Improving ATLAS Particle Flow with deep learning. (arXiv : 2003.08863)

Francesco Di Bello[2] Sanmay Ganguly[1] Michael Pitt[3]
Jonathan Shlomi[1] Eilam Gross[1] Marumi Kado[2] Lorenzo Santi[2]
Weizmann Institute Of Science[1] Roma Sapienza[2] CERN[3]

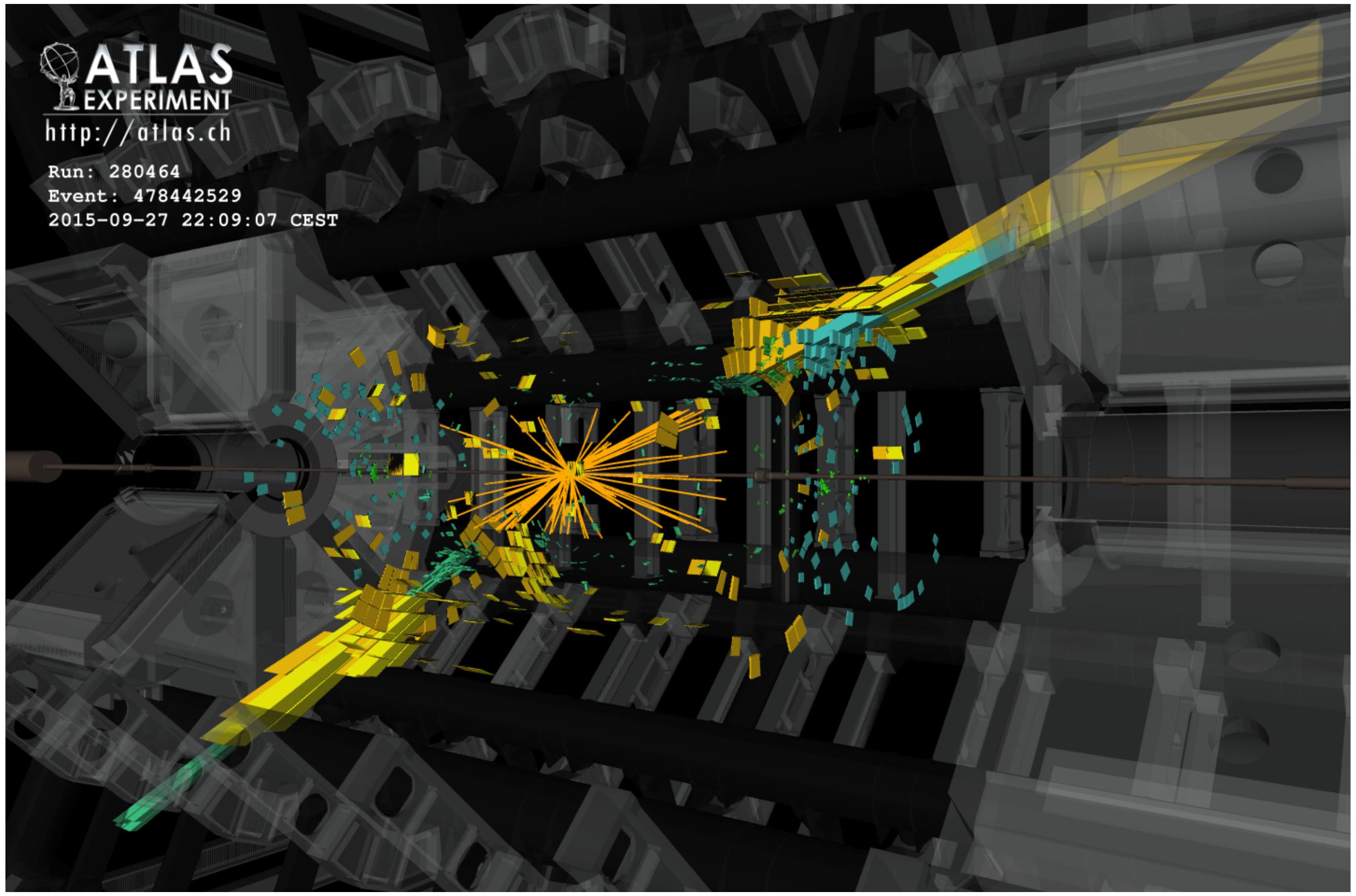
4-th IML Workshop

23/10/2020

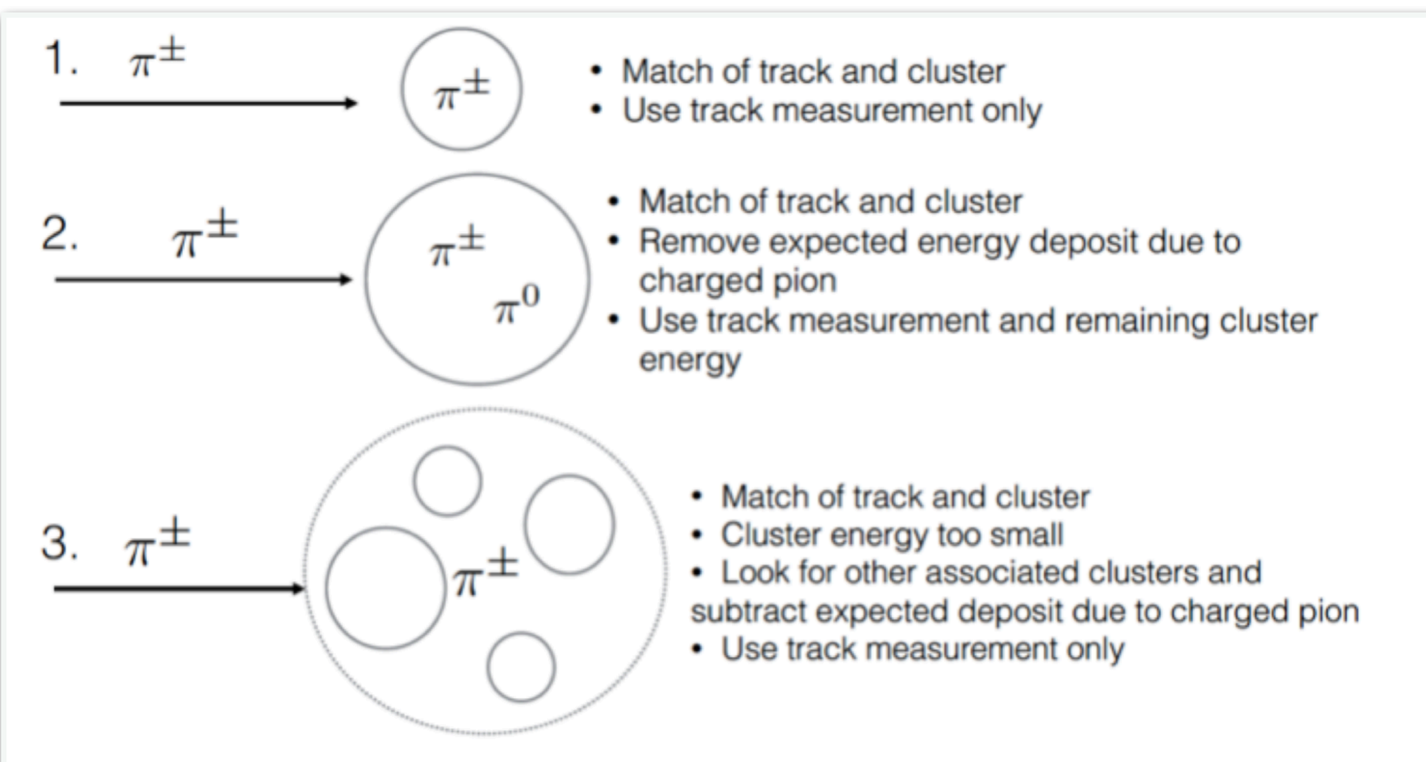
Outline

- Motivation & Baseline energy flow
- Detector Simulation
- The ML Architecture
- Implementation of PFlow
- Comparative Performance
- Summary

A real collision event reconstruction

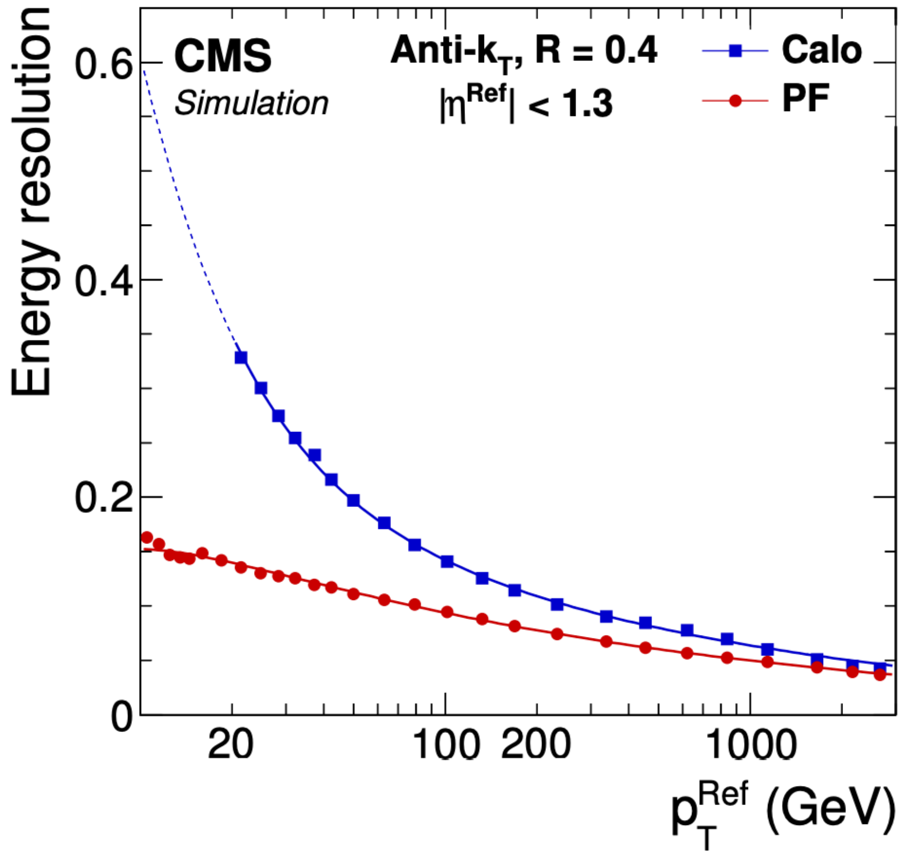


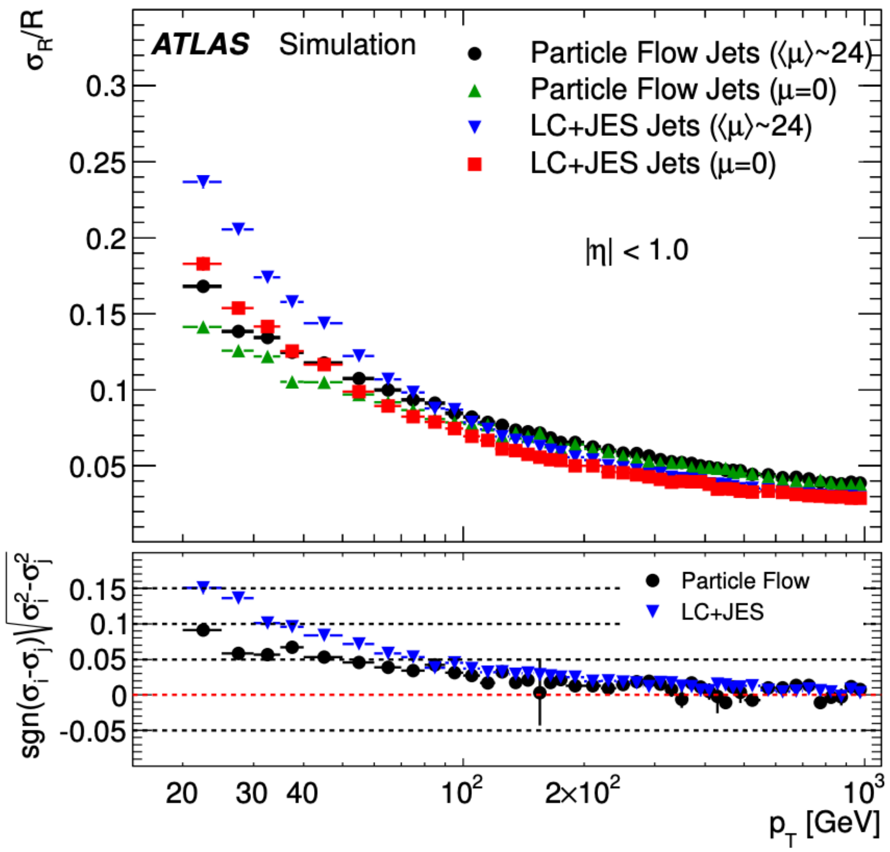
An accurate global event reconstruction (determining the 4-momenta of all the stable objects, combining the information from all sub-detector components) is crucial for understanding the underlying dynamics.



