中央研究院物理研究所
INSTITUTE OF PHYSICS, ACADEMIA SINICA

Status report

2025/04/09
ZDC ML
WAI YUEN CHAN

Highlight of this month

- Confirmed the effect of carbon fiber in G4 sim output
- Transferring the G4 sim output into common ML readable format
- Plot the ML input as histogram to show the energy distribution in ECAL

ZDC Sim: Material test

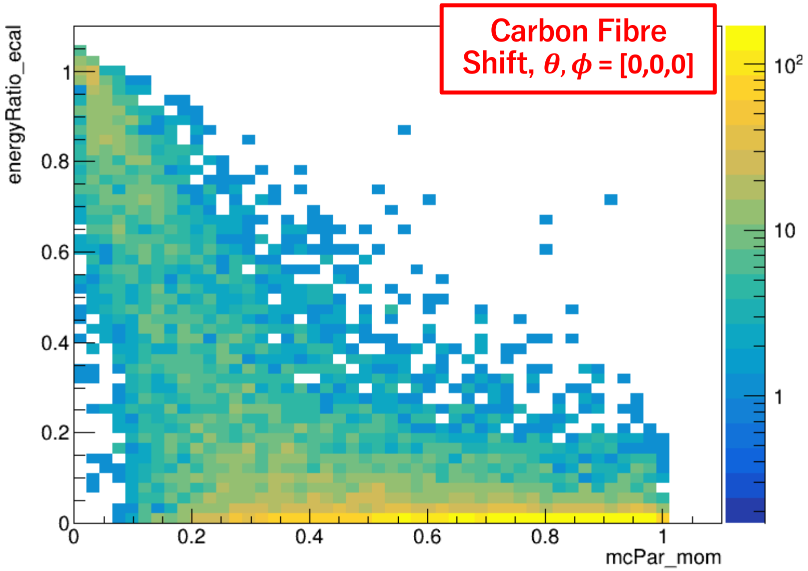
- Checking the effect of beam position and spread in the ZDC Sim
- Comparing samples with the following conditions:

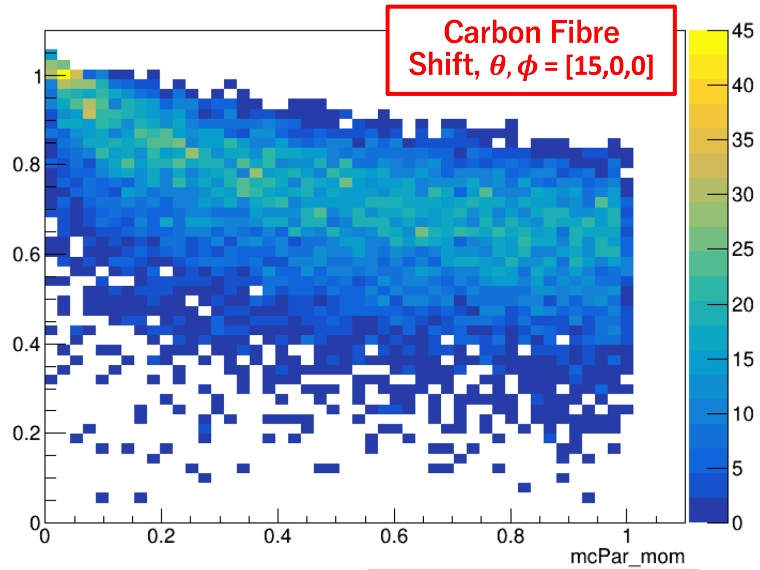
Material	θ, ϕ	Shift [mm]
Carbon fiber	0,0	0
Carbon fiber	0,0	15
LYSO	0,0	0
Carbon fiber	20,180	0
Carbon fiber	20,180	15
LYSO	20,180	0

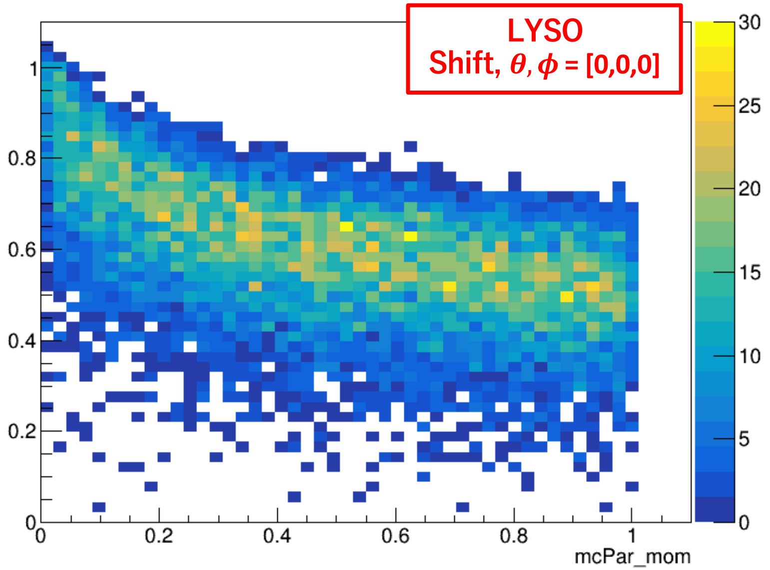
- 10k events within 0.01 GeV – 1 GeV positron beam are generated per sample.
- Energy dump in ECAL and HCAL are separately plotted.

