CEPC

Jets, samples and Wednesday working meeting

Kaili Zhang

zhangkl@ihep.ac.cn

NaN issue solved

Now training results can predict.

```
    - confusion_matrix:
[[0.73624665 0.15256558 0.03817973 0.04011828 0.00385521 0.00157713
  0.00193856 0.00315426 0.00226713 0.00139094 0.01870653]
 [0.13643385 0.75958368 0.03271432 0.0379622  0.00164339 0.00387839
  0.00436045 0.00124897 0.0013914  0.00303478 0.01774856]
 [0.0088283  0.02205431 0.71456615 0.04852825 0.03377786 0.03685954
  0.04519433 0.006284   0.00943148 0.02762546 0.04685032]
 [0.01422383 0.01668038 0.03591599 0.72146735 0.03470966 0.04223282
  0.00911334 0.03592696 0.03005977 0.01095575 0.04871415]
 [0.00215073 0.00283852 0.02557944 0.02370165 0.50521305 0.08571241
  0.02929135 0.08798323 0.09359477 0.04942302 0.09451183]
 [0.00129923 0.00375575 0.02286201 0.02024172 0.06878255 0.53297741
  0.11532541 0.0183966  0.04953435 0.08119616 0.08562881]
 [0.00106888 0.00303212 0.02275181 0.01093963 0.03029939 0.12441512
  0.43260075 0.02721274 0.08220538 0.16920979 0.09626438]
 [0.00219195 0.00231191 0.01137417 0.02365347 0.12105912 0.03432971
  0.05412273 0.32365675 0.2476581  0.07156021 0.10808188]
 [0.00194135 0.00278114 0.01369848 0.02190012 0.0983215  0.08385957
  0.07341121 0.176477   0.34151316 0.07596331 0.11013317]
 [0.00146183 0.00288002 0.02325835 0.012131   0.07471691 0.10363712
  0.26983833 0.04602579 0.09357886 0.27157288 0.10089892]
 [0.01225046 0.01678854 0.0336829  0.03244845 0.06609608 0.07951045
  0.09167862 0.05502128 0.06615486 0.05685532 0.48951304]]
[2025-04-11 11:10:55,425] INFO: ^[[1mEpoch #0: Current validation metric: 0.52981 (best: 0.52981)^[[0m
[2025-04-11 11:10:55,425] INFO: ──────────────────────────────────────────────────────────────────────
```

Results with ParT under preparing.

By adding protection like, then retraining:

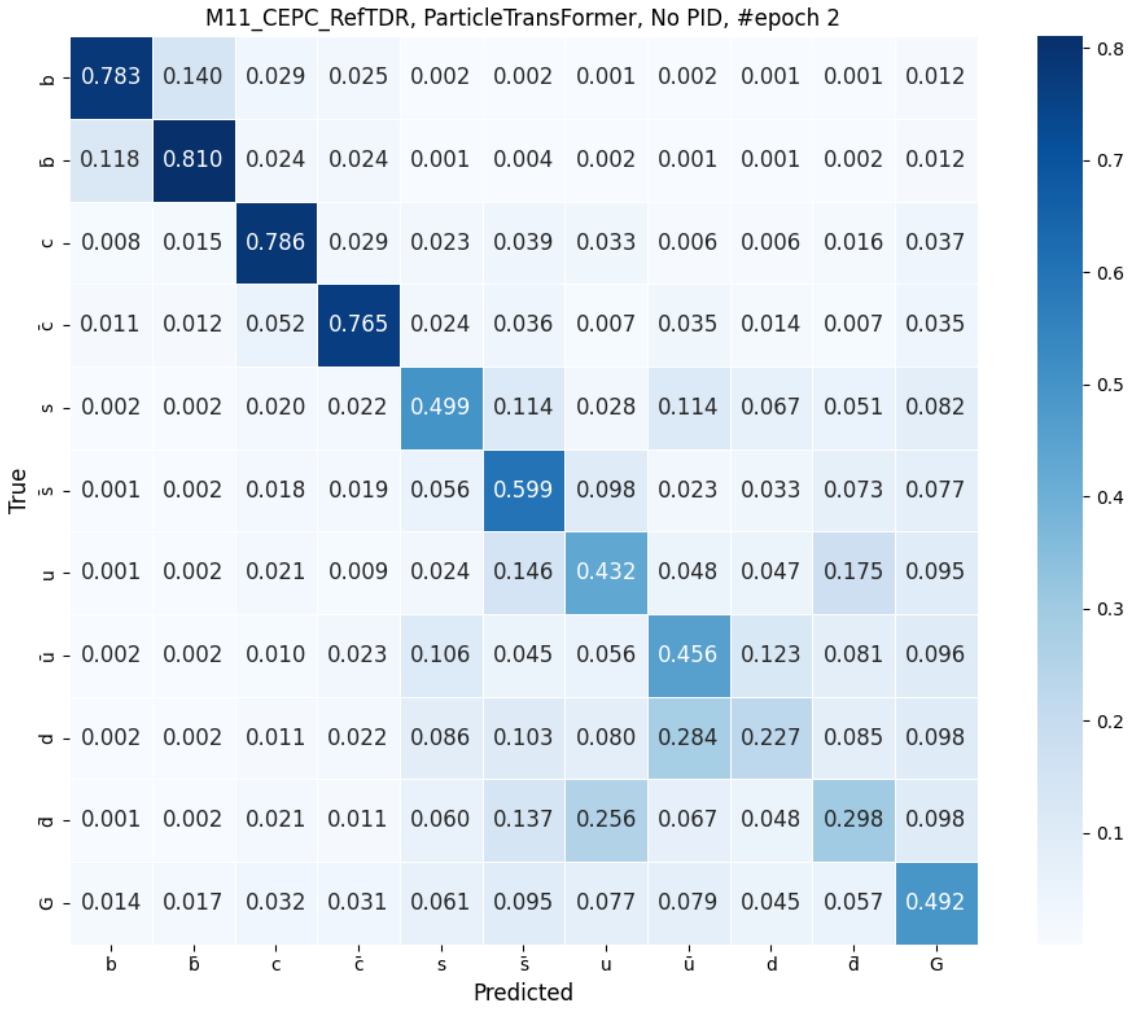
```python
def to_ptrapphim(x, return_mass=True, eps=1e-8):
    # x: (N, 4, ...), dim1 : (px, py, pz, E)
    px, py, pz, energy = x.split((1, 1, 1, 1), dim=1)
    pt = torch.sqrt(to_pt2(x, eps=eps))
    # rapidity = 0.5 * torch.log((energy + pz) / (energy - pz))
    # rapidity = 0.5 * torch.log(1 + (2 * pz) / (energy - pz).clamp(min=1e-20))
    rapidity = 0.5 * torch.log((energy + pz).clamp(min=1e-20) / (energy - pz).clamp(min=1e-20))
    phi = torch.atan2(py, px)
    if not return_mass:
        return torch.cat((pt, rapidity, phi), dim=1)
    else:
        m = torch.sqrt(to_m2(x, eps=eps))
        return torch.cat((pt, rapidity, phi, m), dim=1)

