CEPC

Jets, samples and Wednesday working meeting

Kaili Zhang

zhangkl@ihep.ac.cn

AFB: Practice on fit cos spectrum

@Han Shuo, Jiawei

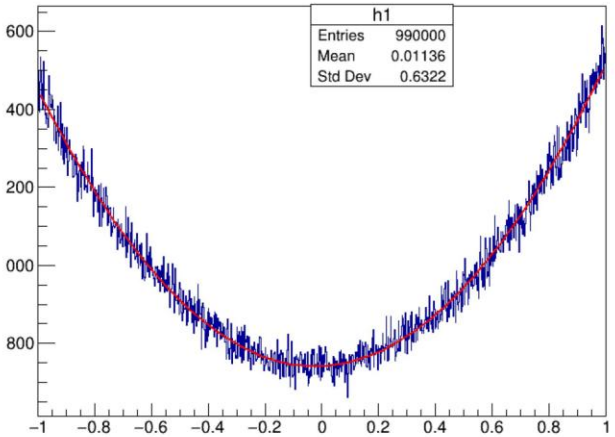
Fitting the costheta distribution

- Set costheta function = [0]*(1 + [1]*x + x*x), where [1] = 8/3 * A_FB
- With MCP, fitted [1] = 0.04463 ± 0.00251, so AFB = 0.01674 ± 0.00094
- Counting AFB = (503202-486798)/(503202+486798) = 0.01657 ± 0.00100
- Compatible within statistical uncertainty
- **Present it as an alternative method?**
  - **Statistical uncertainty of each fit is large (only 1M events), so it's hard to estimate syst. Un. compared with the counting method**

$$\sigma_F = \int_0^1 \frac{d\sigma}{dcos\theta} dcos\theta \quad \sigma_B = \int_{-1}^0 \frac{d\sigma}{dcos\theta} dcos\theta \quad (9)$$

and $cos\theta$ is the angle of the outgoing fermion measured relative to the incident electron direction. The experiments determine A_{FB} from fits to the angular distribution which can be written as

$$\frac{1}{\sigma}\frac{d\sigma}{dcos\theta} = \frac{3}{8}(1 + cos^2\theta) + A_{FB}\, cos\theta \quad (10)$$



9

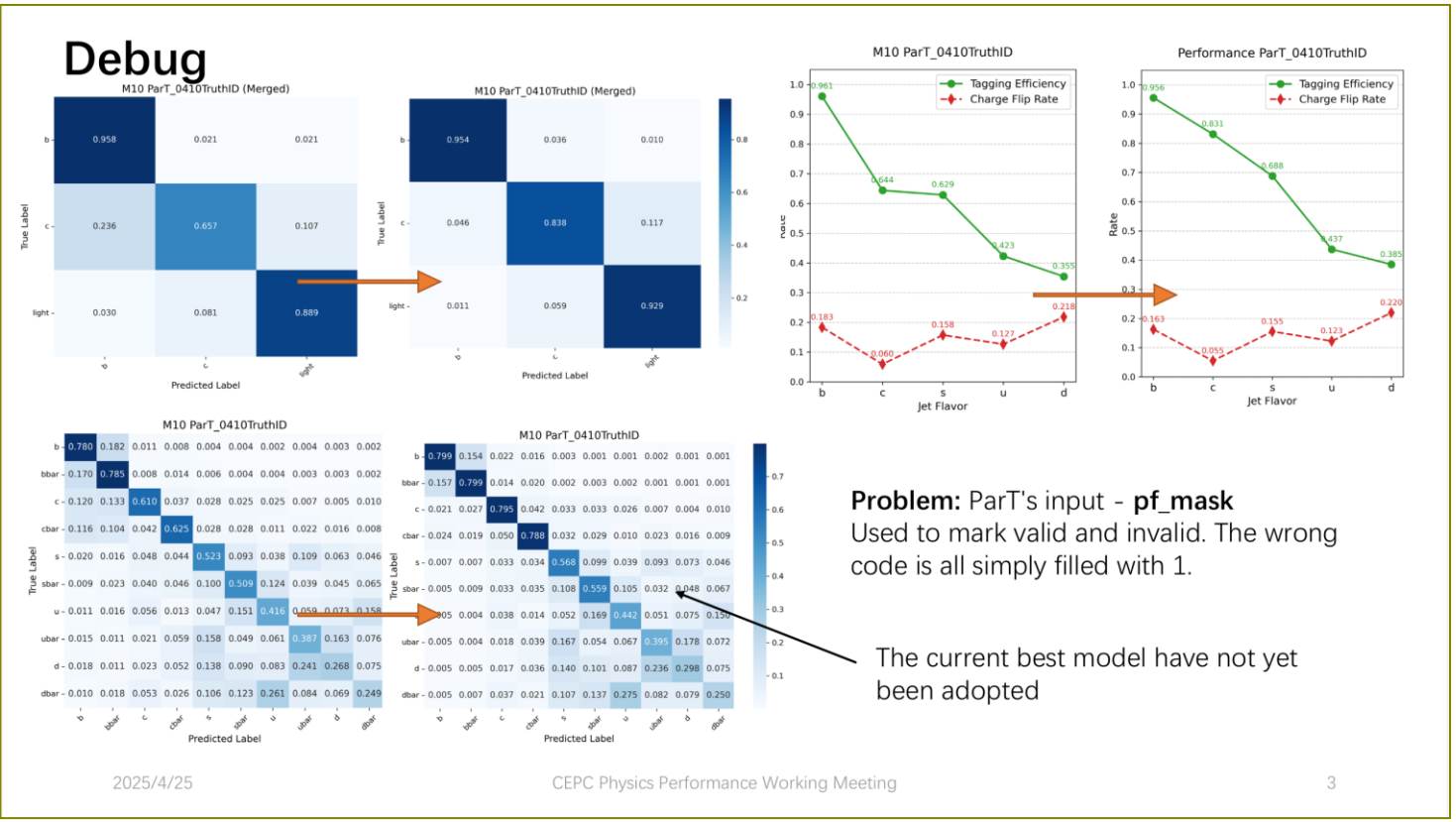
Uncertainty consistent with number counting.

For counting experiment: Including energy spread as uncertainty source. $2*10^{-5}$.

Need better understanding of spectrum shape.

Rb

@Hancen, Mohan



Fixed the bug in model application and least squares.
Now the results consistent.

Error:

Original result (5×10^6 events):
$R_b = 22.808 \pm 0.733\%$
$R_c = 15.434 \pm 0.939\%$

Scaled result (10^{11} events):
$R_b = 22.808 \pm 0.00518\%$
$R_c = 15.434 \pm 0.00664\%$

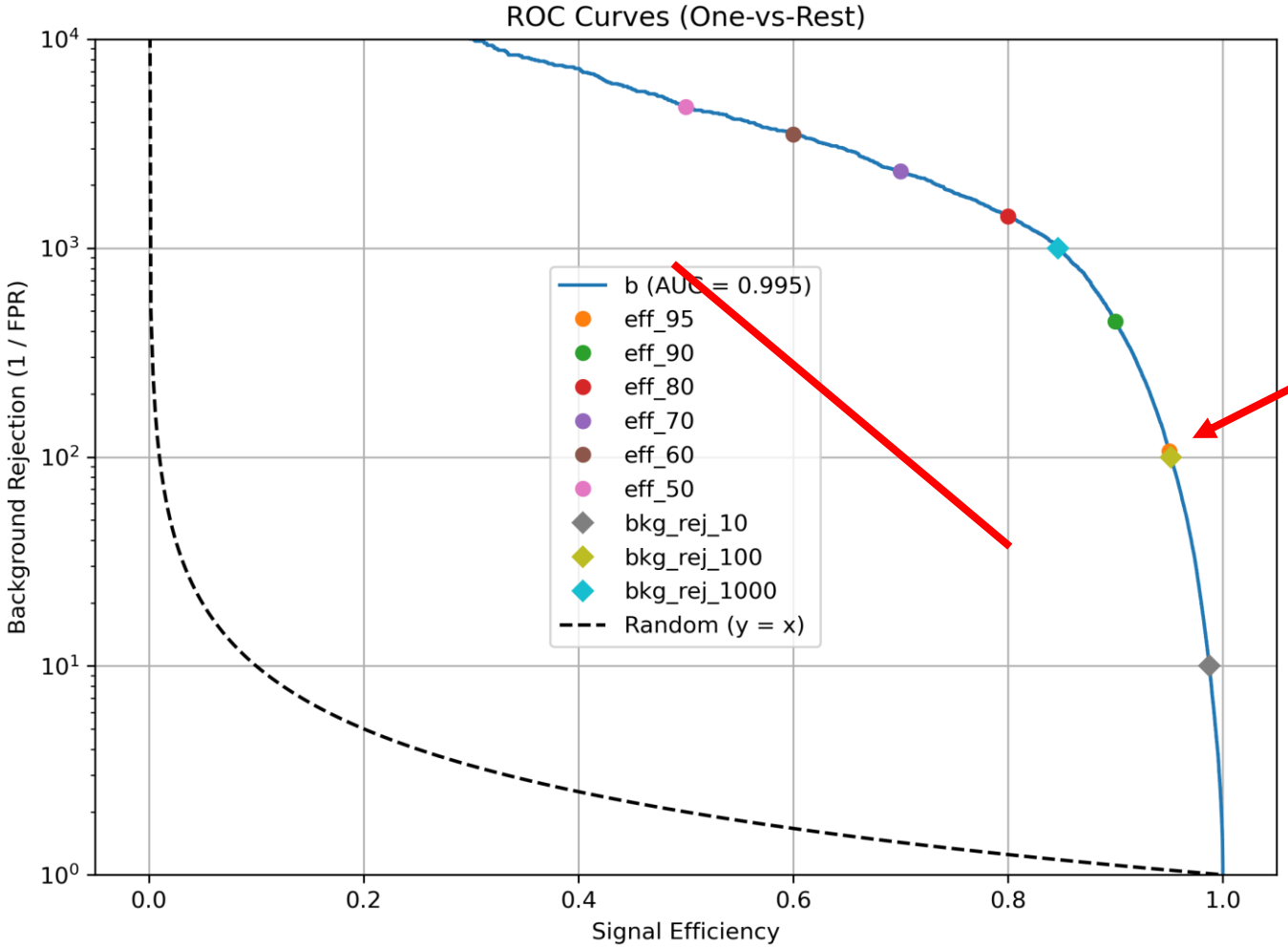
Corrected:

