GeV (expected at SND@LHC)

- Inputs:
  - ν_e neutrino with elastic scattering and charged-current deep inelastic scattering
  - no simulation of readout electronics;
  - hits are defined as a Yes/No signal in each pixel (no amplitude information)
  - true (X, Y) positions of the simulated hits



Example of inputs: here the incoming particle is a 767GeV ν_e.The shower induced by this particle produced hits (white dots) on the 5 SciFi planes

Neutrino energy reconstruction

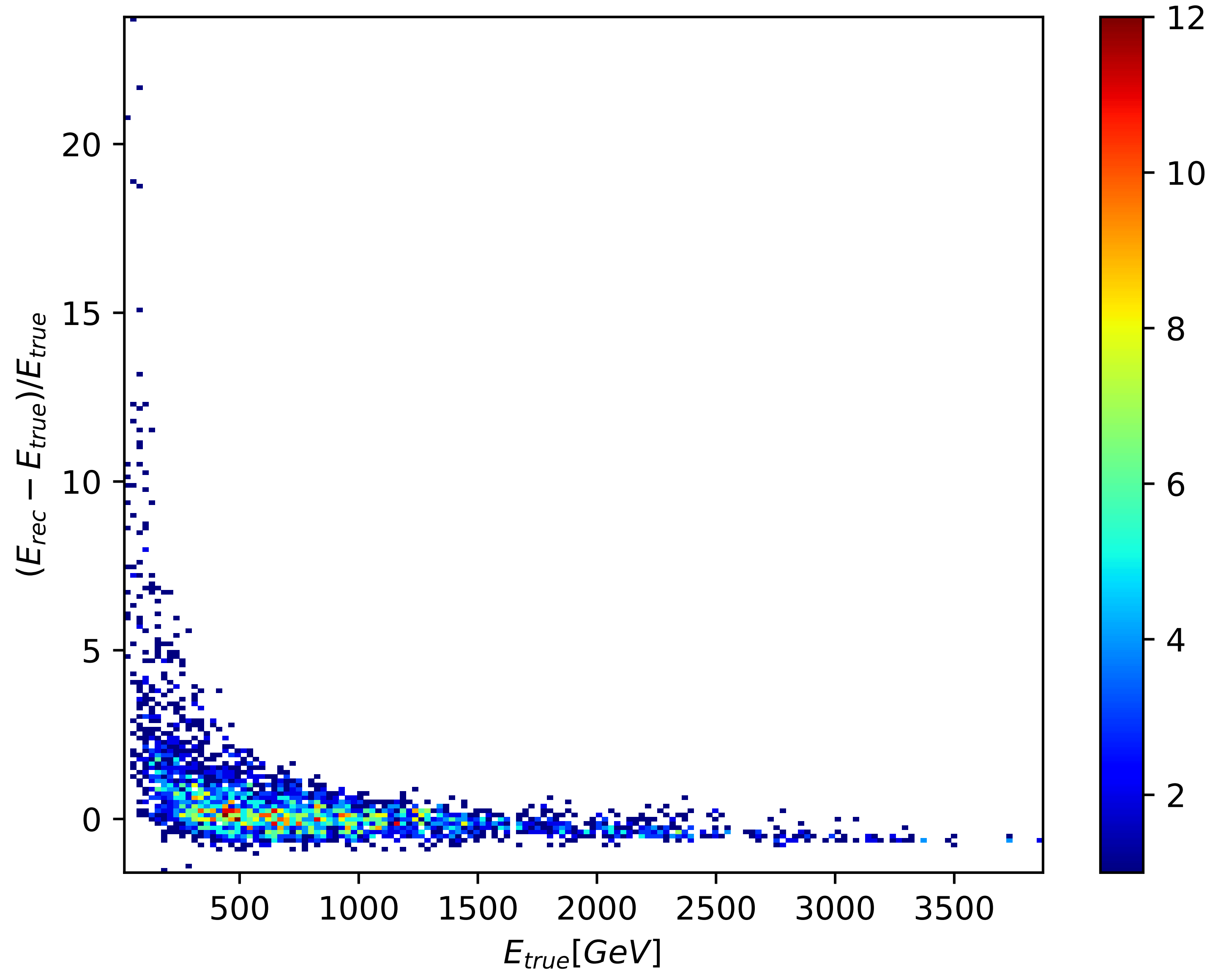
- Bias of ~9 GeV compared to 100-5000 GeV range