```

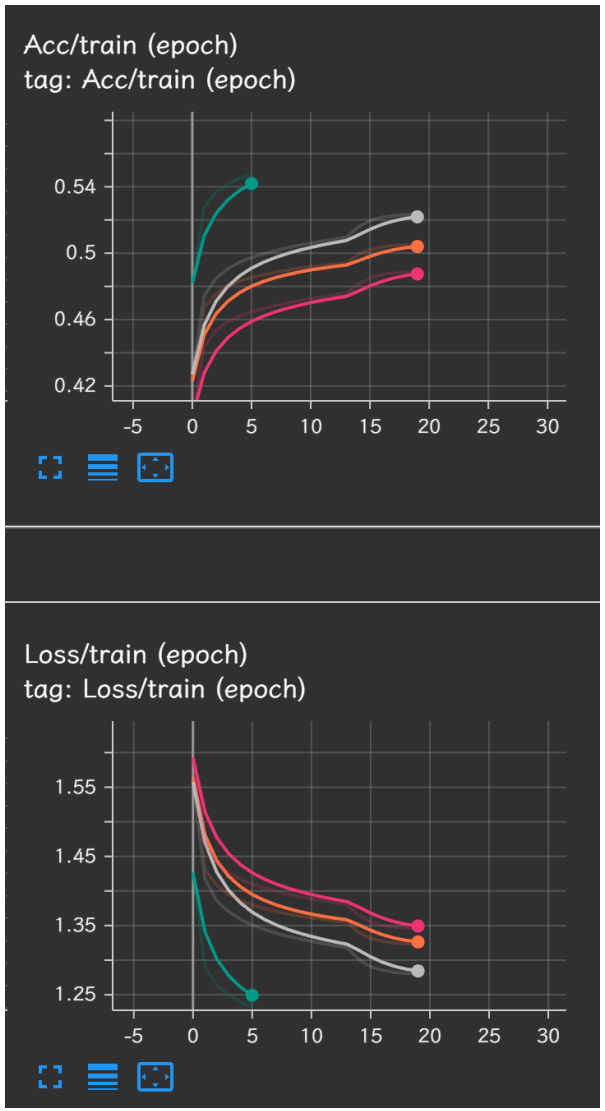
```
[2025-04-11 11:10:37,766] INFO: torch.Size([512, 11]), tensor([[-0.4589, -1.4792,  0.9454,  ...,  0.0225, -0.6776,  0.1235],
        [ 5.0874,  5.9901, -2.0360,  ..., -1.5450, -1.1884, -0.3939],
        [-2.8019, -3.0526, -0.5501,  ...,  1.7558,  0.5533,  1.5081],
        ...,
        [-3.0114, -2.7739, -1.2208,  ...,  0.6870,  1.3693,  1.5613],
        [ 0.6610,  0.1524,  6.8966,  ..., -1.5974, -1.2169,  1.1553],
        [ 2.2230,  4.0747,  5.6043,  ..., -2.4179, -1.1996, -0.1980]],
       device='cuda:0')
[2025-04-11 11:10:37,774] INFO: torch.Size([512]), tensor([ 3,  1,  6,  2,  7,  1,  3,  8,  9,  6,  4,  1,  1,  3,  3, 10,  3,  3,
         9,  7, 10,  9,  7,  0,  6,  5,  5,  8, 10,  5,  9,  8,  3, 10,  3,  1,
         0,  2,  3,  1,  1,  4,  8,  7,  2, 10,  7,  8,  9,  3,  5,  7,  5,  1,
         5,  4,  3,  0,  2,  9,  9, 10, 10,  3,  7,  8,  6,  0,  8,  6,  8,  2,
         4,  1,  0,  6,  0, 10,  9, 10,  3,  8,  2,  1,  7,  1,  7,  9,  8,  7,
         9,  5,  5,  4,  7,  2,  0,  8,  8,  5,  8,  0,  4,  6,  7,  6,  9,  1,
         3, 10,  8,  2,  7,  8,  3,  6,  6,  0,  7, 10,  5,  8,  0,  8,  7,  7,
         4,  5,  5,  7,  1,  1,  0,  3,  6, 10,  6,  9,  9,  0,  1,  2,  3,  3,
         6,  2,  0,  9,  8,  6,  9,  6,  2,  2,  2,  9,  8,  0,  4, 10,  4,  8,
         4,  8,  5,  2,  0, 10,  0,  2,  6,  3,  8,  7, 10,  7,  5,  5,  7,  7,
         2,  1,  4,  2,  6,  9,  4, 10,  4,  5, 10,  3,  2,  1,  5,  5,  8,  9,