In general a PFlow algorithm tries to use all sub-detector information (track momenta, calorimeter cell energies etc) to reconstruct and identify the energy four-momenta of individual particles.

Difference between PFlow in ATLAS & CMS





	ATLAS	CMS
Tracking	arXiv : 1803.06991	
$1/p_T$ resolution	$0.05\% \times p_T/\mathrm{GeV} \oplus 1\%$ [47]	$0.02\% \times p_T/\mathrm{GeV} \oplus 0.8\%$ [48]
d_0 resolution (μm)	20 [49]	20 [48]
ECAL		
E resolution	$10\%/\sqrt{E} \oplus 0.2\%$ [45]	$3\%/\sqrt{E} \oplus 12\%/E \oplus 0.3\%$ [46]
granularity	0.025×0.025	0.017×0.017
HCAL		
E resolution	$50\%/\sqrt{E} \oplus 5\%$ [45]	$100\%/\sqrt{E} \oplus 5\%$ [50]
granularity	0.1×0.1	0.087×0.087

CMS combines the track & calorimeter information into unified PFlow object and forms PFlow jets.

ATLAS used calojets by default until now.

For CMS, the gain from using PFlow is large.
- **CMS used PFlow from Run-1**

ATLAS benefits less from PFlow :
- **better HCAL resolution**
- **smaller magnetic field**
- **longitudinal segmentation of calorimeter**

Motivation

Jet reconstruction and performance using particle flow with the ATLAS Detector (1703.10485)

Particle-flow algorithm is a generic event reconstruction technique. Its performance strongly depends on detector design

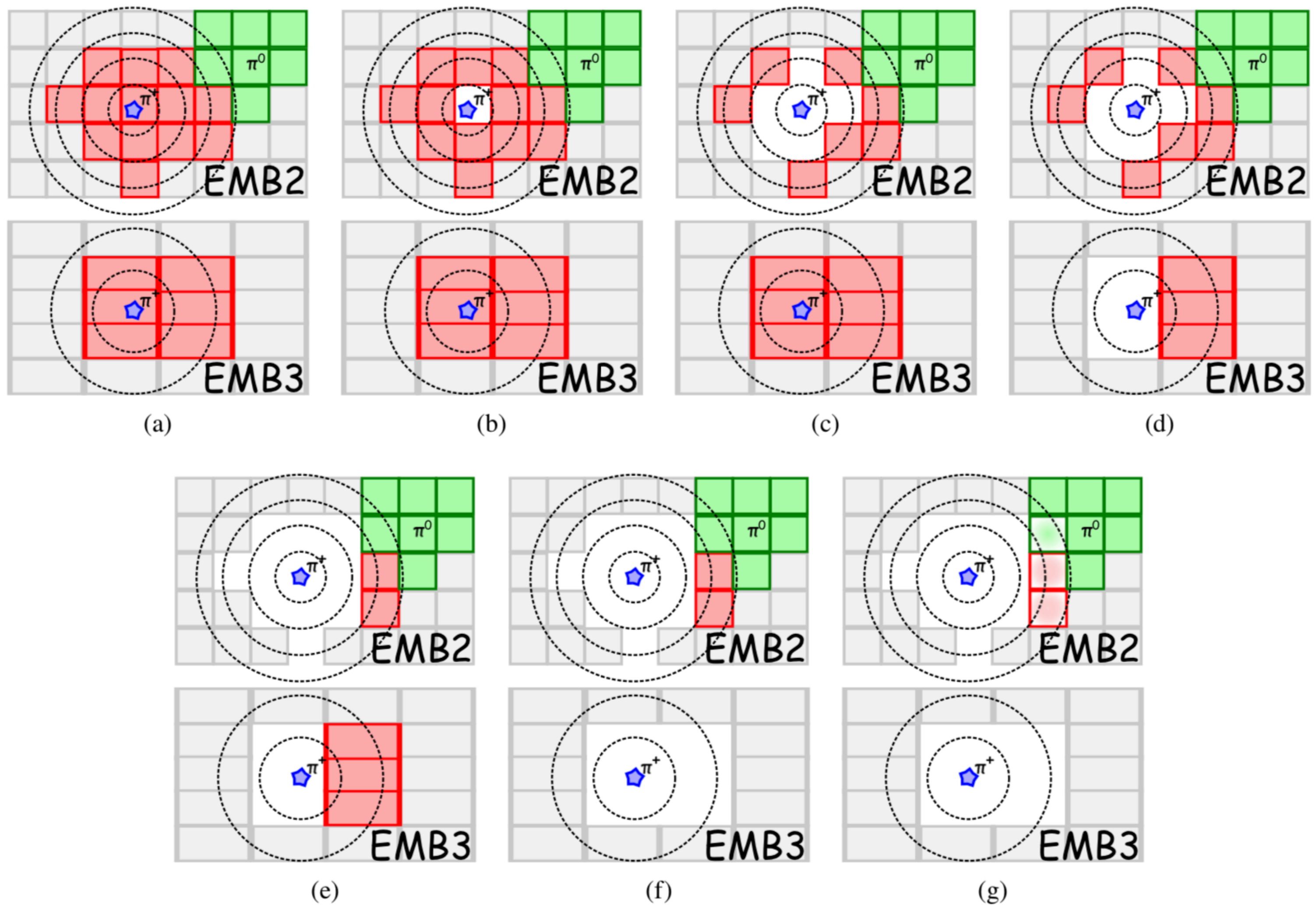
Energyflow algorithm :
Cluster the cells into calorimeter topoclusters and subtract the (statistically) expected energy deposit by charged tracks matched to the topocluster.

The task : predict fraction of neutral energy per cell of the topocluster

The Challenge : Different calorimeter layers have different resolutions.

Our proposal :
Implement a deep learning based method to extract the fraction of neutral energy for each cell in each layer of the ECAL and HCAL calorimeter layers.

Example of existing ATLAS algorithm as of now :



This work is inspired from ATLAS PFlow technique.

ATLAS PFlow has smaller gain compared to CMS.

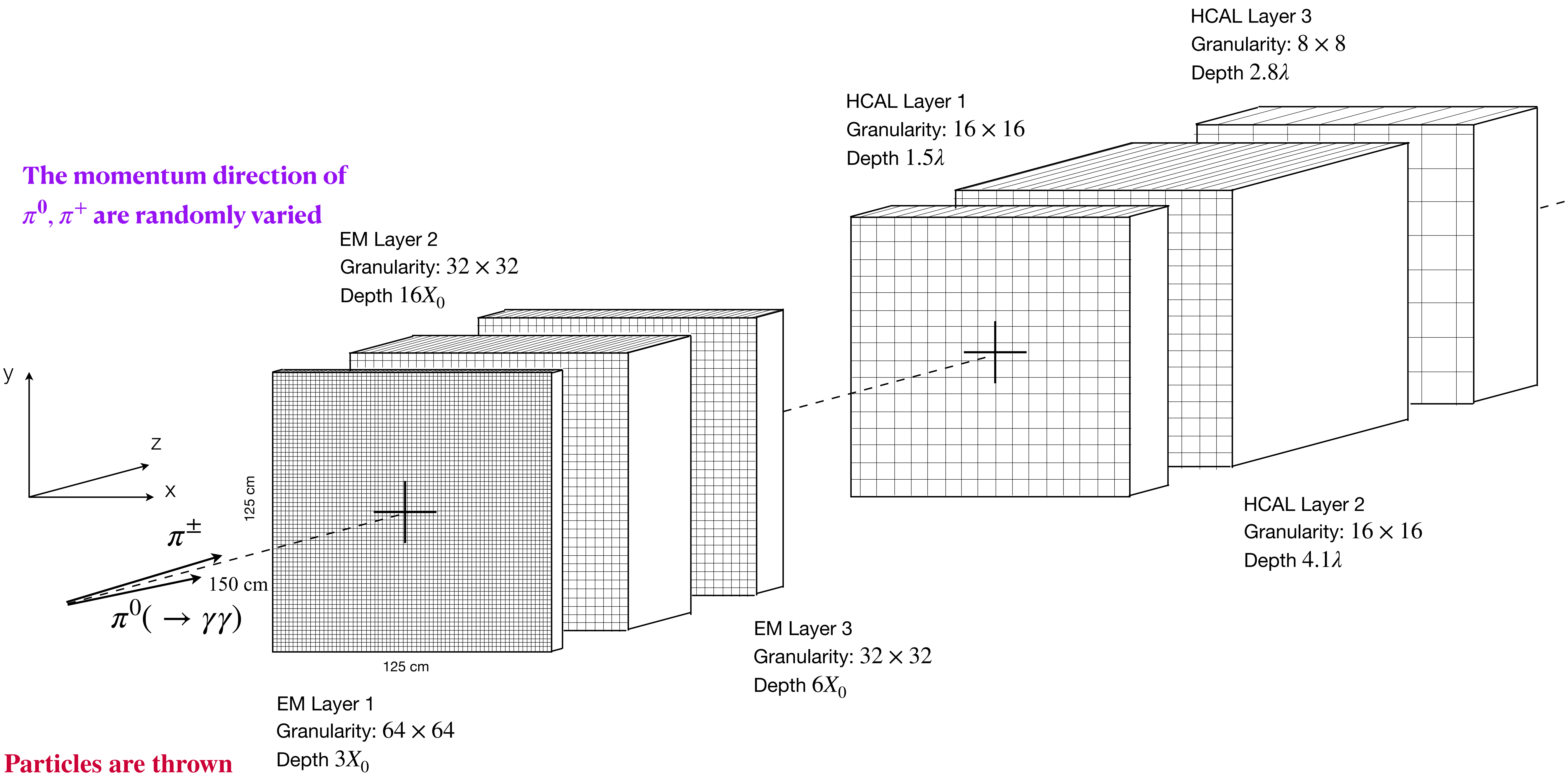
Outline

- ☑ Motivation & Baseline energy flow
- ▷ Detector Simulation
- ☐ The ML Architecture
- ☐ Implementation of PFlow
- ☐ Comparative Performance
- ☐ Summary

Calorimeter with GEANT

The impact of the PFlow highly depends on the granularity of the designed calorimeter.

The momentum direction of π^0, π^+ are randomly varied



Particles are thrown randomly from a corner of 20 cm X 20 cm square.

We make sure that there is a significant energy overlap per cell.