Original result **lstsq** (N_h,pro = 5.0e+06):
R_b = 0.21891525 ± 0.00125486
R_c = 0.17104597 ± 0.00174900

Original result **solve** (N_h,pro = 5.0e+06):
R_b = 0.21939958
R_c = 0.17284379

ROC Curve - b

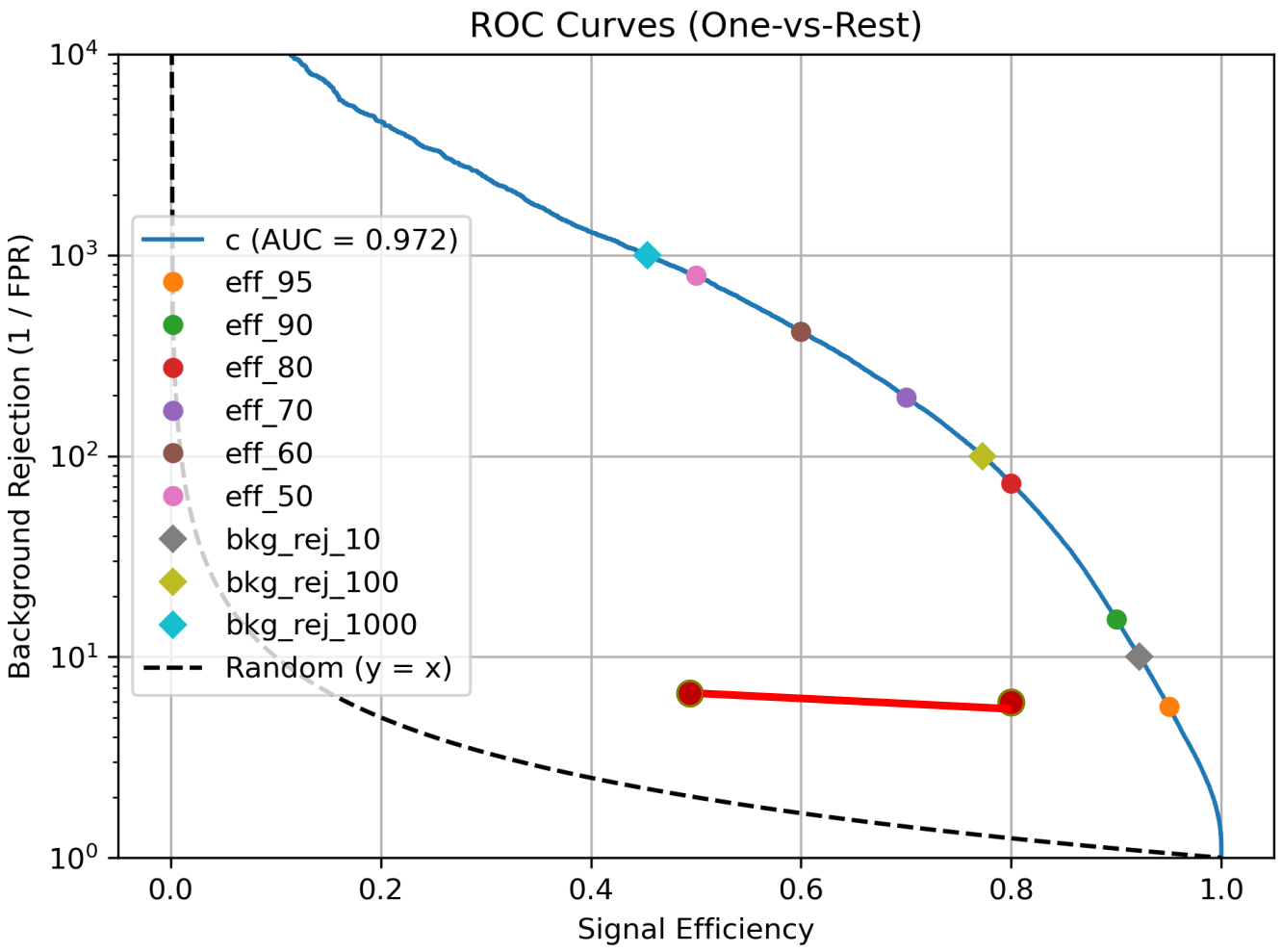
LCFI Plus: 80% b Eff, 2.23% misID to c; (0.05% for ParT)
50% b Eff 0.11% misID to c; (0.02% for ParT)



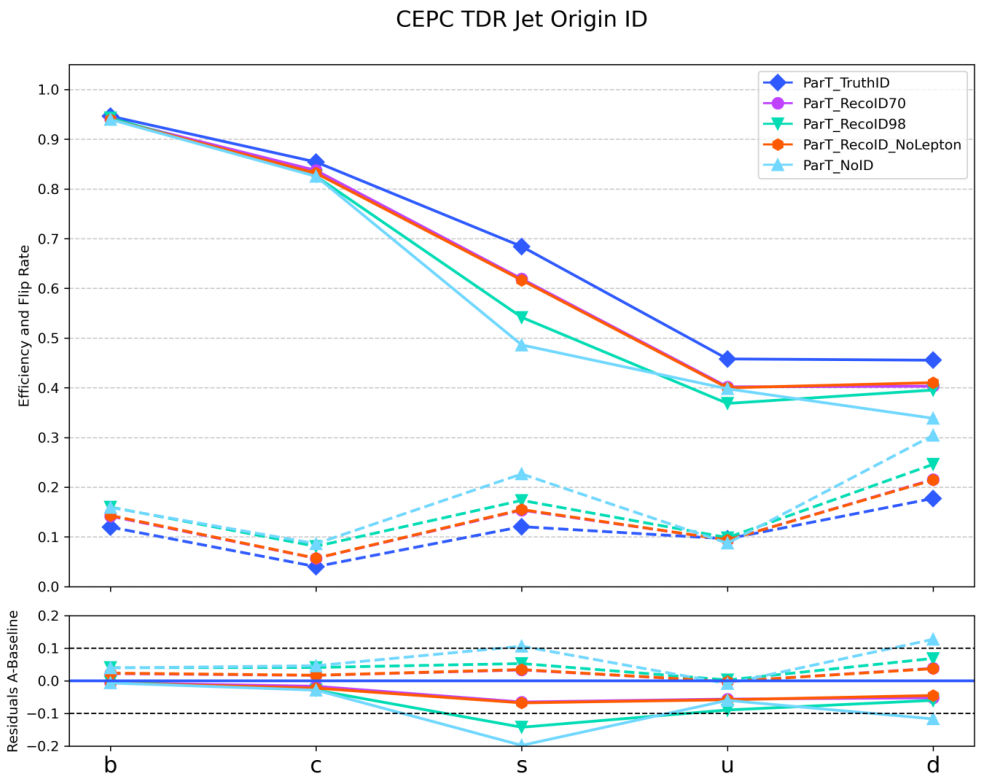
Score= 0.378
Eff = 95%
Bkg_rej =106

Roc curve - c

LCFI Plus: 80% c Eff, 13.58% misID to b; (1% for ParT)
50% c Eff 13.9% misID to b; (0.1% for ParT)