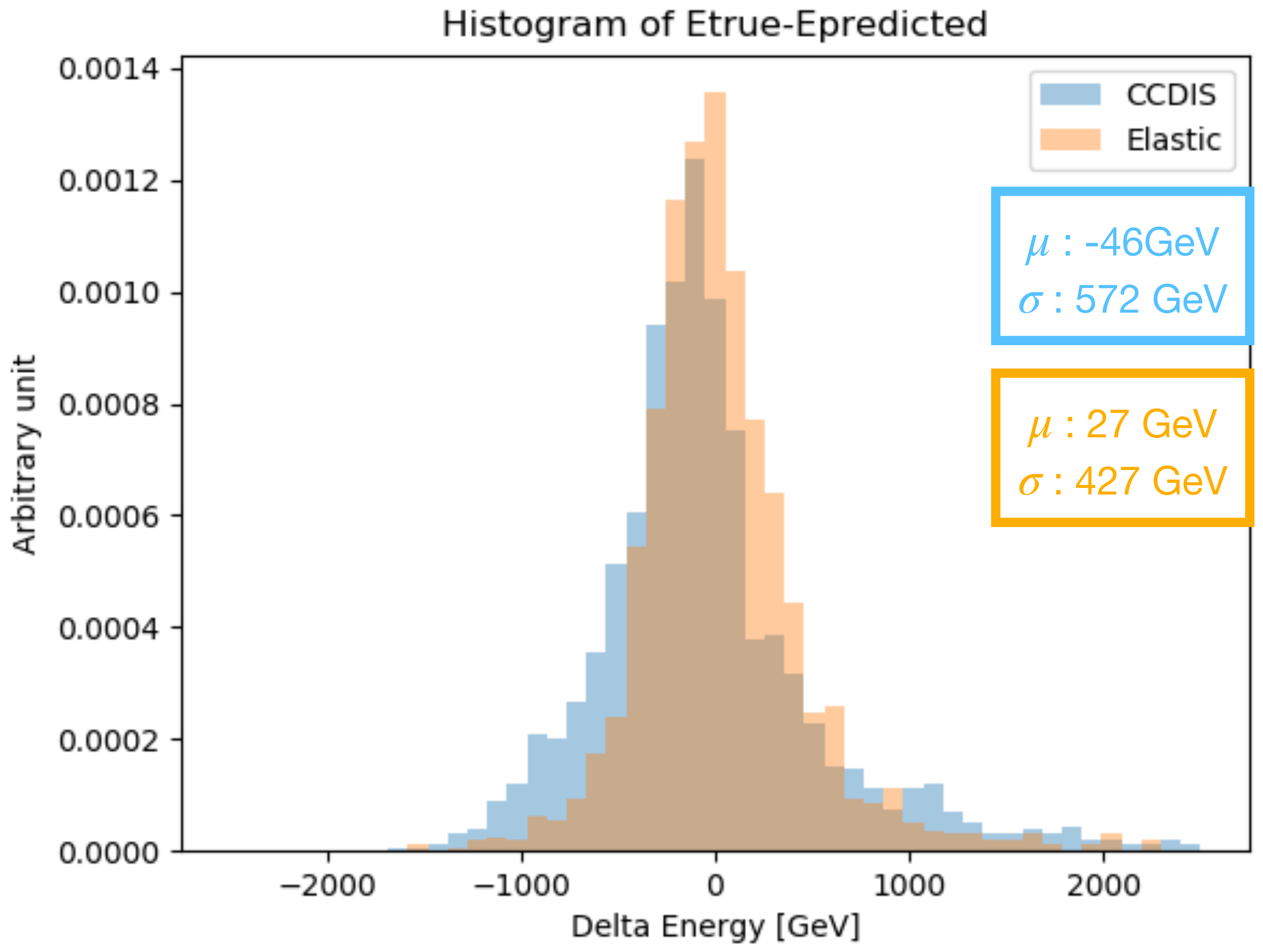
- Energy resolution around 500 GeV

- Degradation in performance is due to:
  -Shower generated at any depth in the detector
  -Shower sampling goes from $10X_0$ to $15X_0$
  –> only 2 planes have a significant amount of hits on average
  -Analysis has to be improved

- We separate the data samples (elastic scattering vs charged-current deep inelastic scattering) during the test of the CNN accuracy

