Particle Identification with the Belle II Calorimeter using Machine Learning.

DPG-Frühjahrstagung, Dortmund

Abtin Narimani Charan

18.03.2021

Belle II

HELMHOLTZ RESEARCH FOR GRAND CHALLENGES

Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

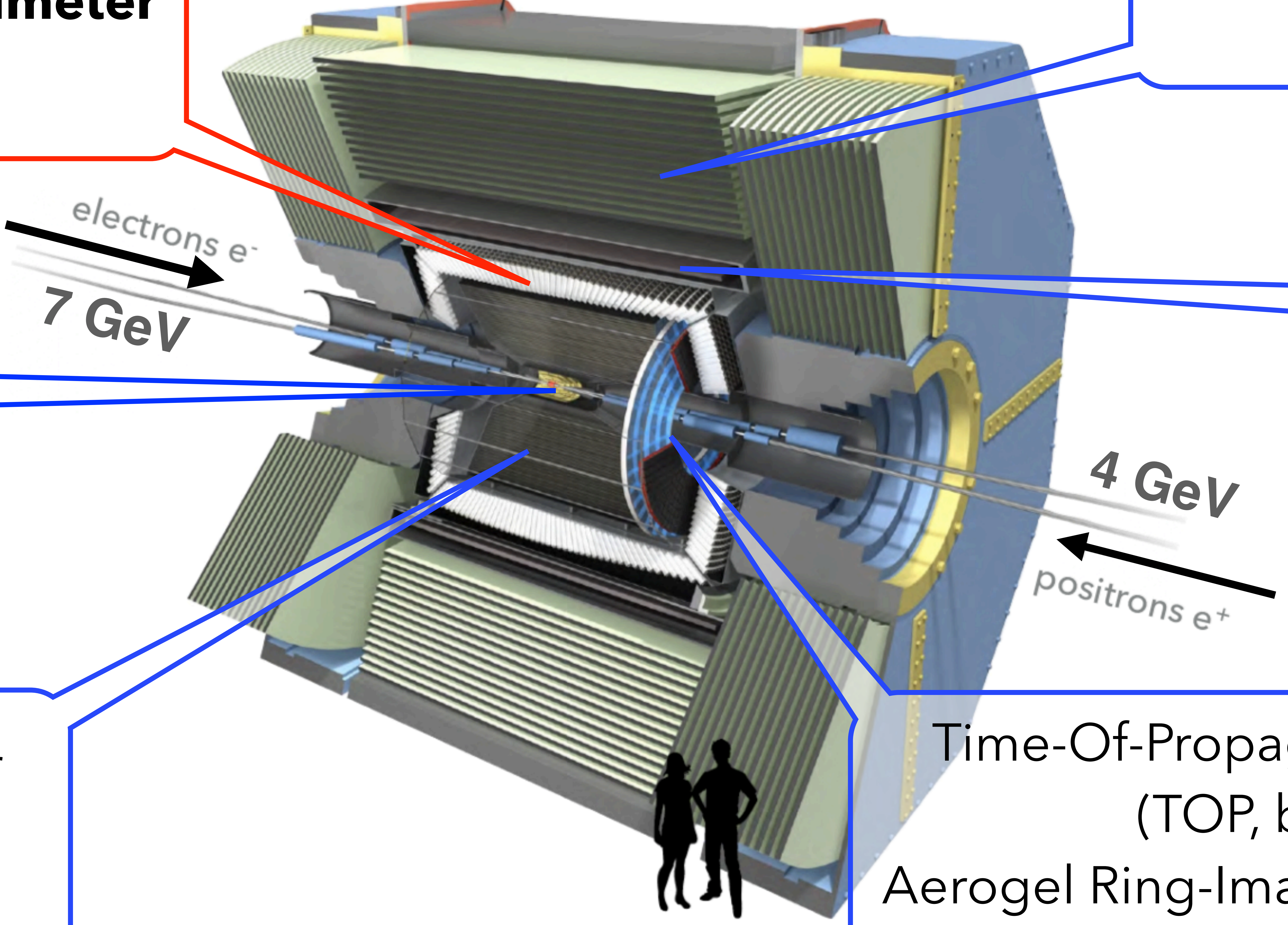
DESY.

Belle II & Particle Identification (PID)

$$e^+e^- \rightarrow \Upsilon(4S) \rightarrow B\bar{B}$$

Electromagnetic Calorimeter (ECL)

K_L and μ detector (KLM)



Vertex Detectors (VXD)