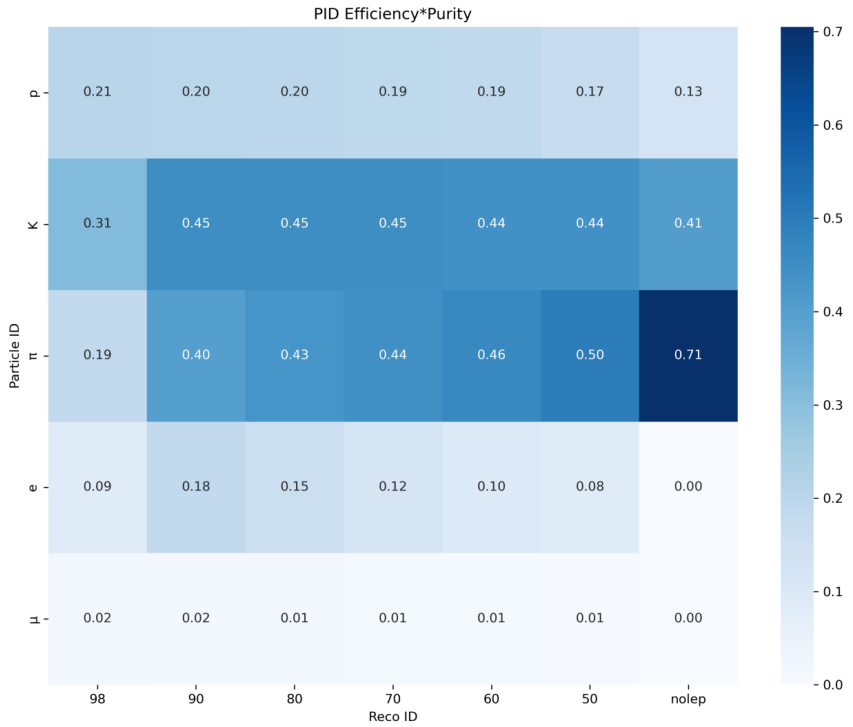


JOI Conclusion



- JOI Performance consistent when:
  - Test event size @bootstrap method
  - Extrapolate to different energy and process
  - Specialized to M10 even tagging:
    - Which M11 can reduce to M2 like b & not b tagging. Still sharp.
    - B tagging WP 95% and bkg 0.01 rejection at ~0.35.

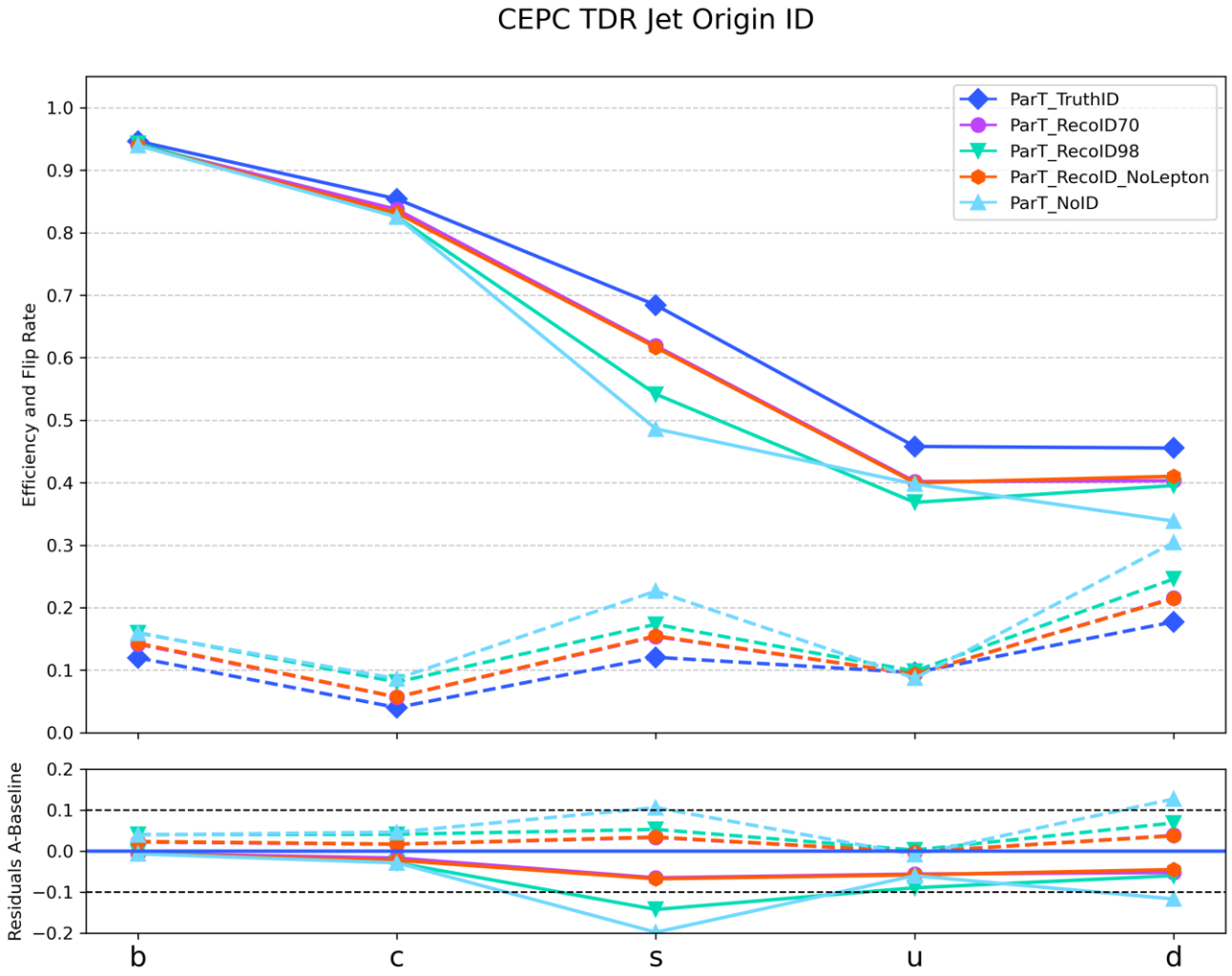
PID Conclusion



- In current lepton ID with chi2 method
  - Lepton eff well with the WP description while the purity too bad.
  - Hardware limitation and high pion ratio
  - Expecting Xgboost method
- With smaller lepton eff, better purity and eff*purity result, better JOI.
  - Also, no lepton WP better than Reco98 WP.

