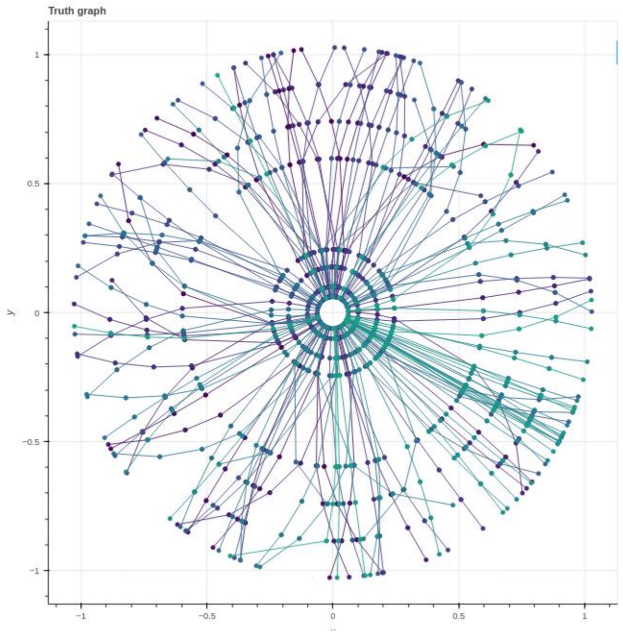
Data

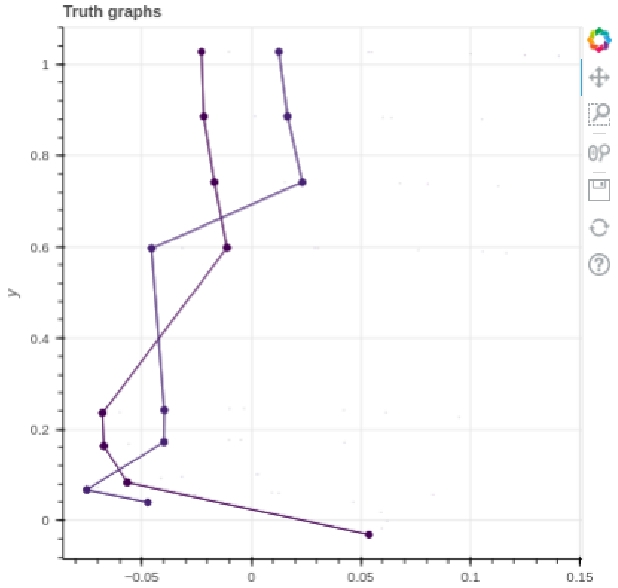
- Use MG5 generate samples through $\mathrm{pp} \rightarrow Q\bar{Q} + \mathrm{j}$
  - Quirk： Collect Quirk and through a simplistic model of the ATLAS detector which consists of 8 layers of trackers.
    - A 500 GeV quirk pair with the string tension (Lambda) = 500 eV (The small Lambda don’t have non-helical tracker)
  - Bkg： jet （~100 particles for one event）

We use pure bkg as training dataset(800 events), pure quirk as validation and testing dataset (100+100 events)

Training dataset:



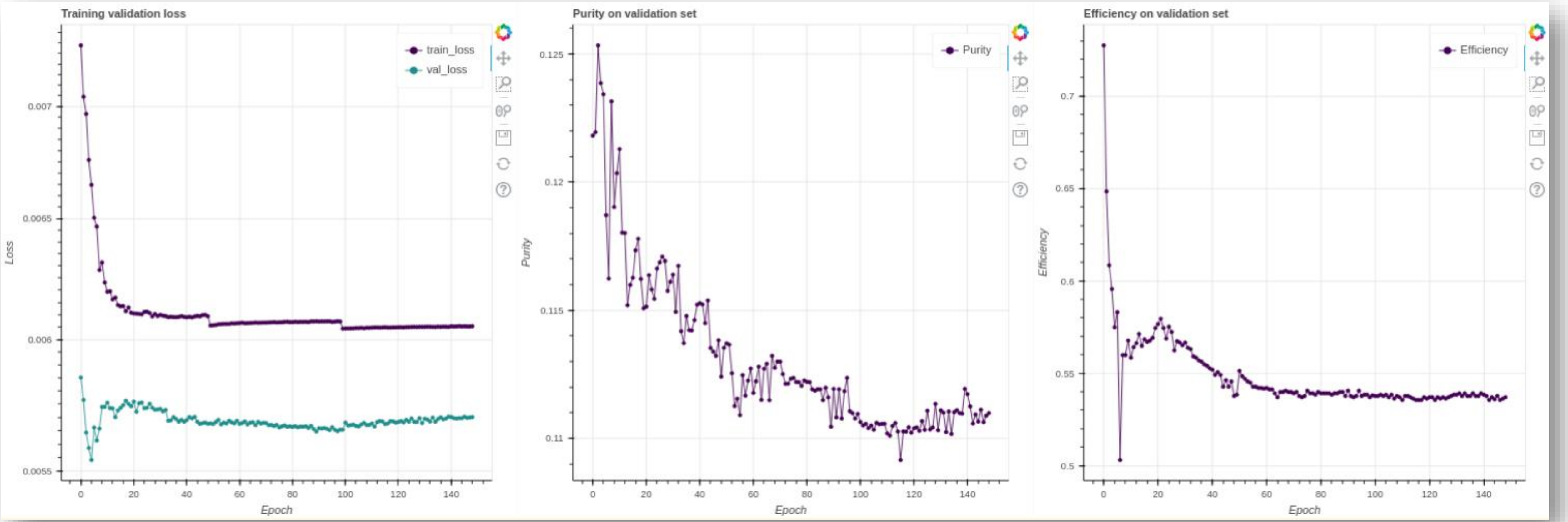
Validation dataset:



Qiyu Sha qsha@cern.cn

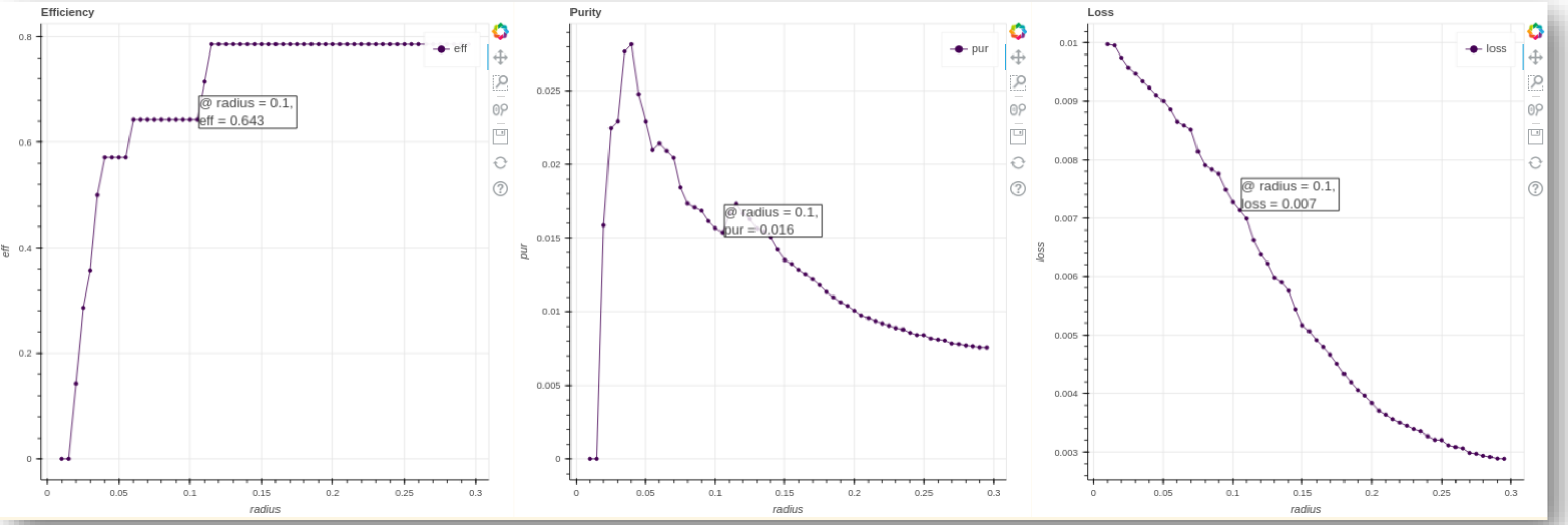
Metric Learning

Use metric learning to reduce the dimension: Embedding the space points on to graphs.