Detector parameters & noise

Detector	Absorber	Scintillator	Subdetector (Legth)
ECAL	Lead **1.2**	Liquid Argon **4.5**	ECAL1 (**3 X_0**) ECAL2 (**16 X_0**) ECAL3 (**6 X_0**)
HCAL	Iron **4.7**	Plastic organic **1.0**	HCAL1 (**1.5 λ_{int}**) HCAL2 (**4.1 λ_{int}**) HCAL3 (**1.8 λ_{int}**)

$X_0 = 3.9$ cm

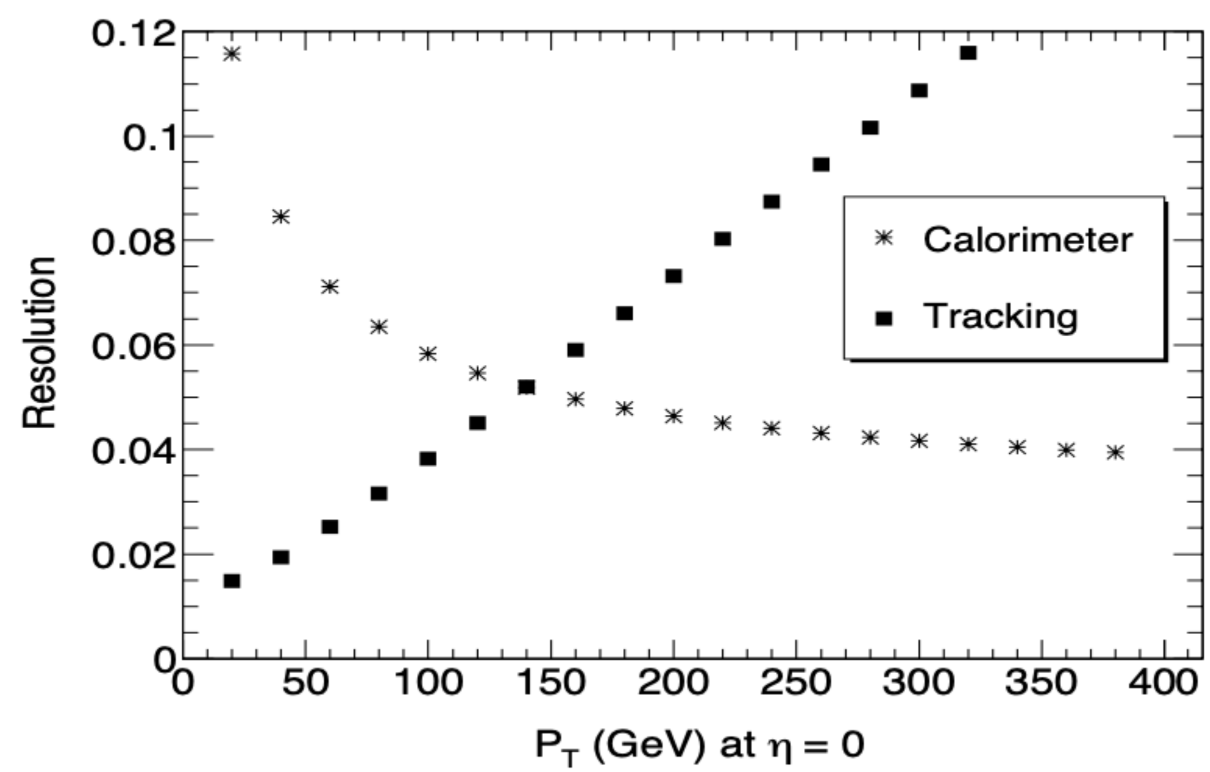
$\lambda_{int} = 17.4$ cm

Particles are generated in four different energy ranges [2-5], [5-10], [10-15] & [15-20] GeV.

Noise is added at per cell level
(used current ~ ATLAS values which includes PU)

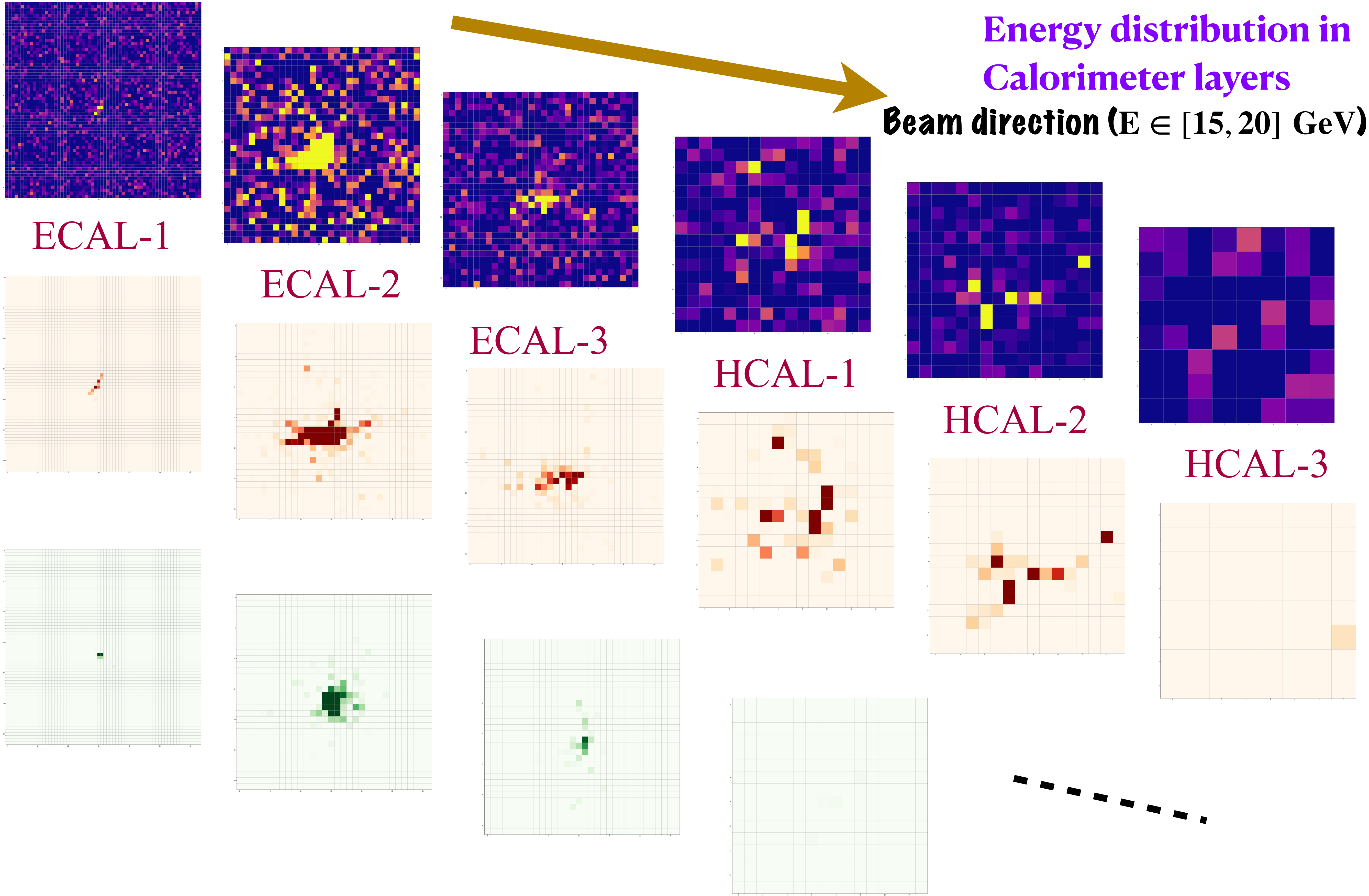
Detector Layer	Res. (HG)	Res. (LG)	Noise [MeV] (cf)
ECAL1	64×64	32×32	13 (4)
ECAL2	32×32	8×8	34 (16)
ECAL3	32×32	8×8	17 (16)
HCAL1	16×16	8×8	14 (4)
HCAL2	16×16	8×8	8 (4)
HCAL3	8×8	8×8	14 (1)

In these energy regimes tracker has better resolution compared to calorimeters.

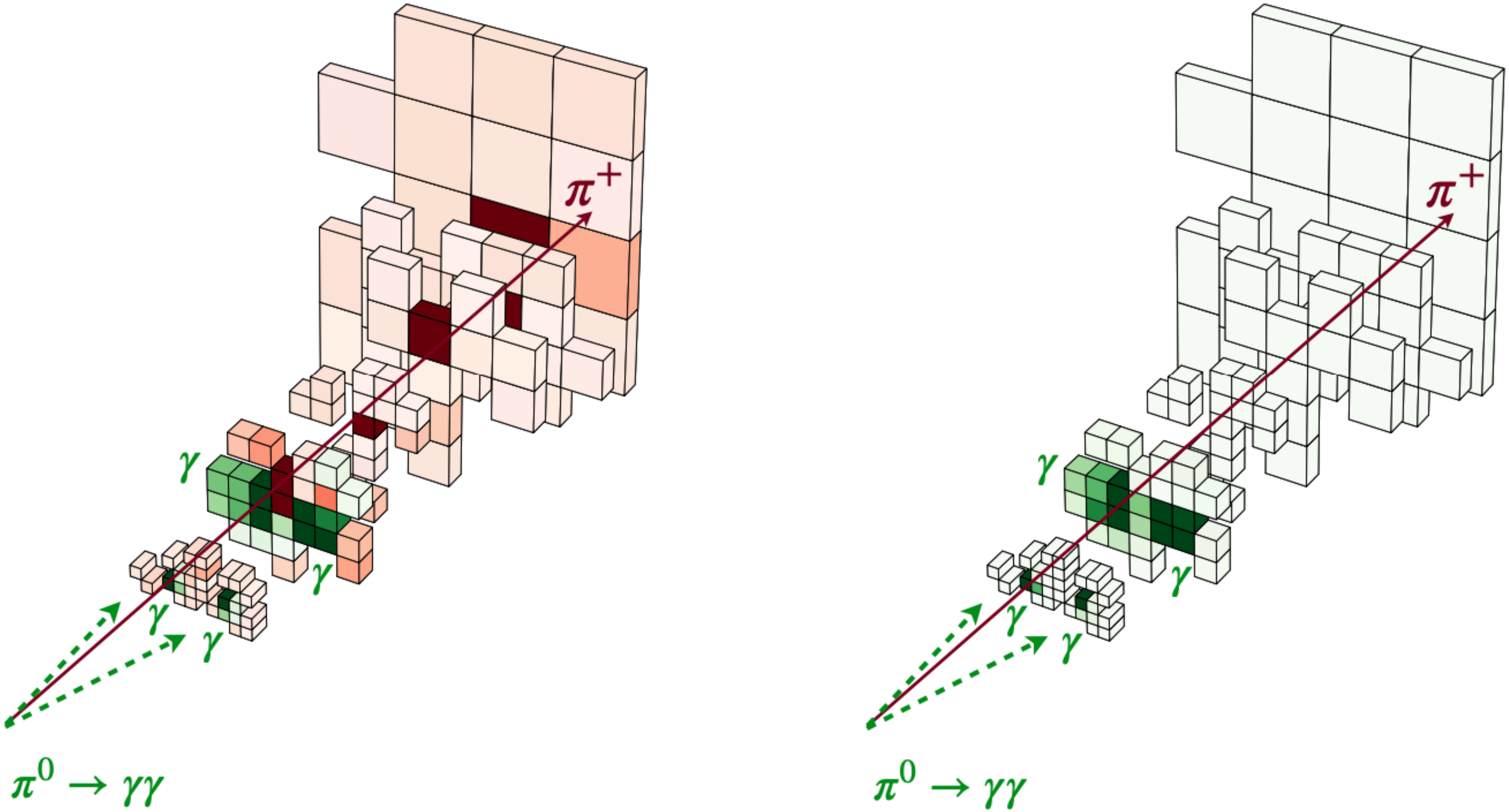


The LG detector configuration is used for superresolution studies (see Francesco's talk for details)

What does an event look like ?