Magnet

Central Drift Chamber (CDC) tracking system

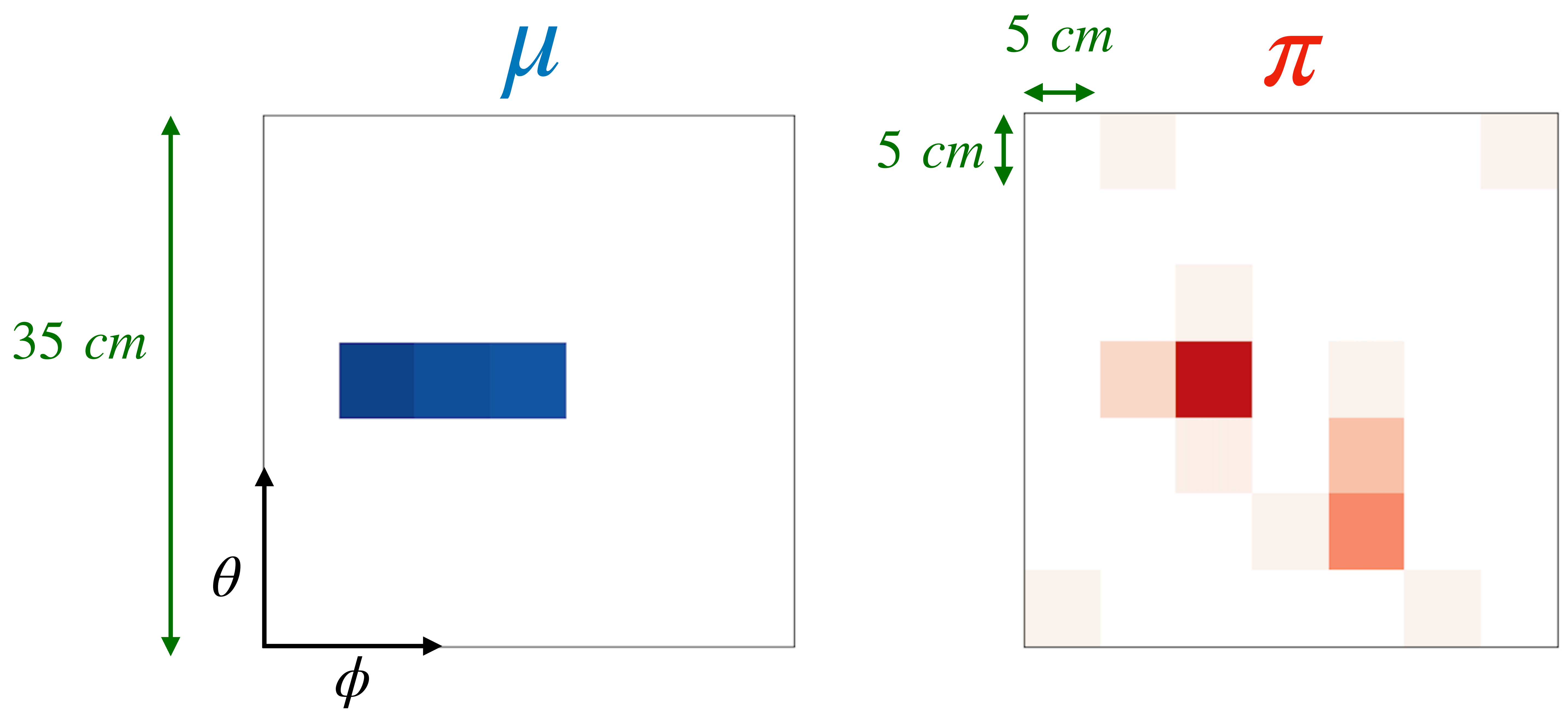
Time-Of-Propagation counter (TOP, barrel)
Aerogel Ring-Imaging Cherenkov counter (ARICH, end-caps)

