Study of individual particle components within hadronic shower using graph neural networks

First steps and preliminary performance results

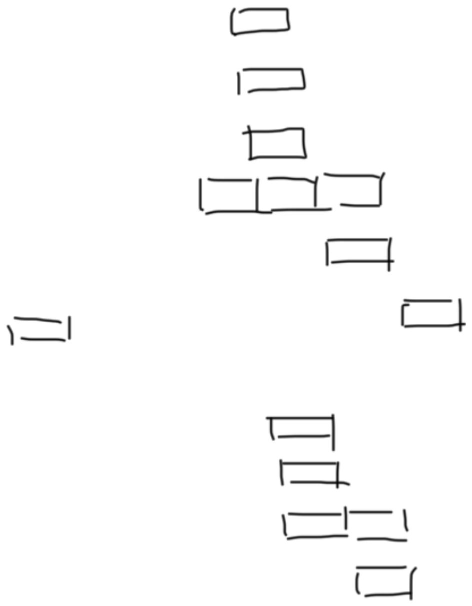
CALICE Collaboration meeting

Vladimir Bocharnikov[1],
[1]DESY, Hamburg
Mar 26 2021

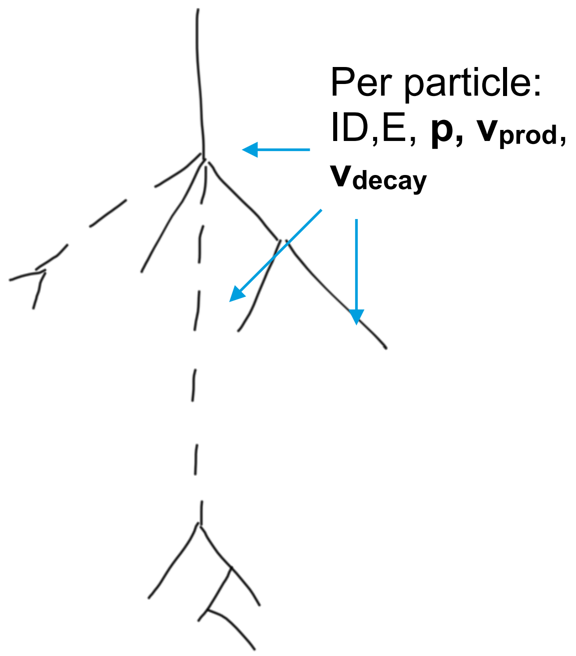
Calorimeter vision for hadronic showers

Ultimate goal and general approach

Set of hits in highly granular calorimeter



Particle interaction tree



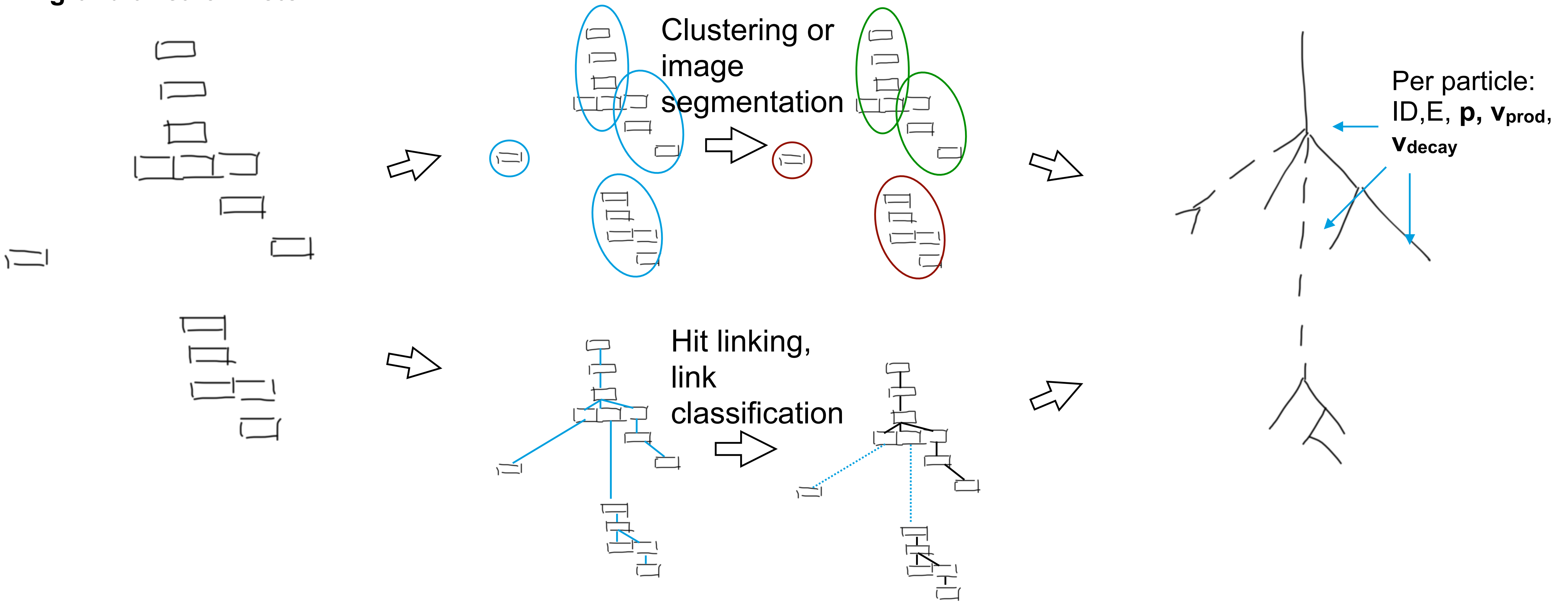
Calorimeter vision for hadronic showers

Ultimate goal and general approach

Set of hits in highly granular calorimeter

Typical ML tasks

Particle interaction tree

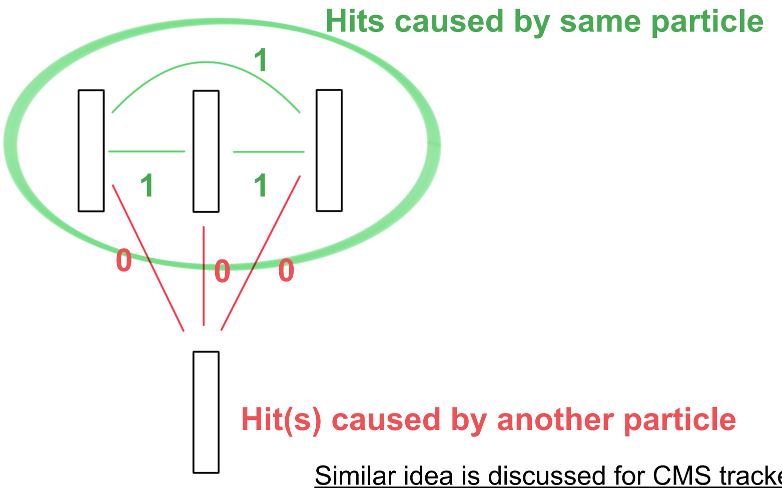


Graph representation of calorimeter event

First steps

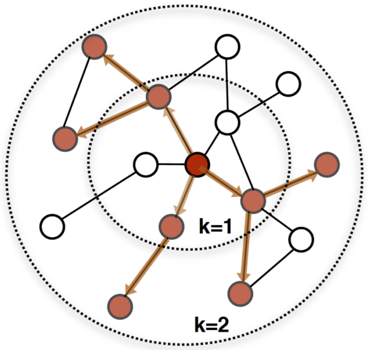
Event graph:

- Nodes - calorimeter hits
- Node features - position, energy, (time)
  - — Edges - neighbours within distance < Rmax (Radius graph)
  - — Edge weights - 1 if pair of hits belong to same **fundamental object** (e/m sub-shower, track), otherwise 0
- ML **objective** - **predict edge weights** given the radius graph of event

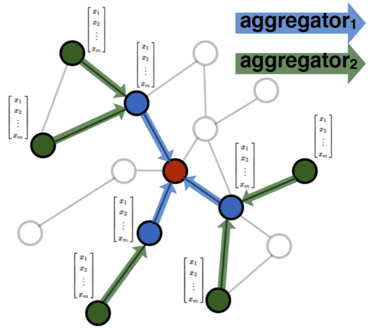


Similar idea is discussed for CMS tracker and HGCAL

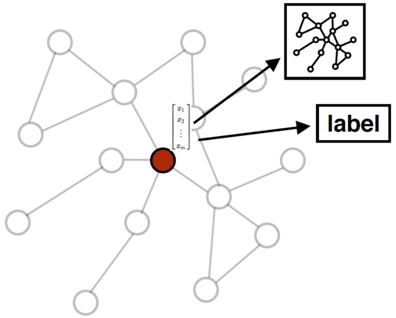
GraphSAGE (**SA**mple and aggre**G**at**E**) architecture (Graph neural network model (GNN)):



Sample neighbourhood of graph nodes



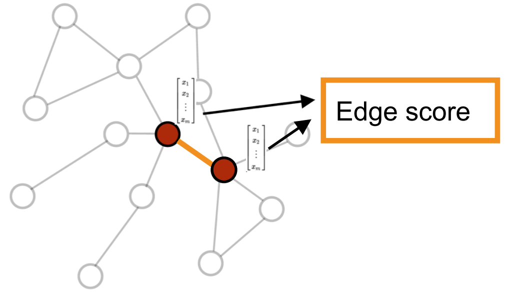
Aggregate feature information from neighbours



Get graph context embeddings for node using aggregated information

+

Custom output layer



Predict edge score for each pair of connected nodes using embedded features

Fundamental objects in hadronic showers

Shower components

<table>
<tr><td colspan="8">Hadronic shower</td></tr>
<tr><td colspan="4">MIP particles