- Elastic events seems to be better reconstructed



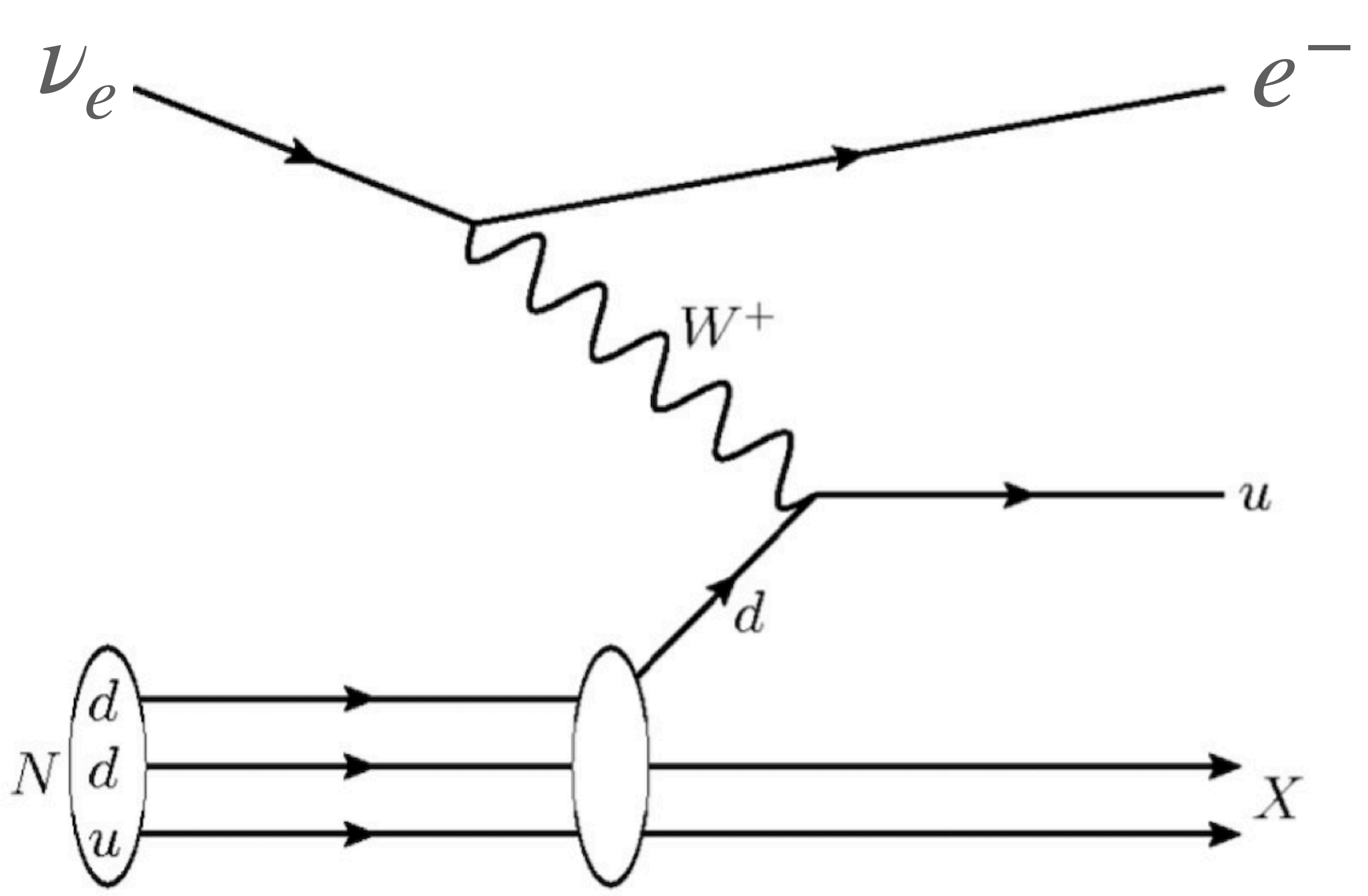


Classification of elastic and inelastic events

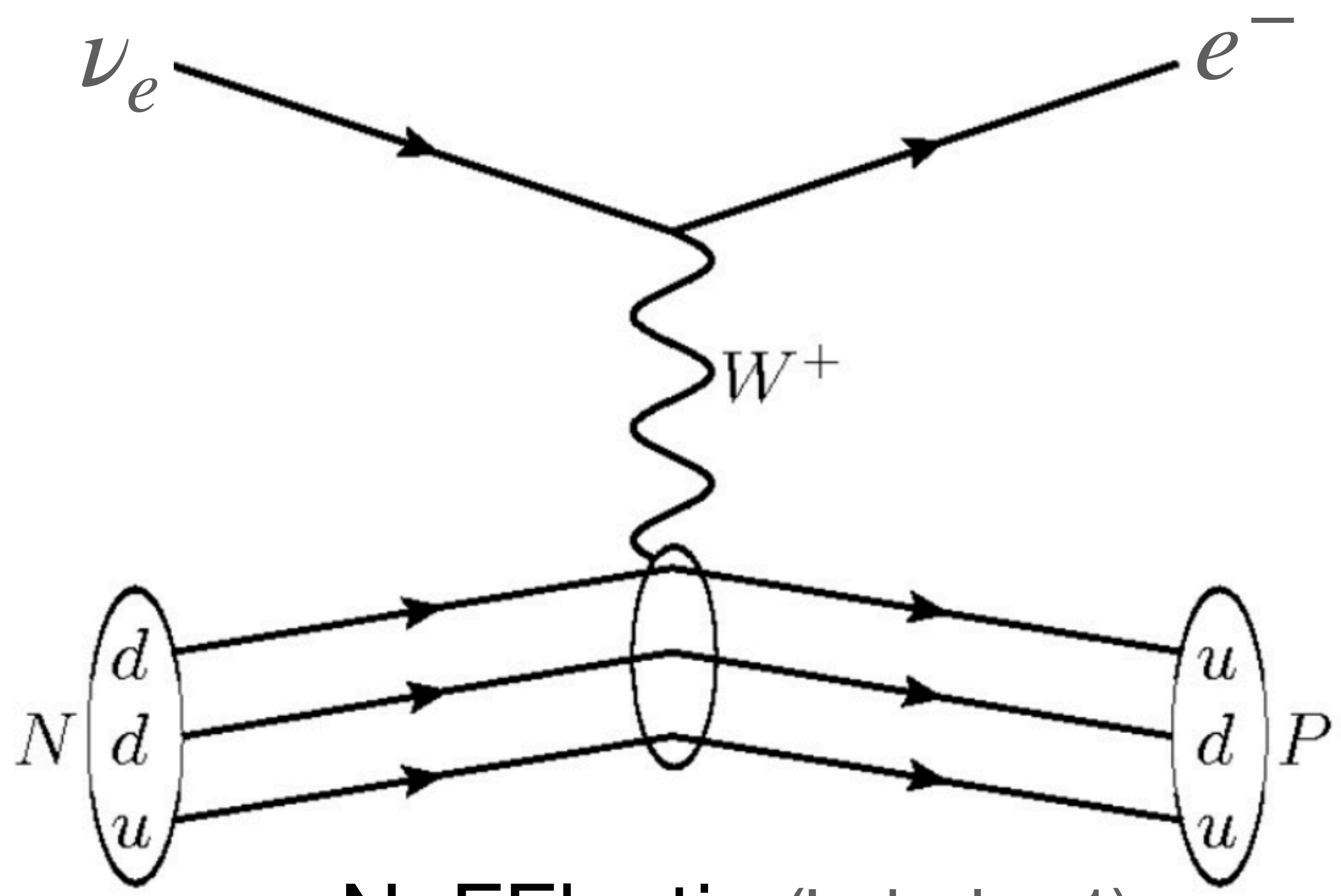
- Objective:
  -as a first step towards flavour tagging we try to performed classification of elastic and inelastic events

- The spatial distribution of hits can be used as a way of discriminating between elastic and inelastic scattering events.

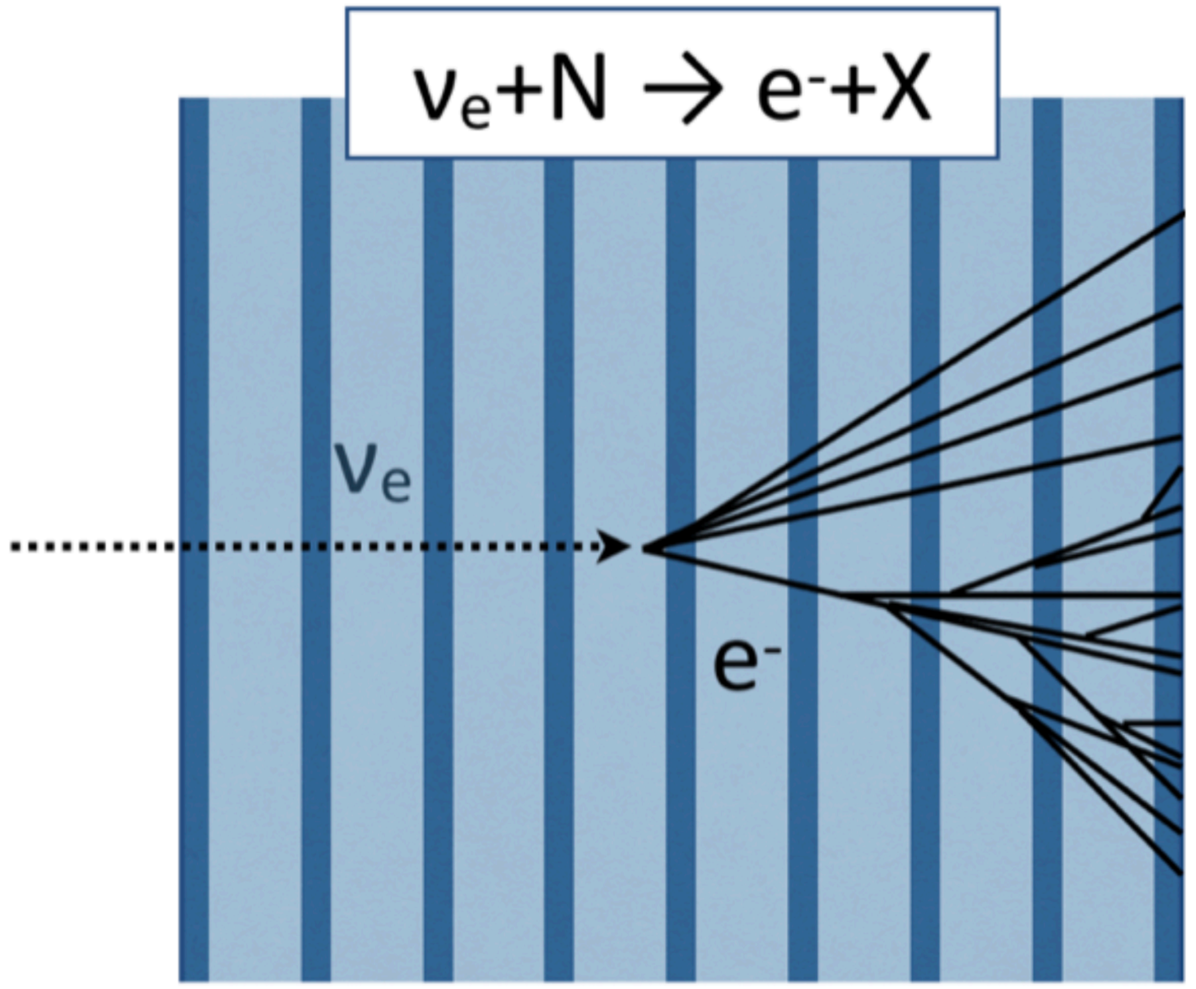
- The CNN used for the first study was modified to provide output label probabilities rather than a predicted energy value.