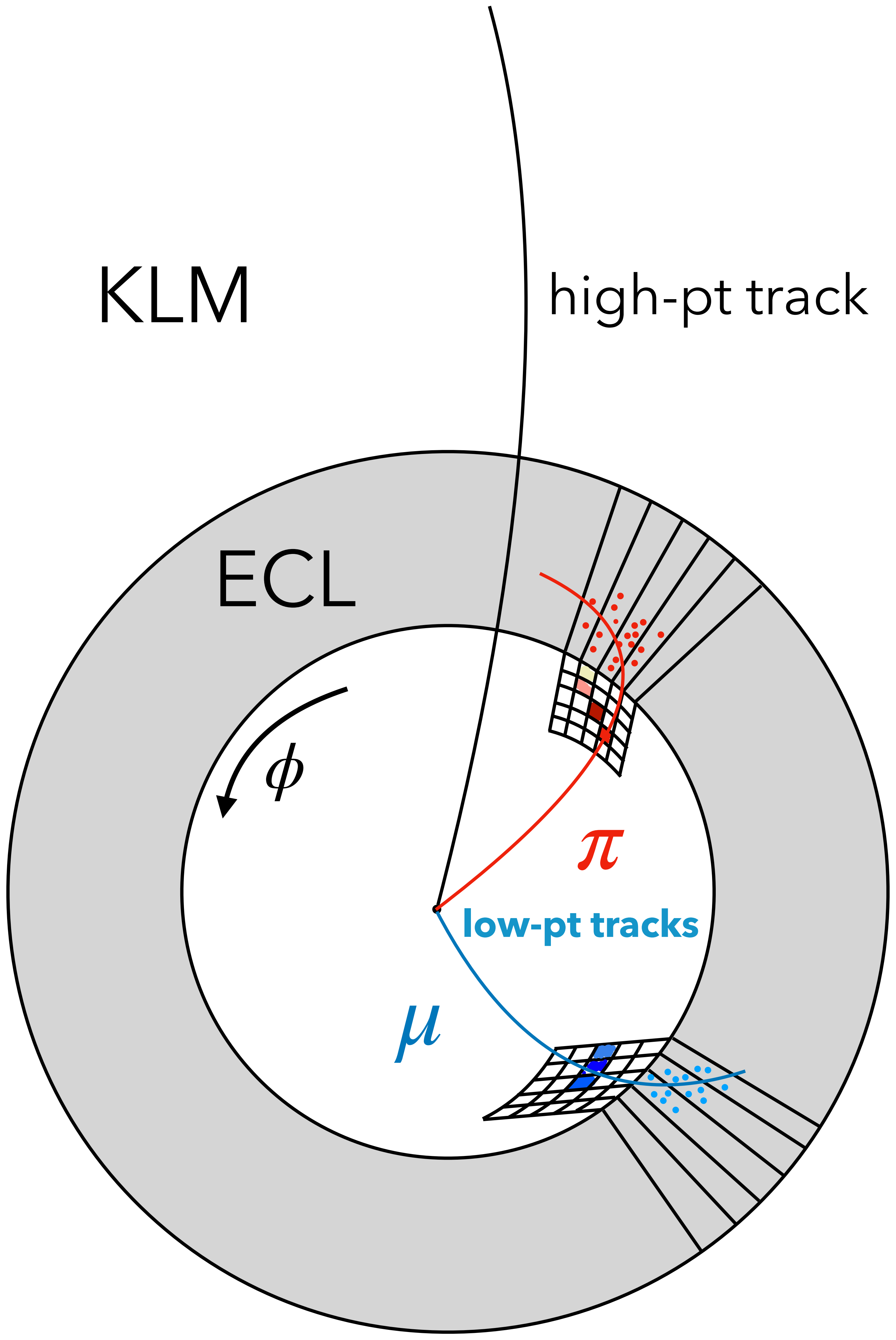
Motivation for improving ECL PID

- Low-P_T muons cannot reach the KLM:

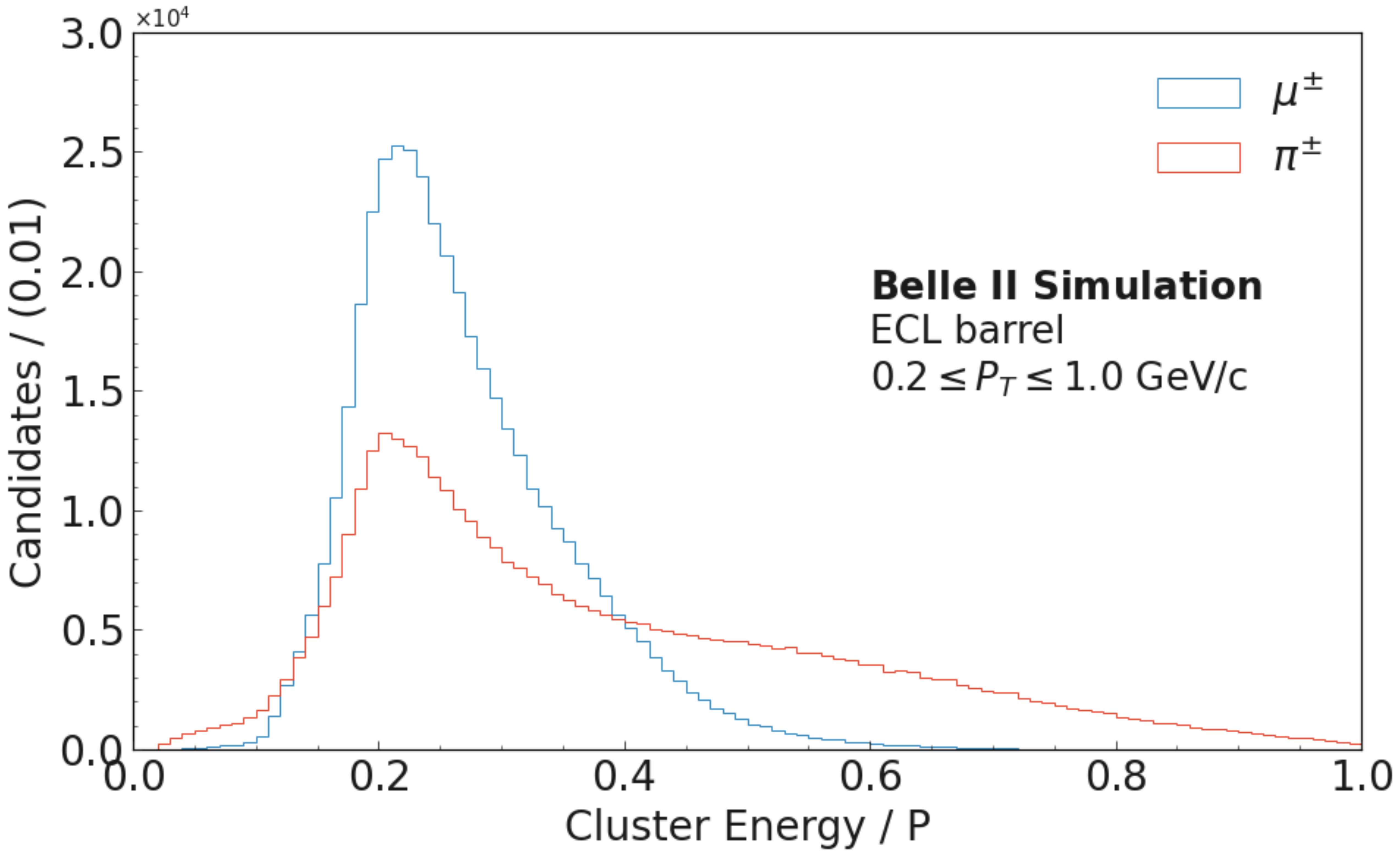
  $0.28 \lesssim P_T \lesssim 0.7\ GeV/c$