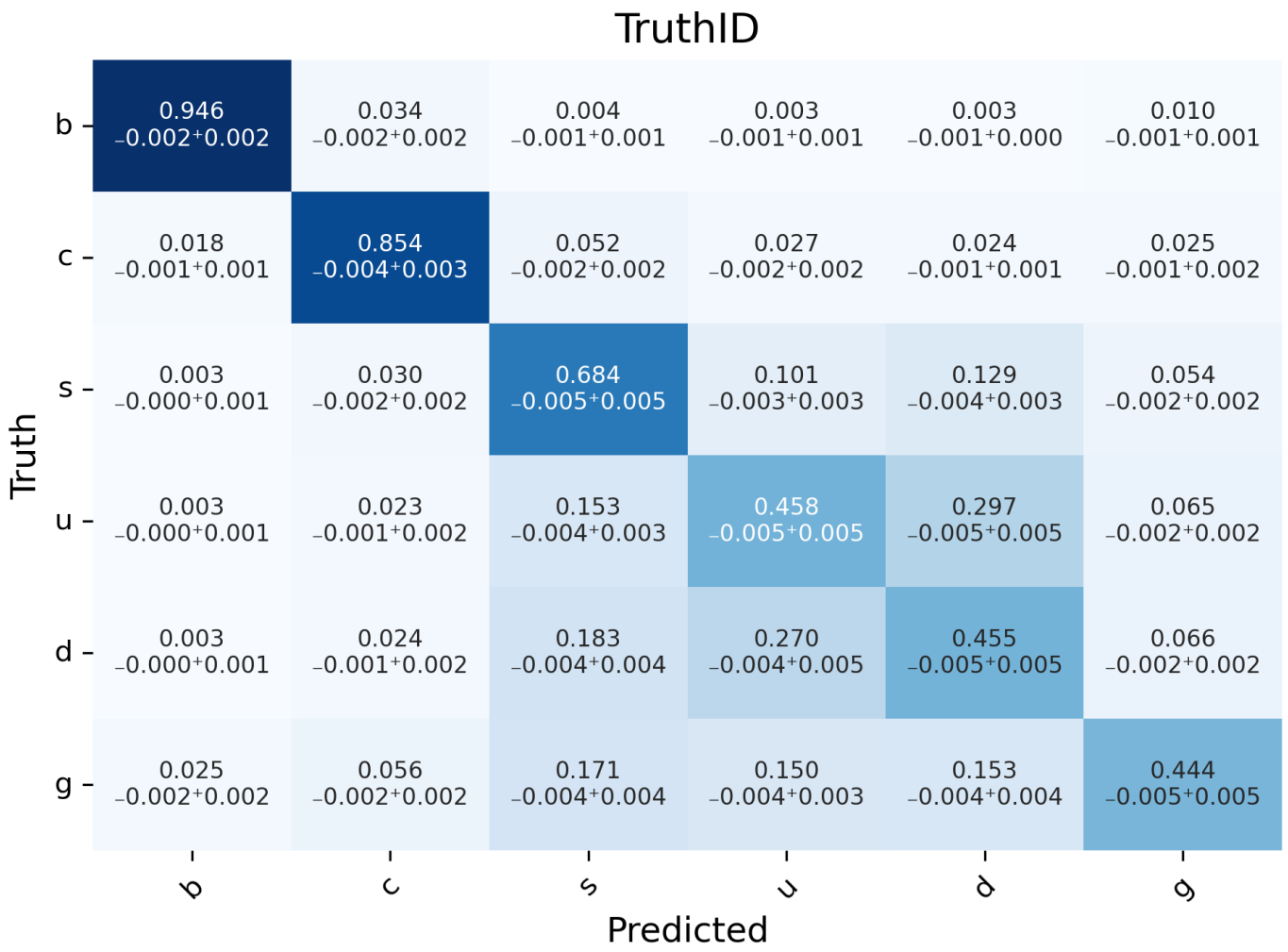
backups

Hqq JOI



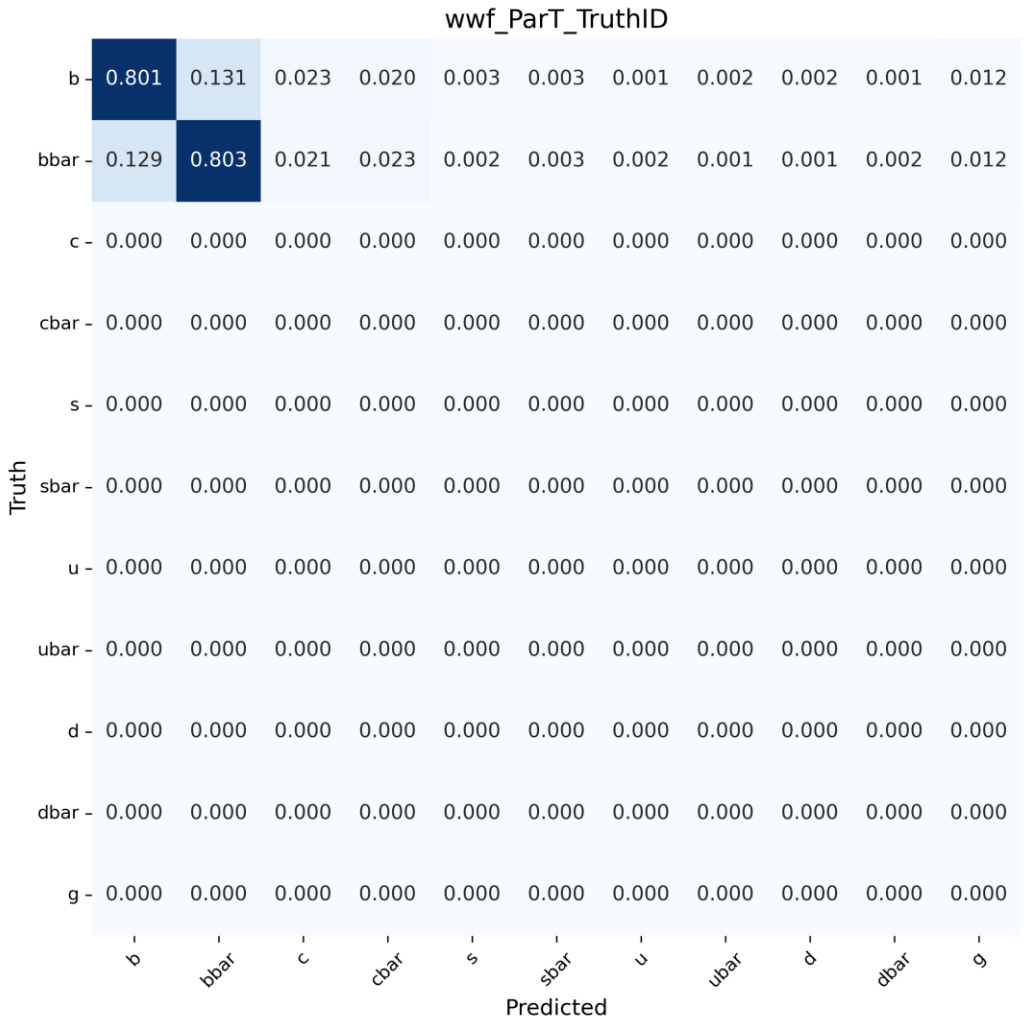
Truth:	0.5936
Reco70:	0.5510
Nolep:	0.5494
Reco98:	0.5153
NoID:	0.4825