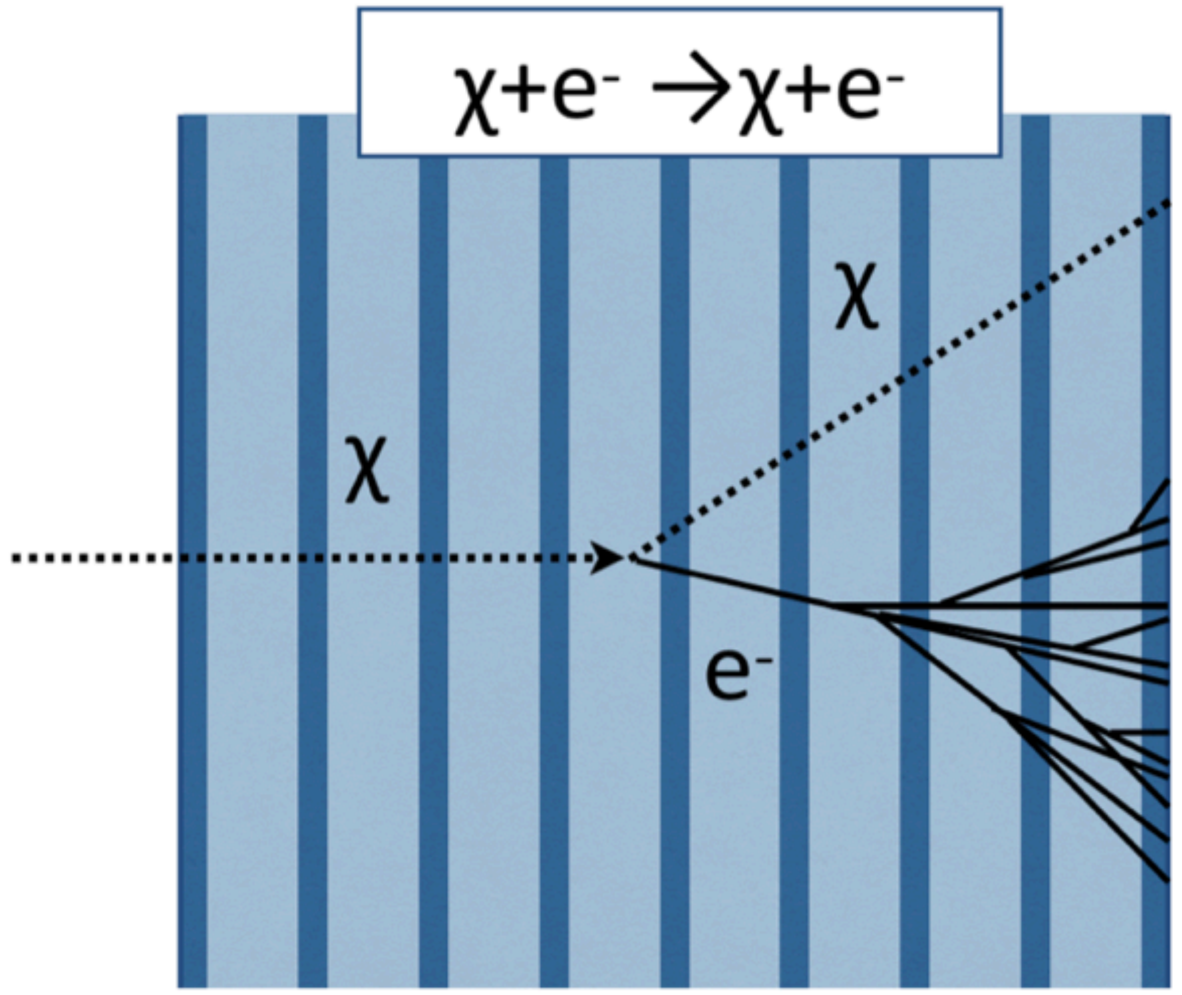


CCDIS (Label: -1)



NuEElastic (Label: +1)