Mostly charged hadrons</td><td>E/m showering particles
$e^{\pm}, \gamma, \pi^0, \eta$</td><td colspan="3">Challenging particles
n, K_0L, ν, …</td></tr>
<tr><td colspan="3">With shower</td><td>W/o shower</td><td rowspan="4">Topology:
e/m shower</td><td colspan="3" rowspan="2">Topology:
Hadronic shower/escape</td></tr>
<tr><td colspan="3">Topology:
MIP track (before shower) +
Hadronic shower</td><td rowspan="3">Topology:
MIP track</td></tr>
<tr><td>MIP particles</td><td>E/m showering particles</td><td>Challenging particles</td><td>MIP particles</td><td>E/m showering particles</td><td>Challenging particles</td></tr>
<tr><td>(…)</td><td>(…)</td><td>(…)</td><td>(…)</td><td>(…)</td><td>(…)</td></tr>
</table>

Fundamental objects in hadronic showers

Shower components

<table>
<tr><td colspan="8">Hadronic shower</td></tr>
<tr><td colspan="4">MIP particles
Mostly charged hadrons</td><td>E/m showering particles
$e^{\pm}, \gamma, \pi^0, \eta$</td><td colspan="3">Challenging particles
n, K_0L, ν, …</td></tr>
<tr><td colspan="3">With shower</td><td>W/o shower</td><td rowspan="4">Topology:
e/m shower</td><td colspan="3" rowspan="2">Topology:
Hadronic shower/escape</td></tr>
<tr><td colspan="3">Topology:
MIP track (before shower) +
Hadronic shower</td><td rowspan="3">Topology:
MIP track</td></tr>
<tr><td>MIP particles</td><td>E/m showering particles</td><td>Challenging particles</td><td>MIP particles</td><td>E/m showering particles</td><td>Challenging particles</td></tr>
<tr><td>(…)</td><td>(…)</td><td>(…)</td><td>(…)</td><td>(…)</td><td>(…)</td></tr>
</table>

Truth information from Monte-Carlo

Algorithm to find truth e/m objects

Simulations

Geant4 (v10.03.p02) QGSP_BERT_HP using CALICE AHCAL geometry

Pure energy deposition in cells (before digitalisation and reconstruction)

Truth electromagnetic sub-shower definition:

"**Electromagnetic**" particles: $e^{\pm}$,γ, π^0, η

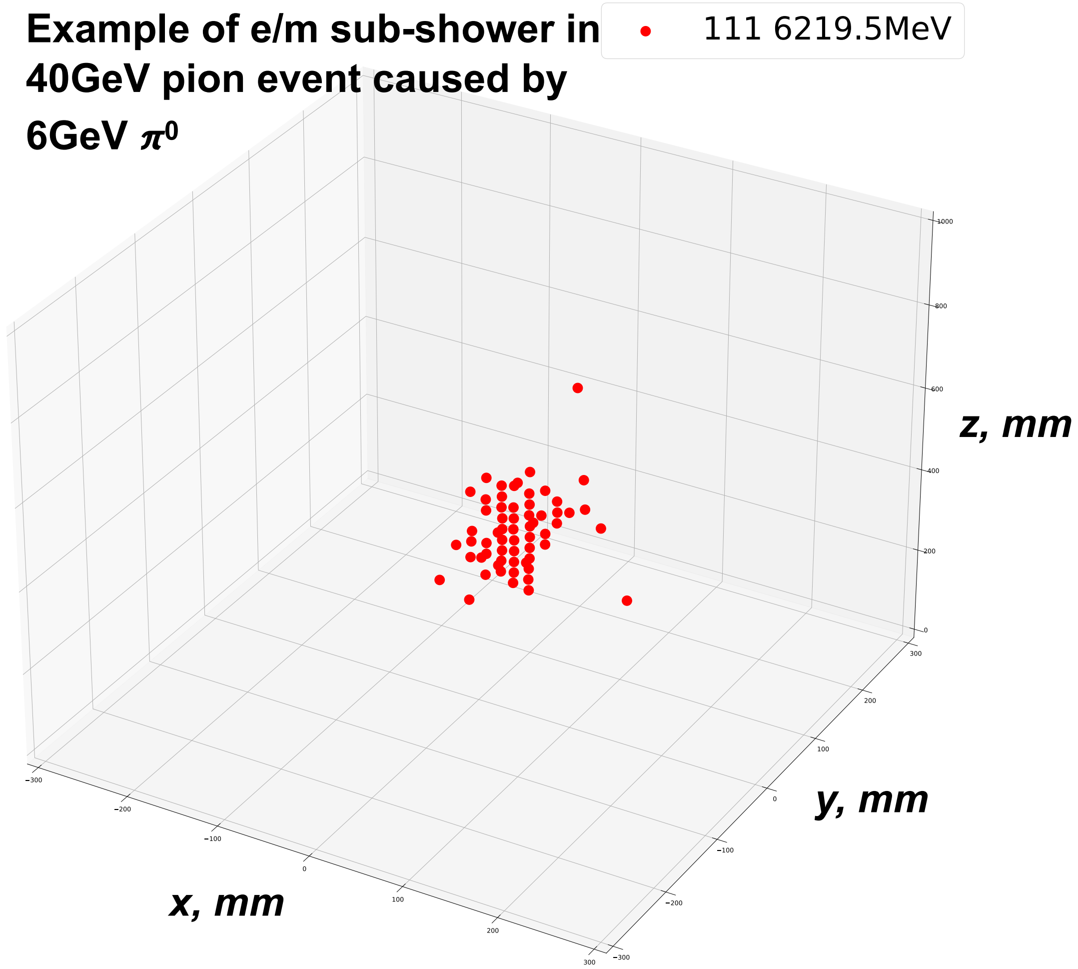
Energy threshold - *0.1GeV* (arbitrary now)

If MC particle is "electromagnetic", all it's "electromagnetic" daughters compose e/m shower and removed from further consideration

Corresponding simulated hits compose sub-shower,

0.5MIP cut: E_{hit}>0.25MeV

Example of e/m sub-shower in 40GeV pion event caused by 6GeV π^0



MC history for **ionising particles** is more complicated to easily define individual objects (tracks). Work in progress

Constructing e/m link graph

Input and training target

Input graph

- Radius graph with calorimeter hit nodes (x,y,z,E) *Rmax=45mm*
- Links - matrix of shape (Nlinks, 2) represent whether his are closer than Rmax, have no attribute

Hit_i	…	…	…	…
Hit_j	…	…	…	…

N_{links}

Target for e/m sub-showers

- For all e/m objects construct it's own radius graph and link attribute equal 1

0	1	…	1	0

N_{links}

Datasets and model parameters

Train & test

Train dataset

- 100 event graphs
  - Pure energy deposition in calorimeter cells (before digitalisation and reconstruction)
  - **10,20,40 GeV pion** samples
  - 88702 edges to process
    - 39994 “truth” electromagnetic links

Test dataset

- ~19K events
  - Pure energy deposition in calorimeter cells (before digitalisation and reconstruction)
  - **10,20,40 GeV pion** samples