Event display for topocluster

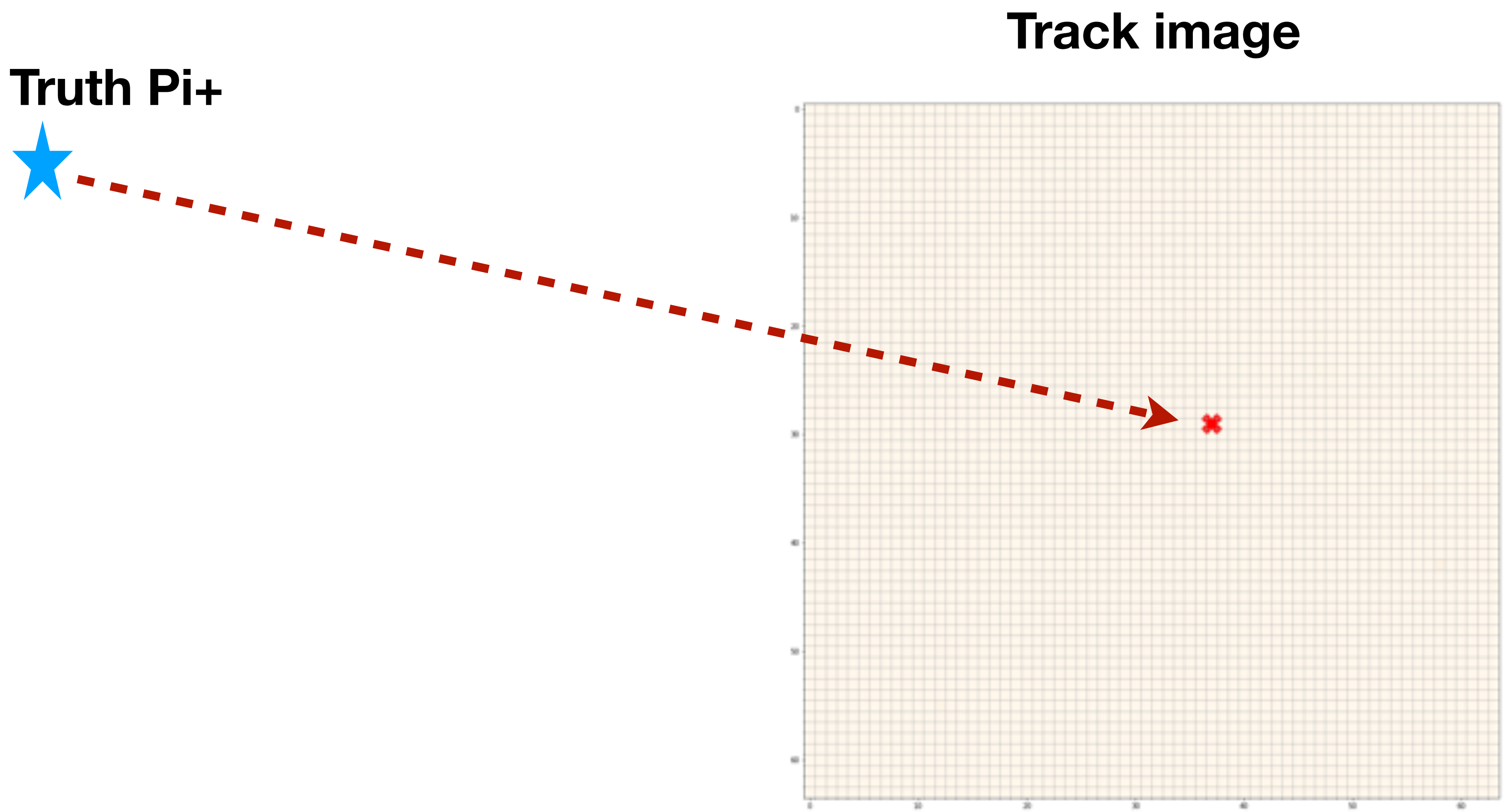


Introducing track image

Forming the track :

smear the truth Pi+ momenta : $\frac{\sigma(p_T)}{p_T} = 5 \times 10^{-4} \times p_T$ **[GeV]**

Keep the direction of Pi+ fixed (no magnetic field)



Outline

- ☑ **Motivation & Baseline energy flow**
- ☑ **Detector Simulation**
- ▶ **The ML Architecture**
- ☐ **Implementation of PFlow**
- ☐ **Comparative Performance**
- ☐ **Summary**

Machine learning task

We want to regress the neutral energy per cell

Input :

6 channel image

(Signal + Noise)

Layer1 —> 64 X 64

Layer2 —> 32 X 32

Layer3 —> 32 X 32

Layer4 —> 16 X 16

Layer5 —> 16 X 16

Layer6 —> 8 X 8

+

Track Layer

→ **NN model** →

Output :

6 channel image

(neutral energy frac)

Layer1 —> 64 X 64

Layer2 —> 32 X 32

Layer3 —> 32 X 32

Layer4 —> 16 X 16

Layer5 —> 16 X 16

Layer6 —> 8 X 8

$$L_{event} = \frac{1}{E_{tot}} \sum_c E_c (f_t^c - f_d^c)^2$$

E_c : Energy of a cell, $E_{tot} = \sum E_c$

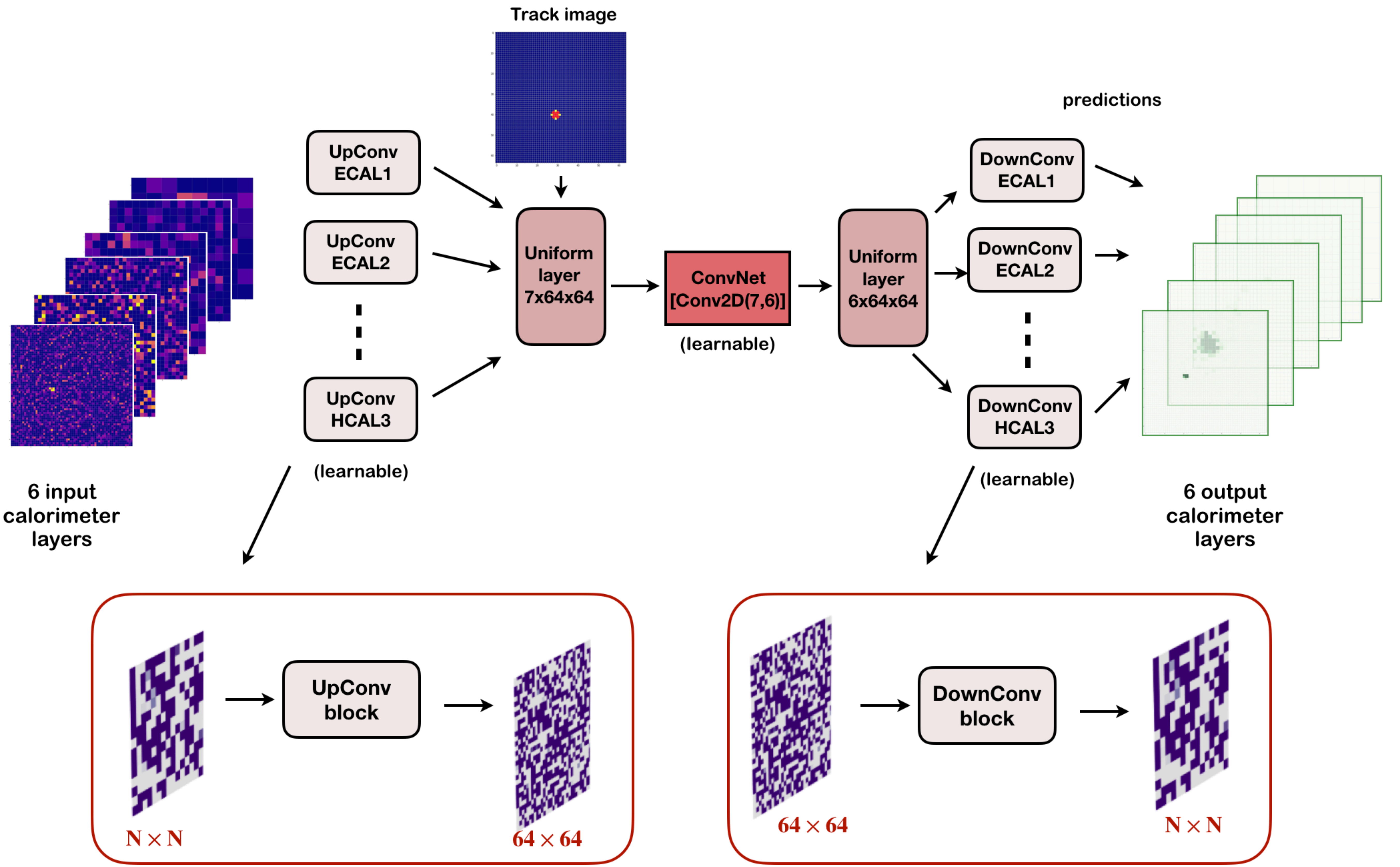
f_t^c : target neutral energy fraction

f_c^d : predicted neutral energy fraction

A simple L2 loss function doesn't serve the purpose, we need to put extra weights on highest seed cells inside a topocluster

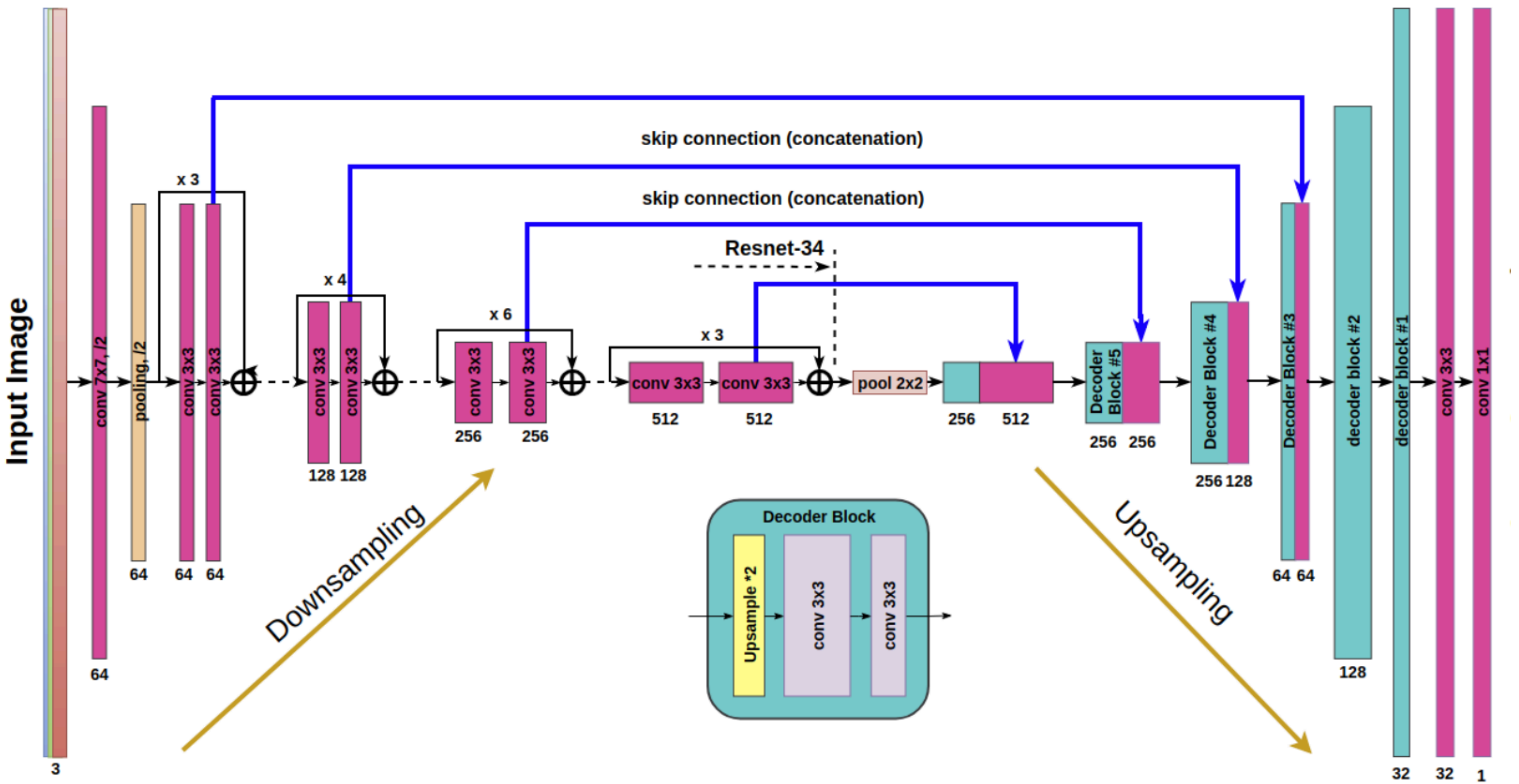
A weight factor of $\sqrt{E}$ won't work either .