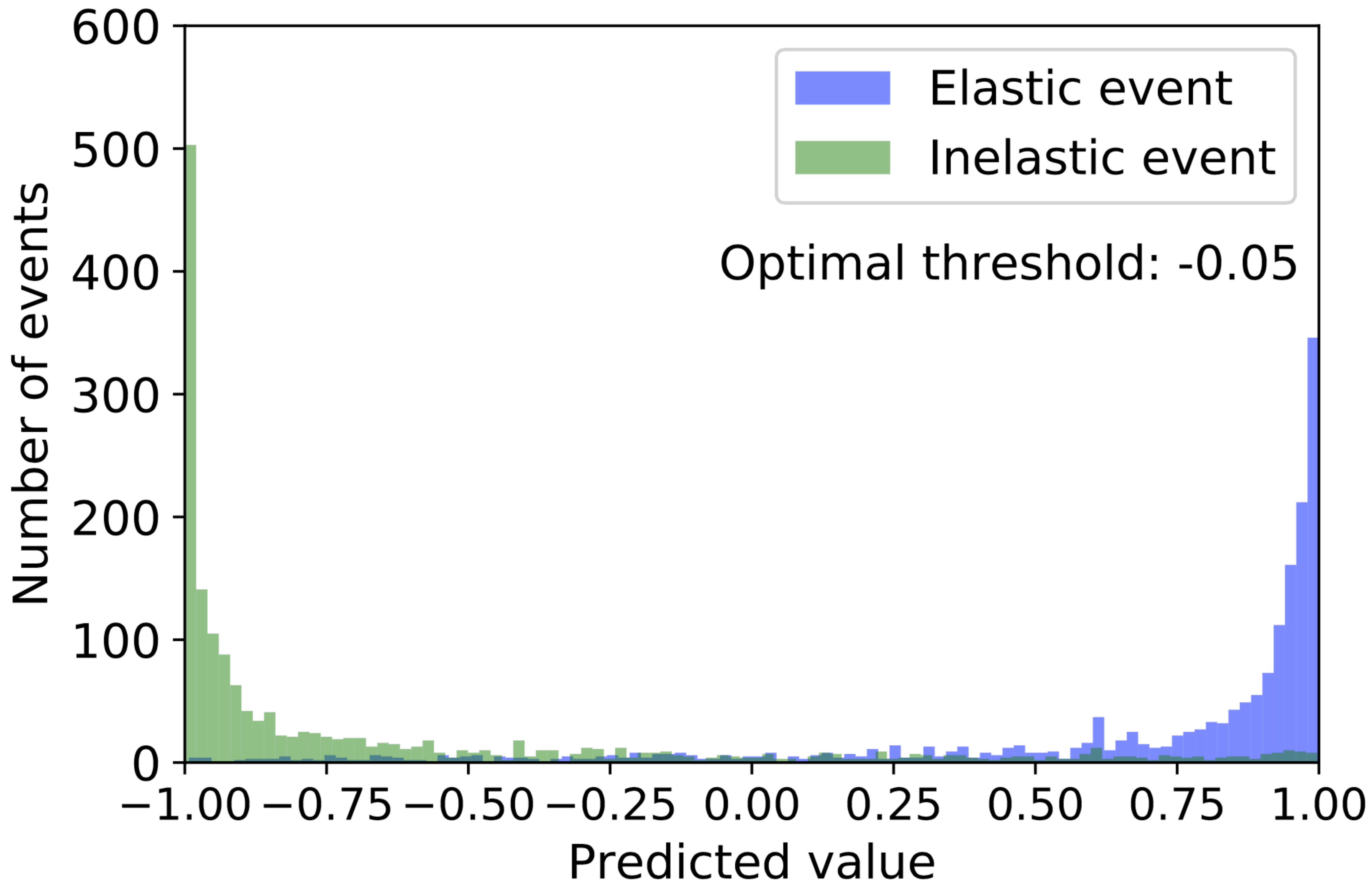


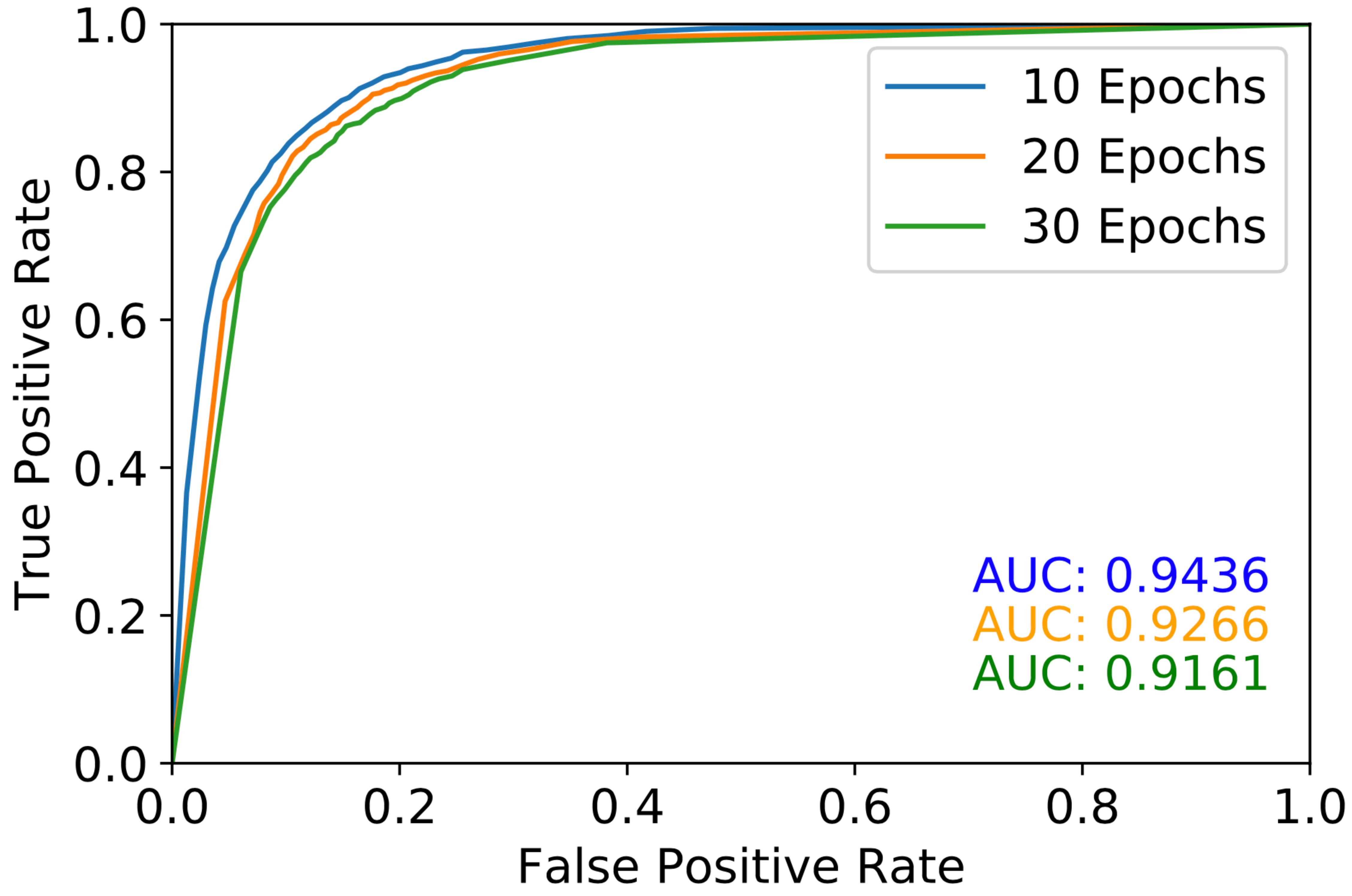


Classification of elastic and inelastic events

- This time the loss function will evaluate the difference between the true label, and the probability for the event to be elastic or inelastic predicted by the CNN