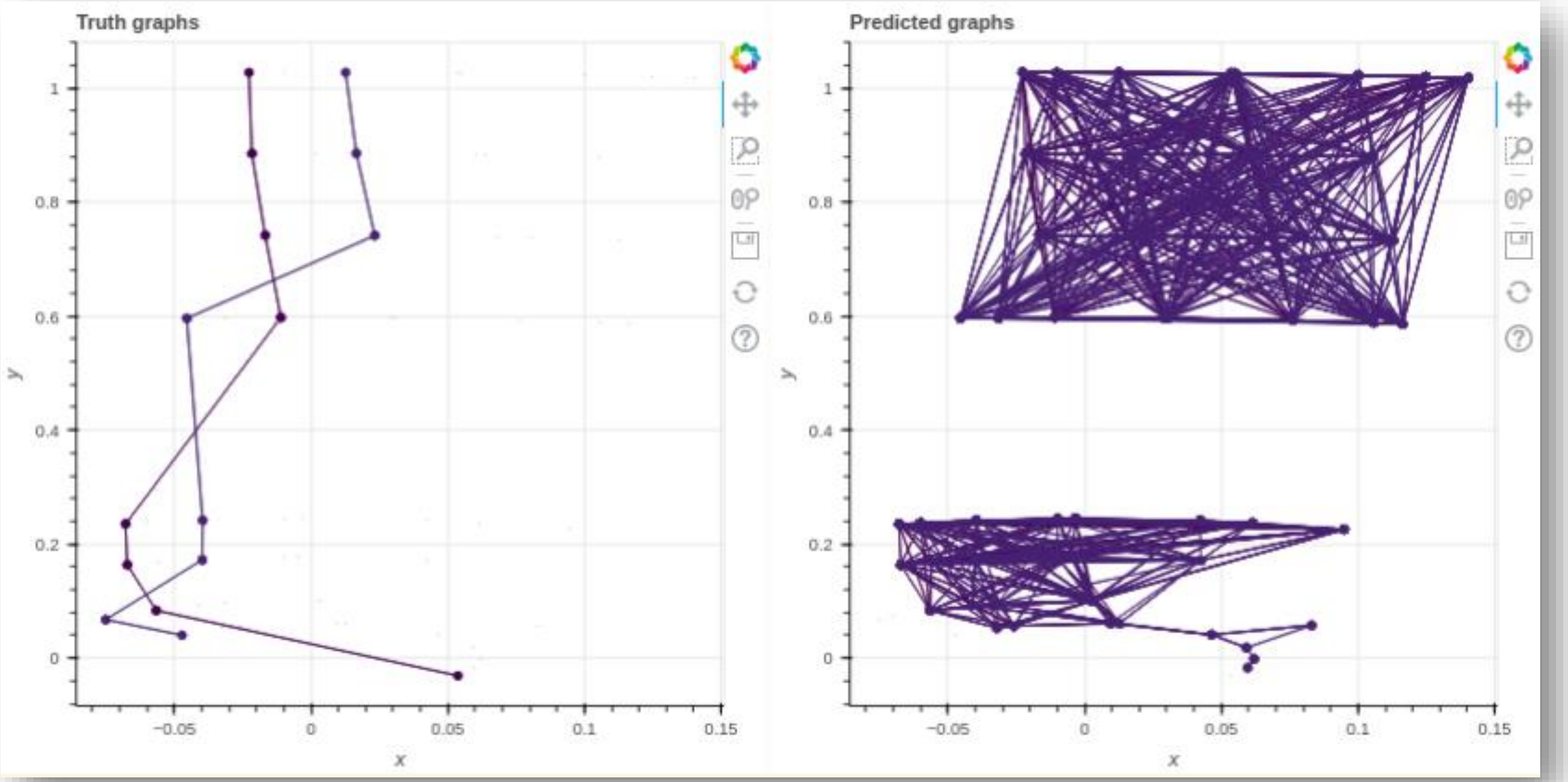
Metric Learning

Evaluate the model performance on one test data sample to see how the efficiency and purity change with the embedding radius.

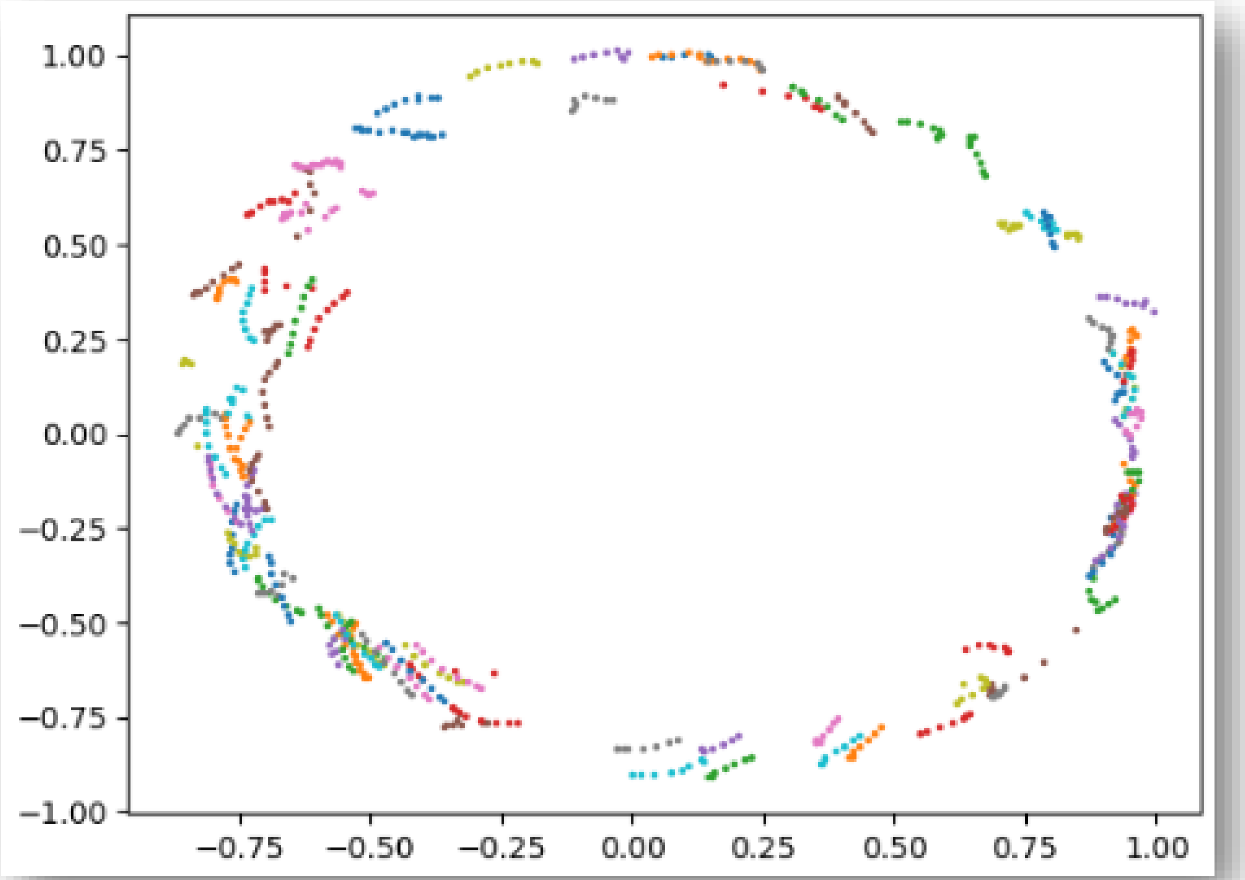


Metric Learning

Example: Truth and predicted track.
Draw few tracks to take a look:

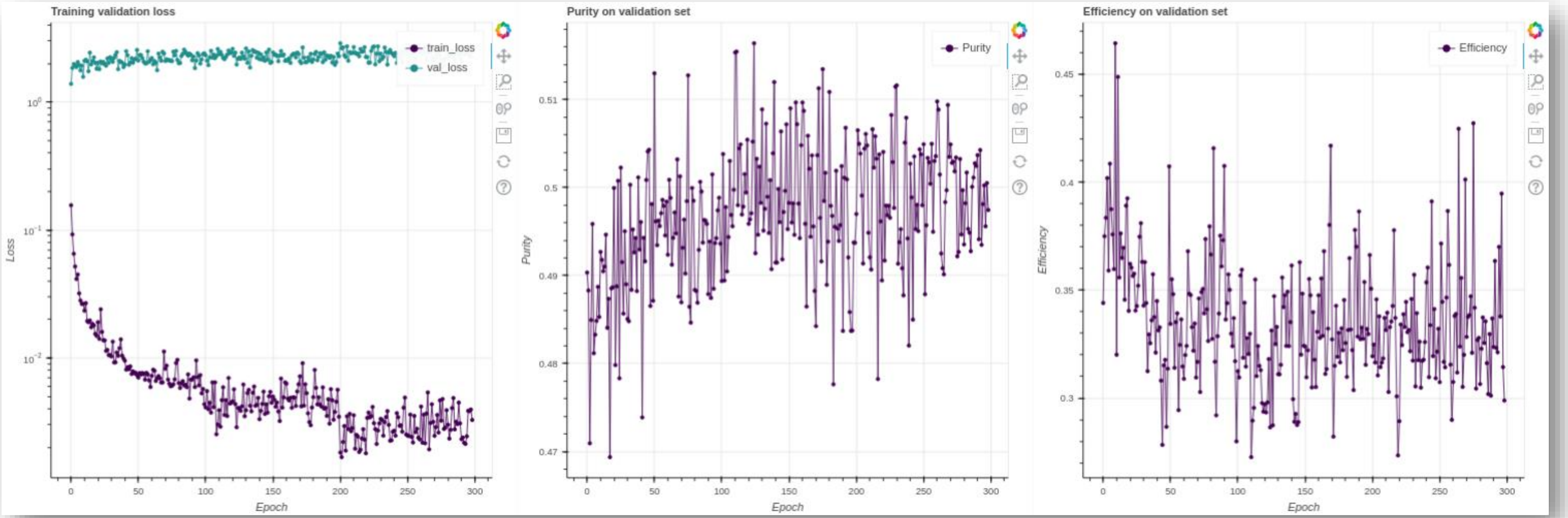


Metric Learning



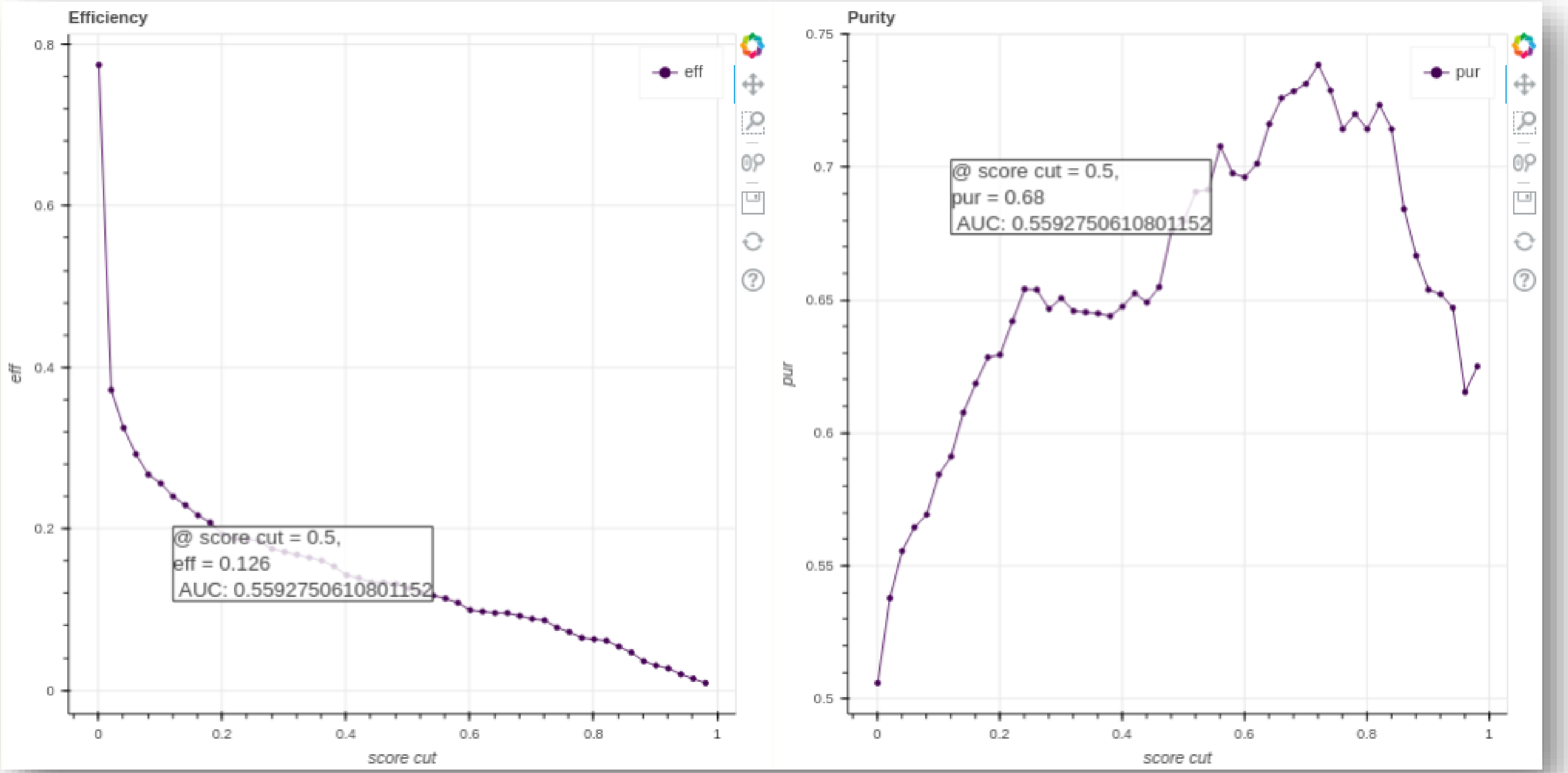
GNN

Train GNN to classify edges as either “true” (belonging to the same track) or “false” (not belonging to the same track)



GNN

Train GNN to classify edges as either “true” (belonging to the same track) or “false” (not belonging to the same track)



GNN

Based on pure quirk sample: ~7%

Based on pure bkg sample: 97%

Some selection for reconstructed particles:

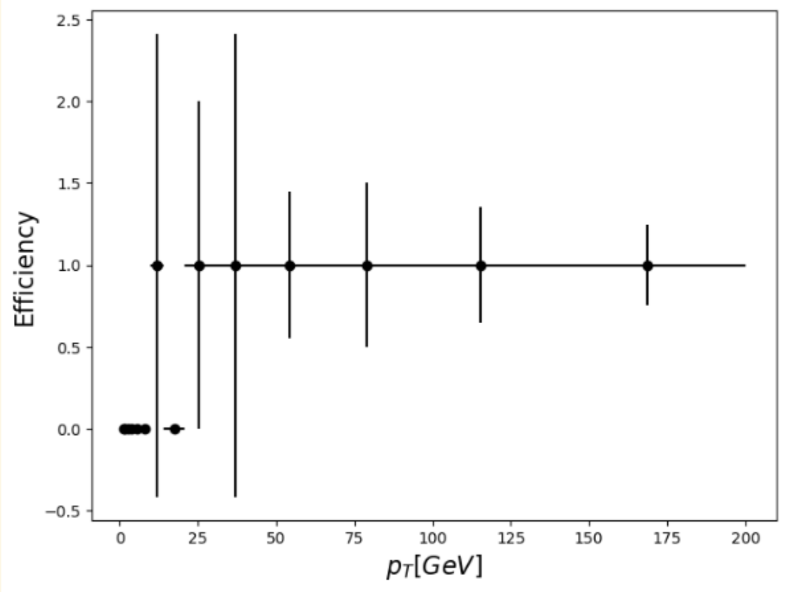
For bkg, we have 8 true hits for each particles, for quirk, we have ≥ 8 true hits.

- min_track_length: 5
- min_particle_length: 7
- ["is_matchable"] = spacepoint_matching.n_reco_hits >= min_track_length
- ["is_reconstructable"] = spacepoint_matching.n_true_hits >= min_particle_length
- ["is_catchable"] = spacepoint_matching.n_true_hits - spacepoint_matching.n_reco_hits <= 5

GNN

min_track_length: 5
min_particle_length: 7
["is_matchable"] = spacepoint_matching.n_reco_hits >= min_track_length
["is_reconstructable"] = spacepoint_matching.n_true_hits >= min_particle_length
["is_catchable"] = spacepoint_matching.n_true_hits - spacepoint_matching.n_reco_hits <= 5

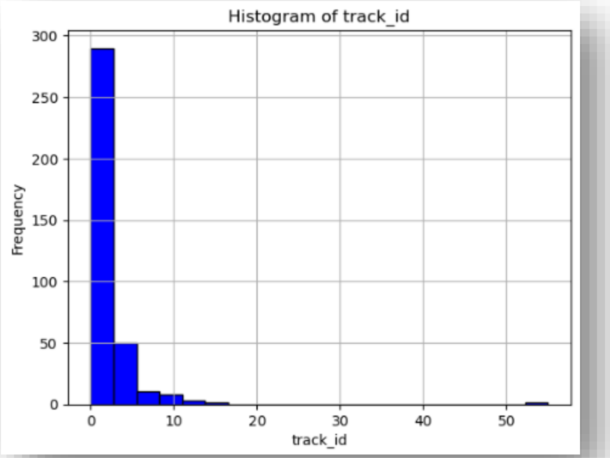
```
--------------------- b) Calculating the performance metrics ---------------------
Number of reconstructed particles: 30
Number of particles: 400
Number of matched tracks: 180
Number of tracks: 217
Number of duplicate reconstructed particles: 2
Efficiency: 0.075
Fake rate: 0.171
Duplication rate: 0.067
------------------------------ c) Plotting results ------------------------------
```

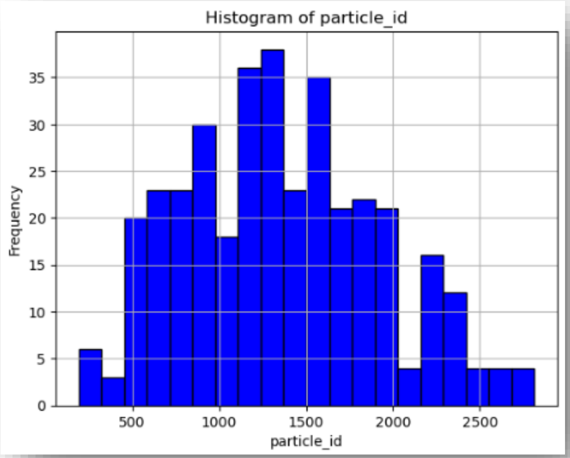


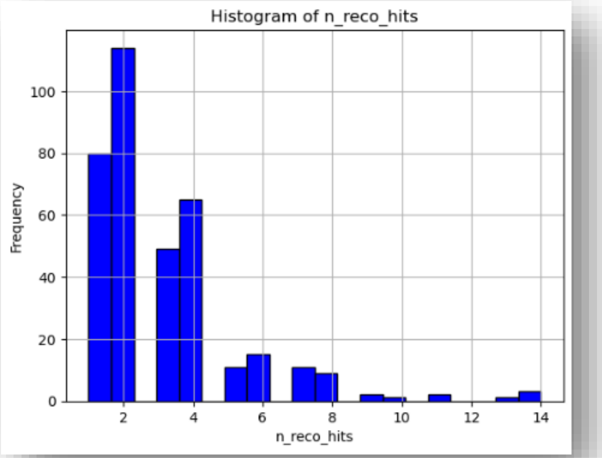
Qiyu Sha qsha@cern.cn

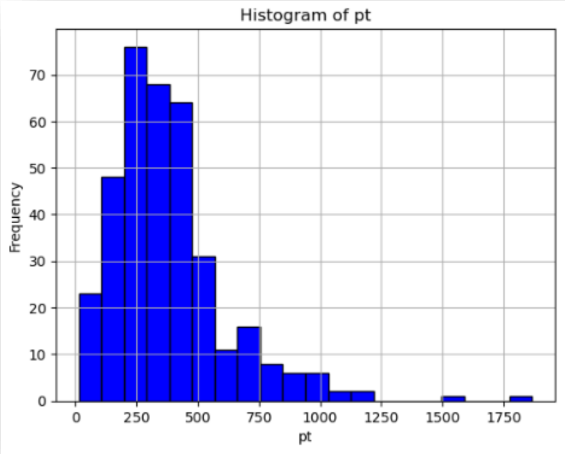
GNN

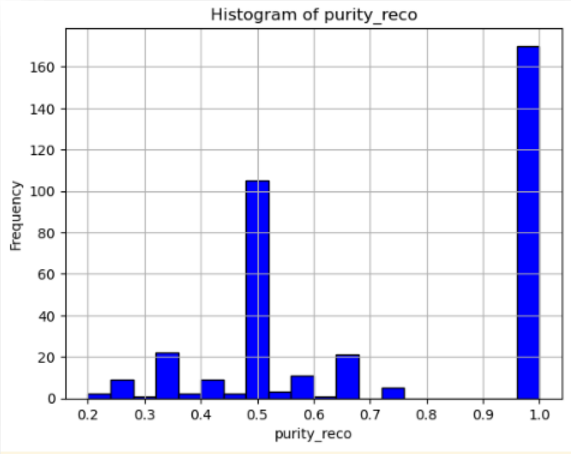
The distribution of whole quirks' track paramters (400 particles)