The neural network architecture (pPFlow)



The architecture (cPFlow + Res-UNet)

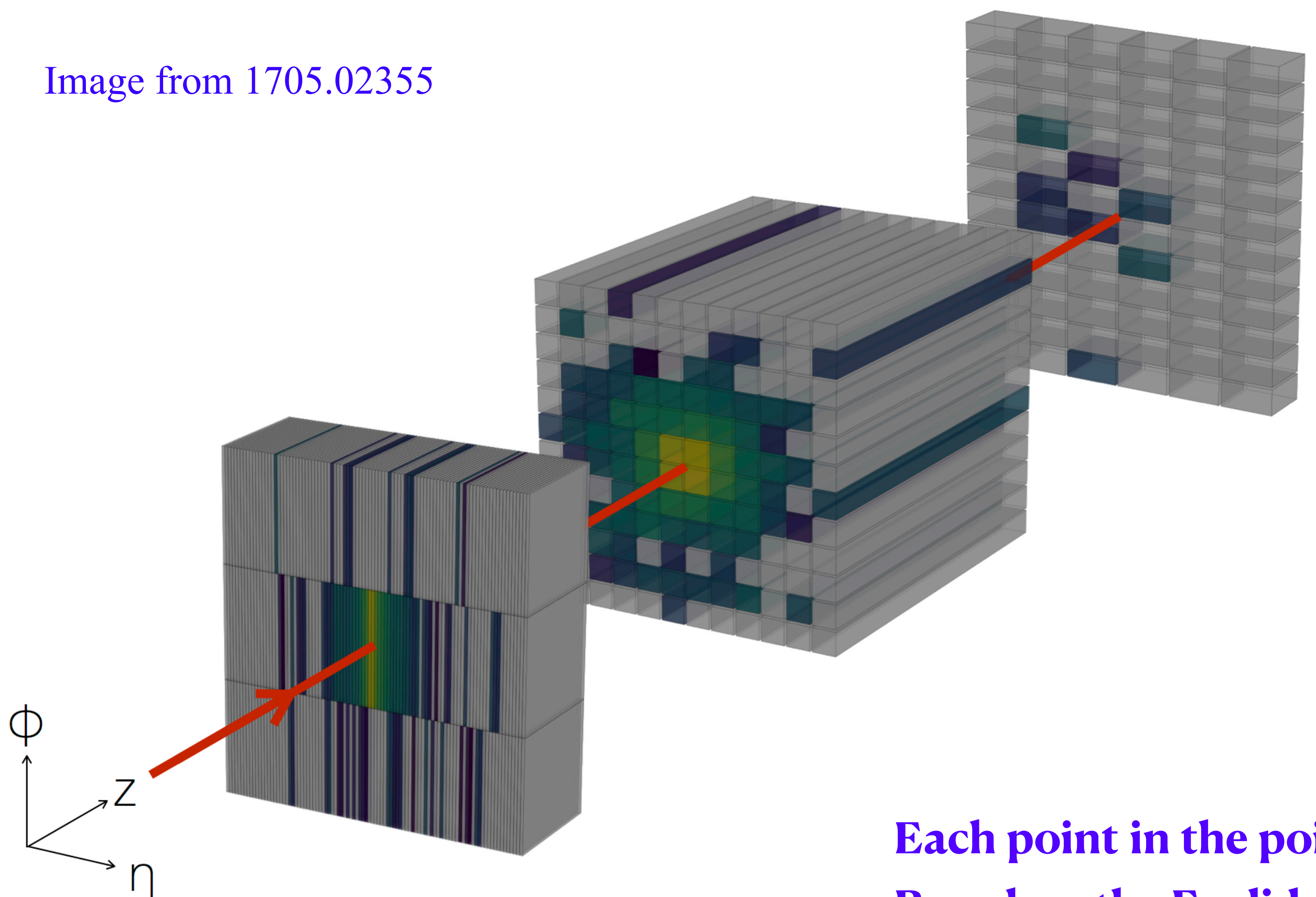
Image from : arXiv 1806.05182

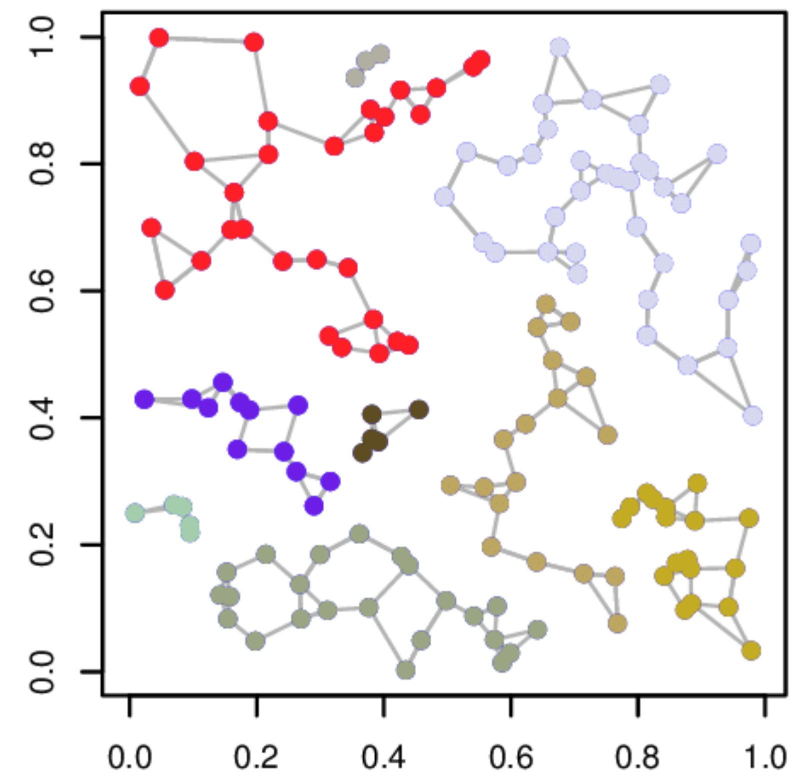


The graph network

Calorimeter showers have natural representation of a point cloud.

Image from 1705.02355





Each point in the point-cloud has 4 features (x, y, z, E). Based on the Euclidean distance among the points, one can form a K-nearest-neighbor graph

The graph network

https://arxiv.org/pdf/1801.07829.pdf

$$(x')_i^{l+1} = max_{j \in \mathcal{N}(i)} \, \Theta_x(x_j^l - x_i^l) + \Phi_x(x_i^l)$$

$$(e')_i^{l+1} = mean_{j \in \mathcal{N}(i)} \, \Theta_e(e_j^l - e_i^l) + \Phi_e(e_i^l)$$

Learned “coordinates”

$(\overrightarrow{x_1}, E_1)$

$(\overrightarrow{x_2}, E_2)$

$(\overrightarrow{x_3}, E_3)$

.

.

.

$(\overrightarrow{x_N}, E_N)$

$\longrightarrow$ $\mathrm{EC_{L1}(d_{in} = 3,\ d_{out} = M)}$ $\longrightarrow$

Forms a K-NN graph
and performs
the message passing

(z_1^M, E_1^{L1})

(z_2^M, E_2^{L1})

(z_3^M, E_3^{L1})

.

.

.

(z_N^M, E_N^{L1})

Learned “energy”

The deepset network

$$f(\mathbf{x}) = \sigma\left(\gamma\ \mathrm{I}\ \mathbf{x} - \lambda\ mean(\mathbf{x})\ \mathrm{I}\ \right),\ x = (\overrightarrow{x}, E)$$

Learned "features"

$$\begin{matrix} (\overrightarrow{x_1}, E_1) & & & & (z_1^M, E_1^{L1}) \\ (\overrightarrow{x_2}, E_2) & & & & (z_2^M, E_2^{L1}) \\ (\overrightarrow{x_3}, E_3) & \longrightarrow & \mathbf{DS_{L1}(d_{in} = 4,\ d_{out} = M)} & \longrightarrow & (z_3^M, E_3^{L1}) \\ \vdots & & & & \vdots \\ (\overrightarrow{x_N}, E_N) & & & & (z_N^M, E_N^{L1}) \end{matrix}$$

Outline

- ☑ Motivation & Baseline energy flow
- ☑ Detector Simulation
- ☑ The ML Architecture
- ➢ Implementation of PFlow
- ☐ Comparative Performance
- ☐ Summary

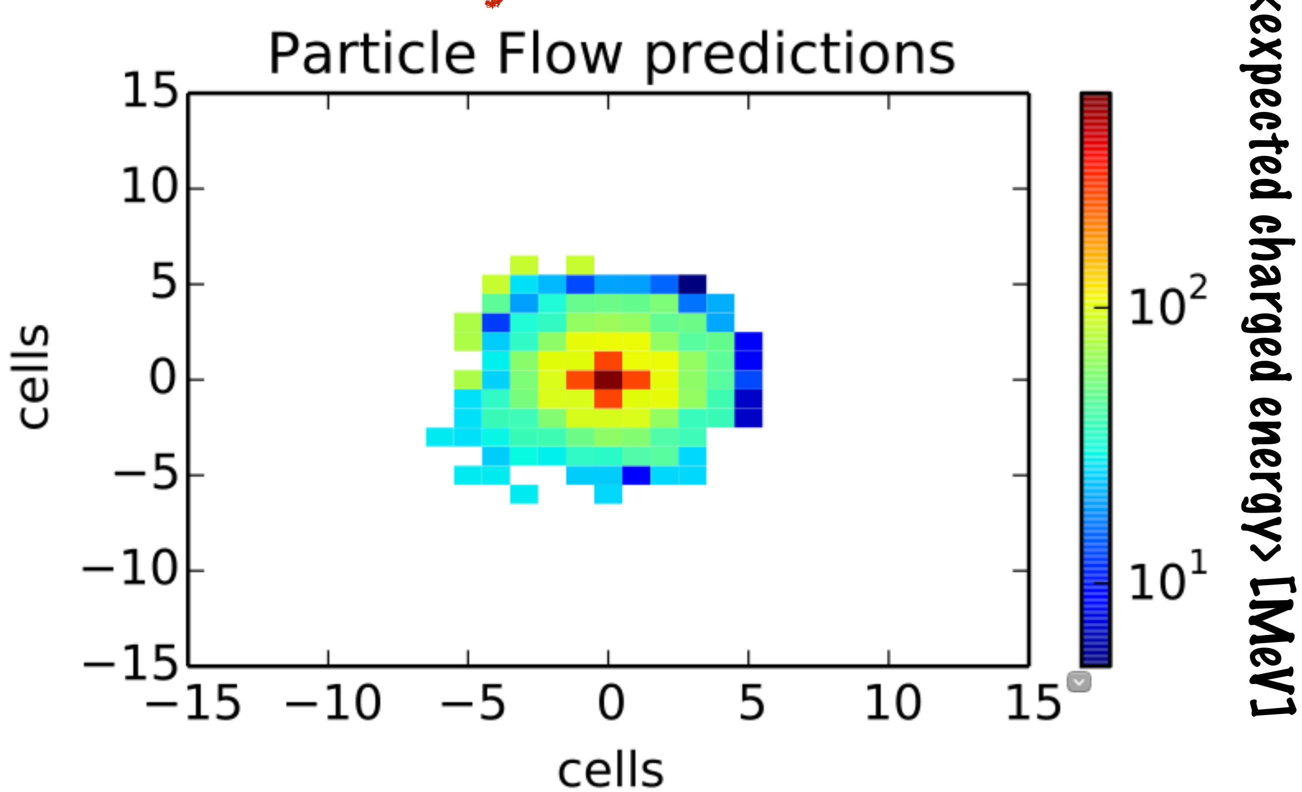
The Parametrized ParticleFlow (pPFlow)