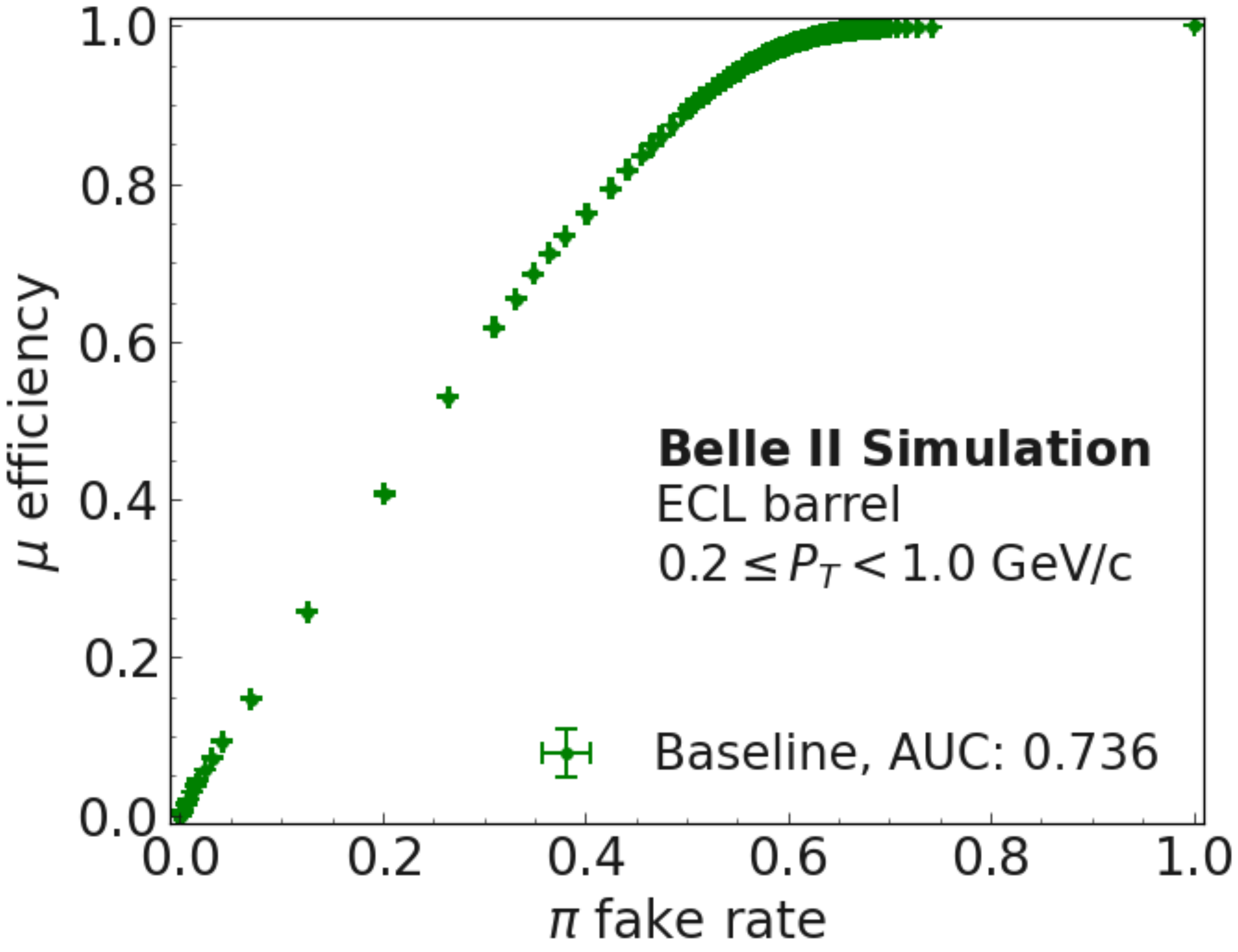
- We need to rely on the information in the ECL.