Eff with uncertainty from bootstrap

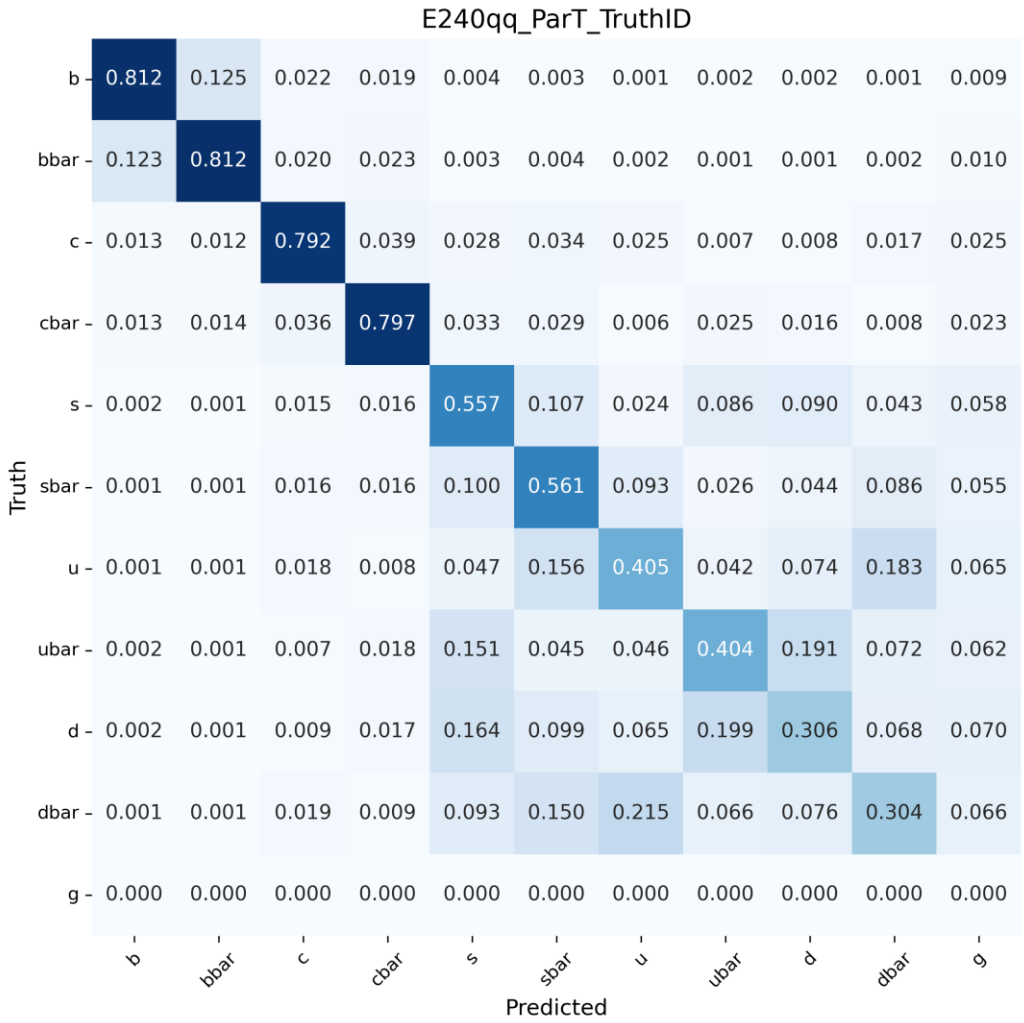


Extrapolate to: E360 n1n1_bb

B tagging eff ~93%



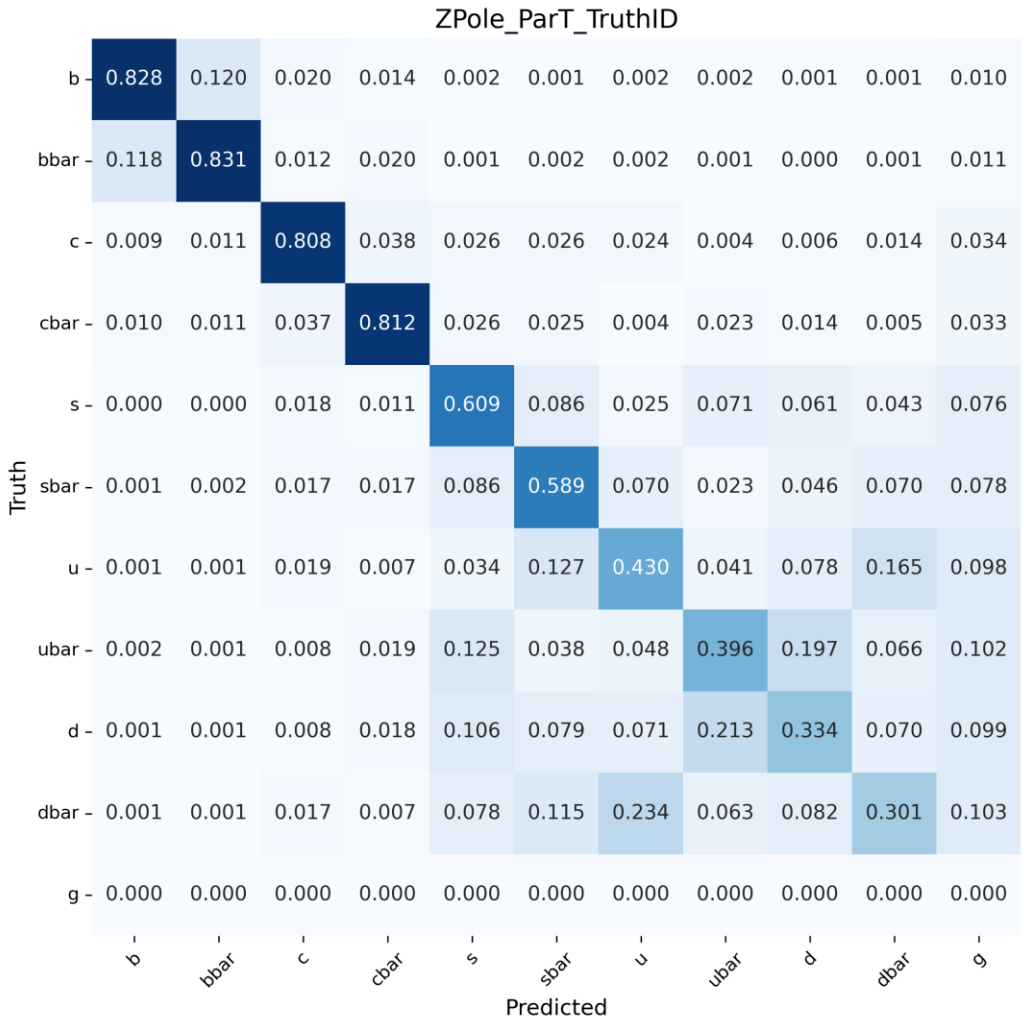
Extrapolate to: E240 2f qq



0.5750
(With no gluon tag)

E91.2 Zpole b/c/u/d/s

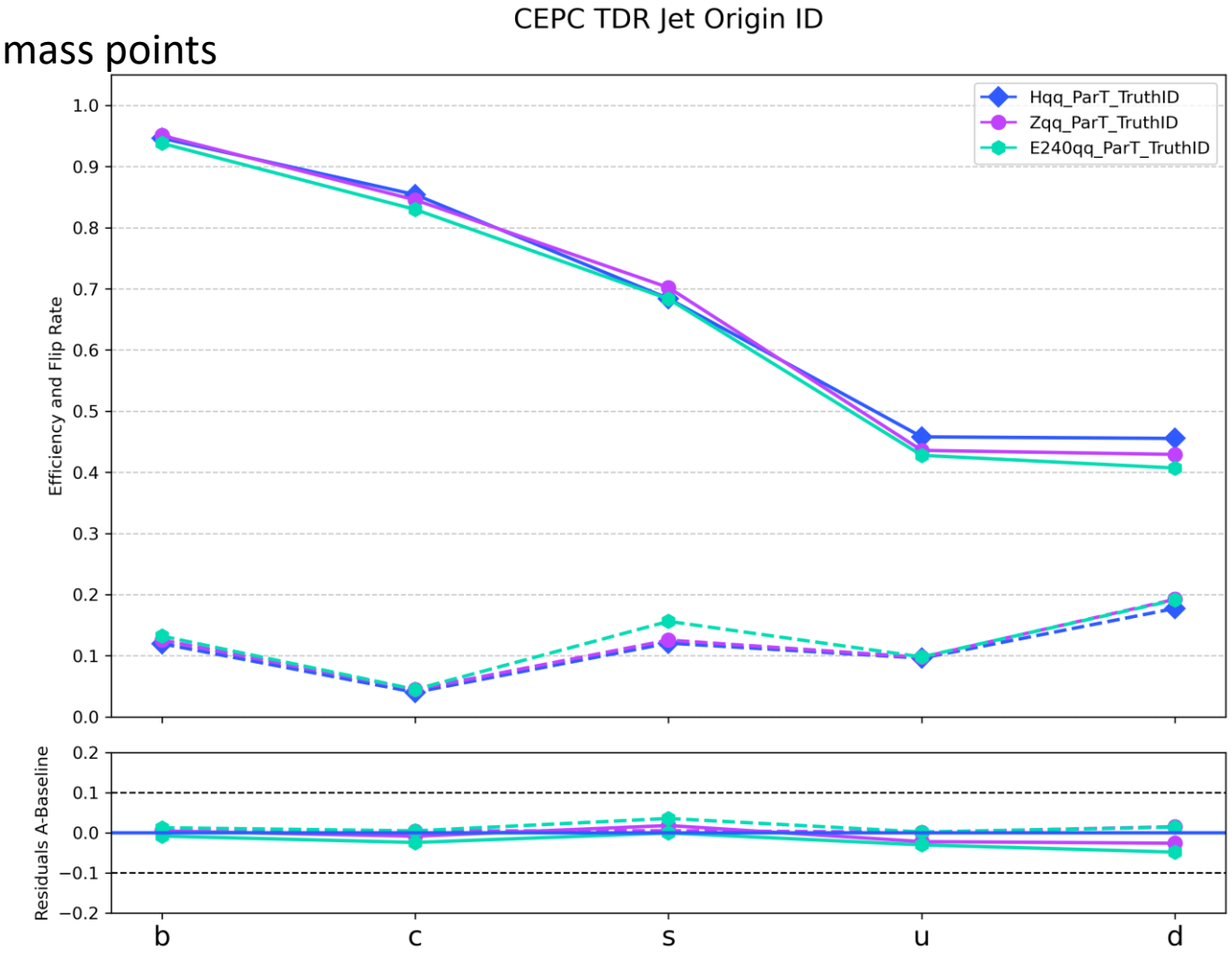
No performance downgrading.



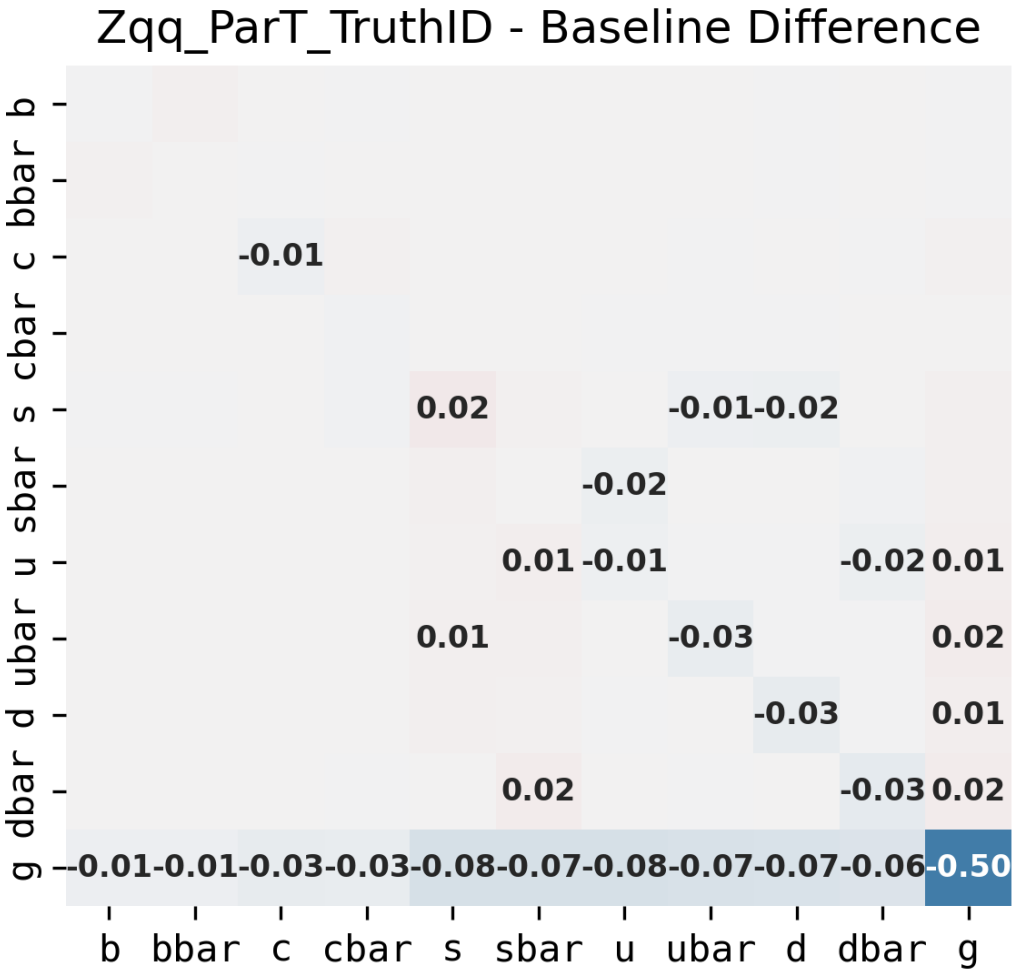
0.5938
(With no gluon tag)