- After training, the prediction accuracy was found to be 94.5%





Outlook

- Add information from the scintillating pads of the Muon detector
  –> greatly improve efficiency for not fully contained showers
- Study effect of non-binary SciFi channels response
- Move on to full detector simulation and higher stat samples
- Optimize network structure & parameters to achieve best energy resolution
- Add the possibility of adding tracking for isolated tracks (muons) to aid flavour tagging
- SND@LHC finishing installation -> Use in the LHC Run 3

Conclusion

- Using CNN algorithms for SND@LHC prompt analysis has many advantages:
  - once trained, it is fast and straightforward to compute the energy
  - It is flexible, as the same architecture can be used for multiple tasks (type of interaction, flavour tagging)
  - It can be very accurate
- Allow to perform **real time** physic
- It is very greedy: we need feed it a lot with labelled data to achieve acceptable accuracy. For this, an accurate description of the detector geometry and digitalisation of the detector signal is essential.
- It would be a crucial asset for particle detector operating at a high luminosity collider or beam dump
- In case of SND@LHC, after HL-LHC upgrade, emulsions cannot operate and a ML approach would become even more important
- New approach to neutrino physics with a tracking detector

감사 해요
(Thank you)

Contact: paul.debryas@epfl.ch, elena.graverini@epfl.ch

BACKUP

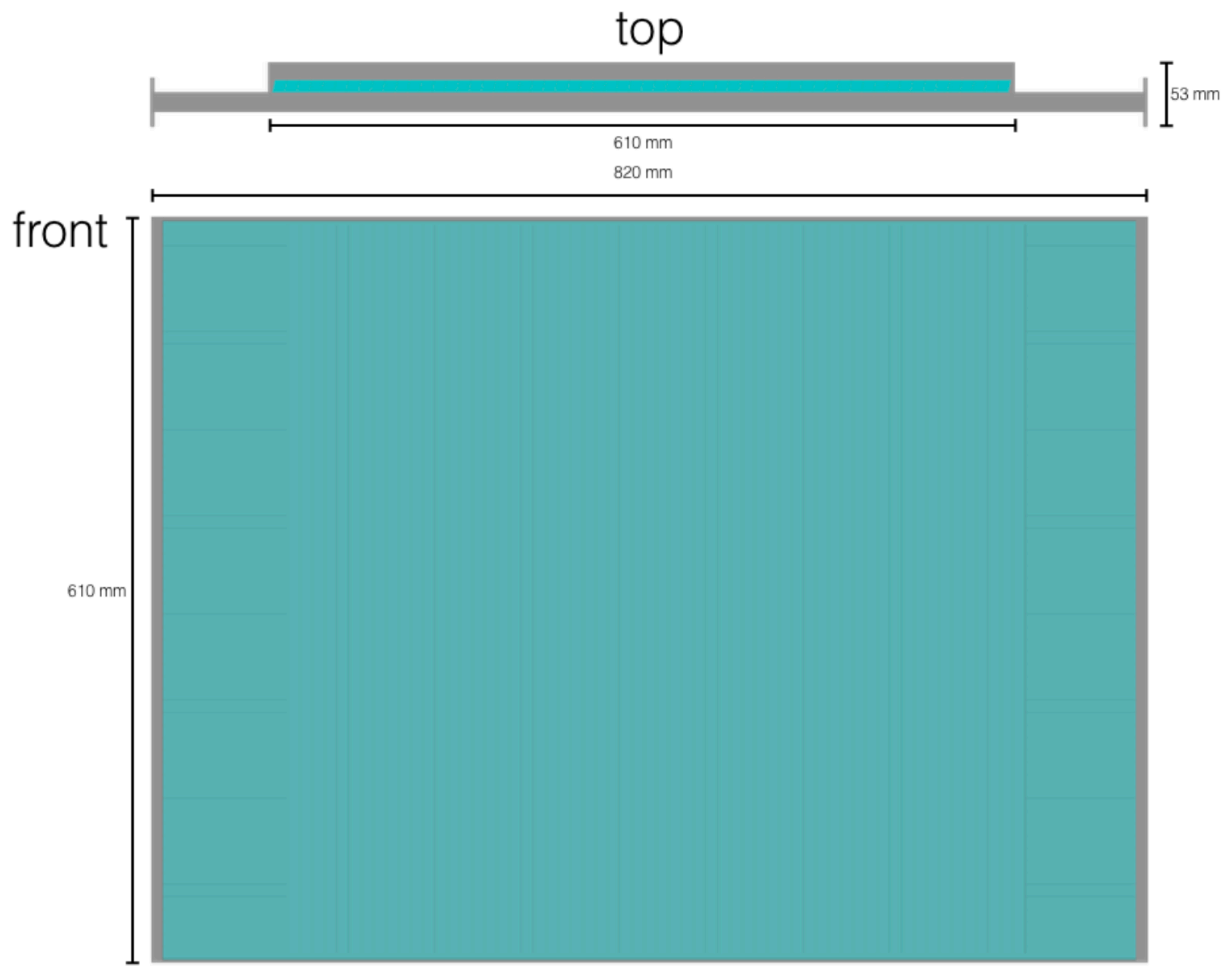
Electronic detectors

Muon upstream

5 planes of 82x61 cm made of 10 horizontal bars

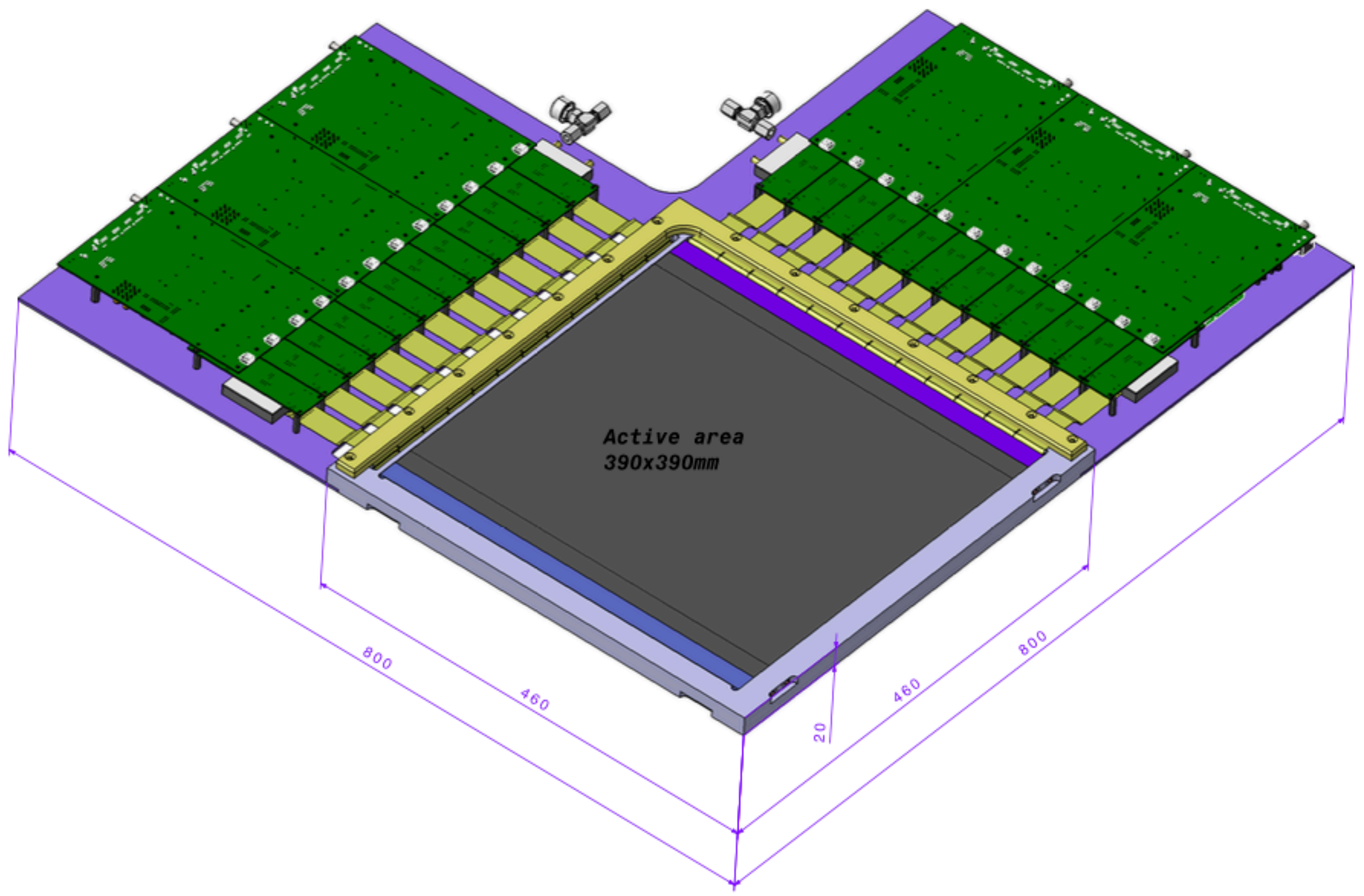
Muon downstream

3 planes of 82x61cm made of 2 layers (X and Y) of 60 bars each



SciFi tracker

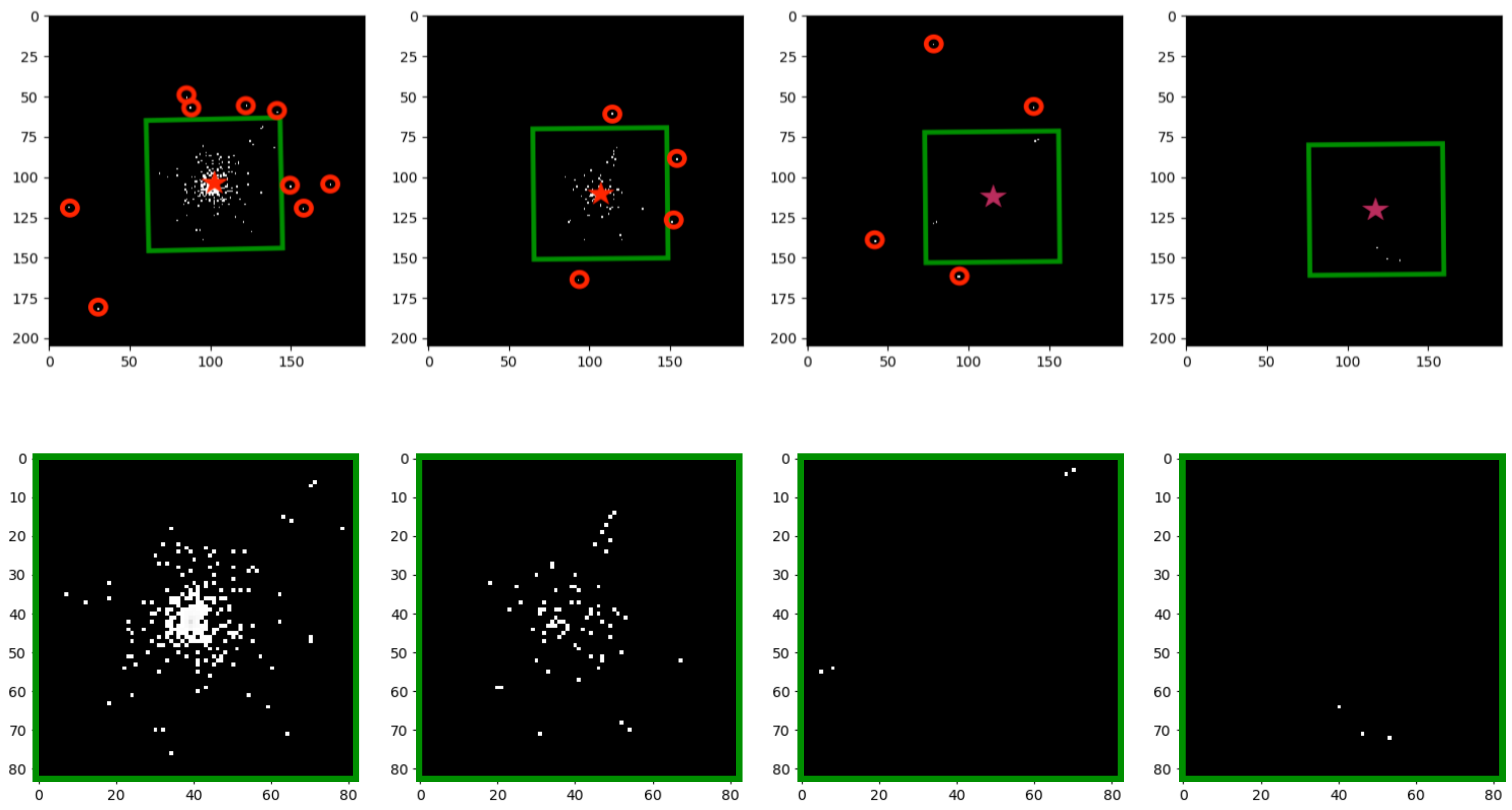
5 planes of 41x40cm made of 2 layers (X and Y) with a 250μm resolution