Default PID in Belle II





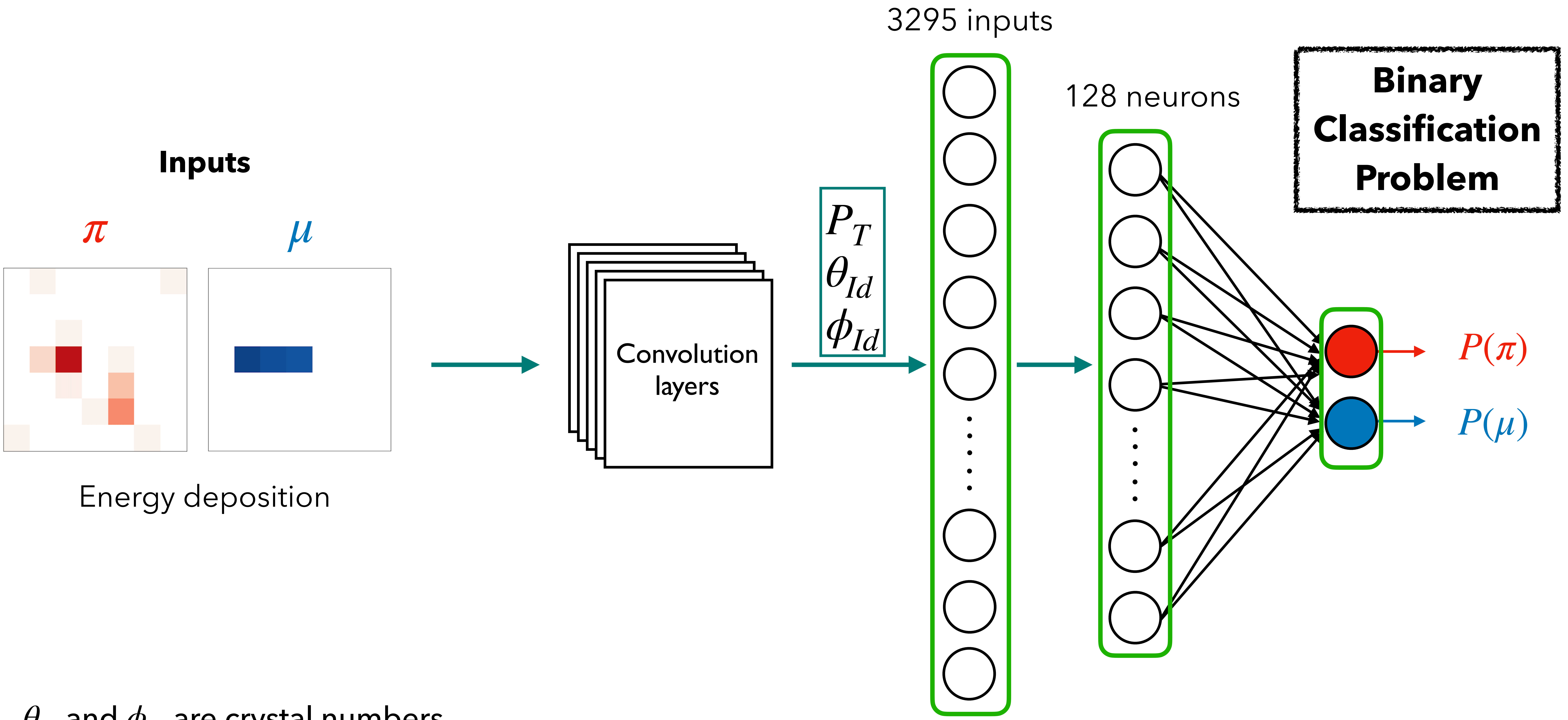
Boosted decision trees (BDT) for PID

- It is based on multi-variate classification algorithm.
- BDTs are trained combining measurements from the ECL and the tracking system.
- There are observables that are defined to describe **lateral shower shape** development in the ECL which differs for e, μ, and π.

Variable	Range	Description
E/p [c]	–	Ratio of cluster energy over track momentum.
$E_{cluster}$ [GeV]	–	Cluster energy.
E_1/E_9	–	Ratio of the energy of the seed crystal over the energy sum of the 9 surrounding crystals.
E_9/E_{21}	–	Ratio of the energy sum of 9 crystals surrounding the seed over the energy sum of the 25 surrounding crystals (minus 4 corners).
$\lvert Z_{40}\rvert$	–	Zernike moment $n = 4$, $m = 0$, calculated in a plane orthogonal to the EM shower direction.
$\lvert Z_{51}\rvert$	–	Zernike moment $n = 5$, $m = 1$, calculated in a plane orthogonal to the EM shower direction.
Z_{MVA}	–	Score of BDT trained on 11 Zernike moments.
ΔL [mm]	–	Projection on the extrapolated track direction of the distance between the track entry point in the ECL and the cluster centroid.