         0,  1,  6,  8,  9,  5,  4,  3,  2,  2,  2,  6,  0,  9,  3,  1,  1, 10,
         3,  5,  6,  3,  2,  0,  7,  3, 10,  5,  8,  6,  2,  9,  9,  0,  3,  9,
         0,  8,  0,  5,  2, 10,  0,  3,  5,  2,  8,  1,  2,  6,  9,  3,  1,  2,
         6,  5,  8, 10,  0,  4,  2,  0,  3,  9,  2,  5,  9,  0,  5,  4,  6,  9,
         2,  5,  9,  0,  6,  8,  2,  9, 10,  7, 10,  2,  9,  8,  7,  5,  4,  3,
         0,  6,  5,  7, 10,  1,  9,  3,  1,  0,  8,  4,  9, 10,  6,  8,  5,  4,
         4,  5,  9,  2,  4,  8,  1,  7, 10,  5,  6,  6,  0,  5,  8,  3,  5,  9,
        10, 10,  6,  0,  8,  4,  0,  3, 10,  6,  2,  9,  5,  9,  7,  2, 10,  5,
         0,  2,  7,  6, 10,  1, 10,  1,  5,  2,  2,  9,  2,  7,  0,  6,  9,  5,
         0,  1,  7,  0,  3, 10,  8,  3,  2, 10,  9,  4,  8,  0,  4,  6,  5,  3,
         1,  1,  6,  3,  5,  0, 10,  0, 10,  5,  9, 10,  3,  1,  7,  9,  1, 10,
         8,  9,  2,  3,  0,  3,  1,  7,  3,  4,  5,  1,  3, 10,  6,  9,  7,  8,
         7,  2,  6,  1,  2,  3,  1,  1,  6,  3,  1,  4,  6,  1,  6, 10,  2,  9,
         4,  8,  8,  5,  0,  0,  5,  7,  4,  8,  1, 10,  5,  3,  6,  5,  3,  8,
         0,  0,  2,  2,  1,  5,  8,  1,  0,  0,  8, 10,  1,  5,  3,  7,  0,  8,
         3,  9,  8,  9,  4,  4,  5,  6,  7,  2,  6,  7,  6,  0,  8, 10,  8,  0,
        10,  7,  0,  0, 10,  6,  8,  3,  5,  0,  4, 10,  2,  6,  9,  0,  8, 10,
         4,  2,  4,  0,  3,  5,  2,  2], device='cuda:0')
```