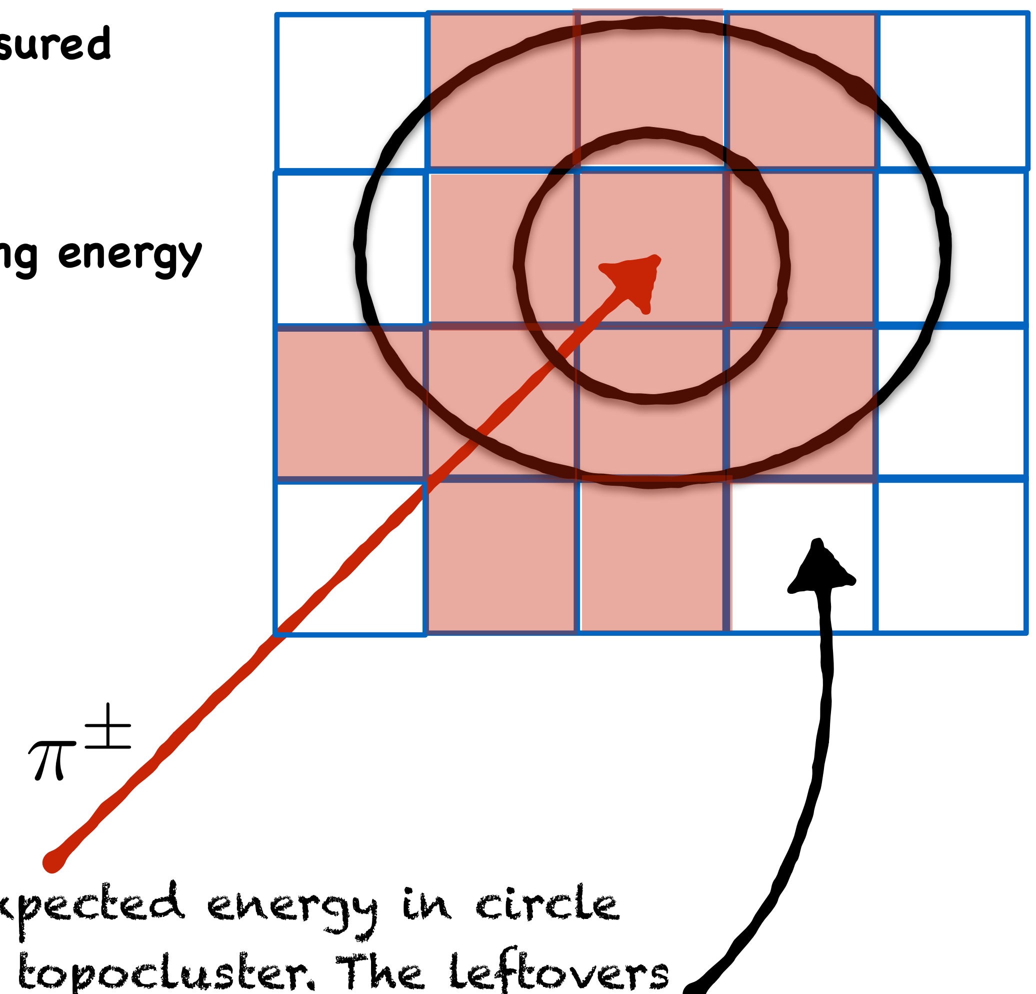
1. Group together the cells from their energy —-> topoclustering.
   First identify the seed cells ($\frac{E}{\sigma} \geq 5$) and merge with it neighboring cells with $\frac{E}{\sigma} \geq 2$
   Needed to remove the noise and cluster together cells fired by a common source (pi+, pi0)

2. Parametrize the expected energy given the momentum of the track (measured from the inner tracker, the layer of the "first" nuclear interaction).

3. The expected energy is subtracted from the topocluster and the remaining energy is considered as neutral energy

Example of the parametrization in the first calorimeter layer





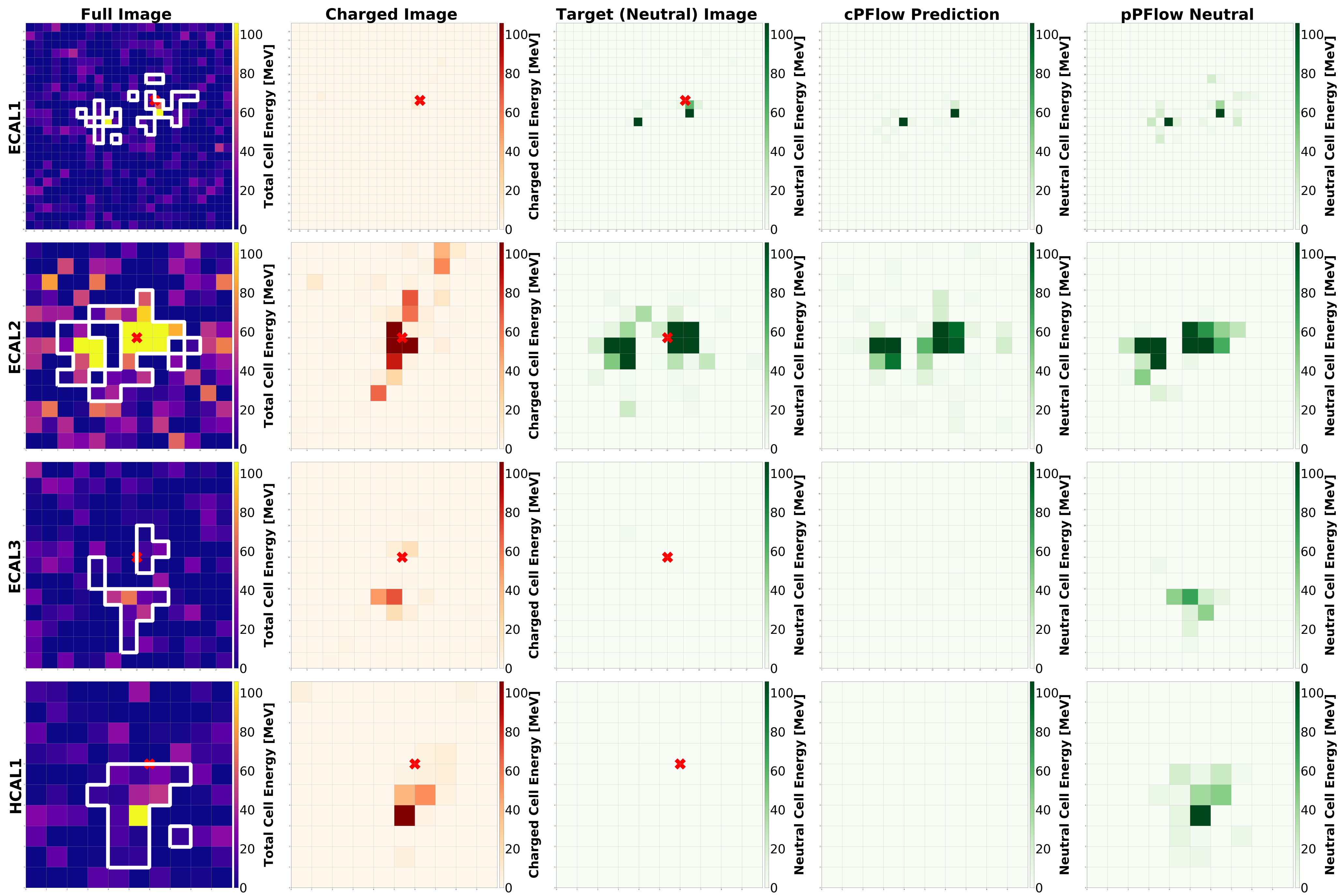
Subtract expected energy in circle within the topocluster. The leftovers energy is considered as the neutral component

This method is used as a baseline for comparison

Outline

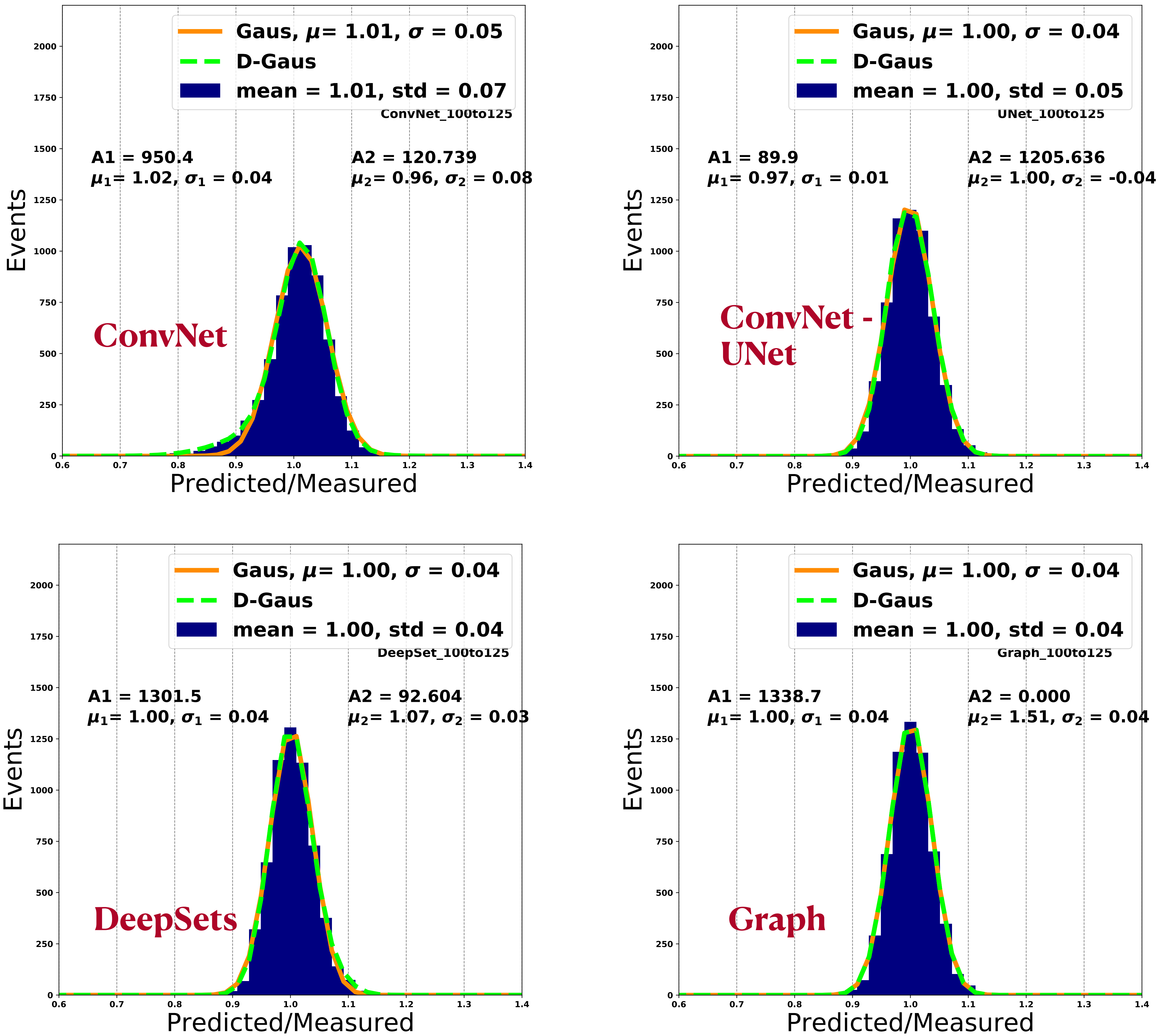
- ☑ Motivation & Baseline energy flow
- ☑ Detector Simulation
- ☑ The ML Architecture
- ☑ Implementation of PFlow
- ▶ Comparative Performance
- ☐ Summary