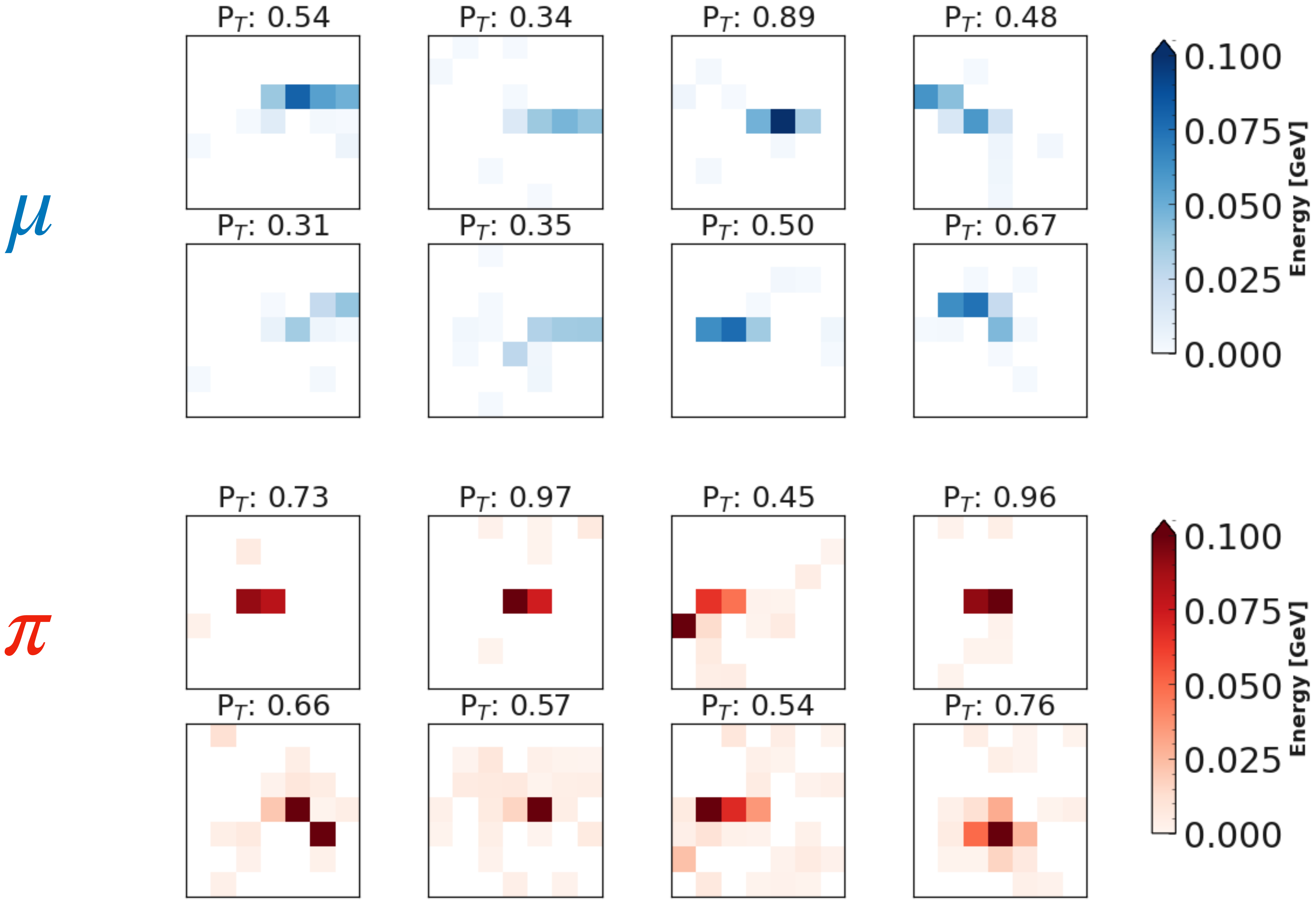
ICHEP 2020 proceeding

Convolutional Neural Network (CNN)



θ_{Id} and ϕ_{Id} are crystal numbers.

Energy deposition in crystals (MC samples)

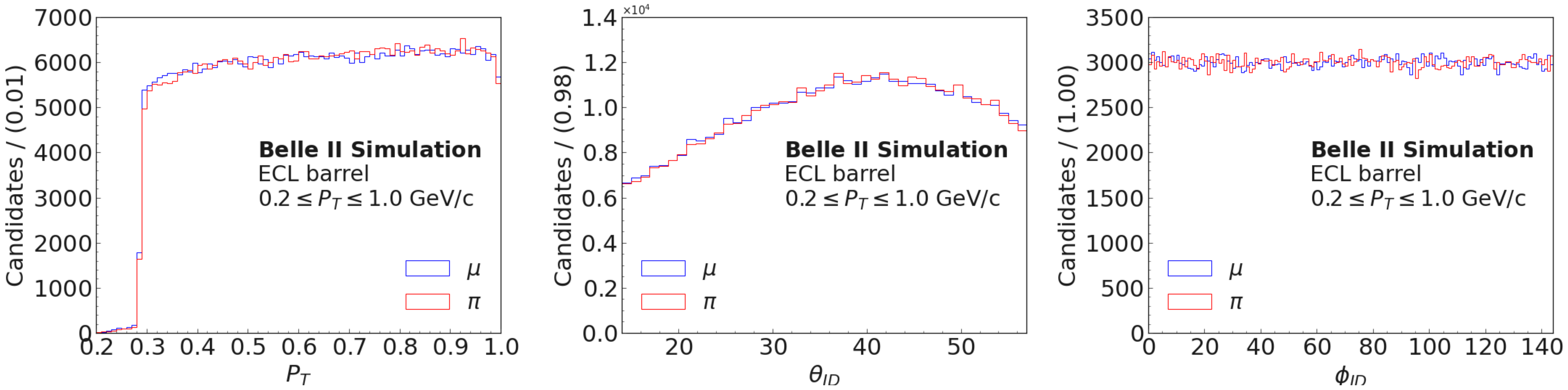


Tracks are extrapolated in the calorimeter and I do not rely on clustering and shower shape variables.

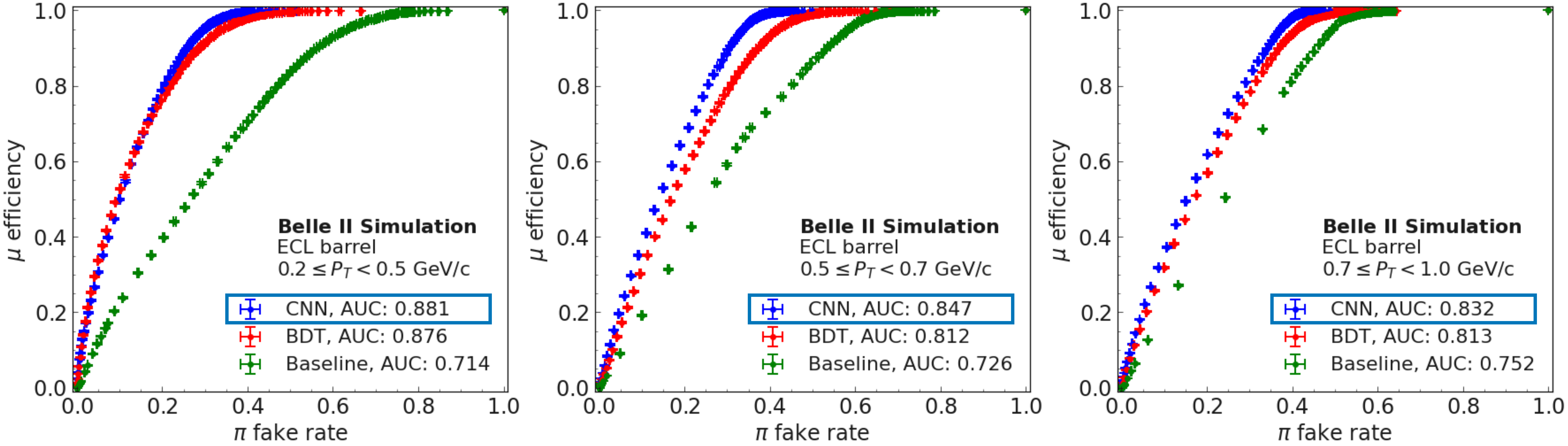
Training samples and inputs

- $\mu^{\pm}$ and $\pi^{\pm}$ samples (1 track per event)
- The samples include beam background.
- Number of training samples: 865k