ParT, no ID result

Intermediate stage, final results can better;
Metric: 0.55868



ParticleNet:



No PID; (No PID information entry)
TruthPID (not finished yet)
RecoPID
RecoPID+Only ChargedTracks(No nertral PFO)

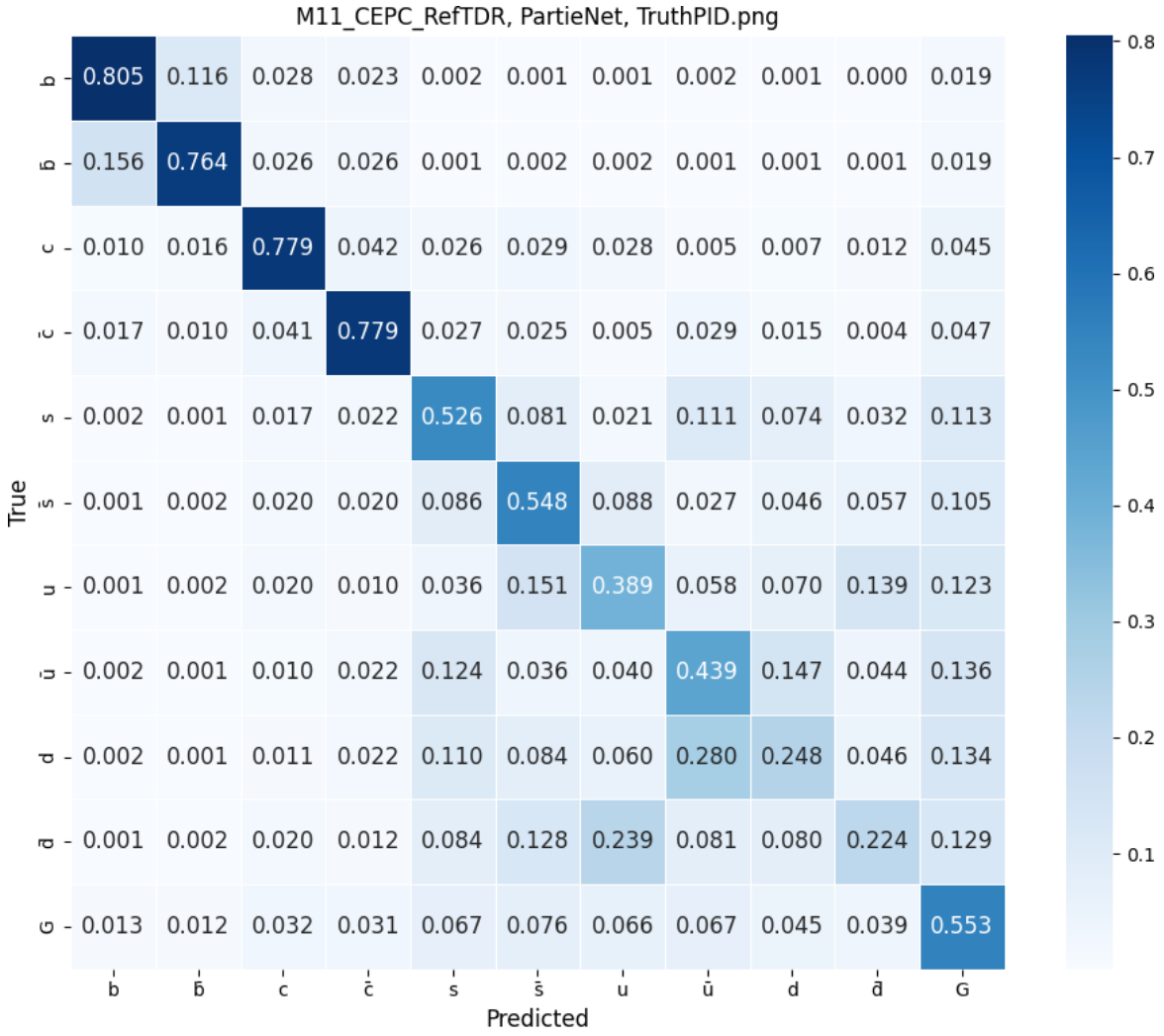
Value	Step
0.4889	19
0.5476	5
0.5234	19
0.5052	19

No overtraining;
Results shows:

TruthPID > RecoPID > RecoPID chargedtracksonly > No PID.