An event display (2-5 GeV)



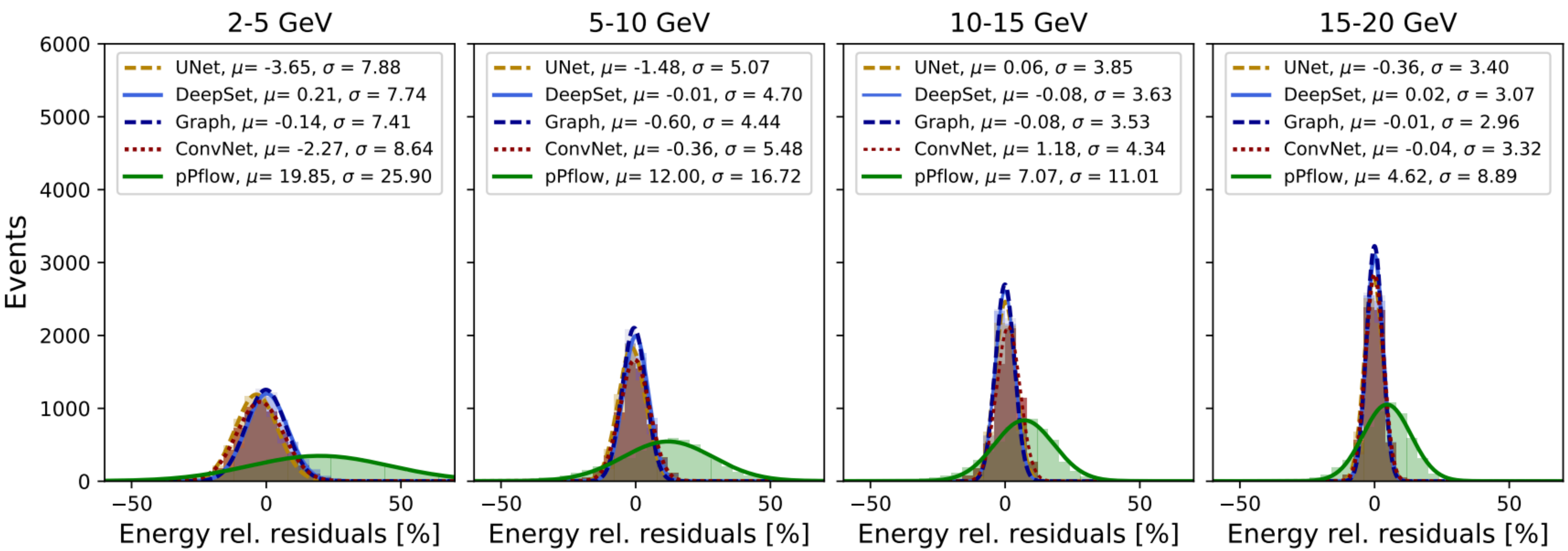
The trained network efficiently suppress the noise contribution

Comparison of training



Energy response comparison

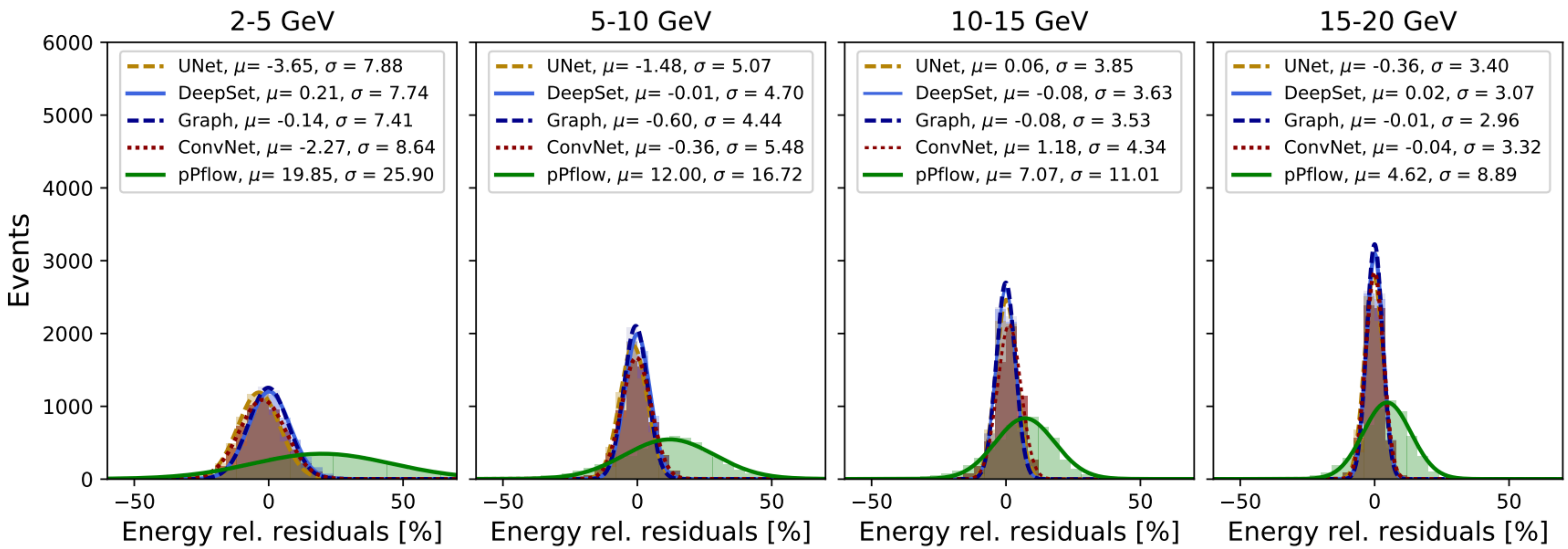
At low energy the GNN has 4X better resolution than traditional pPFlow



$$\textbf{Relative Residual} = \left(\frac{E_{predicted} - E_{neutral}}{E_{neutral}}\right)$$

Energy response comparison

With the current model, we observe a small bias + a Non-Gaussian tail
For image based methods.