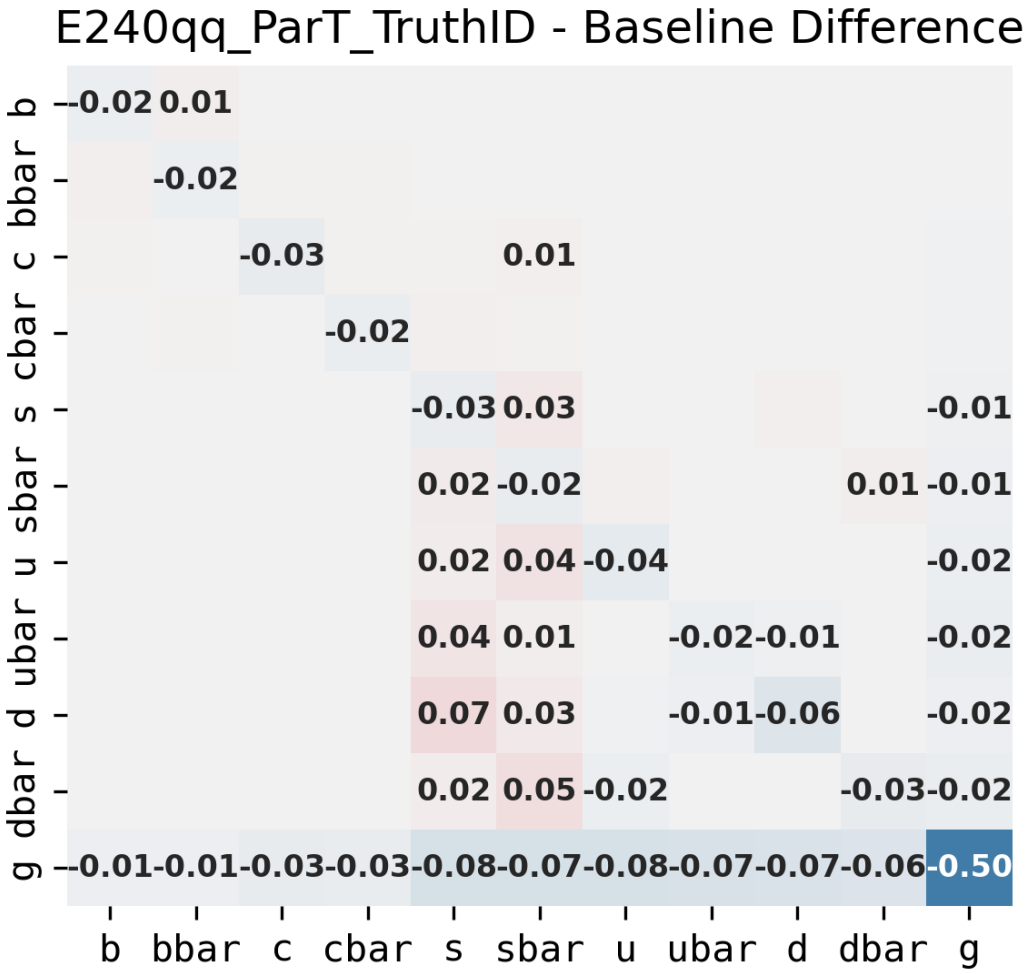
10 points-comparision

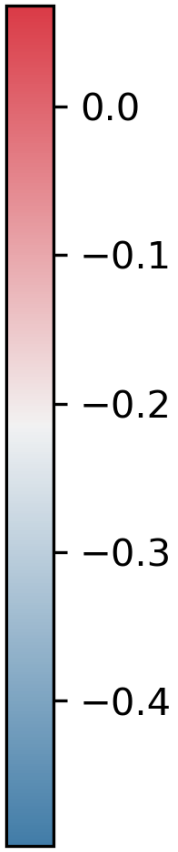
E91.2-E125-E240 qq
Performance consistent.
JOI can extrapolate to other mass points and other processes.