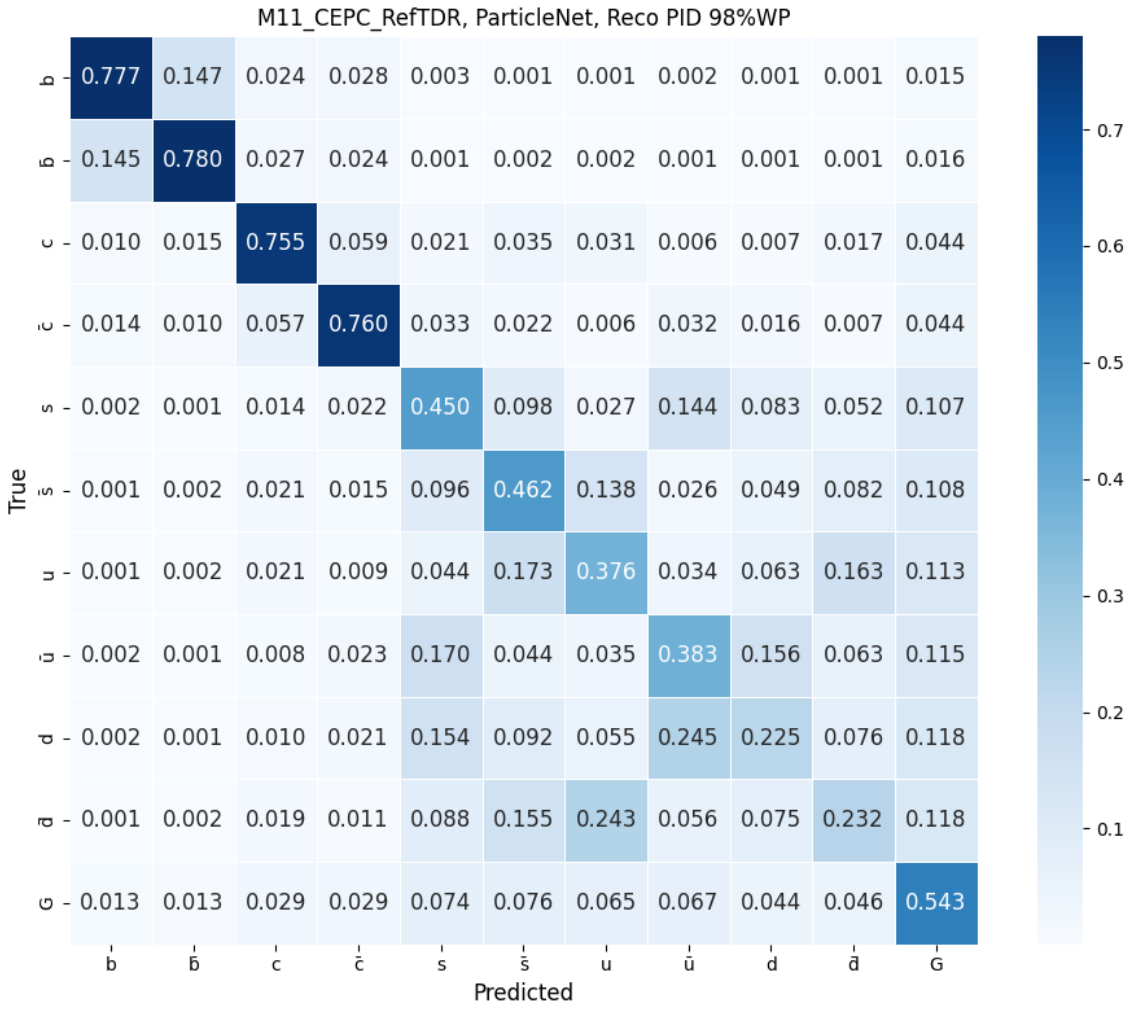
PN, truth PID

Intermediate stage, final results can better;
Issue: charge asymmetry;
Metric: 0.55052