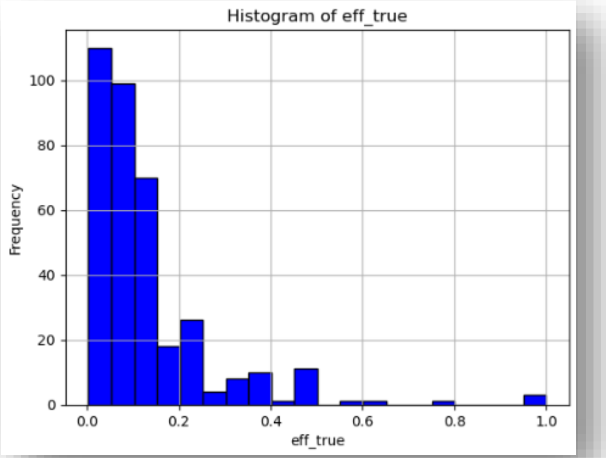






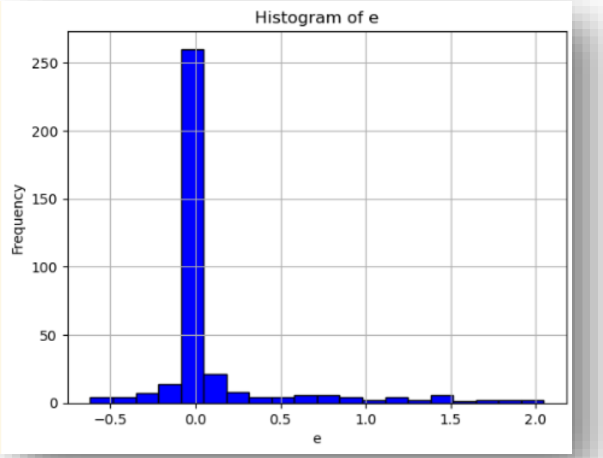


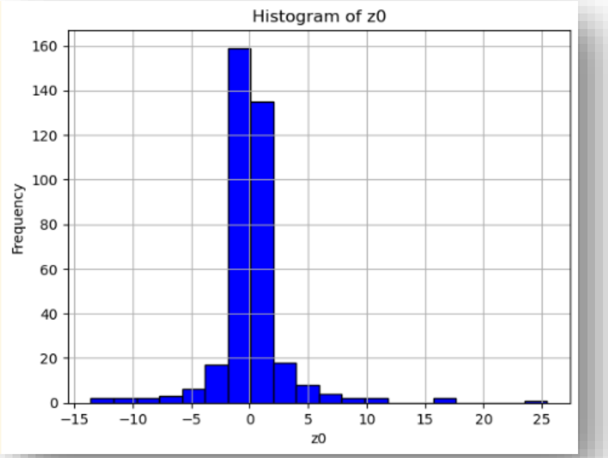


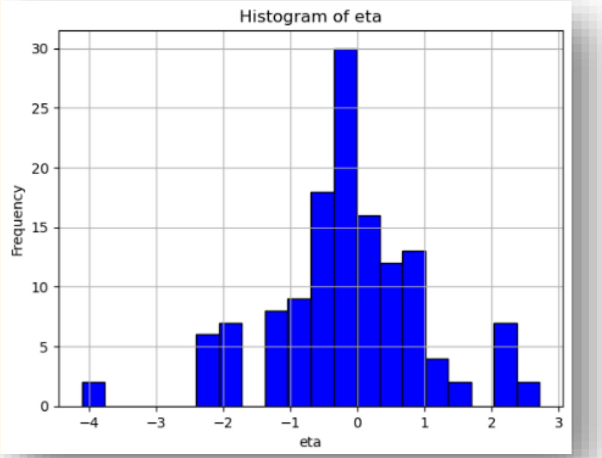


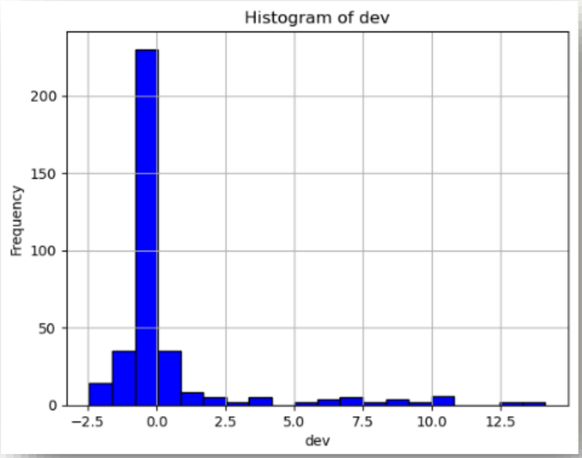
GNN

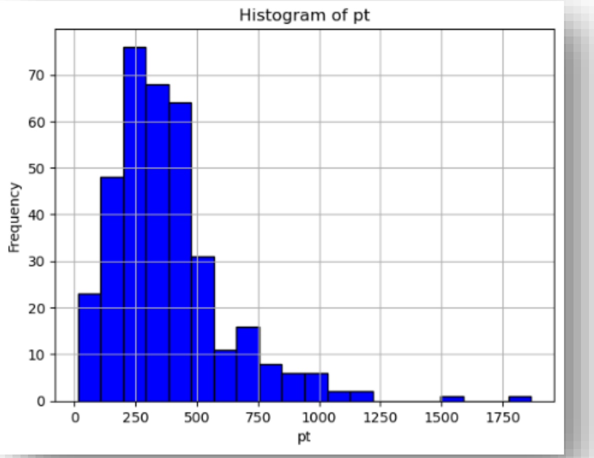
The distribution of whole quirks' track paramters (400 particles)