Difference in M11 Matrix

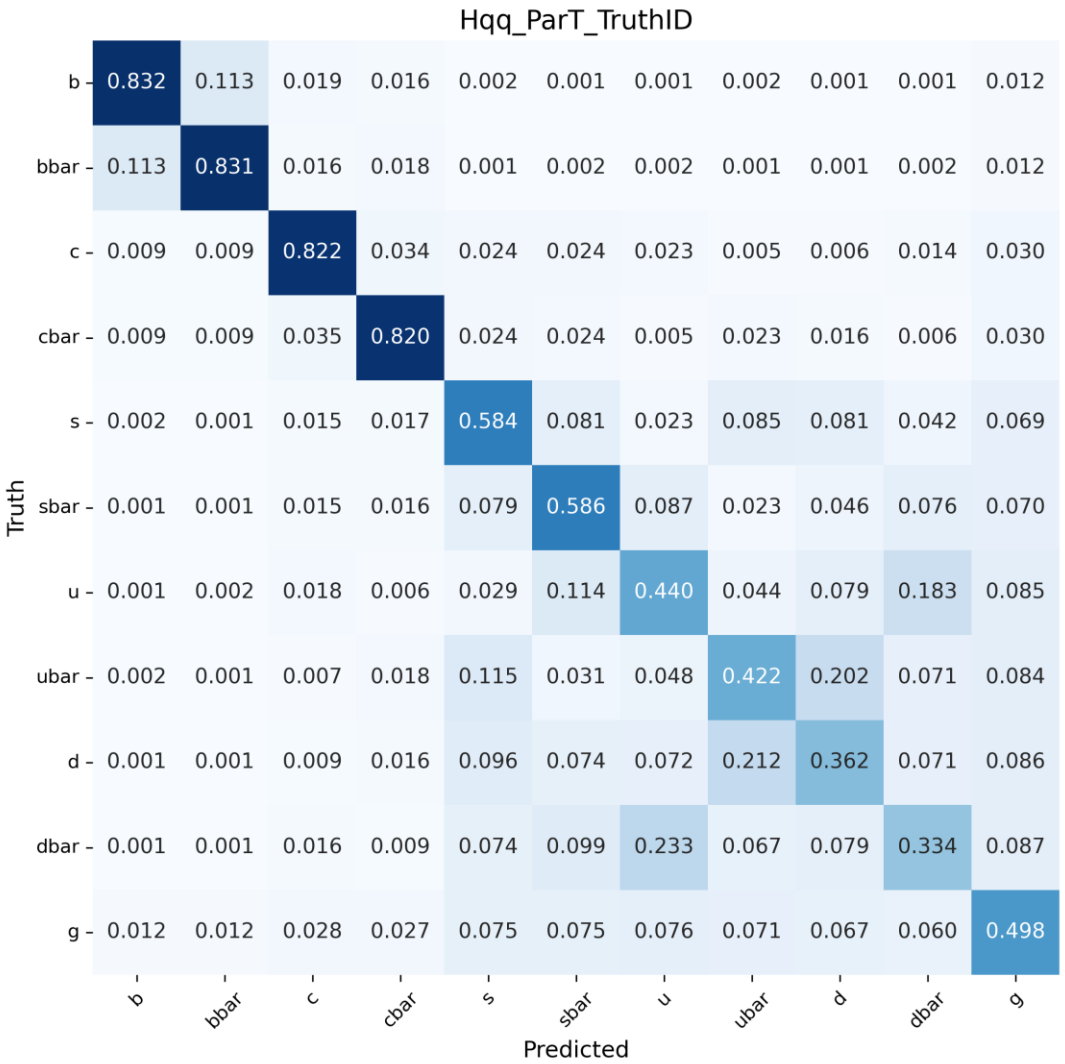




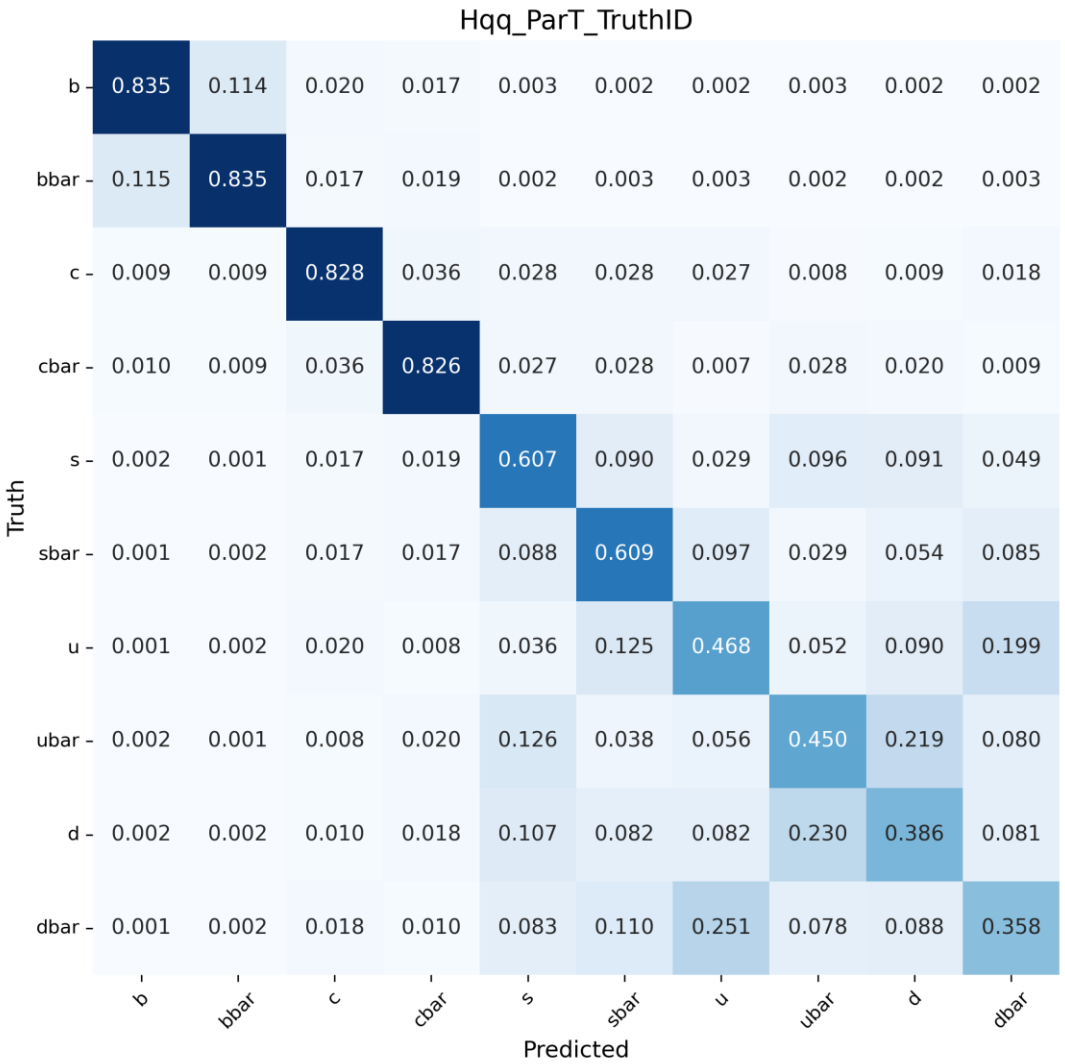


M11 to M10

Ignore gluon score and gluon category.

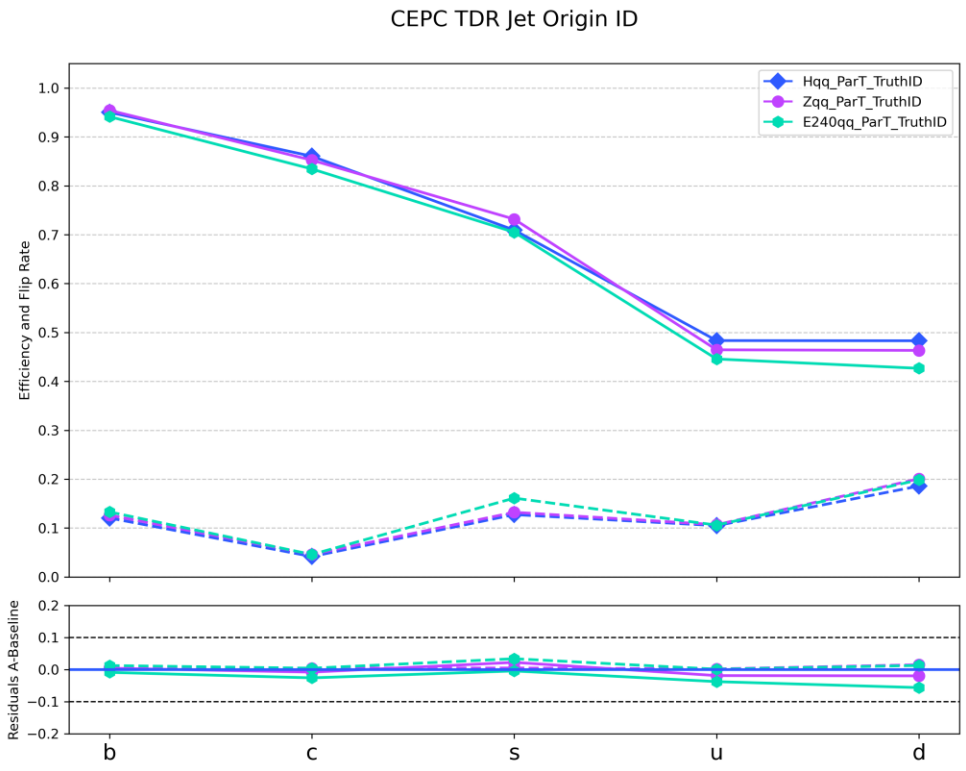


0.6202

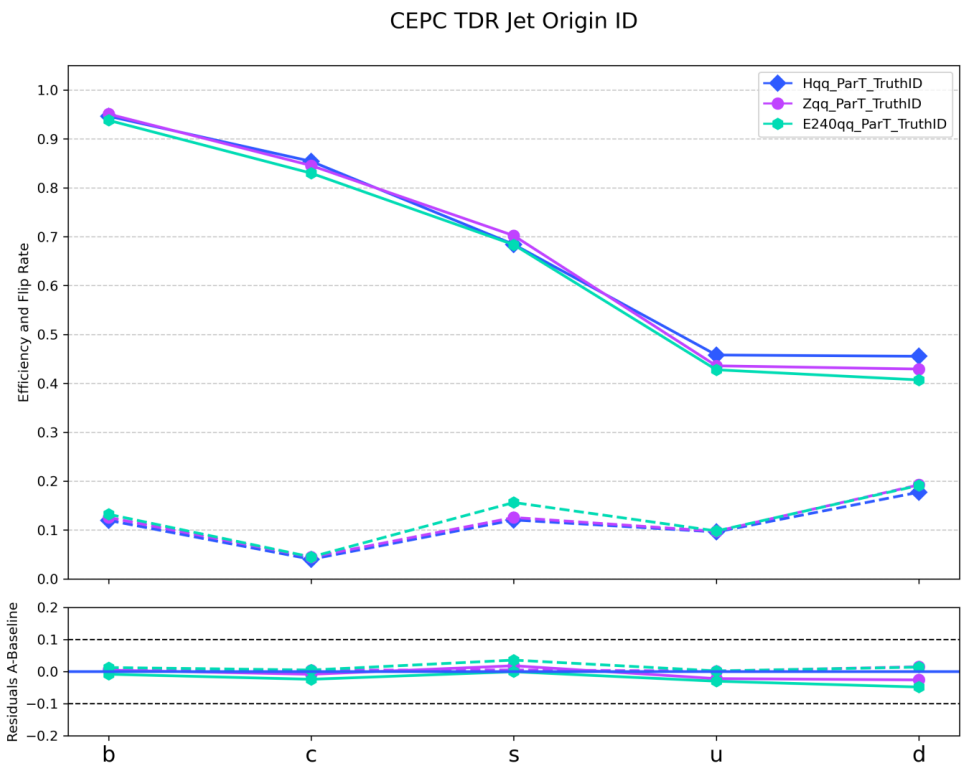


Tiny difference when M11 to M10 in 10 points plot.