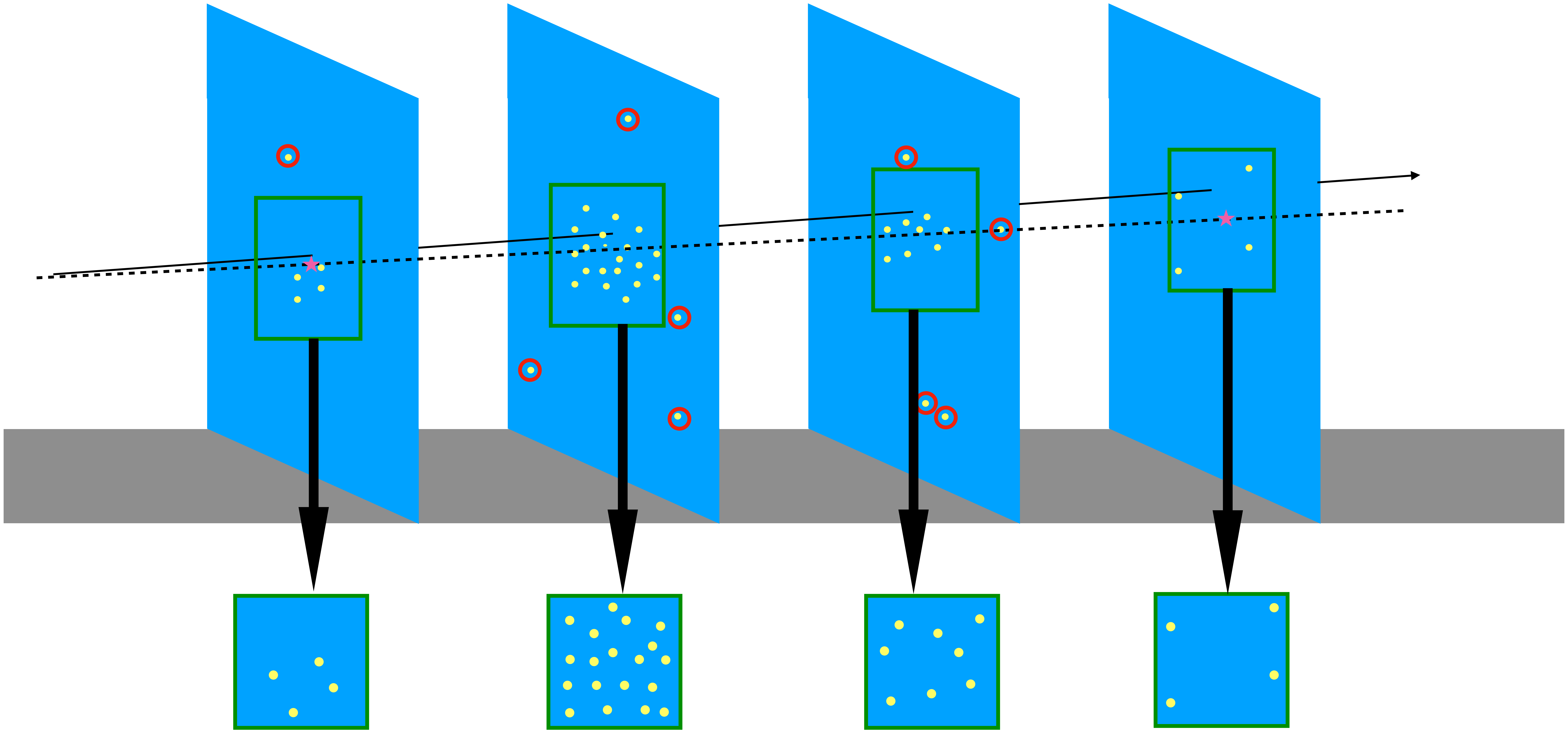
Downsizing process

- To find the optimal CNN architecture, we worked on a simplified dataset with smaller SciFi planes (smaller dimension of the input images)
- Advantages:
  -That avoid polluting the CNN with extra hits so that the shower features stand out better
  -It save a lot of computing time
- Optimal size of the sub-image calculated using whole sample: Take 3 sigma around the barycenter of the hits of the planes

Example of the preprocessing step of downsizing on a single event



Downsizing process



Feasibility study

- Objective:
  - Measure the energy of an electronic shower in the energy range 0-100 GeV.

- Detector geometry used:
  - 4 planes of 41.6x38.7 cm.
  - Tungsten bricks of 5.6cm ($10X_0$ radiation lengths).
  - granularity implemented corresponds to the fibre diameter (250μm)

- Particle gun electron:
  - shot before the first plane though the center with a random angle within plus/minus 10°.

- Input information for the reconstruction:
  - (X, Y) truth positions of the hits used (from MC)
  - no simulation of readout electronics
  - hits are defined as a greyscale signal (amplitude of the hit proportional to the number of hits that fall into the plane pixel).

Neutrino energy reconstruction

- Detector geometry:
  - 5 planes of 41x40 cm.
  - separated by tungsten bricks of 7.8cm (15X0 radiation lengths).
  - granularity implemented correspond to the resolution of the SciFi (250μm)

- Neutrino:
  - produced with Genie using the NuEElastic and CCDIS settings
  - energy range 100-5000 GeV
  - shotted in accordance with the alignment of the IP (off center of the plane) with a random angle.
  - shower starting point is spread uniformly over the detector length

- Input information for the reconstruction:
  - no simulation of readout electronics;
  - hits are defined as a Yes/No signal in each pixel (no amplitude information is used)
  - (X, Y) truth positions of the hits used (from MC)