  Number of validation samples: 288k



CNN performance on MC samples with 1 track / event



$$\mu \text{ Efficiency} = \frac{\text{Number of muons identified as muons}}{\text{Total number of muons}}$$

$$\pi \text{ fake rate} = \frac{\text{Number of pions identified as muons}}{\text{Total number of pions}}$$

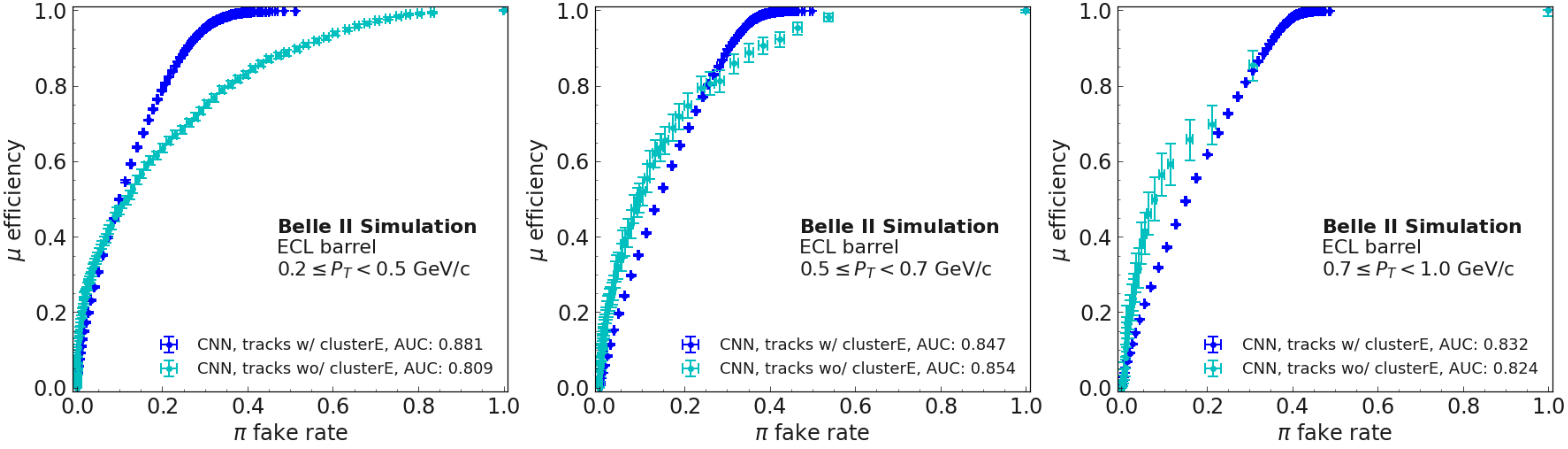
These plots includes tracks which have a cluster in the ECL assigned to it.

If you are interested in tracks without cluster, look at next slide.