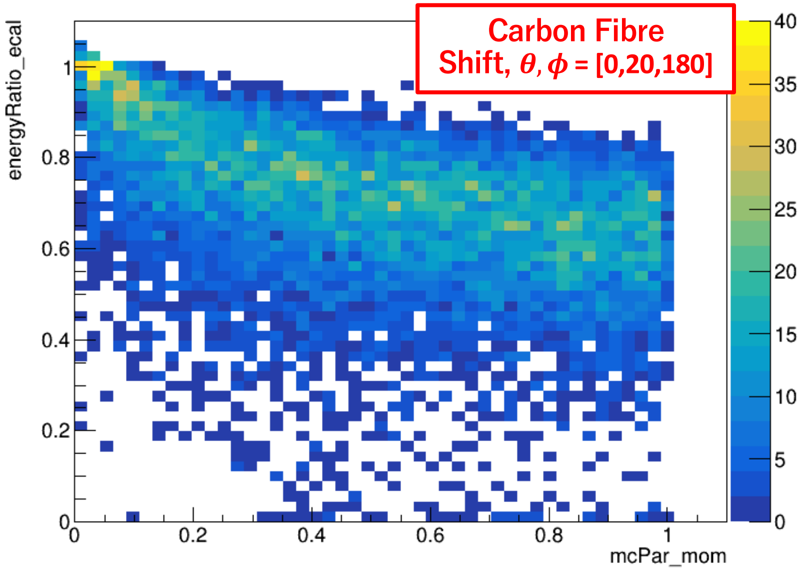
Results (ECal)

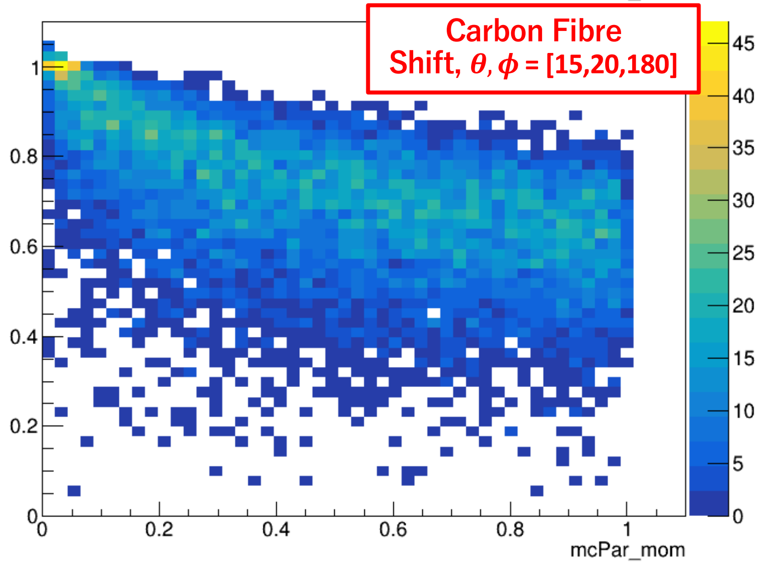
中央研究院物理研究所
INSTITUTE OF PHYSICS, ACADEMIA SINICA