Model

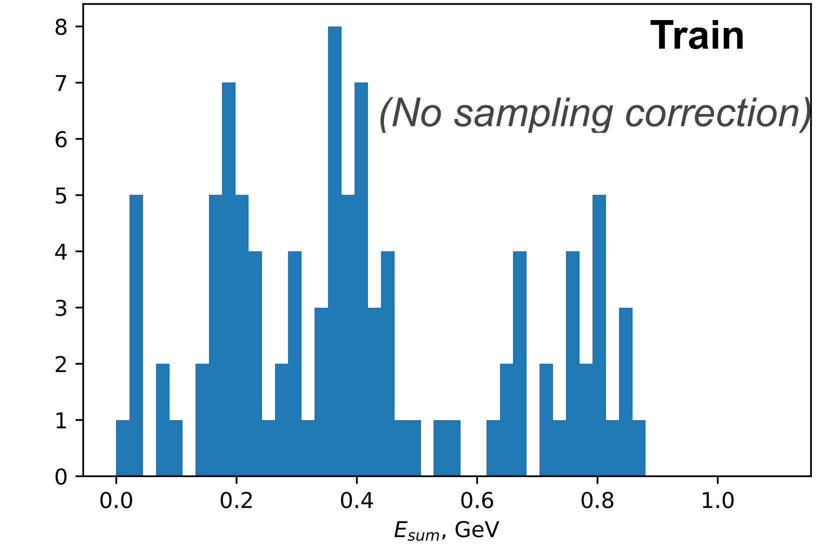
GraphSAGE GNN

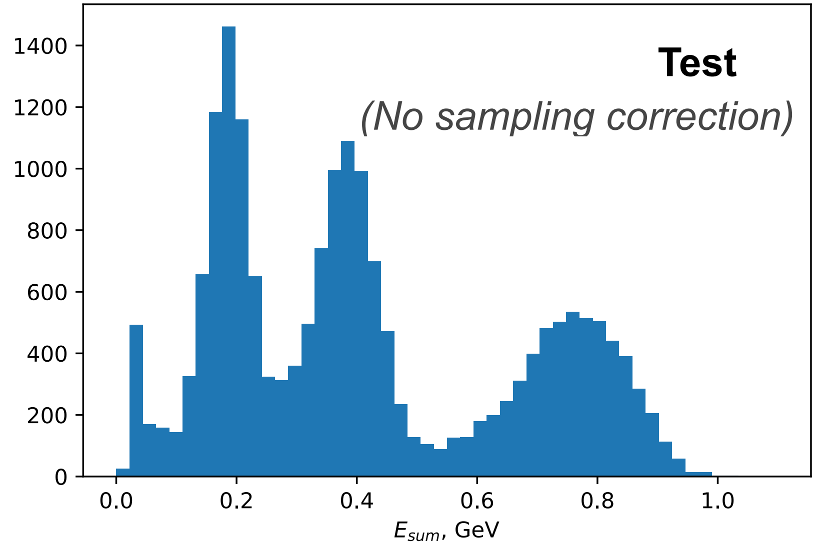
8 layers with 16 hidden channels + 1 linear output layer to convert node embeddings to edge scores