PN, reco PID

Metric: 0.52211

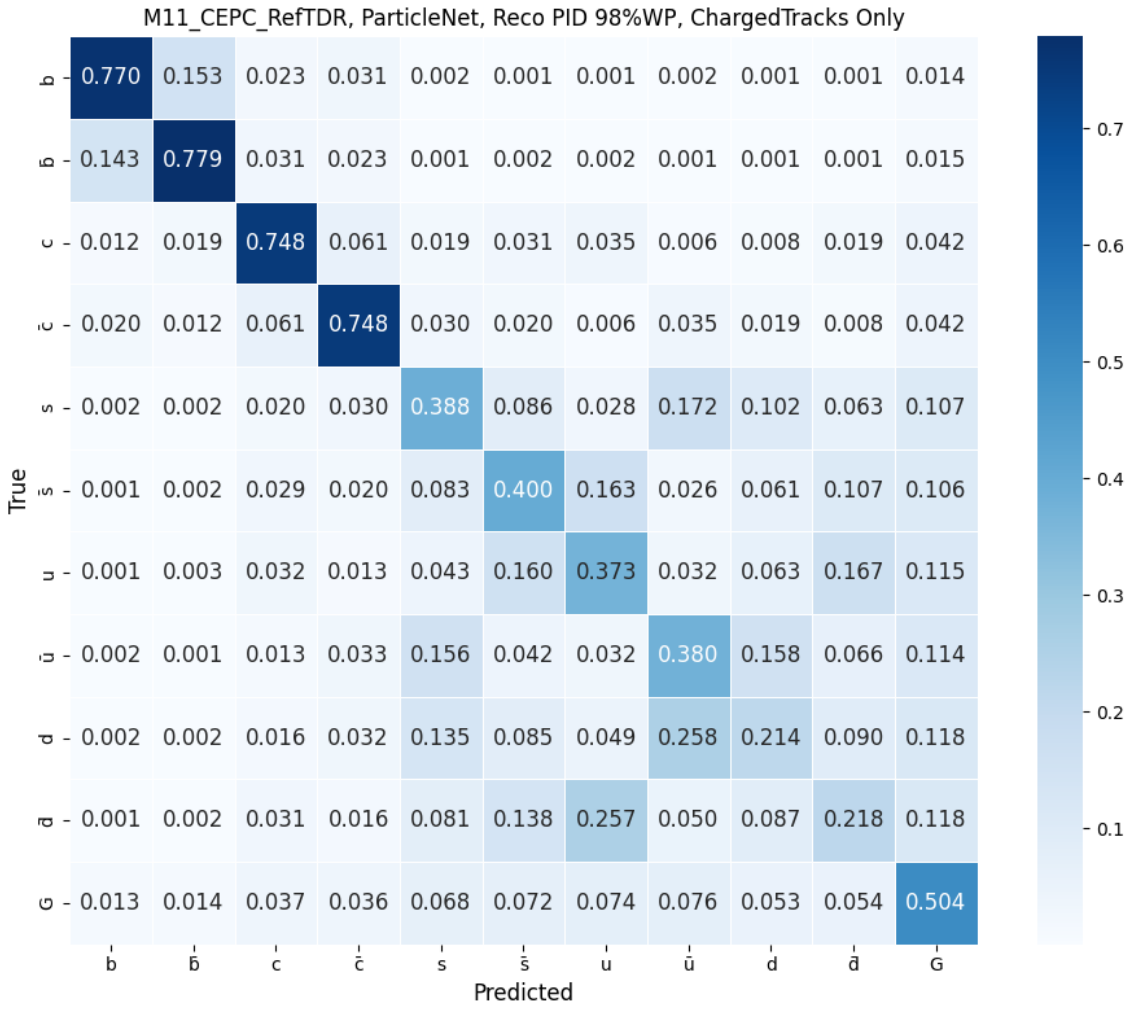


98% Lepton WP used in this JOI.
May not the best option. (high pi to muon migration rate, worse purity)
Impact of different WP Under tuning.

@geliang: 70%WP recommended in Jet. (purity essential).
98%WP for high energy leptons.