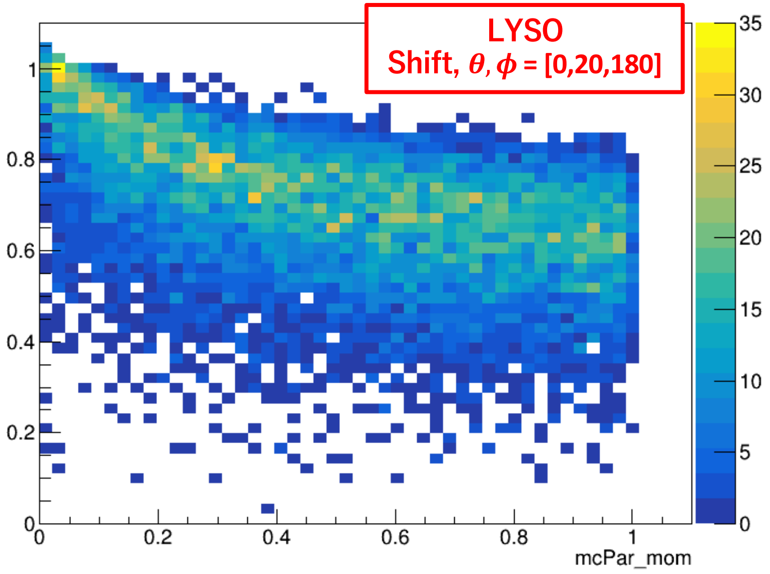






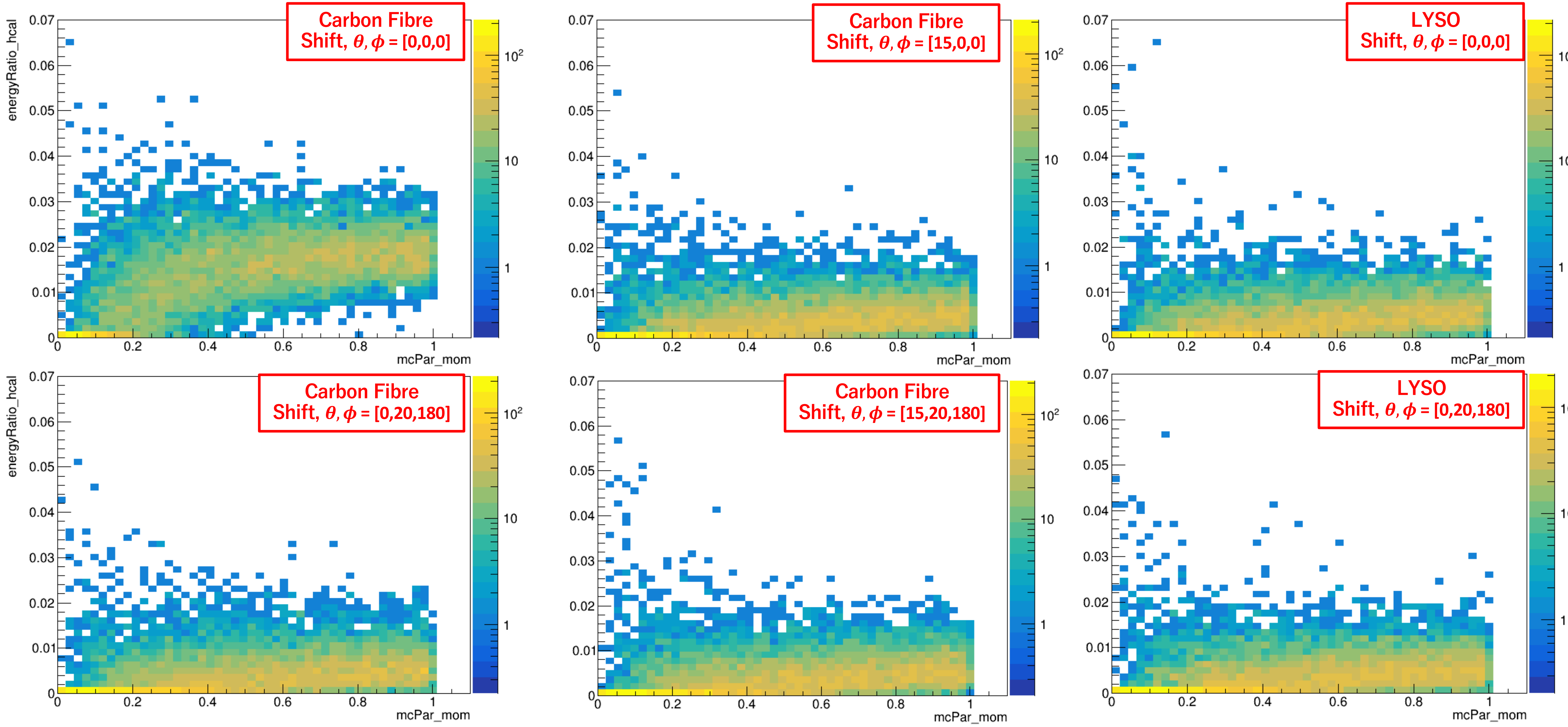






Results (HCal)

中央研究院物理研究所
INSTITUTE OF PHYSICS, ACADEMIA SINICA



ML Sample preparation

中央研究院物理研究所
INSTITUTE OF PHYSICS, ACADEMIA SINICA

- Need a larger sample which contain the following info:

Ecal hit related	Position X	Position Y	Position Z	Cell ID	Energy (Sum)
MC particle related	Momentum X	Momentum Y	Momentum Z		

- Fixed vertex (position fixed at [-902.408,-15.000,35488.907])
- $\boldsymbol{\theta}$, $\boldsymbol{\phi}$,shift = [20,180,15]
- Material at the center: Carbon fiber
- 50k events within 0.01 GeV – 20 GeV positron beam are generated.
- First step: modify the input, link up the energy distribution and the Cell ID.

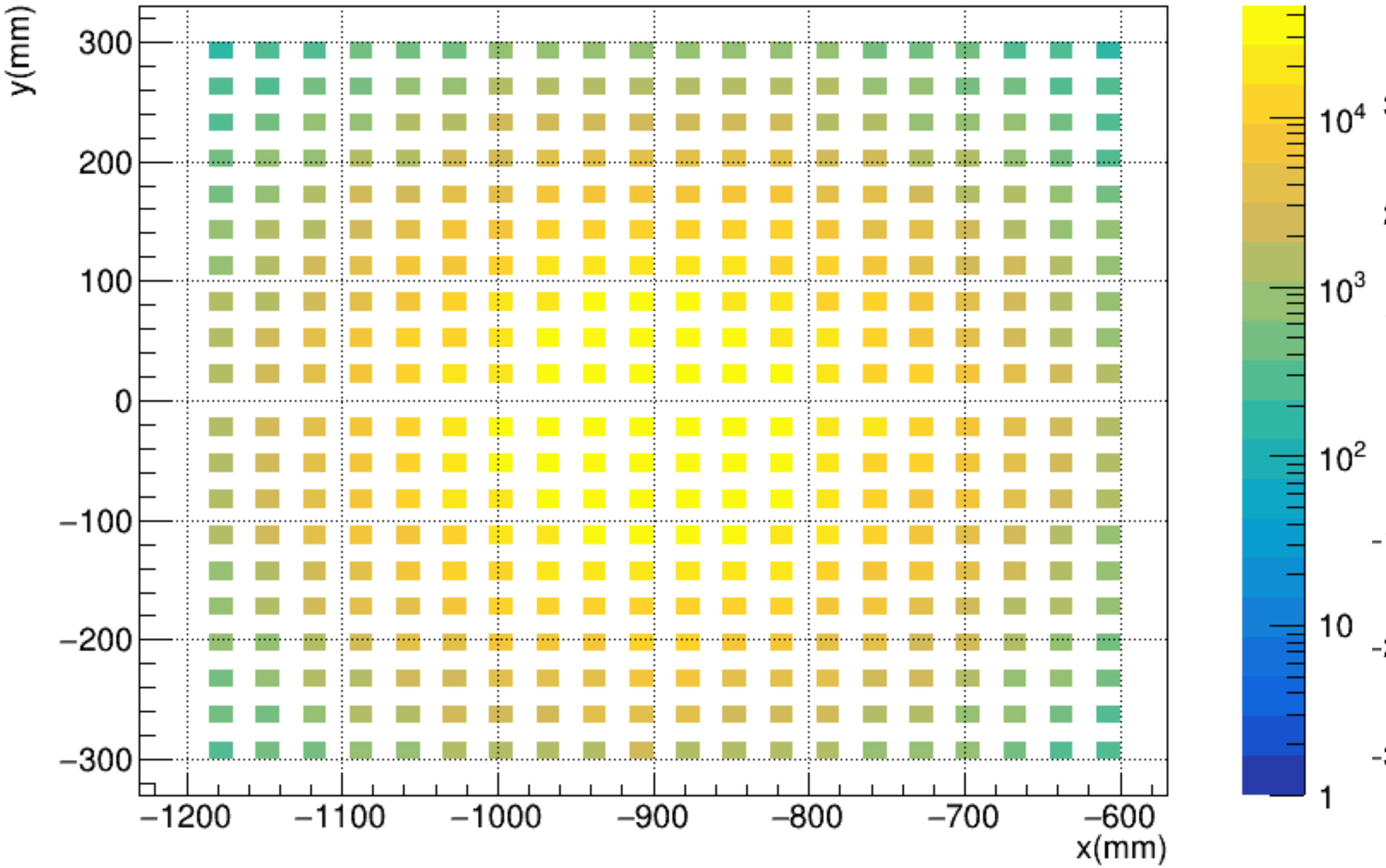
ML input: array

- We make use of the ROOT file structure (looping over entries), we can make a 20x20 2D array to store the energy per cell.
- A print out with energy per cell (amplified by 10000 times for visual purpose) in 1 example event has been shown as a 2D array:

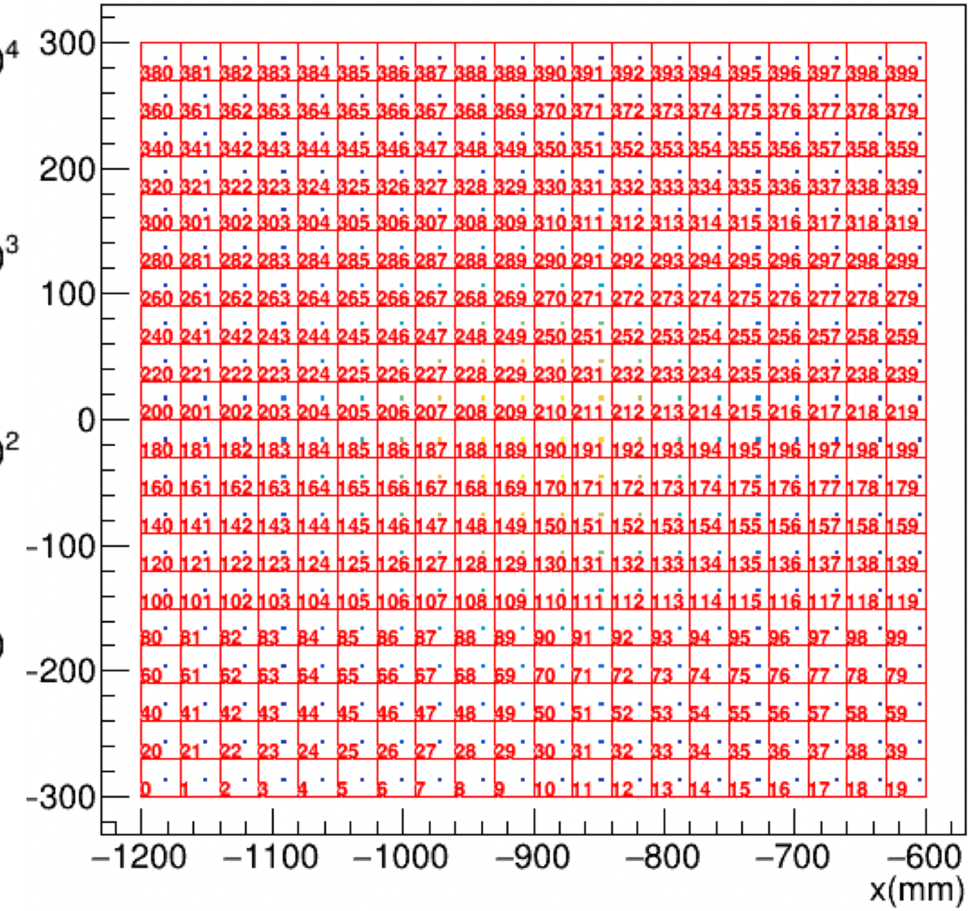
```
[[  0   0   0   0   0   0   0   0   0   0  10   3  12   0   0   0   0   0   0   0]
 [  0   0   0   0   0   0   0   0   0   0   7  37   4   0   0   0   0   0   0   0]
 [  0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   1   0   0]
 [  0   0   0   0   0   0   0   0   0   3   0   0   0   0   0   0   0   0   0   0]
 [  0   0   0   0   0   0   0   0   0   0   5   0   0  14   2   0   0   0   0   0]
 [  0   0   0   0   0   0   5   0   0   0   0   0   0   0   0   0   0   0   0   0]
 [  0   0   0   0   0   0   0   0   0   0   8  25   0   0   5   0   0   6   0   6]
 [  0   0   0   0   0   0   0   0   0   0  12  13  33  14   4   3  18   0   0   0]
 [  0   0   0   0   0   0   0   0   5   1  47 163 141  39  36  12   3   0   0   0]
 [  0   0   0   0   0   0   0   2   0  13 196  33 210  81   8   7   3   0   0   0]
 [  0   0   0   0   0   0   0   2   0  13  60 196 217  47  22   0   0   0   0   0]
 [  0   0   0   0   0   0   0   0   5   2  12  34  29   6   9   7   0   0   0   0]
 [  0   0   0   0   0   0   0   0   0   4   0  89  28  51   6   0   0   0   0   0]
 [  0   0   0   0   0   0   0   0   0   0   1  11   0   5  50   0   0   0   0   0]
 [  0   0   0   0   0   0   0   0   1   0   0   3   0   0   0   0   0   0   0   0]
 [  0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0]
 [  0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0]
 [  0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0]
 [  0   0   0   0   0   0   0   0   0   0   0   0   0   0   6   0   0   0   0   0]
 [  0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0]]
```