Prediction of edge scores

Binary cross entropy loss

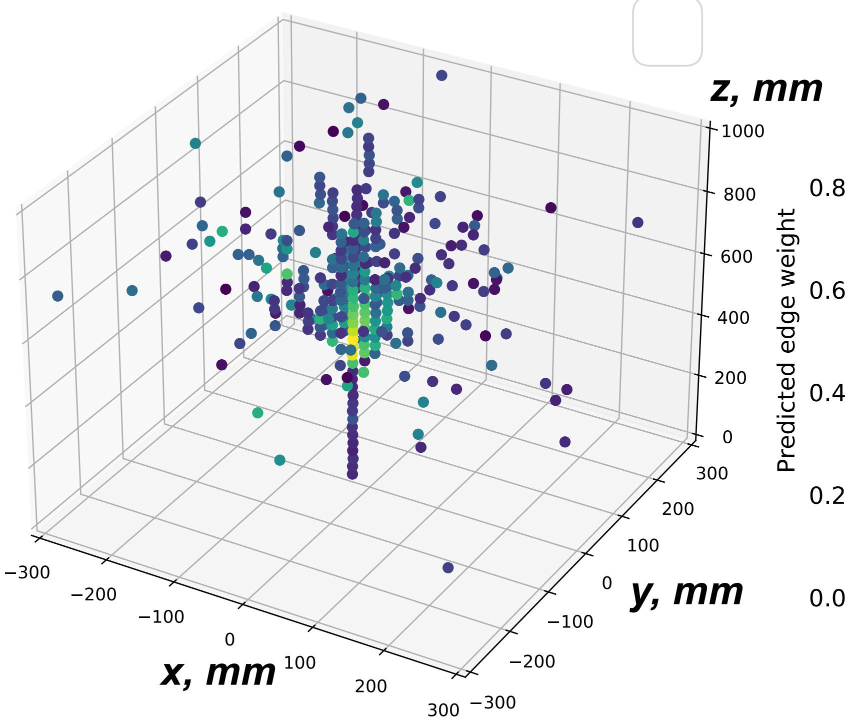
Energy sum distributions for datasets



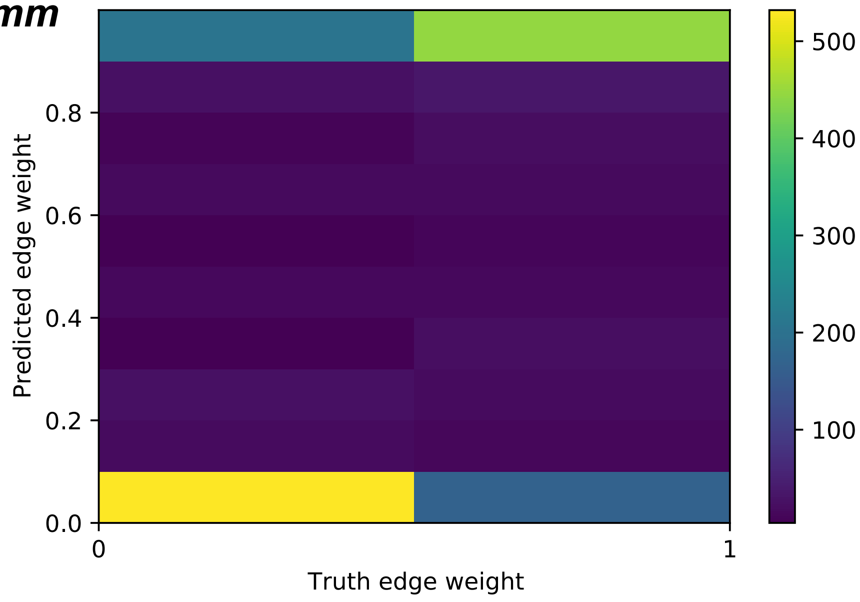


Example of output for test event

Preliminary results for single test event

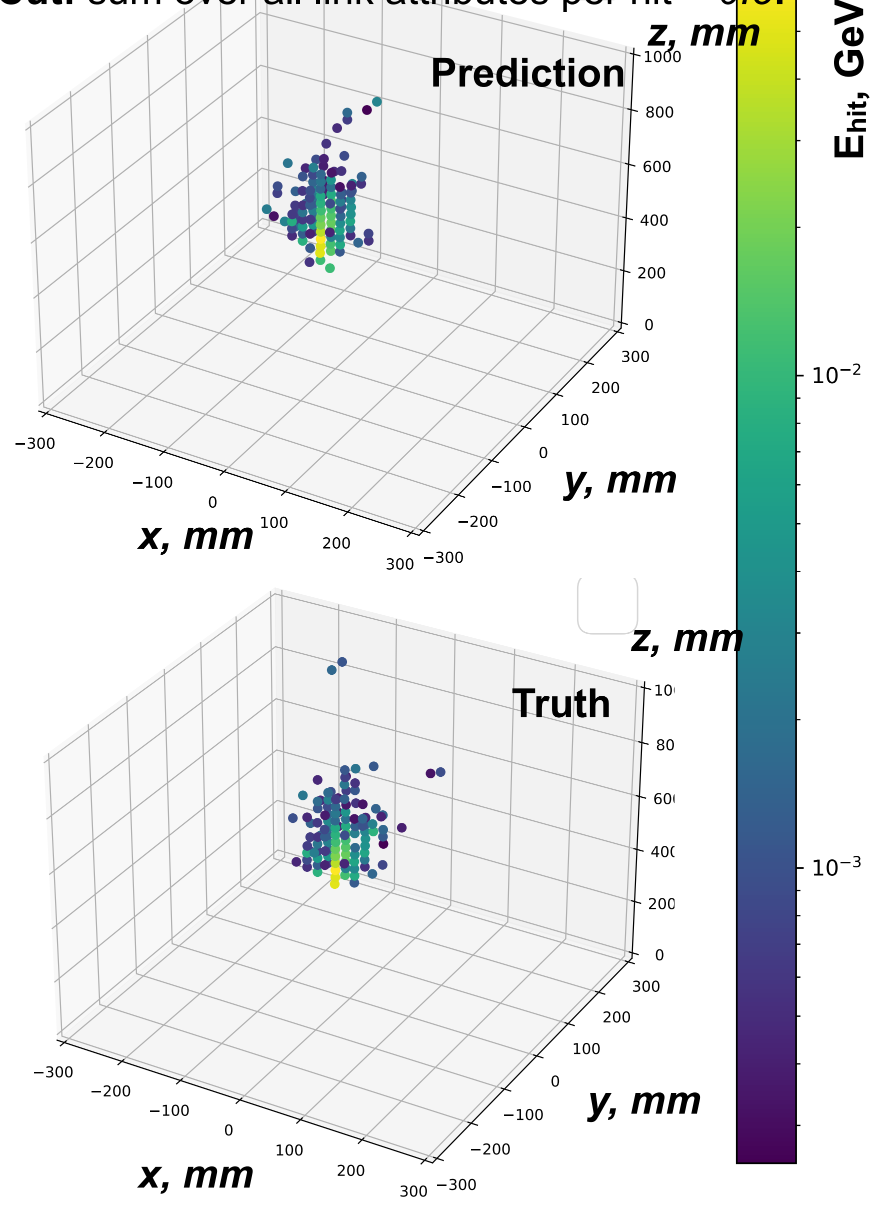


1638 graph edges



Electromagnetic part of the shower

Cut: sum over all link attributes per hit > 0.5:



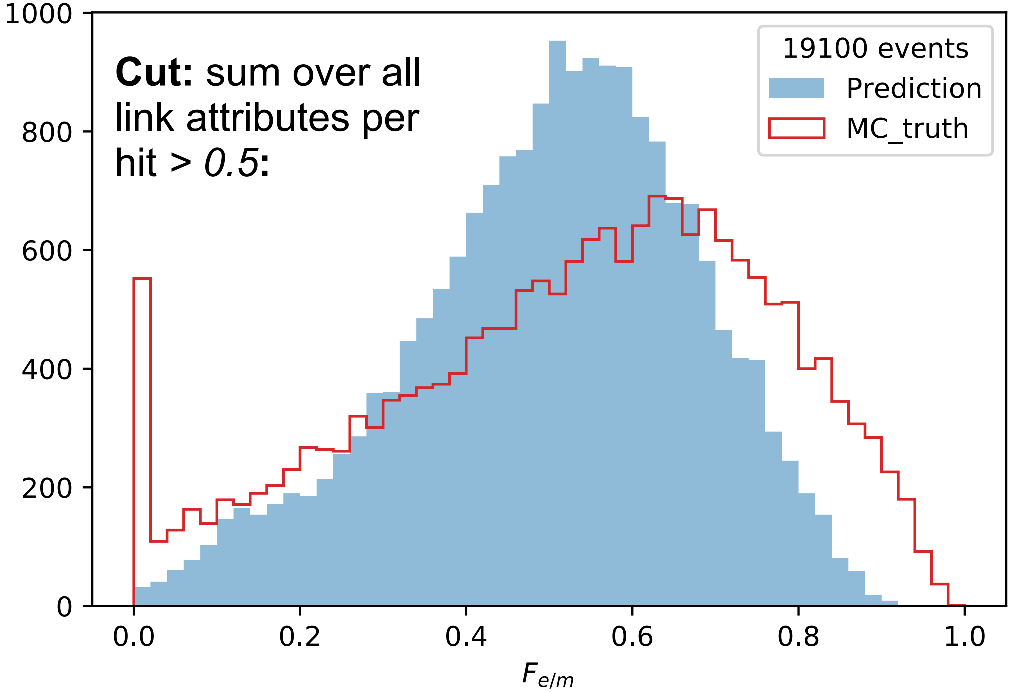
Work in progress…

Electromagnetic fraction of hadronic showers

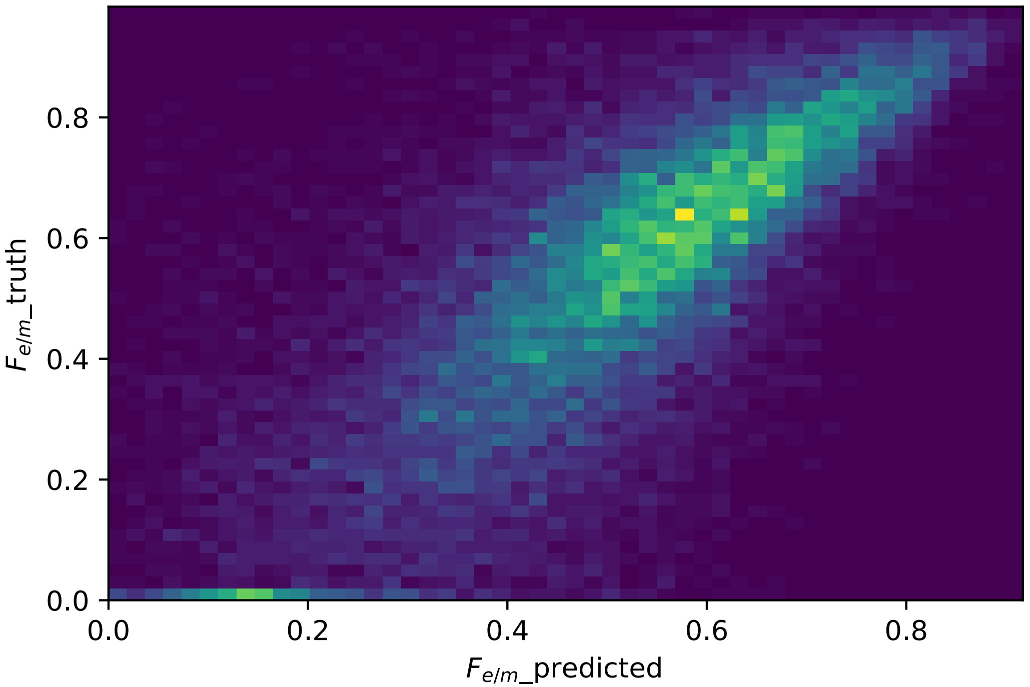
Preliminary results on larger dataset

Work in progress…

Electromagnetic fraction for 10,20,40 GeV pions



Prediction vs truth correlation



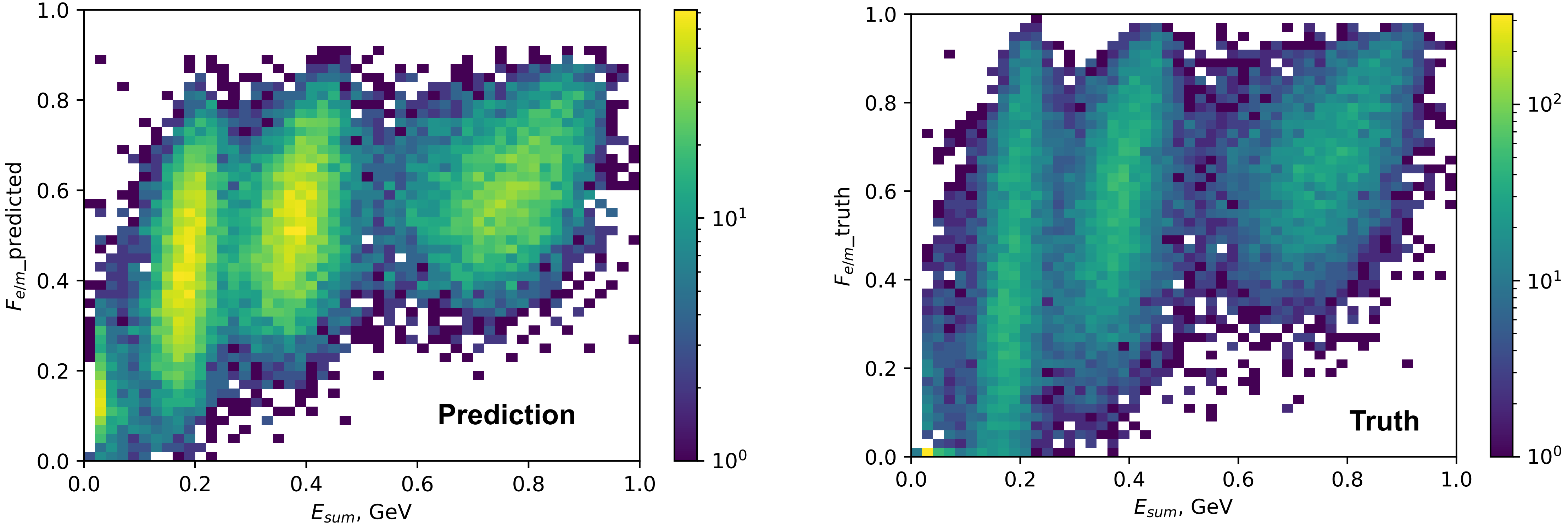
- ☐ Shape of e/m fraction distribution is not described well
  - ☑ Satisfying correlation of prediction vs truth
  - ☐ Slight overestimation of e/m fraction for punch-through events

Electromagnetic fraction of hadronic showers

Preliminary results on larger dataset

W o r k i n p r o g r e s s …

Correlation of Fe/m with energy sum for 10,20,40 GeV MC pions



- ☑ Predicted e/m fraction correlate with energy sum (less pronounced than truth fraction)
  - ☐ No leakage correction
  - ☑ Can be used for global software compensation
- ☐ Resulting electromagnetic structure of hadronic shower can be used for GNN energy reconstruction

Summary & outlook

- ☑ Monte-Carlo simulations provide detailed information about hadronic shower sub-structure
  - ☑ Can be reconstructed exploiting high granularity
- ☑ GNN approach shows promising results to act as tool for reconstructing individual particles within hadronic showers
  - ☑ GNN reconstruction of electromagnetic part within hadronic shower gives physically meaningful output
  - ☐ Benchmark simulations against test beam data
  - ☐ Architecture and model parameters can be optimised for better performance
    - ☐ Include energy contribution of different particles to same hits (in contrast to used “Geiger” mode)
    - ☐ Include timing information
  - ☐ Repeat study for reconstruction of ionising tracks within hadronic showers
  - ☐ Individual particle separation method is under development