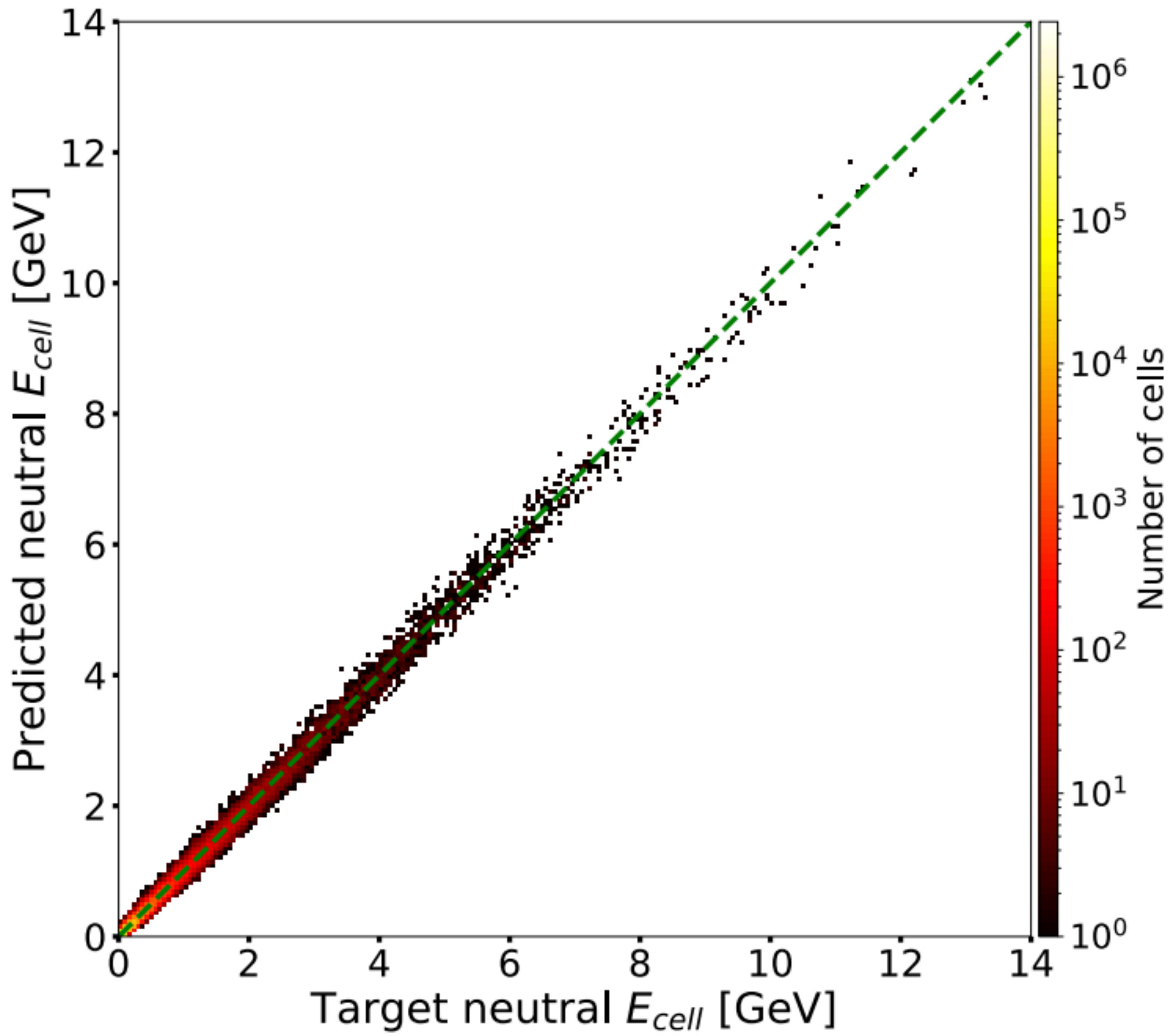


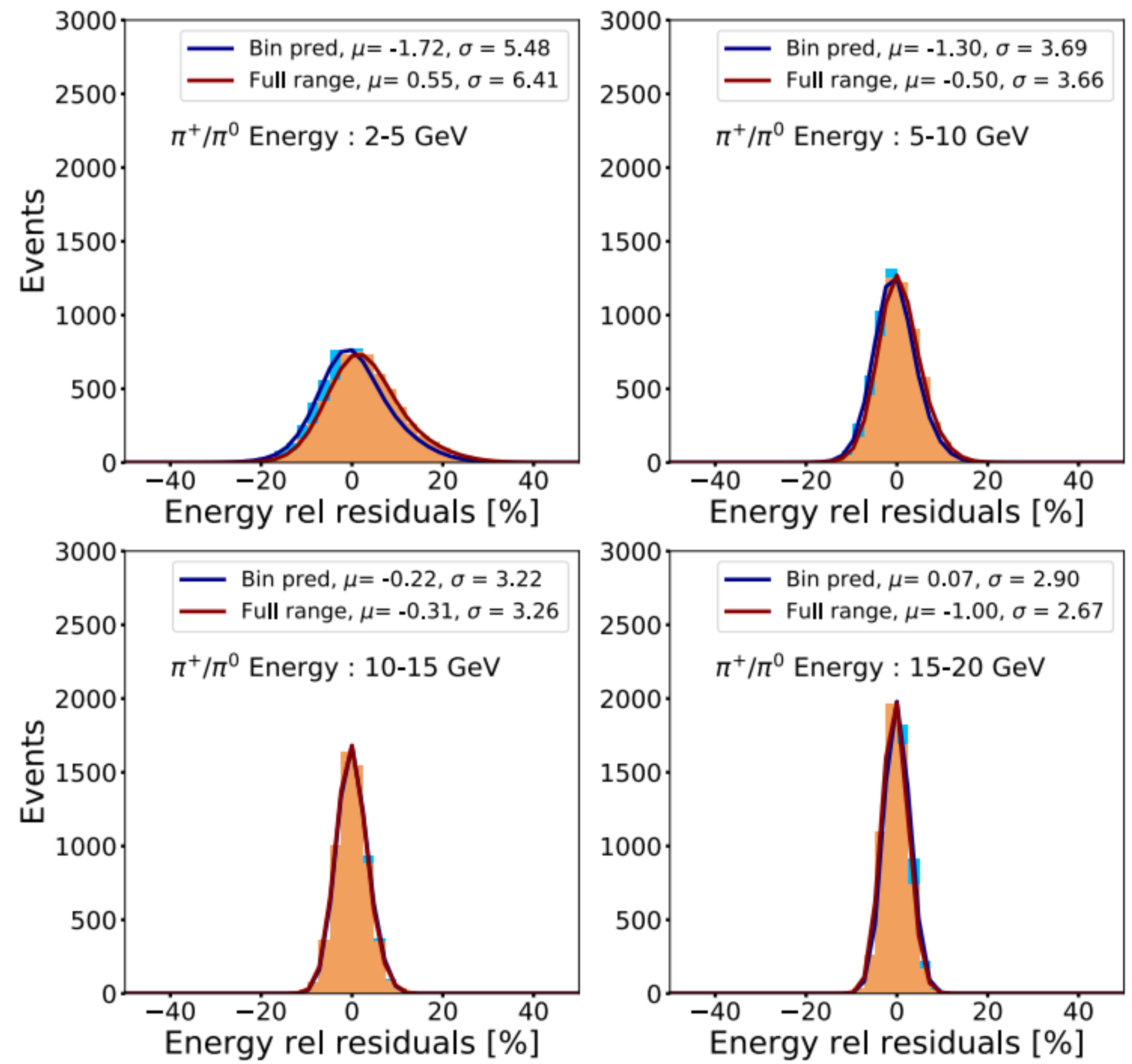
- The current NN trained on full images, the performance is evaluated only within topocluster. Hence an underestimate of the mean residual within topoclusters.

- The soft cells are treated as per noise & hence the NN predicts zero energy for those cells. Compensates by over-predicting the hard cells, leading to the Non-Gaussian tail.

- The bias in the pPFlow originates from systematically lower energy & size of the topocluster of single π^+ samples.

Cell level performance

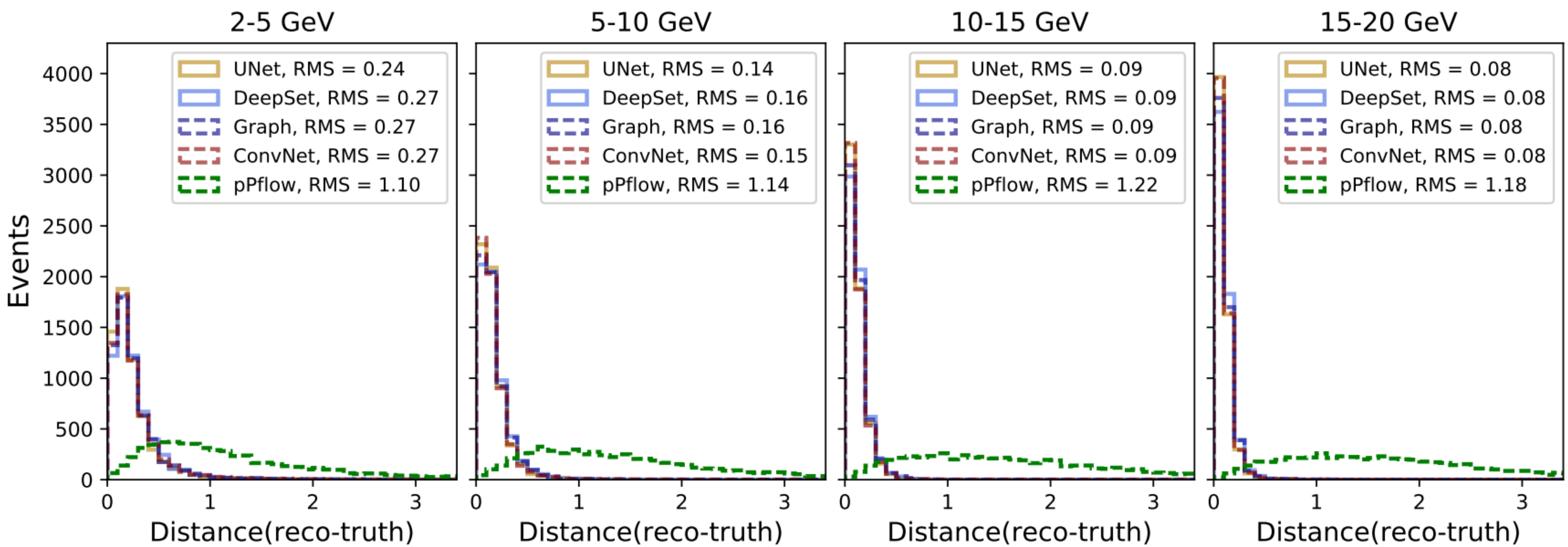




Per cell energy regression over the entire energy range

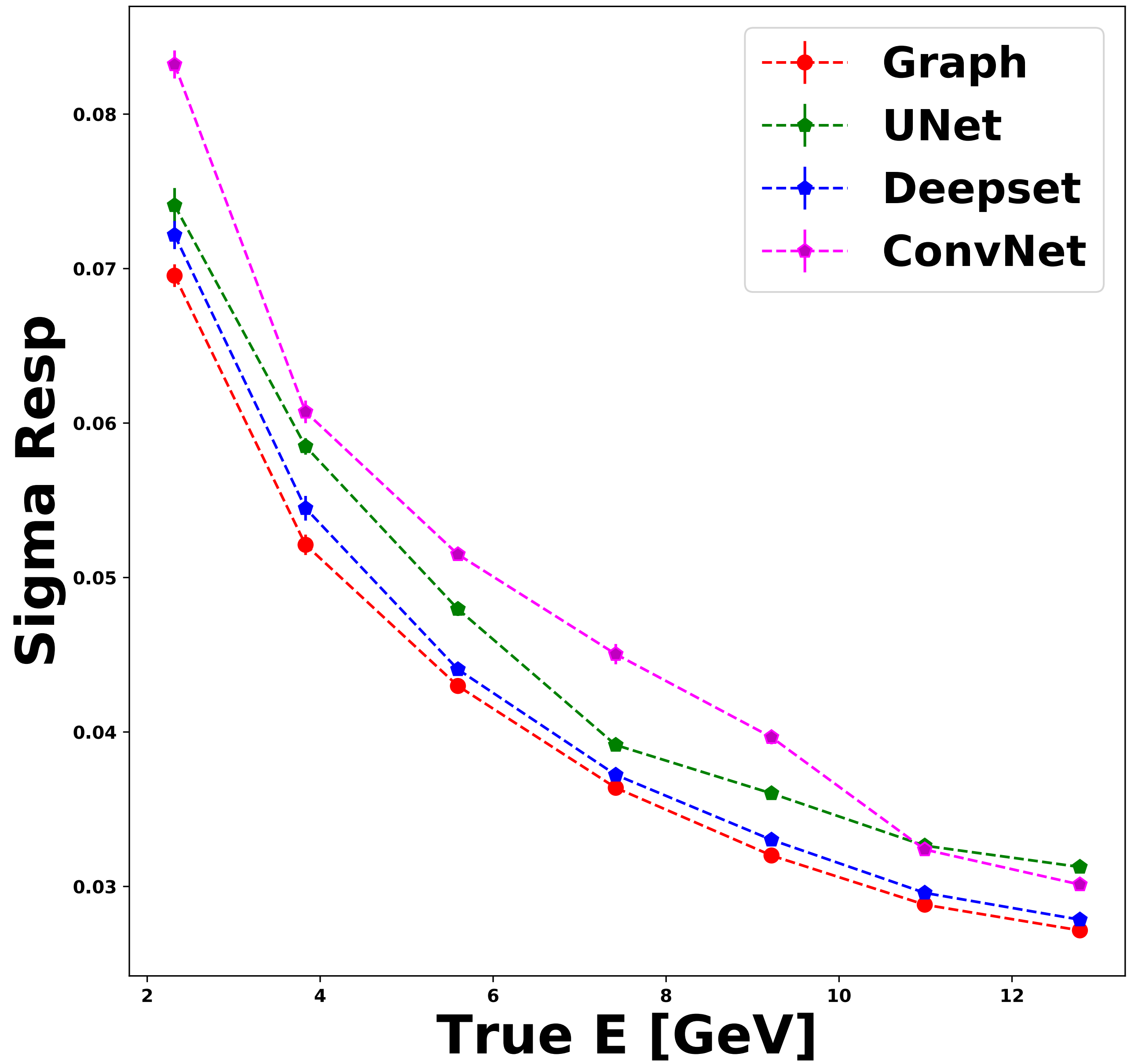
Direction response comparison

The distance computed in number of cells between the barycenter of the predicted and truth neutral energy in the ECAL2 layer.



The pPFlow algorithm has much better (upto 6 X) spatial resolution than traditional PFlow

Stability of different networks



Graph networks have best stability and resolution followed by deepsets.

Outline

- ☑ Motivation & Baseline energy flow
- ☑ Detector Simulation
- ☑ The ML Architecture
- ☑ Implementation of PFlow
- ☑ Comparative Performance
- ▶ Summary

Summary

- ☑ We have demonstrated that a suitable ConvNet, Graph, Deepset architecture gives estimation of the energyfraction for the generalized case :
  Input —> Variable Resolution + Noise + Track, Output —> Real resolution.

- ☑ The algorithm actually succeeds in yielding a complete image of neutral energy profile of the layers.

- ☑ The trained NN is able to learn and predict the noise pattern.
  A network trained on topoclusters has better performance on the topoclusters.

- ☑ These ML based algorithms are shown to improve the energy and direction estimation over existing PFlow algorithm, in case of overlapping charged and neutral pions.
  3 to 5 times resolution improvement obtained at low energy regime.

- ☑ The next target is to extend the work to jets and eventually implement in real experimental analysis.

- ☑ The ML algorithm at its current form is applicable to ATLAS as it is based on topoclusters.
  Will perform further studies in complex environment before implementation.

And what's next in this direction ?

Our detector design allows us to build a low granularity calorimeter by merging the cells but keeping the high-resolution truth information.

This can be used to establish super-resolution techniques for calorimetry :

Stay tuned for the next talk by Francesco Di-Bello

THANK YOU !!