ECAL Cell ID grid system

#Event count per cell



Corresponding cell ID

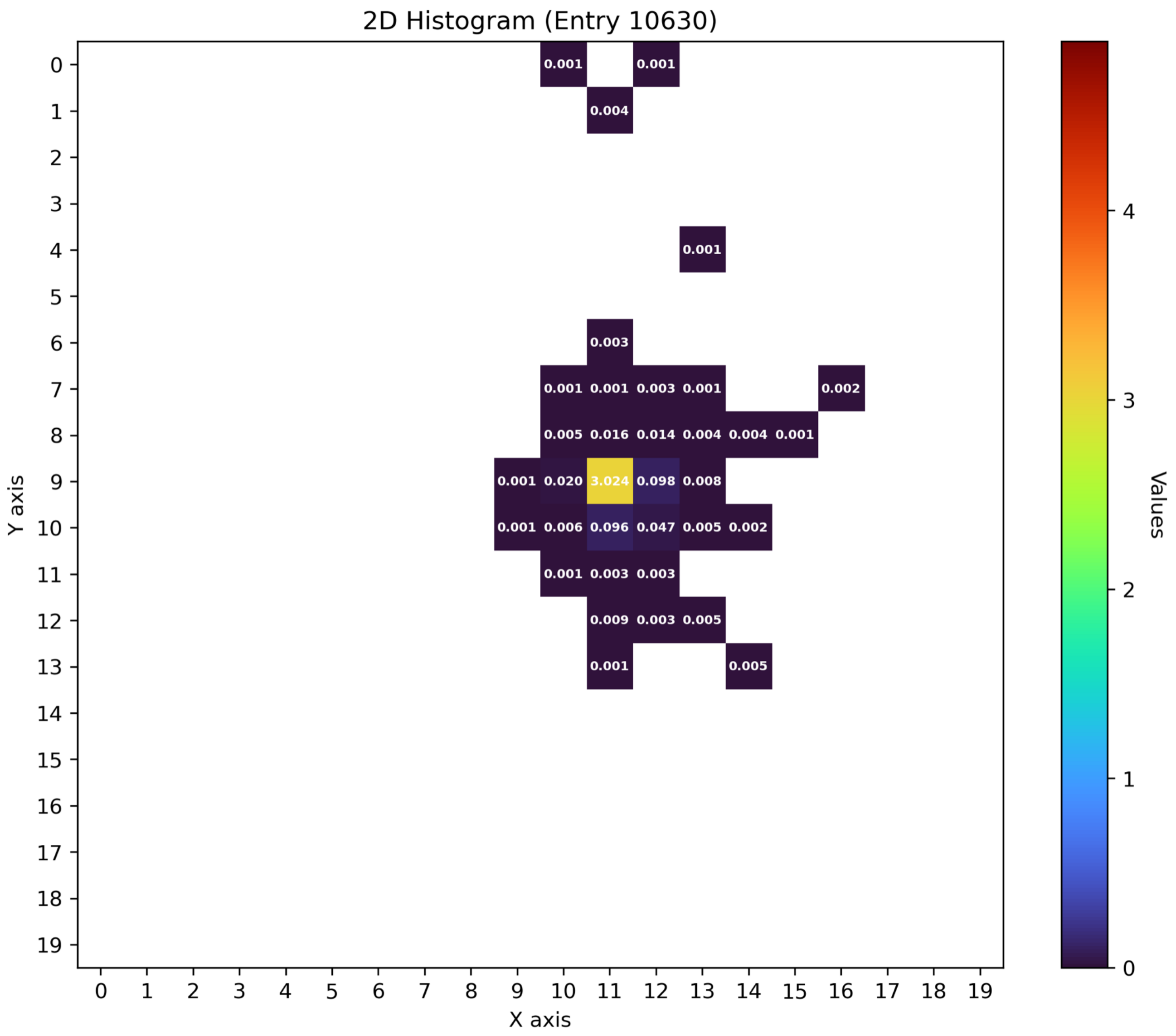


2D Histogram

中央研究院物理研究所
INSTITUTE OF PHYSICS, ACADEMIA SINICA

- We can transfer the array into 2D Histogram for each single event.
- Plotting positron events

```
[[   0   0   0   0   0   0   0   0   0   0  10   3  12   0   0   0   0   0   0   0]
 [   0   0   0   0   0   0   0   0   0   0   7  37   4   0   0   0   0   0   0   0]
 [   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   1   0   0]
 [   0   0   0   0   0   0   0   0   0   3   0   0   0   0   0   0   0   0   0   0]
 [   0   0   0   0   0   0   0   0   0   0   5   0   0  14   2   0   0   0   0   0]
 [   0   0   0   0   0   0   5   0   0   0   0   0   0   0   0   0   0   0   0   0]
 [   0   0   0   0   0   0   0   0   0   0   8  25   0   0   5   0   0   6   0   6]
 [   0   0   0   0   0   0   0   0   0   0  12  13  33  14   4   3  18   0   0   0]
 [   0   0   0   0   0   0   0   0   5   1  47 163 141  39  36  12   3   0   0   0]
 [   0   0   0   0   0   0   0   2   0  13 196  33 210  81   8   7   3   0   0   0]
 [   0   0   0   0   0   0   0   2   0  13  60 196 217  47  22   0   0   0   0   0]
 [   0   0   0   0   0   0   0   0   5   2  12  34  29   6   9   7   0   0   0   0]
 [   0   0   0   0   0   0   0   0   0   4   0  89  28  51   6   0   0   0   0   0]
 [   0   0   0   0   0   0   0   0   0   0   1  11   0   5  50   0   0   0   0   0]
 [   0   0   0   0   0   0   0   0   1   0   0   3   0   0   0   0   0   0   0   0]
 [   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0]
 [   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0]
 [   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0]
 [   0   0   0   0   0   0   0   0   0   0   0   0   0   0   6   0   0   0   0   0]
 [   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0]]
```