Backup

E/m showers

Algorithm to find truth objects

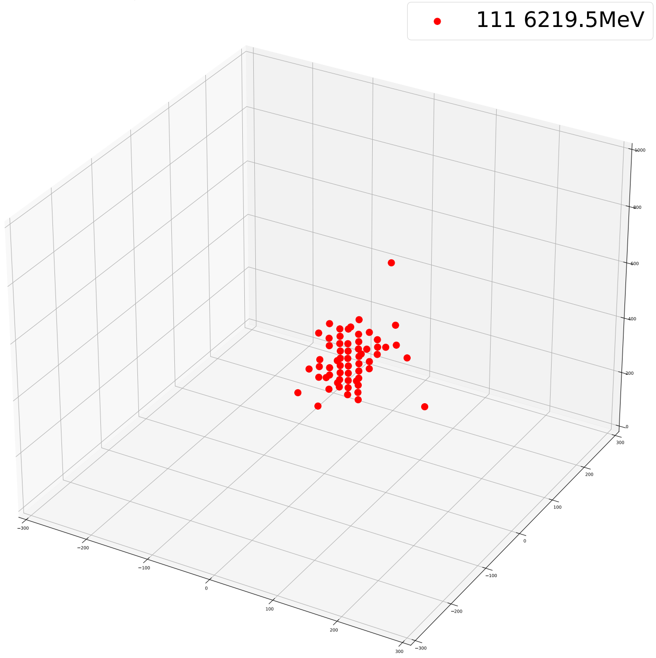
"**Electromagnetic**" particles: $e^{\pm}$,γ, π^0, η

Energy threshold - *0.1GeV* (arbitrary now)

If MC particle is "electromagnetic", all it's "electromagnetic" daughters compose e/m shower and removed from further consideration

Corresponding simulated hits compose sub-shower,

0.5MIP cut: $E_{hit}>0.25MeV$

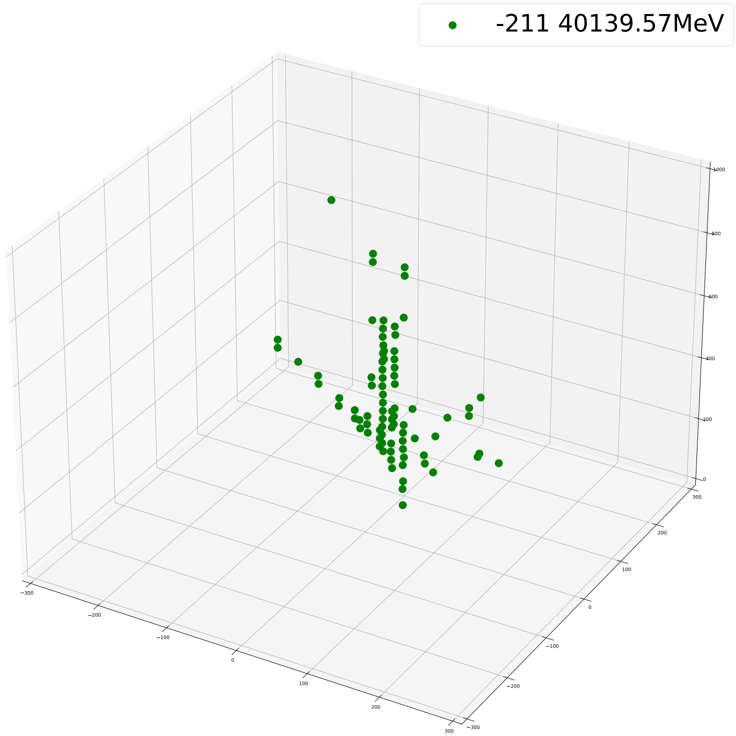


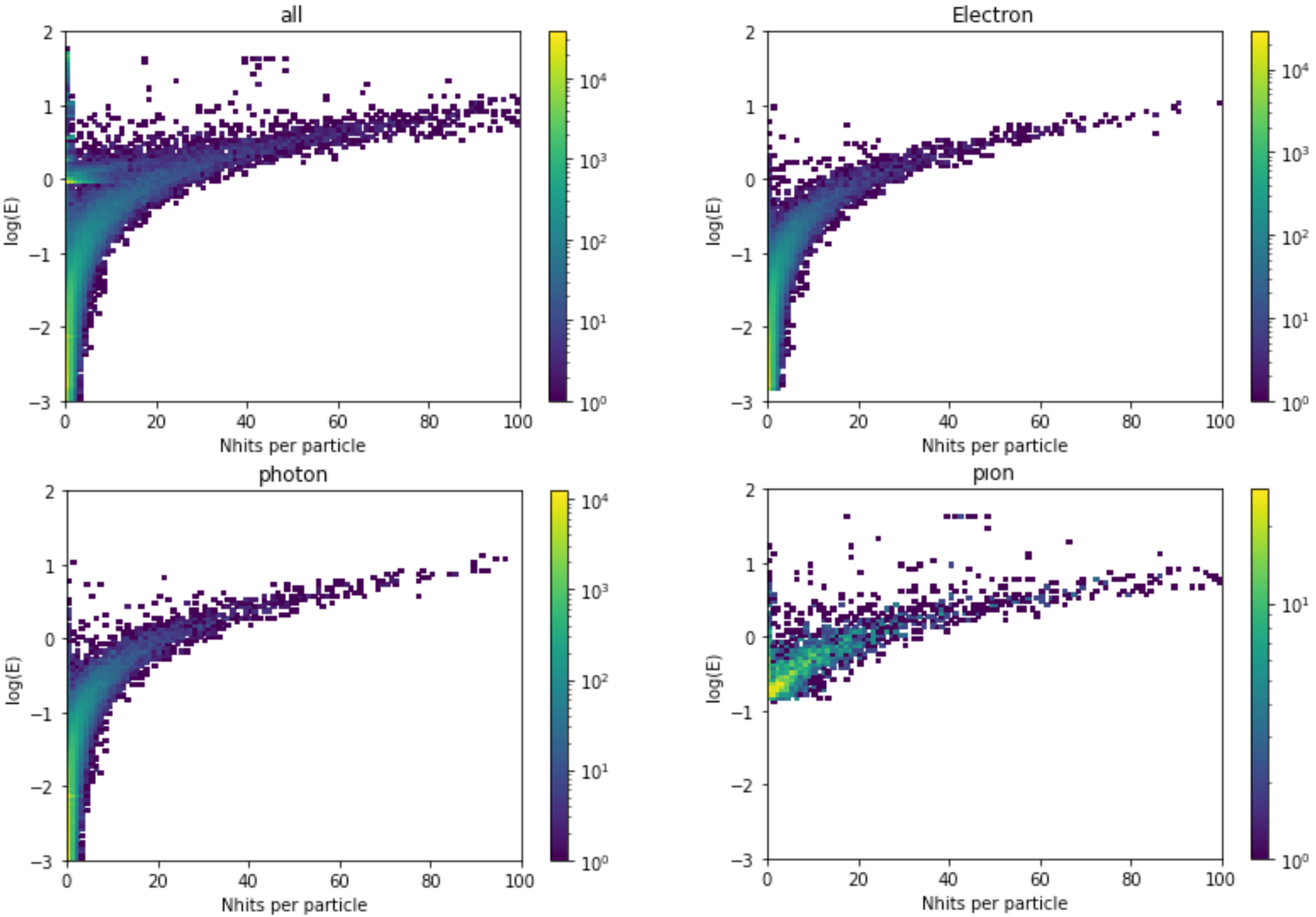
"**Ionising**" particles: $\pi^{\pm}$,$p^{\pm}$, $\mu^{\pm}$