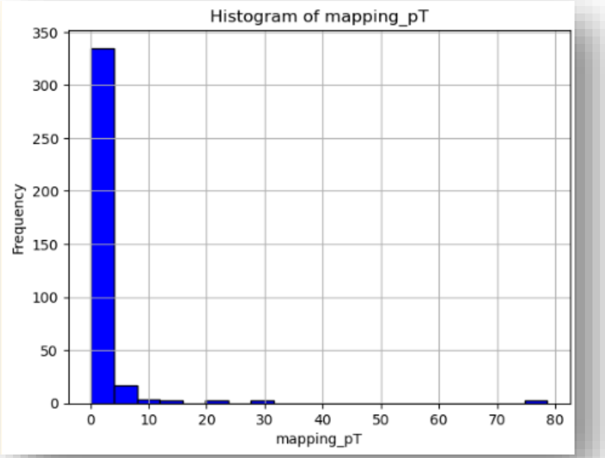


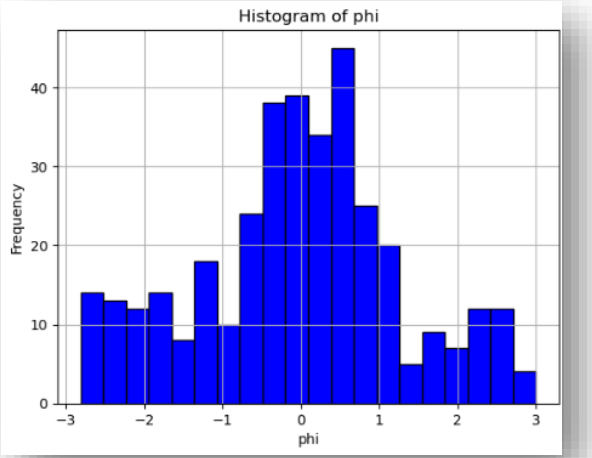


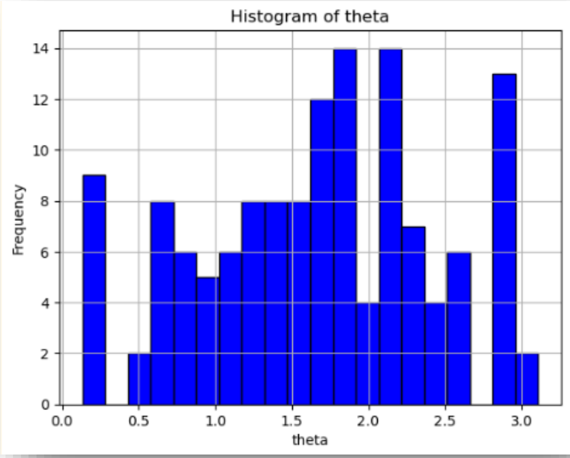






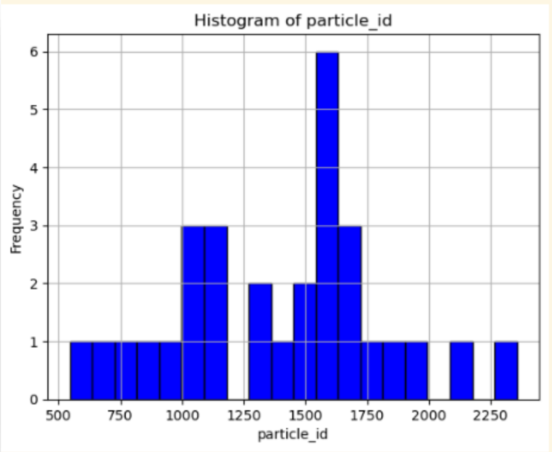


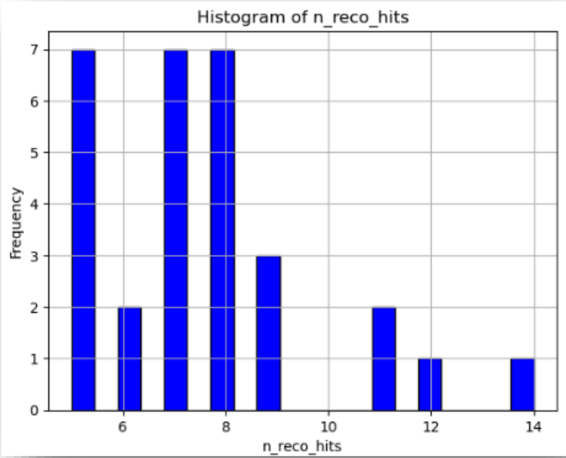


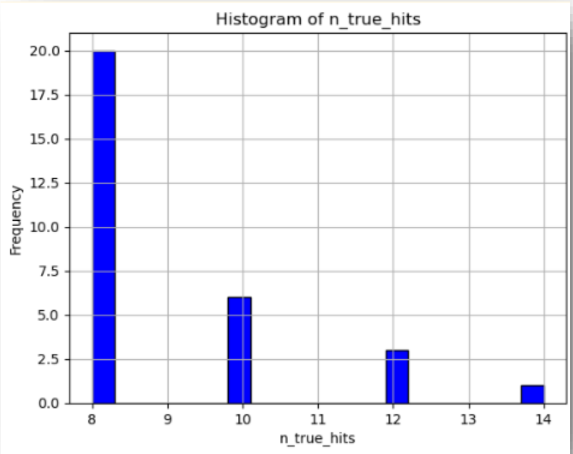


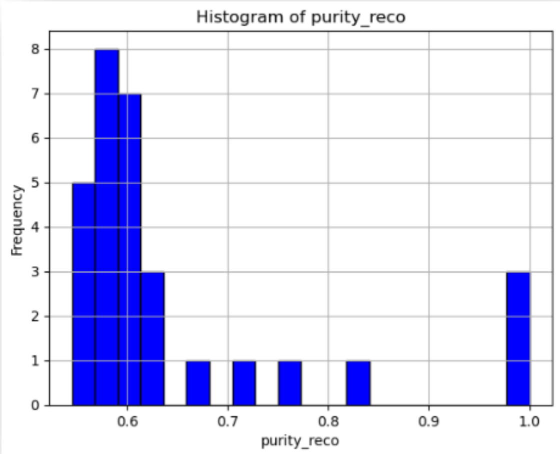
GNN

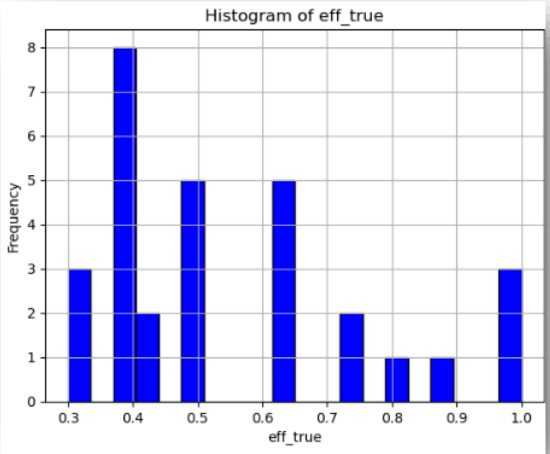
The distribution of reconstructed quirks' track paramters







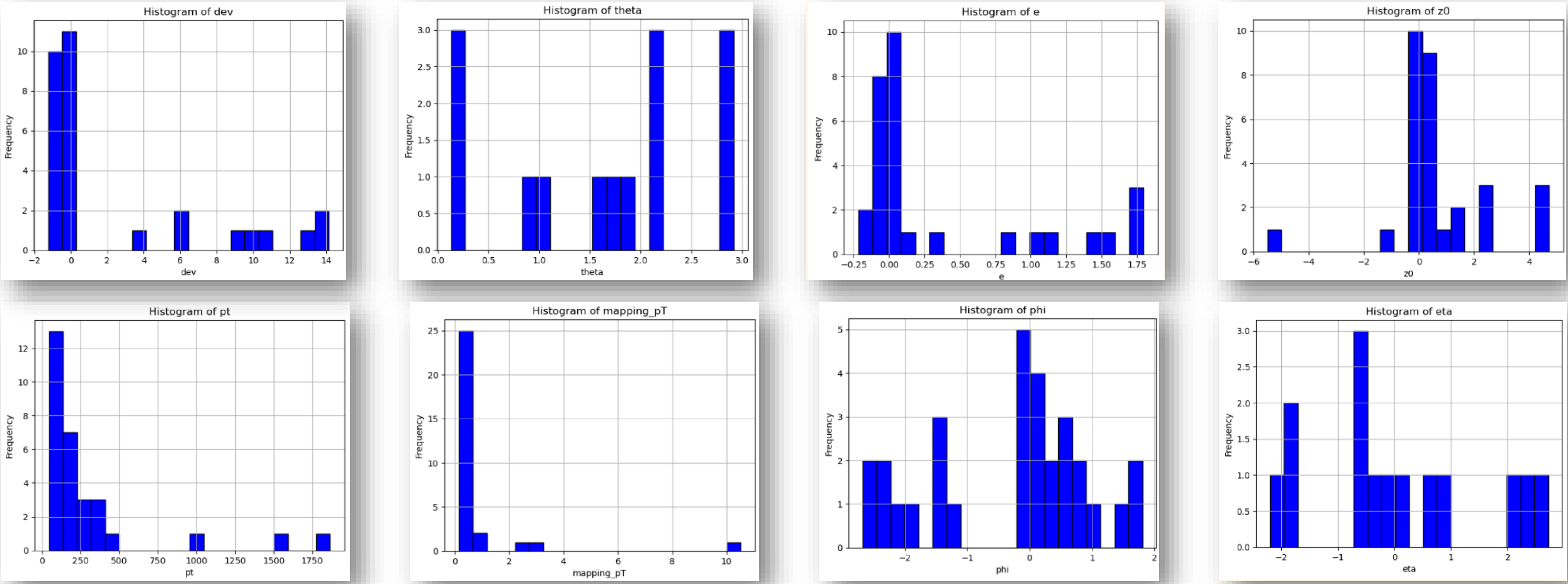




GNN

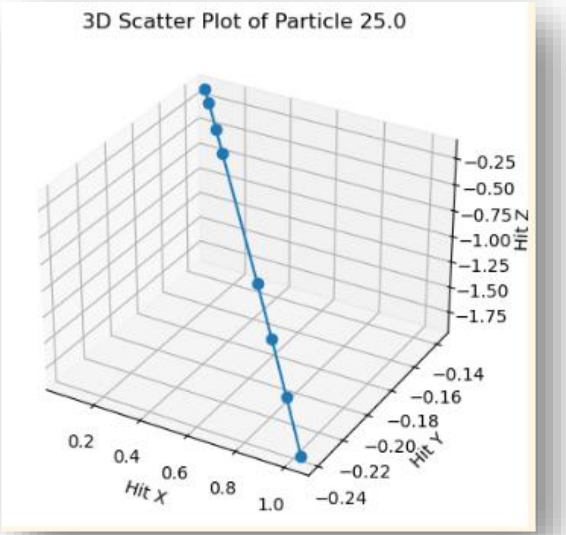
```
pT = 0.3*magnetic_field*R # in MeV
```

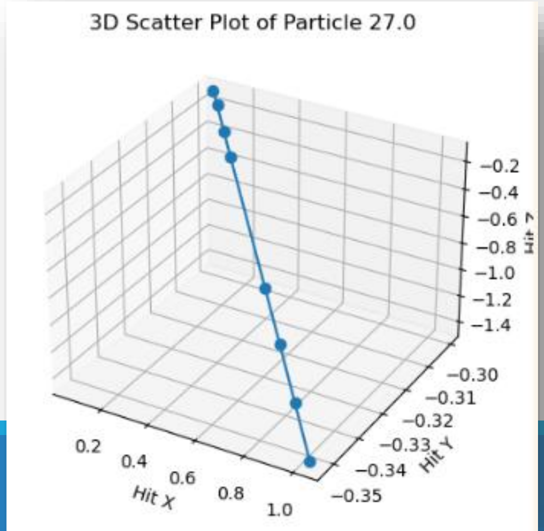
The distribution of reconstructed quirks' track paramters. (mapping_pT means pT comes from conformal_mapping)



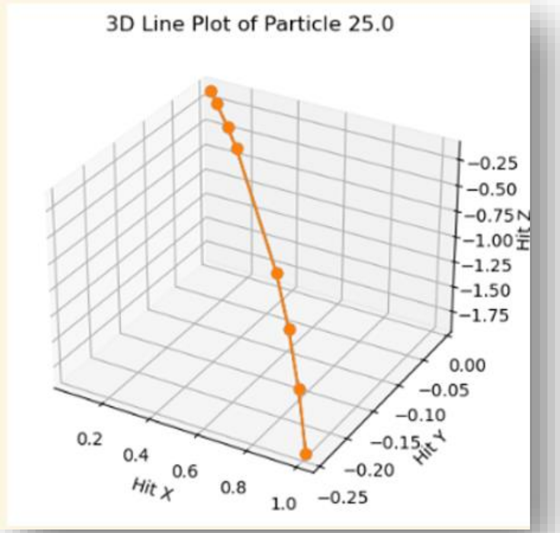
GNN-Bkg

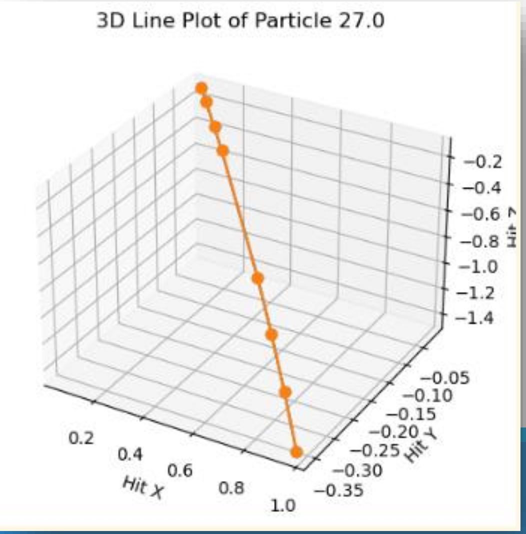
Hits plot for constructed particles(some of them)