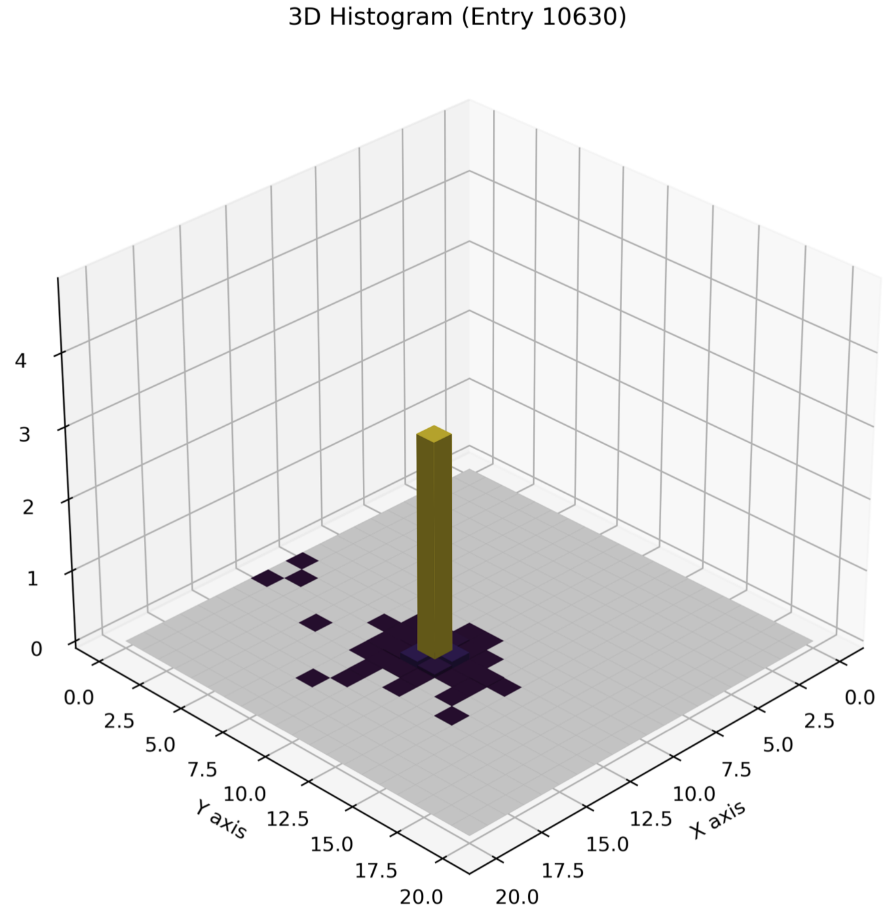


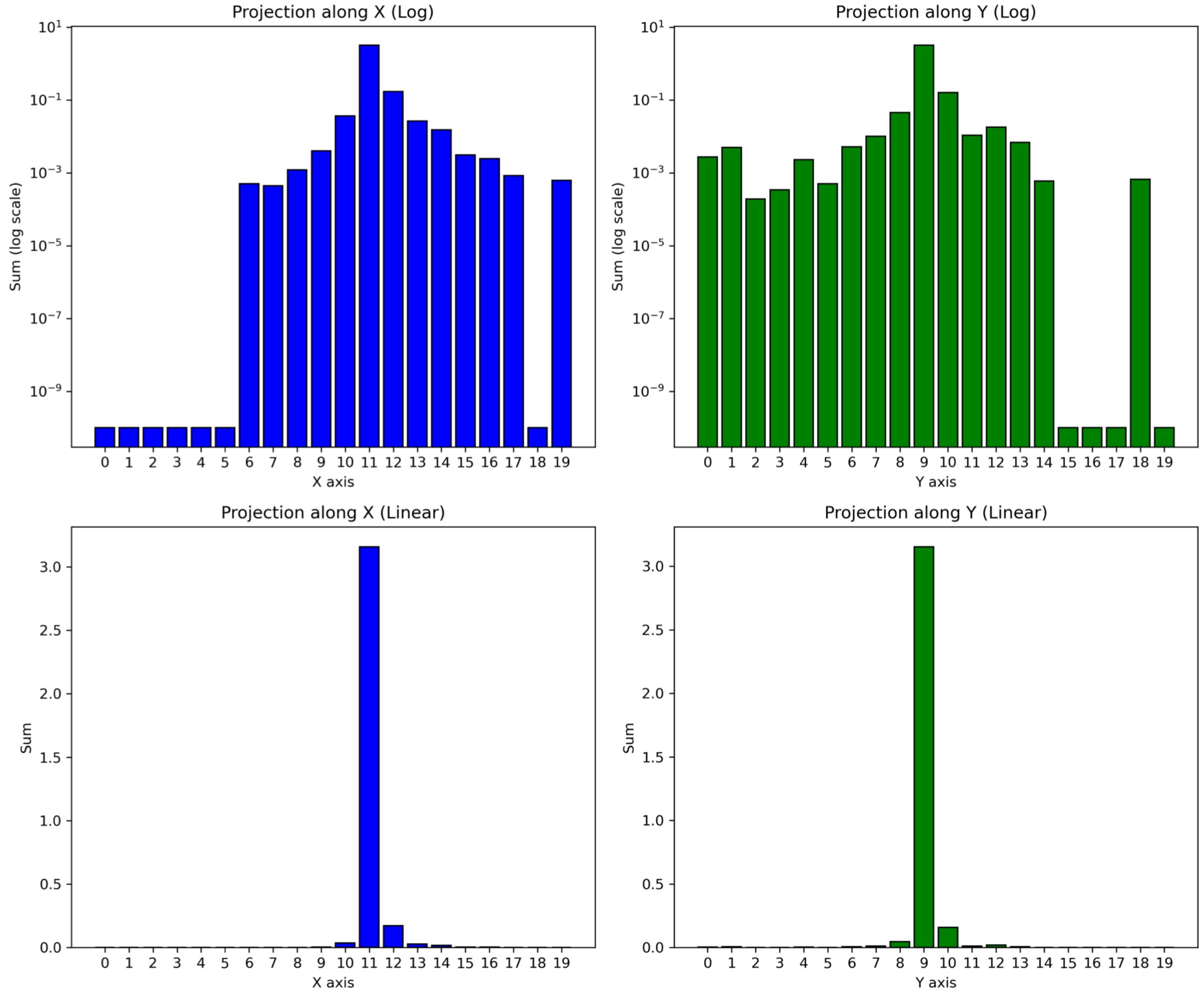
2D Histogram

- For each event we have also record the sum of energy and the number of cell which have energy >= 0.001 GeV

```
Entry 10630 2D array:
[[0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0. ]
 [0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0. ]
 [0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0. ]
 [0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0. ]
 [0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0. ]
 [0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0. ]
 [0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0. ]
 [0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0. ]
 [0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0. ]
 [0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  3.  0.1 0.  0.  0.  0.  0.  0.  0. ]
 [0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.1 0.  0.  0.  0.  0.  0.  0.  0. ]
 [0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0. ]
 [0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0. ]
 [0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0. ]
 [0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0. ]
 [0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0. ]
 [0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0. ]
 [0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0. ]
 [0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0. ]
 [0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0.  0. ]]
Total sum of all elements: 3.4
Number of elements >= 0.001: 35
```