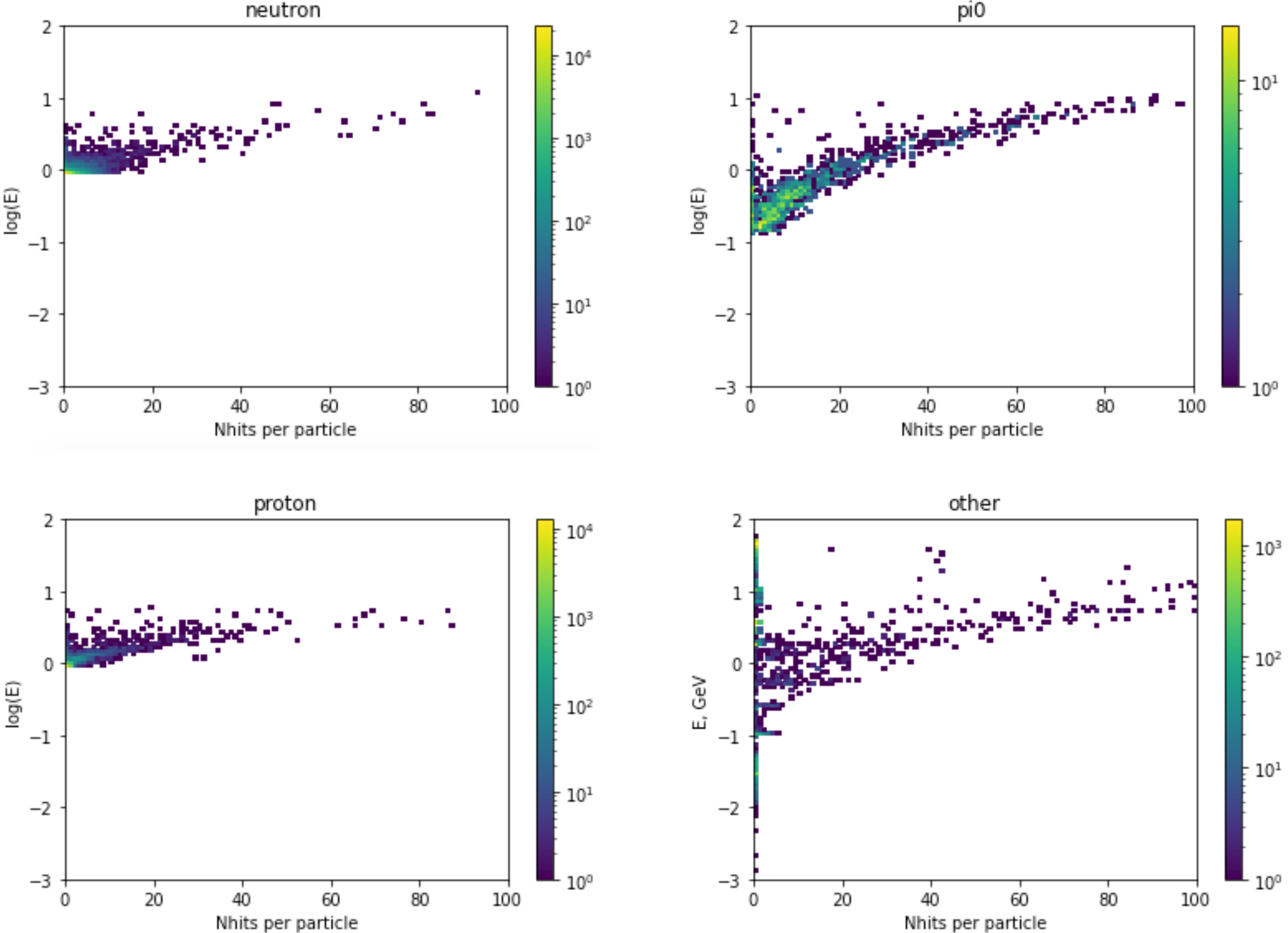
Energy threshold - *0.1GeV* (arbitrary now)

If MC particle is "Ionising", all it's "Ionising" daughters compose ion shower (not individual tracks)

Corresponding hits are collected from SimCalorimeterHit collection ($E_{hit}>0.25MeV$)




all
Electron
photon
pion
log(E)
Nhits per particle


neutron
pi0
proton
other
log(E)
E, GeV
Nhits per particle
2
1
0
−1
−2
−3
0
20
40
60
80
100
10⁰
10¹
10²
10³
10⁴

Example of ~6 GeV neutron within a 40 GeV pion shower

Going down the Monte-Carlo tree

Color coding represent different branches of MC tree corresponding to direct daughters of neutron