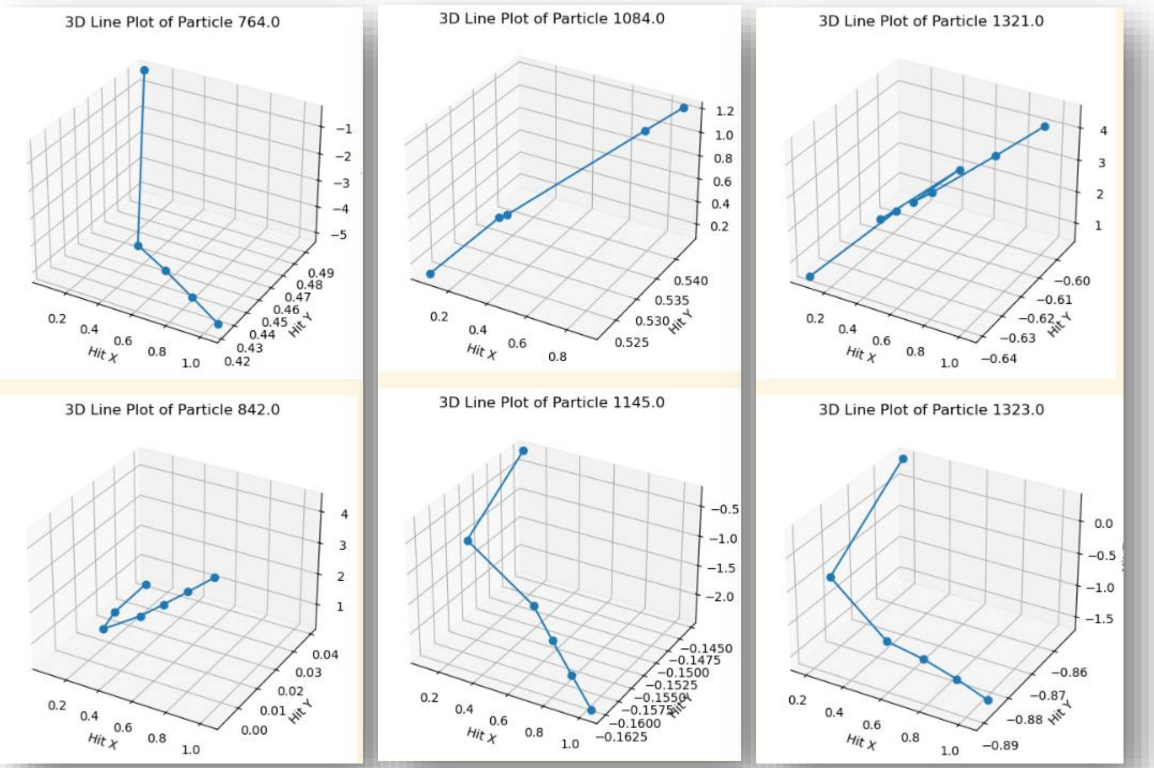
Ture Hits plot for constructed particles:



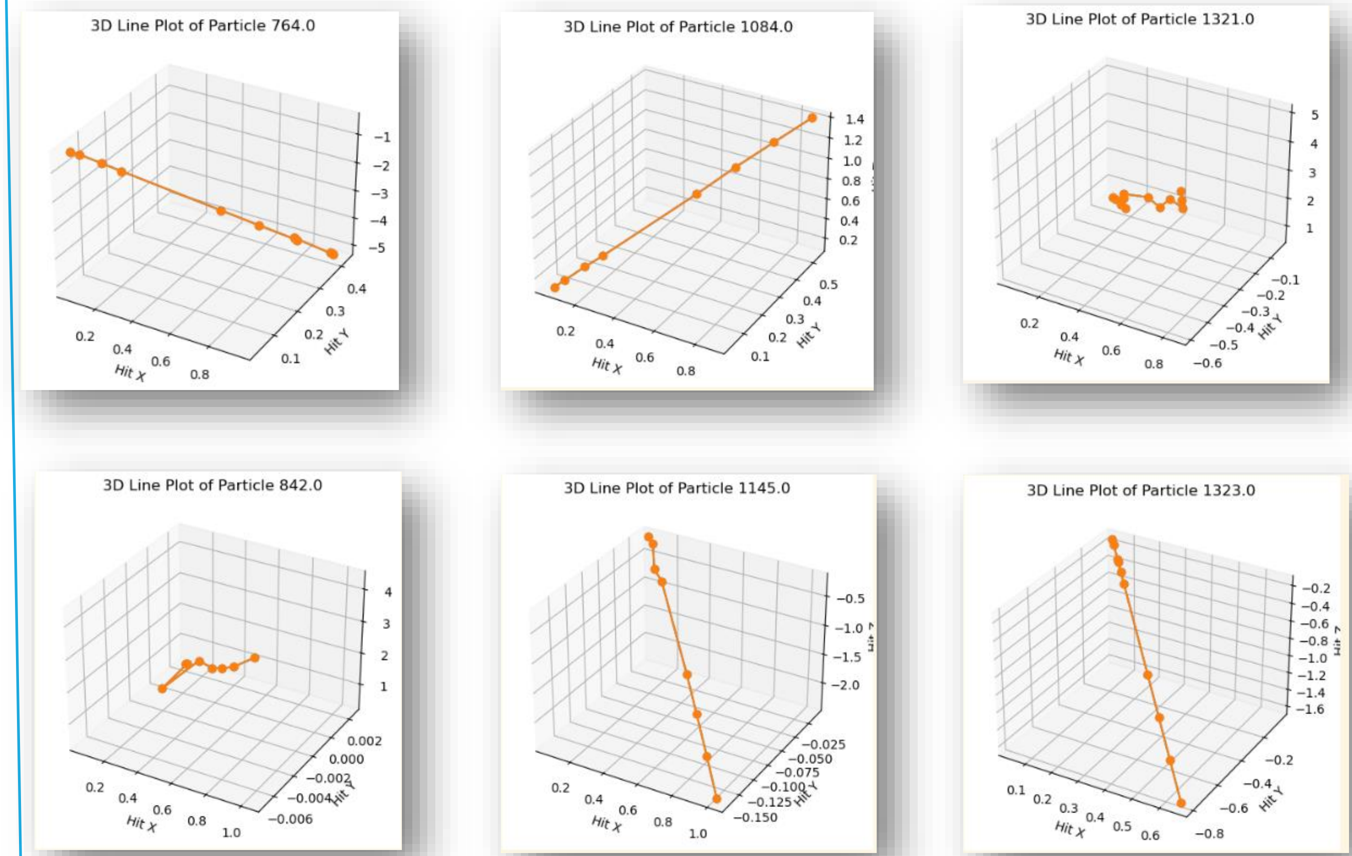


GNN-quirk

Hits plot for constructed particles(some of them)



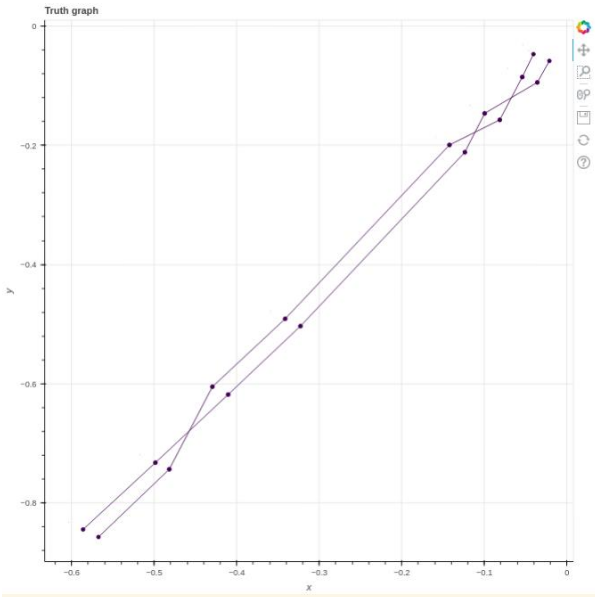
Ture Hits plot for constructed particles:



Data

We use pure quirk as training dataset(600 events), pure quirk as validation and testing dataset (100+100 events)

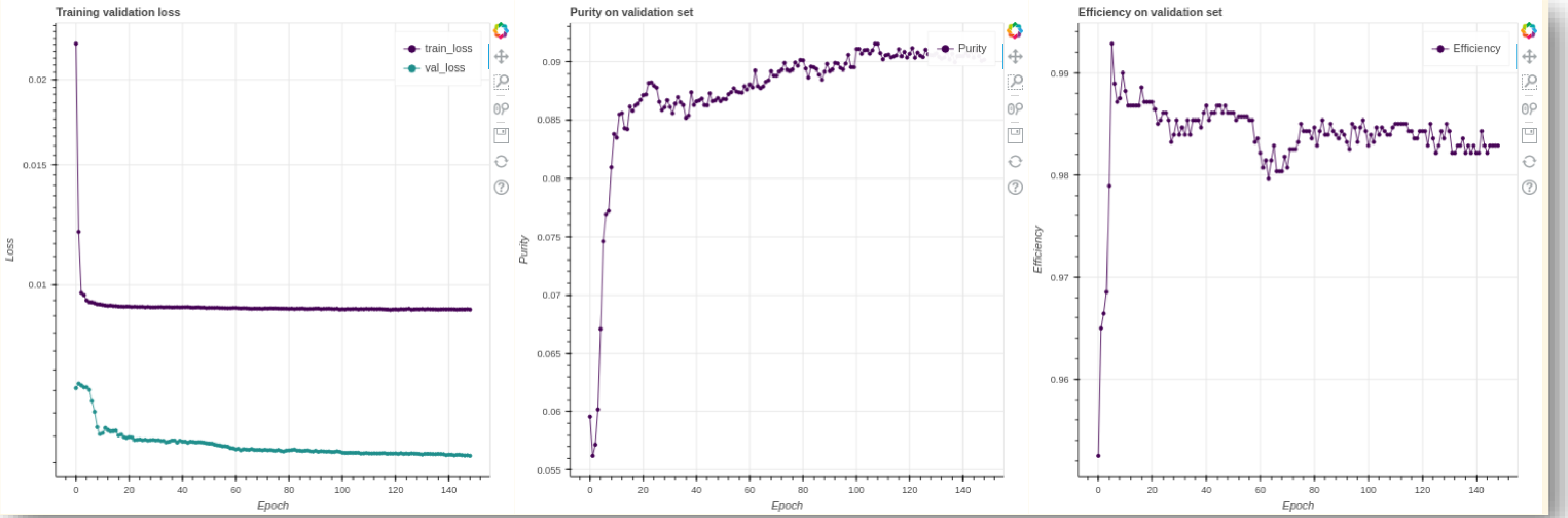
Training dataset:



Validation dataset:

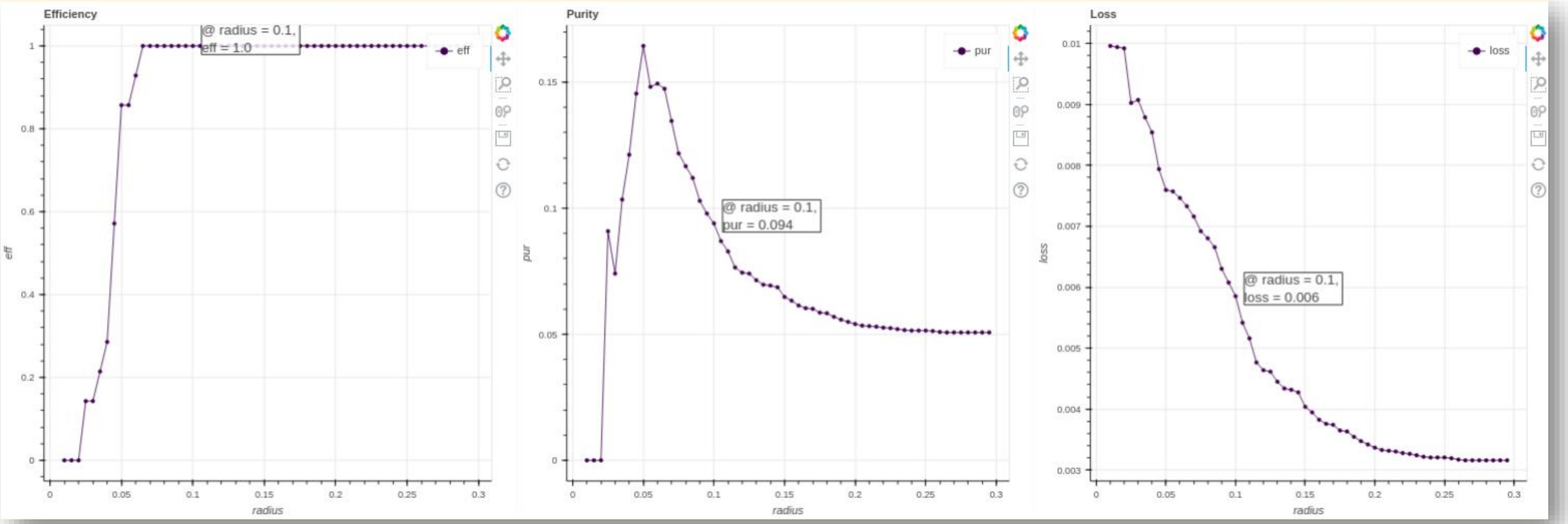
Metric Learning

Use metric learning to reduce the dimension: Embedding the space points on to graphs.



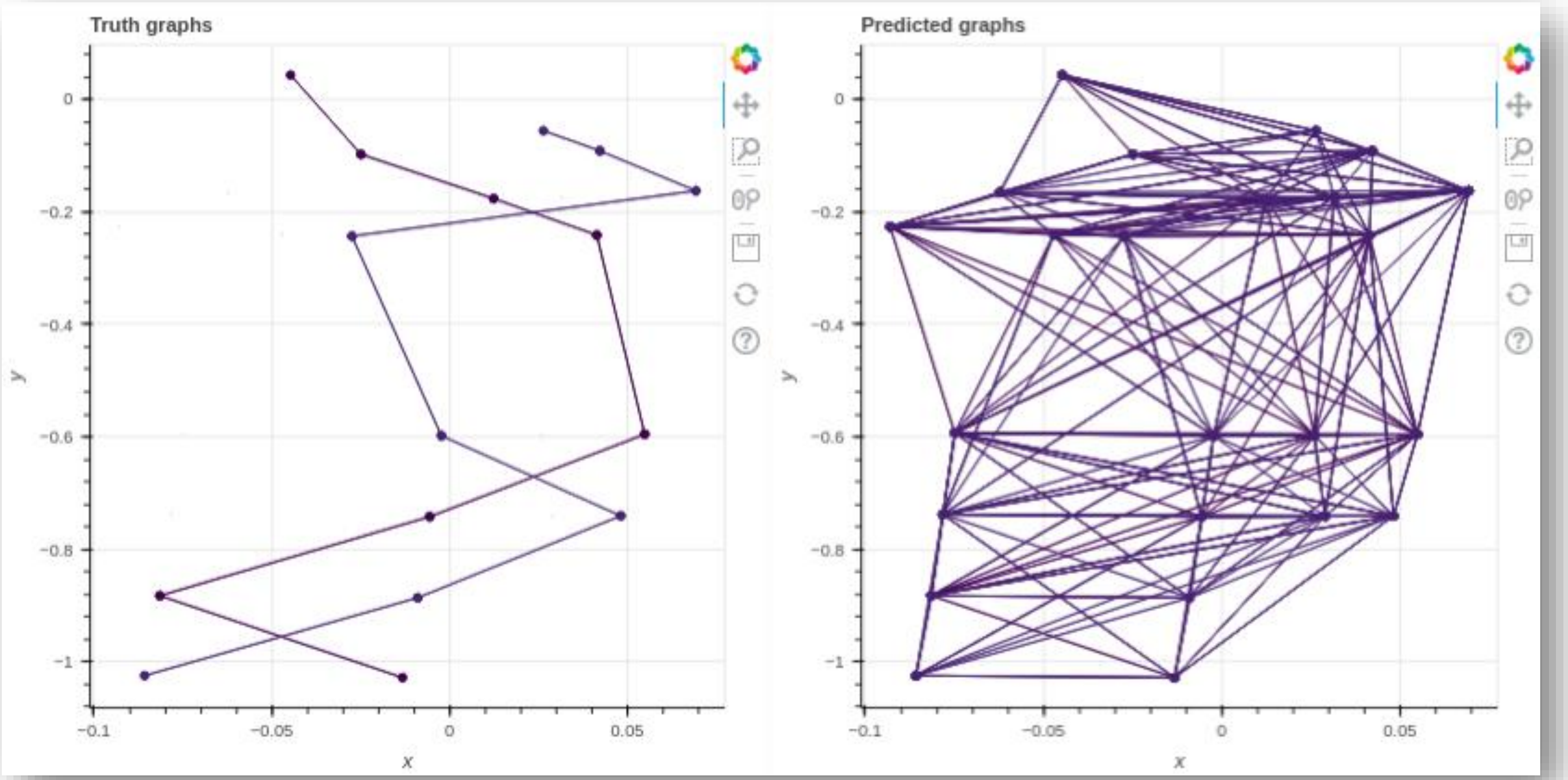
Metric Learning

Evaluate the model performance on one test data sample to see how the efficiency and purity change with the embedding radius.

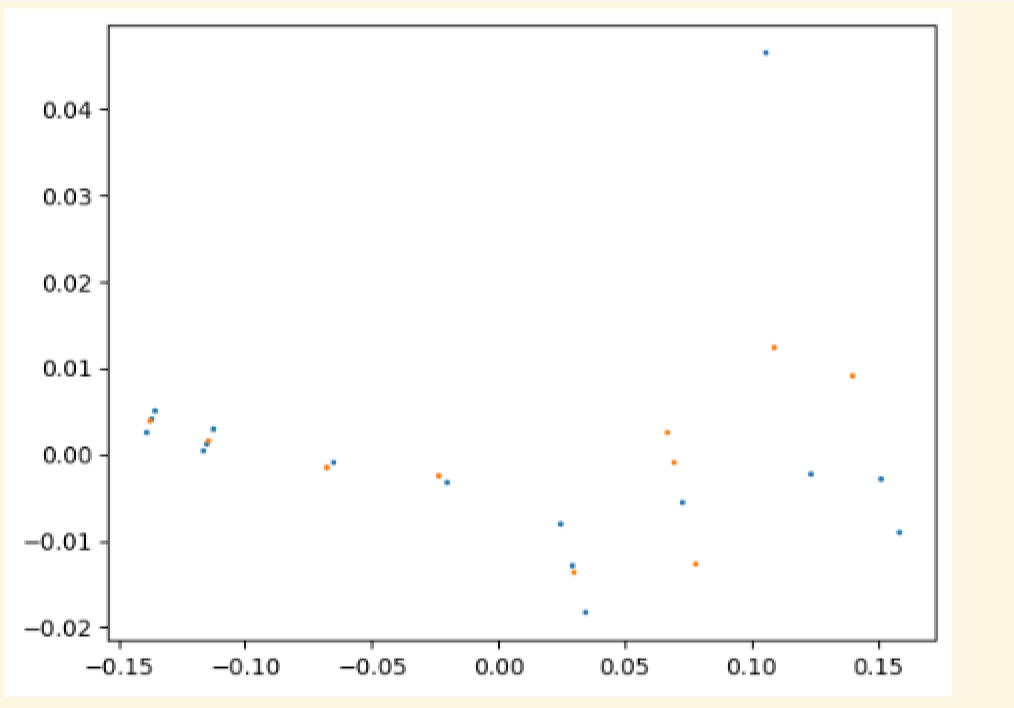


Metric Learning

Example: Truth and predicted track.
Draw few tracks to take a look:

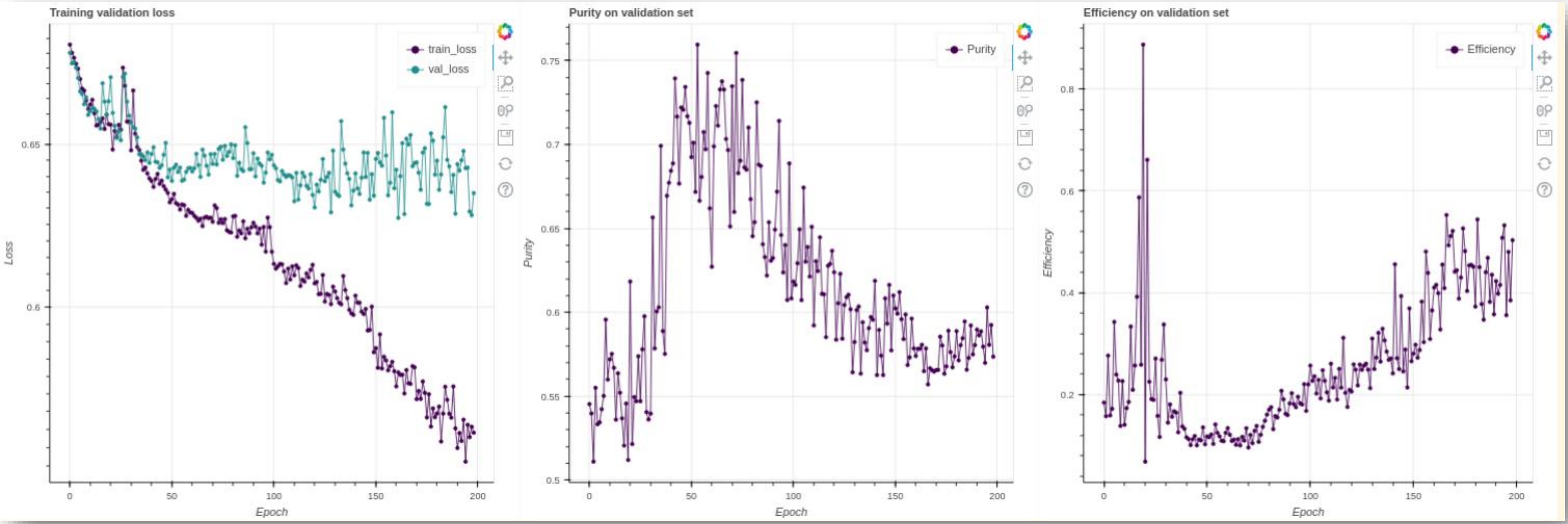


Metric Learning



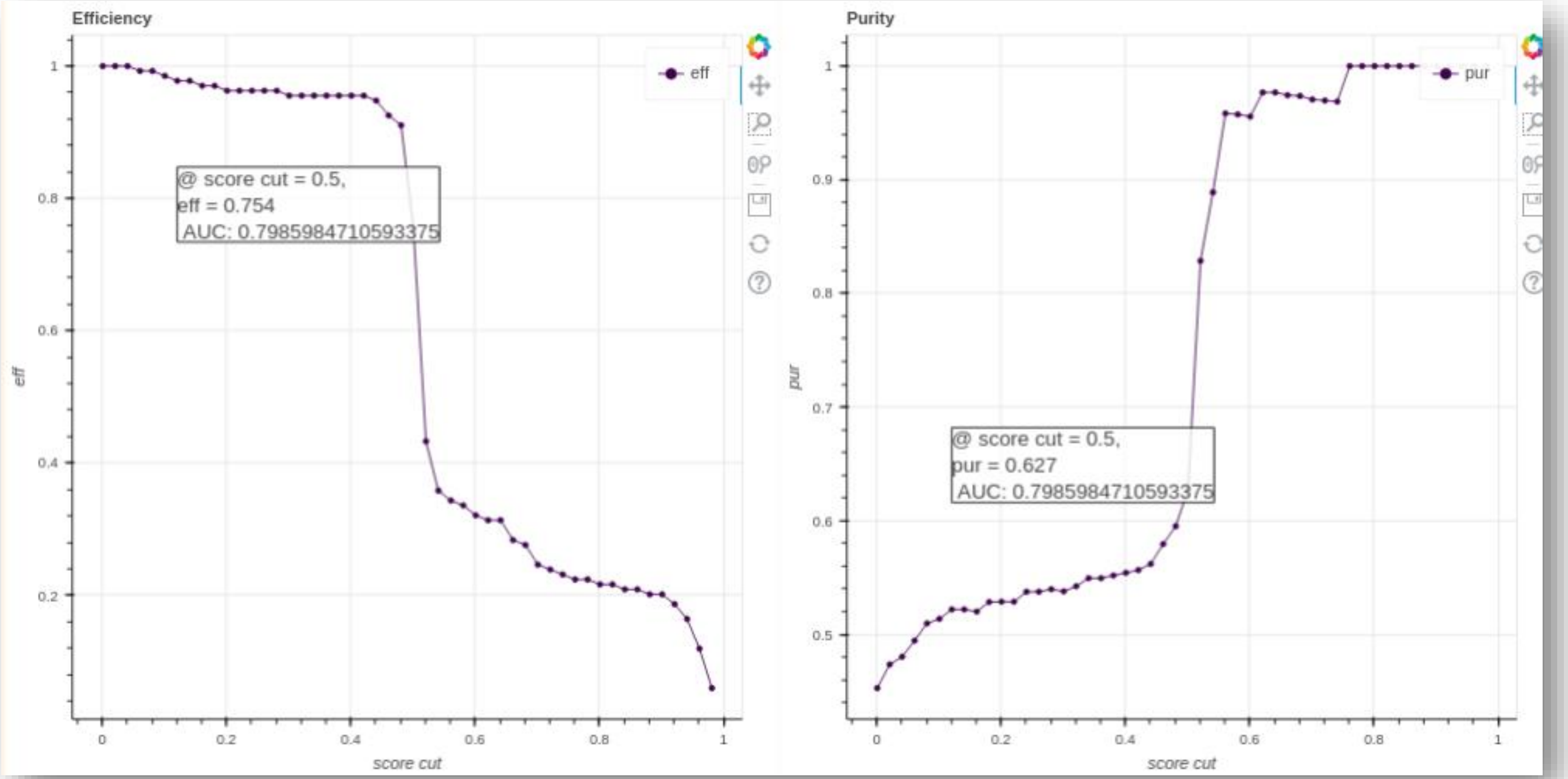
GNN

Train GNN to classify edges as either “true” (belonging to the same track) or “false” (not belonging to the same track)



GNN

Train GNN to classify edges as either “true” (belonging to the same track) or “false” (not belonging to the same track)

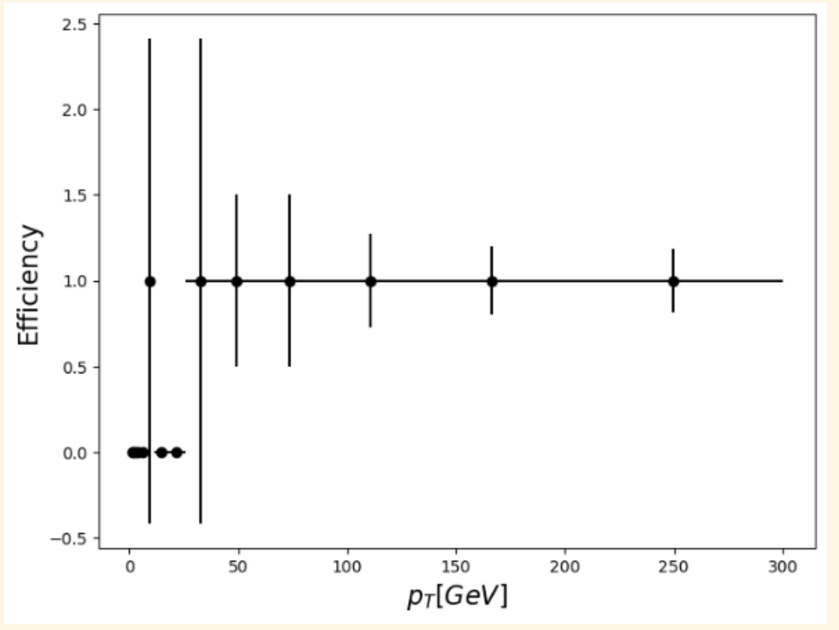


Qiyu Sha qsha@cern.cn

GNN

Based on test/val sample:

```
--------------------- b) Calculating the performance metrics ---------------------
Number of reconstructed particles: 289
Number of particles: 400
Number of matched tracks: 317
Number of tracks: 337
Number of duplicate reconstructed particles: 28
Efficiency: 0.723
Fake rate: 0.059
Duplication rate: 0.097
```



Based on pure training sample:

```
------------ Step 6: Evaluating the track reconstruction performance ------------
-------------------------- a) Loading labelled graphs --------------------------

100%|██████████| 600/600 [00:13<00:00, 44.23it/s]

-------------------- b) Calculating the performance metrics --------------------
Number of reconstructed particles: 932
Number of particles: 1200
Number of matched tracks: 1032
Number of tracks: 1047
Number of duplicate reconstructed particles: 100
Efficiency: 0.777
Fake rate: 0.014
Duplication rate: 0.107
```