1D, 3D Histogram



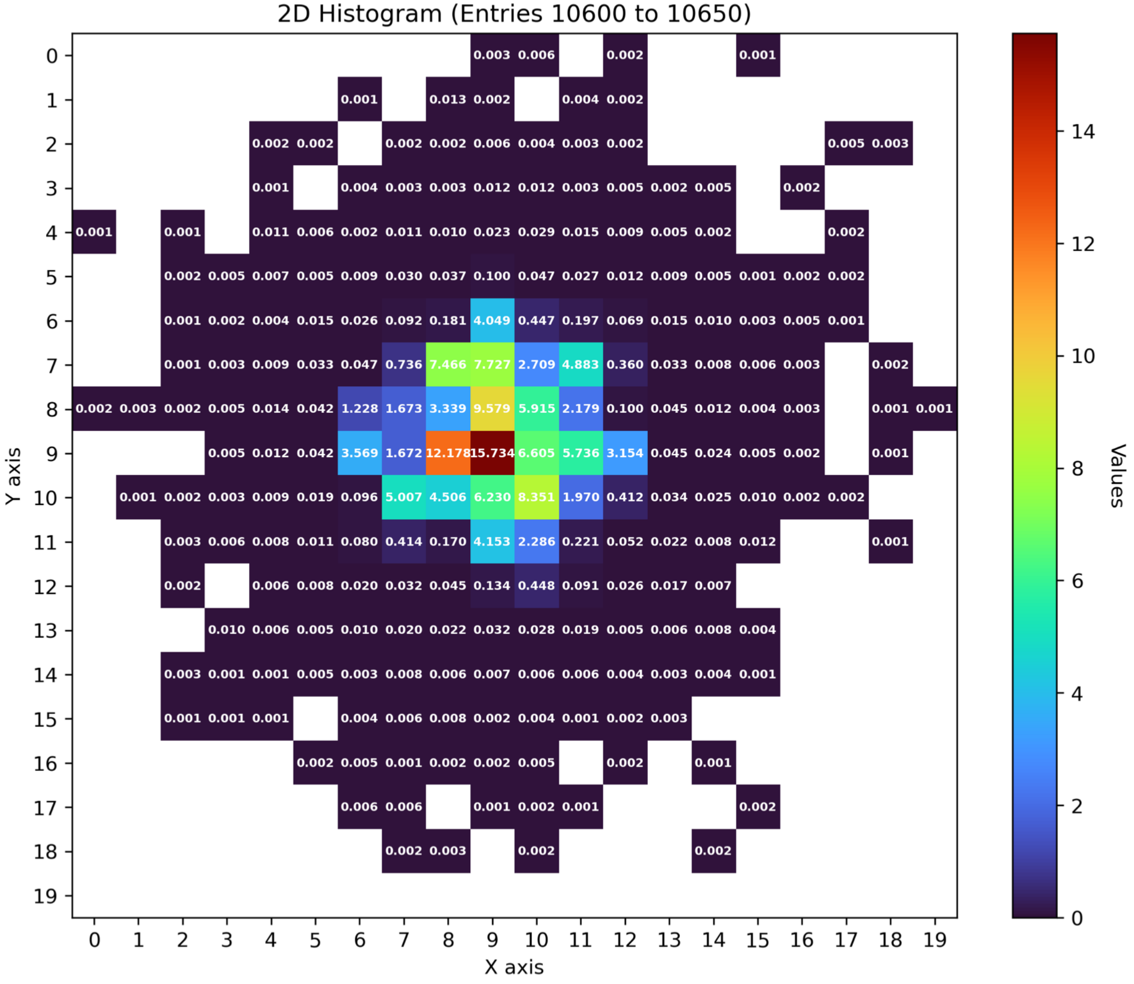


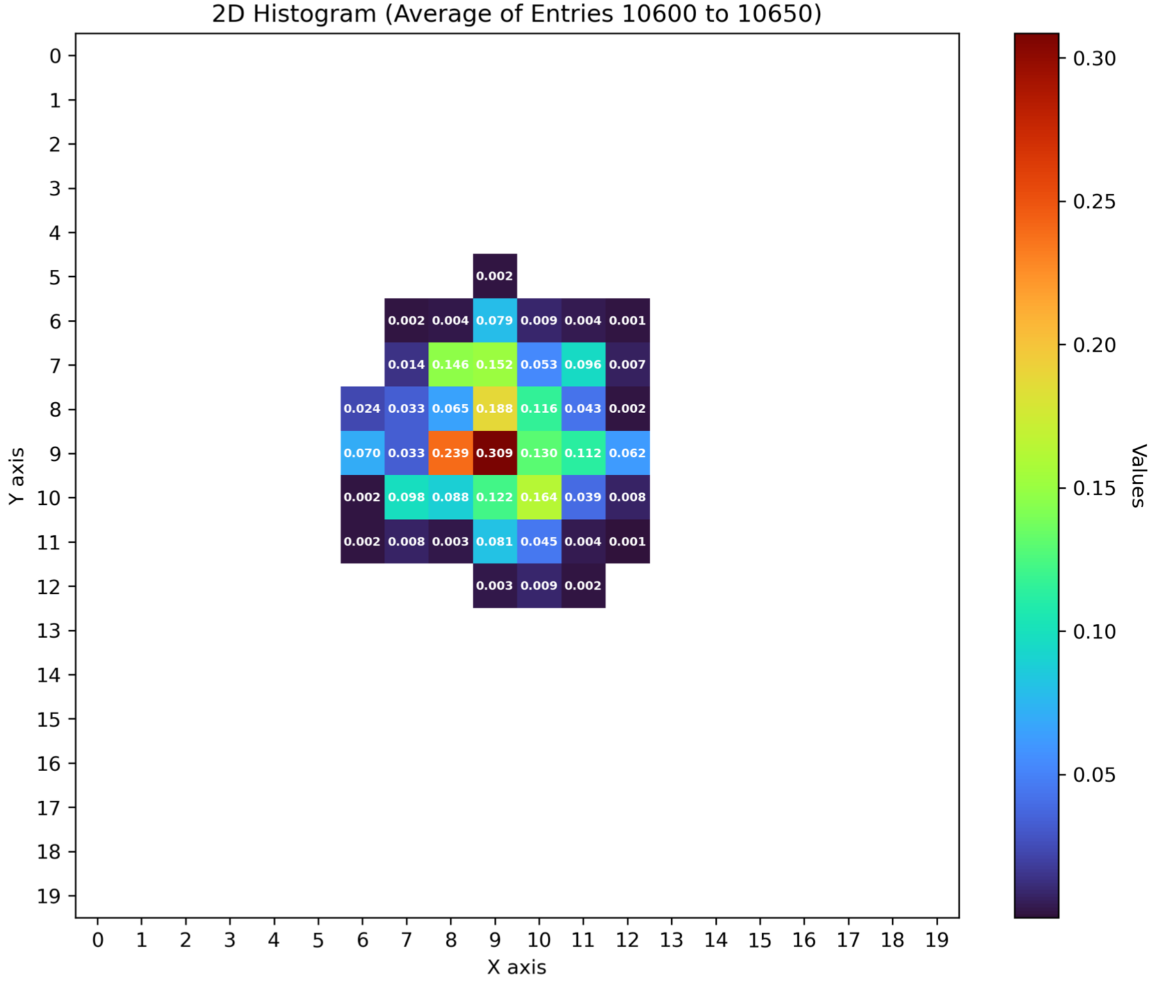
- 1D and 3D histogram will help us to identify any decay product hit the ECAL

2D Histogram (Summed)

中央研究院物理研究所
INSTITUTE OF PHYSICS, ACADEMIA SINICA

- 2D histogram for multiple events are also available.



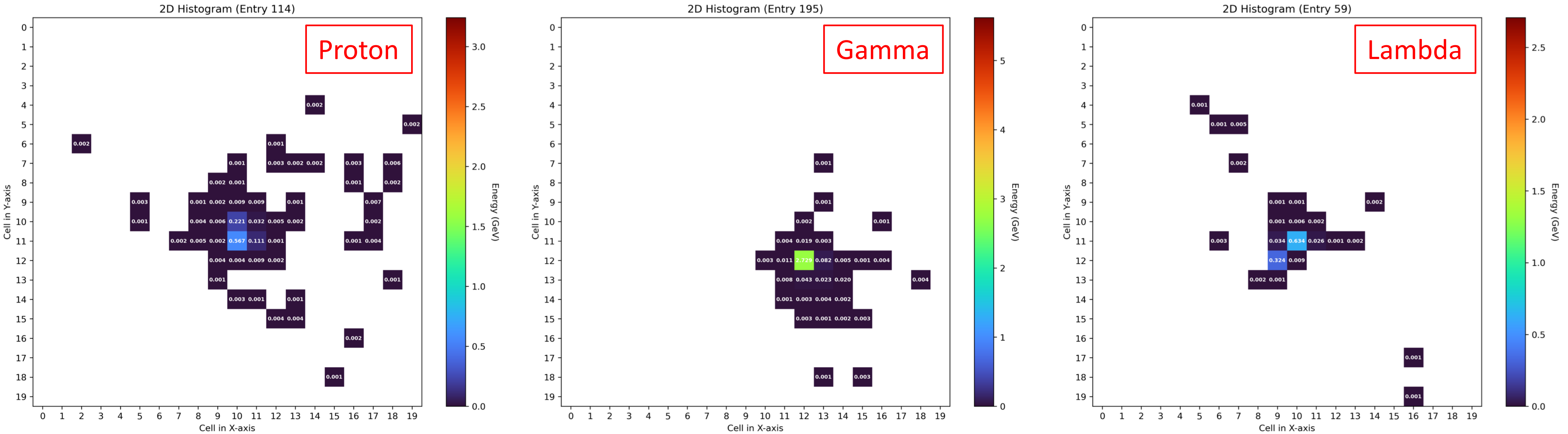


2D Histogram with other particles

- In order to start the ML training, we need samples from other particles (classes) as well.
- Here we generated G4 sim samples with the following conditions:
  - Fixed vertex (position fixed at [-902.408,0.000,35488.907])
  - $\boldsymbol{\theta}, \boldsymbol{\phi}$,shift = [20,180,0]
  - Material at the center: Carbon fiber
  - 5k events within 10 GeV – 15 GeV
  - Particles : $e^+, e^-, \gamma, p, n, \Lambda$

2D Histogram with different particles

中央研究院物理研究所
INSTITUTE OF PHYSICS, ACADEMIA SINICA



- For hadrons, we might better add HCAL info including HCAL cell ID as the input parameter.
- At the moment maybe we can use e^+, γ, p to arrange our CNN structure.

Pending & Outlook

Pending items

- We need to test the job submission pipeline
- In order to show the energy distribution in a 2D array, we need HCAL Cell ID info as well (As HCAL have a depth in z direction, we might eventually need a 3D array instead)

Outlook

- We don’t need all type of particles at the moment, a binary classification is a good and simple starting point
- We can try to run the actually ML code:
    1. Check does it actually read samples as expected
    2. Adjust the CNN structure (for ECAL only)