PN, reco PID, charged tracks only

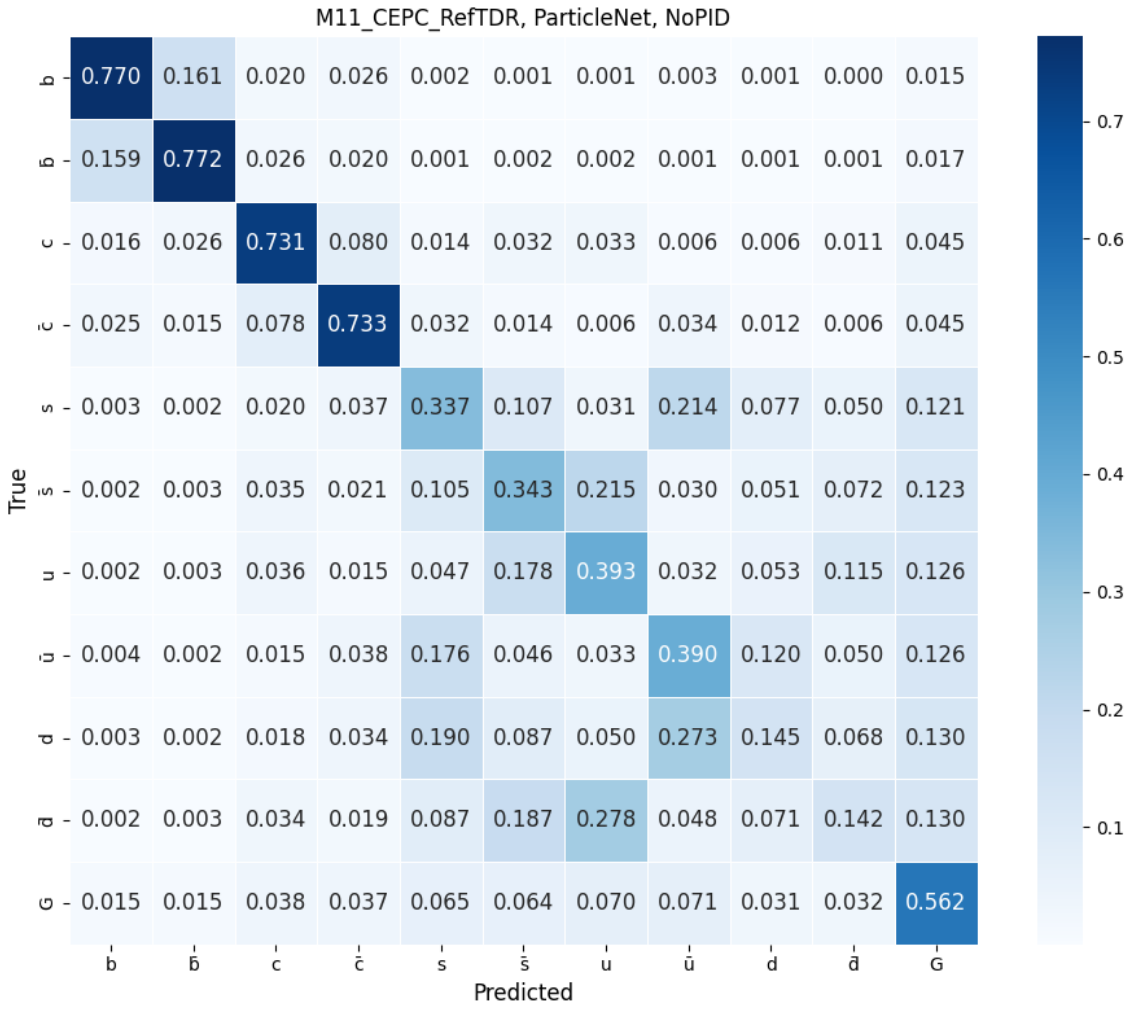
Metric: 0.50198



Only charged track informations in JOI.
Neutral PFOs not included.
This corresponds to committee question “What if no Calorimeter”.

PN, no PID

Metric: 0.48355



Extrapolation

Red: from extrapolation.

	PN		ParT	
Truth PID	0.55052	+5.4%	0.63605	
Reco PID	0.52211	+4.0%	0.60323	
Reco PID, charged tracks only	0.50198	+3.8%	0.57997	
No PID	0.48355		0.55868	+15.5%

10 points plot, ROC curve under preparation.