CNN performance - comparison with tracks that has no cluster

total # of μ: 144162

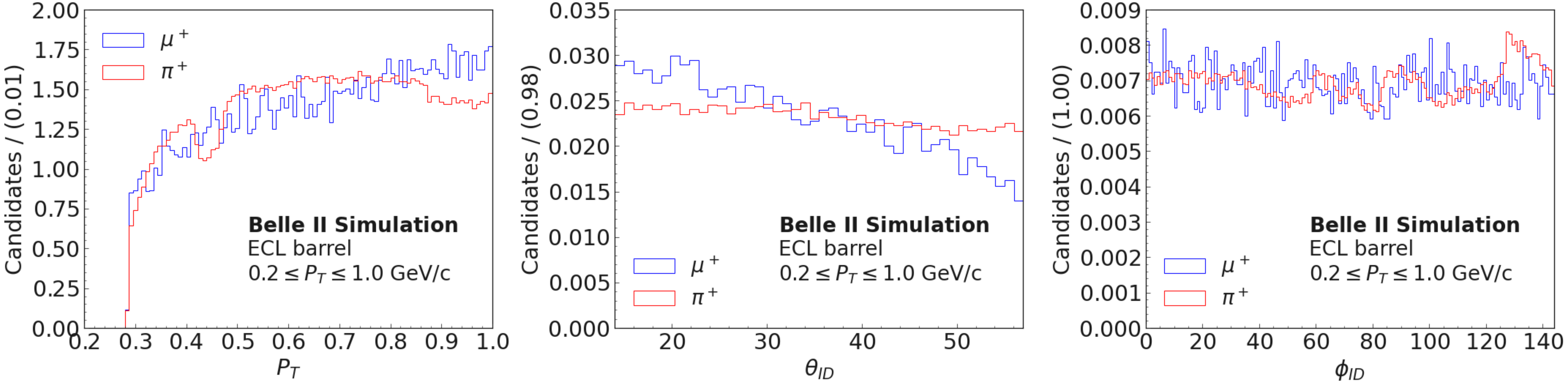
total # of π: 144162



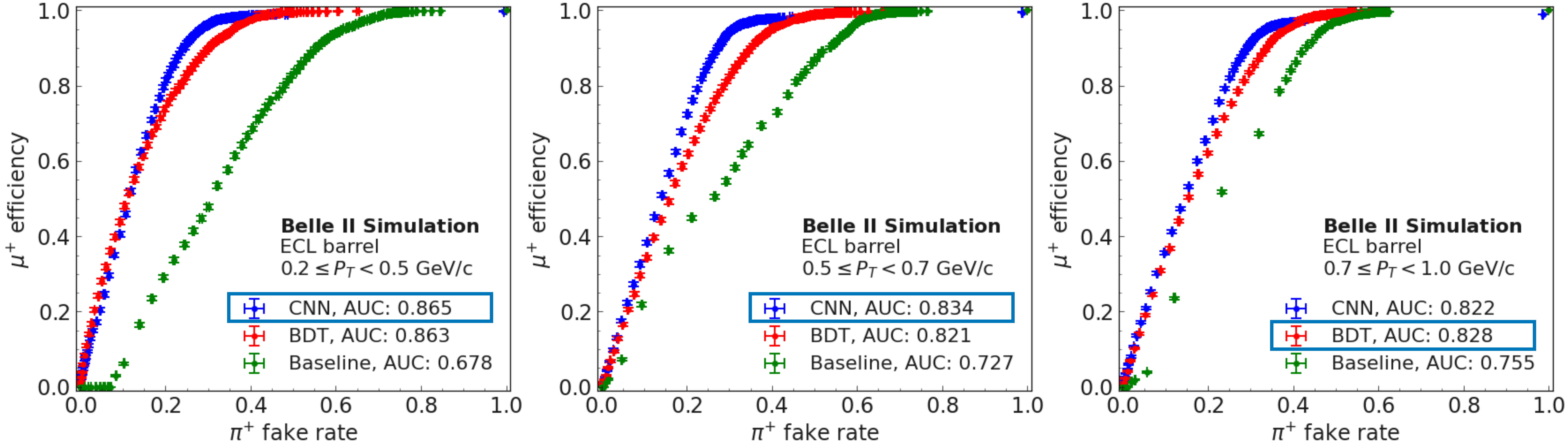
$0.2 \leq P_T < 0.5$ GeV/c	$0.5 \leq P_T < 0.7$ GeV/c	$0.7 \leq P_T < 1.0$ GeV/c
μ with cluster: 95.7 %	μ with cluster: 99.6 %	μ with cluster: 99.9 %
π with cluster: 84.4 %	π with cluster: 94.3 %	π with cluster: 95.8 %
μ without cluster: 4.3 %	μ without cluster: 0.4 %	μ without cluster: 0.1 %
π without cluster: 15.6 %	π without cluster: 5.7 %	π without cluster: 4.2 %

Physics processes with low-P_T particles

- $e^+e^- \to \mu^+\mu^-(\gamma_{ISR})$
- $e^+e^- \to \tau^+[\to \pi^+\pi^-\pi^+]\,\bar{\nu}_\tau \;\; \tau^-[\to 1\ prong]\,\nu_\tau$



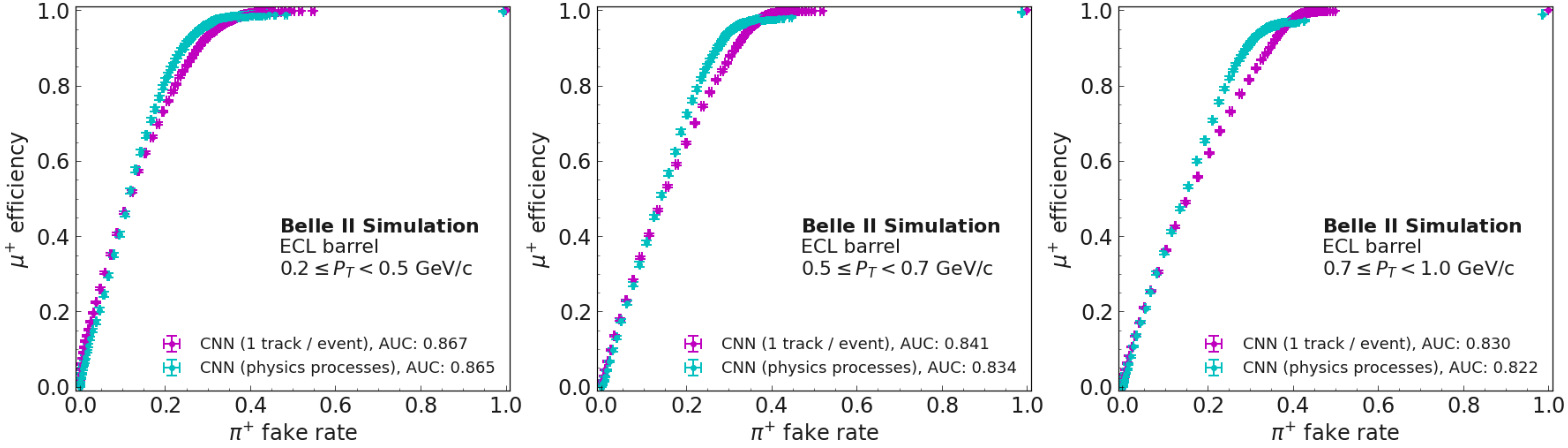
CNN performance on physics processes (MC)



$$\mu^+ \text{ Efficiency} = \frac{\text{Number of muons identified as muons}}{\text{Total number of muons}}$$

$$\pi^+ \text{ fake rate} = \frac{\text{Number of pions identified as muons}}{\text{Total number of pions}}$$

CNN performance - comparison in different MC samples



NOTE: These plots includes only μ^+ and π^+.

Summary

- PID for low-P_T tracks in the Belle II detector can be improved using low-level information in the ECL.
- A Convolutional Neural Network (CNN) is trained using 7x7 pixel images in the ECL, based on energy deposition, P_T, θ_{Id}, and ϕ_{Id} .
- CNN method does not depend on clustering or shower shapes.
- CNN surpasses baseline and BDT method in the ECL:
  - For a π^+ fake rate of 0.2, μ^+ efficiency is:
    - Baseline: 0.42 –> BDT: 0.65 –> **CNN: 0.71**