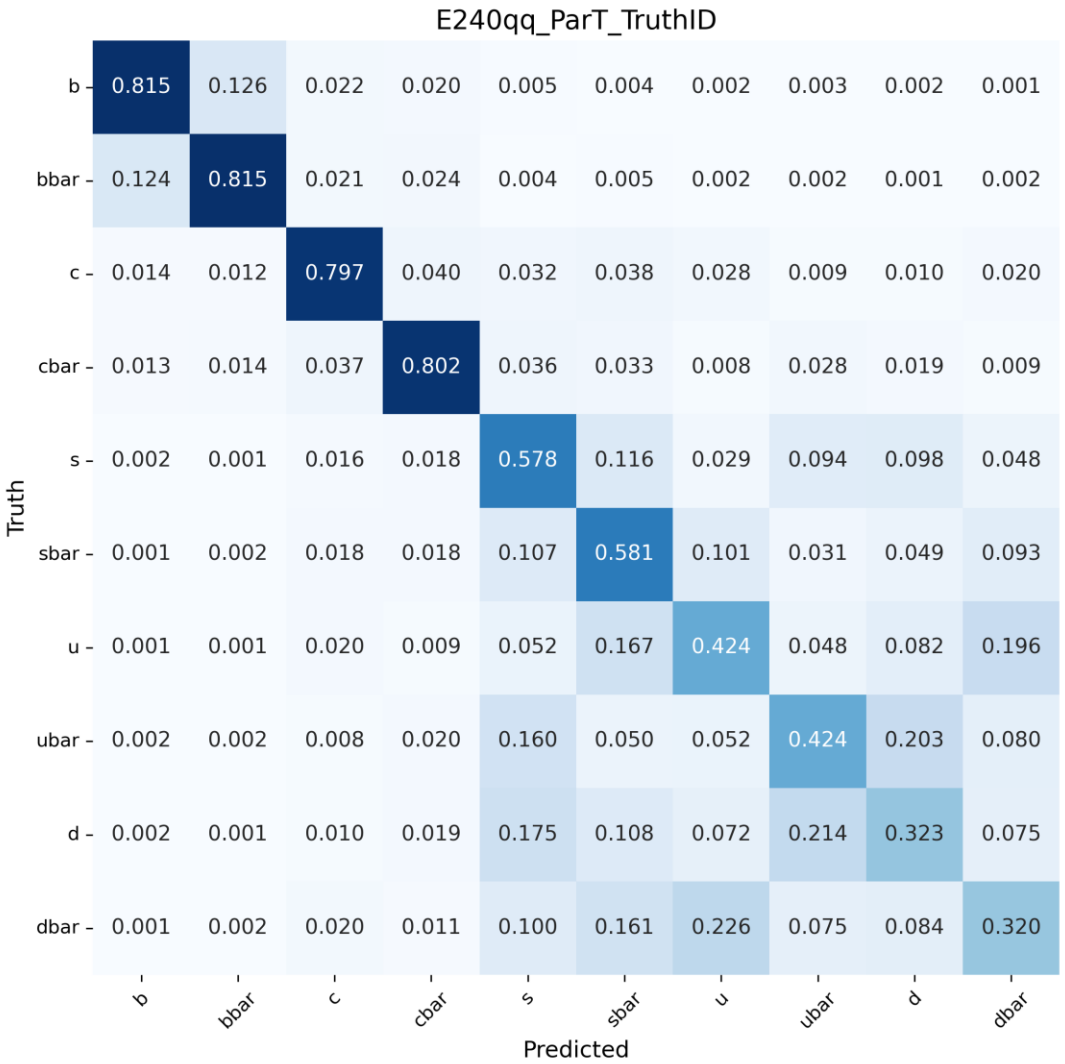
M11

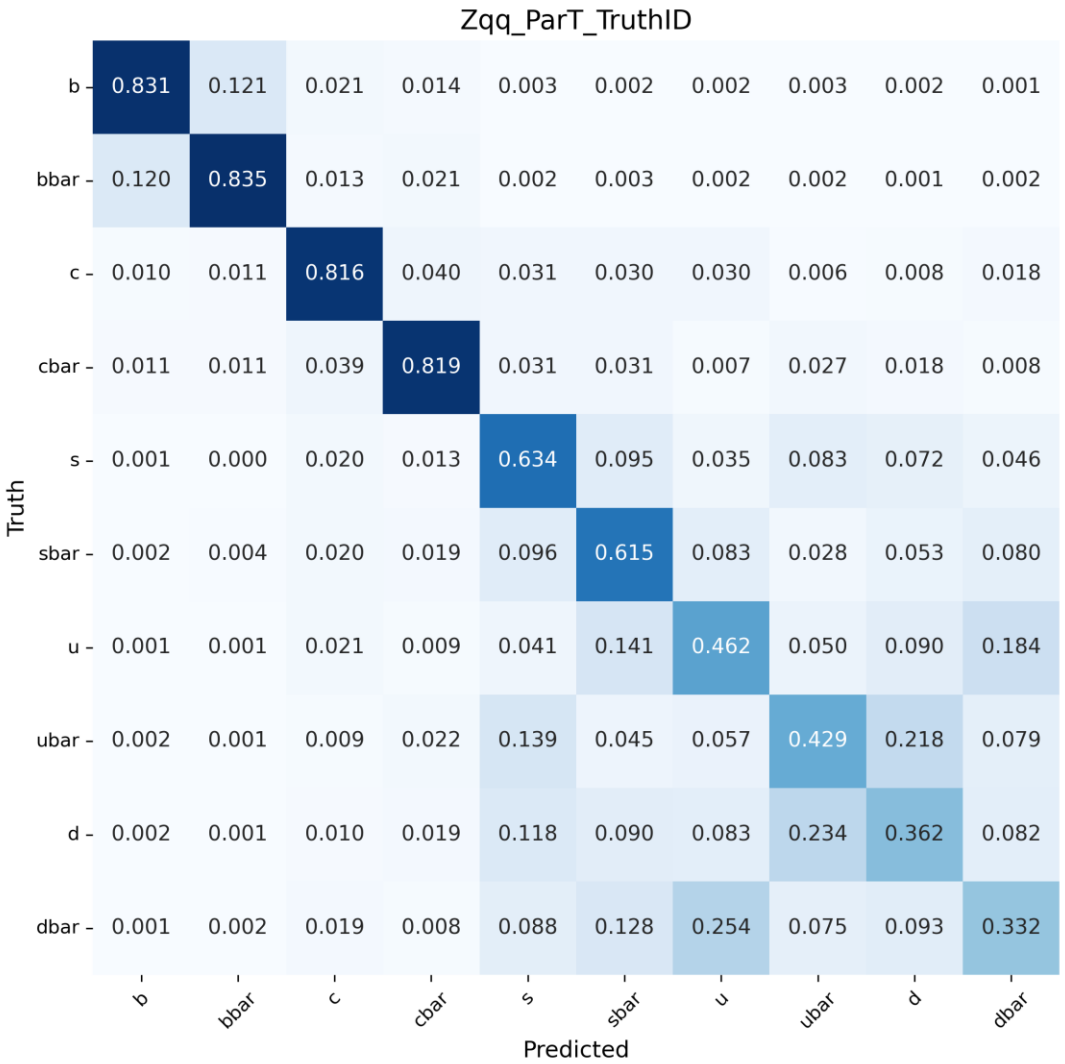


M10

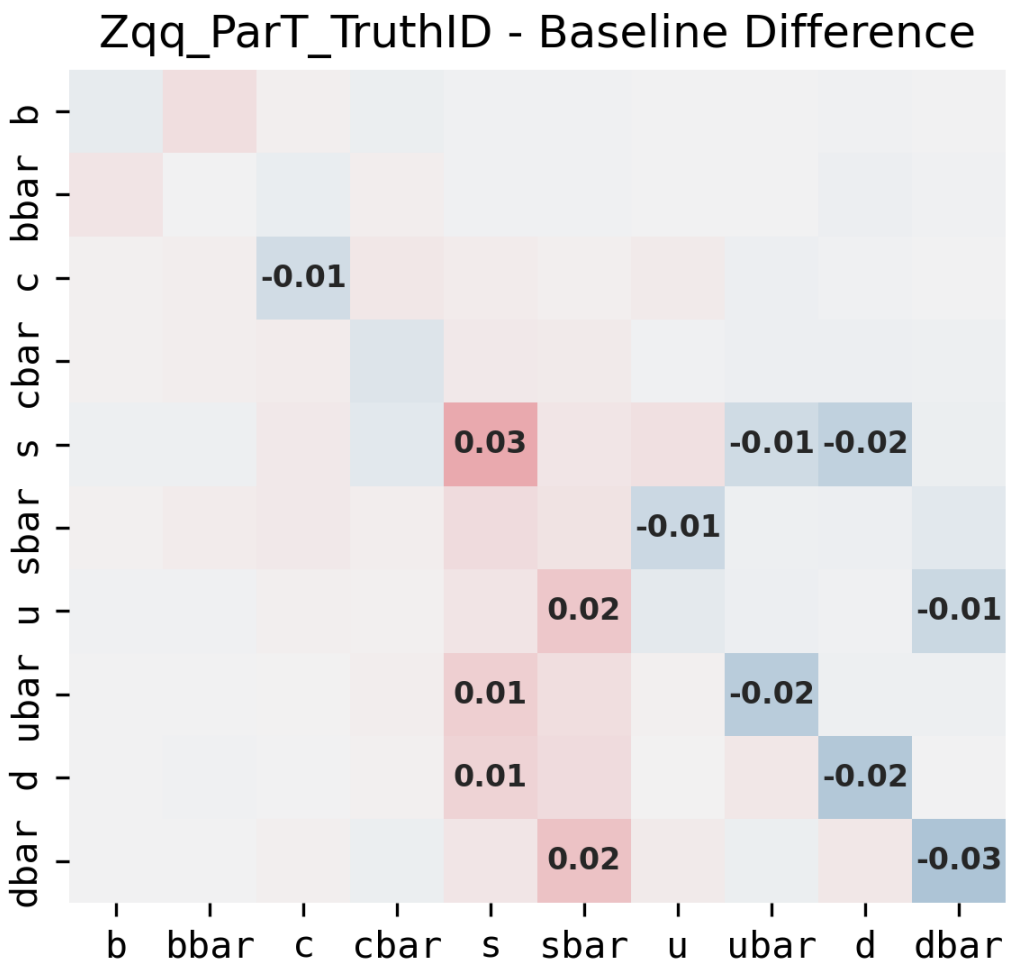


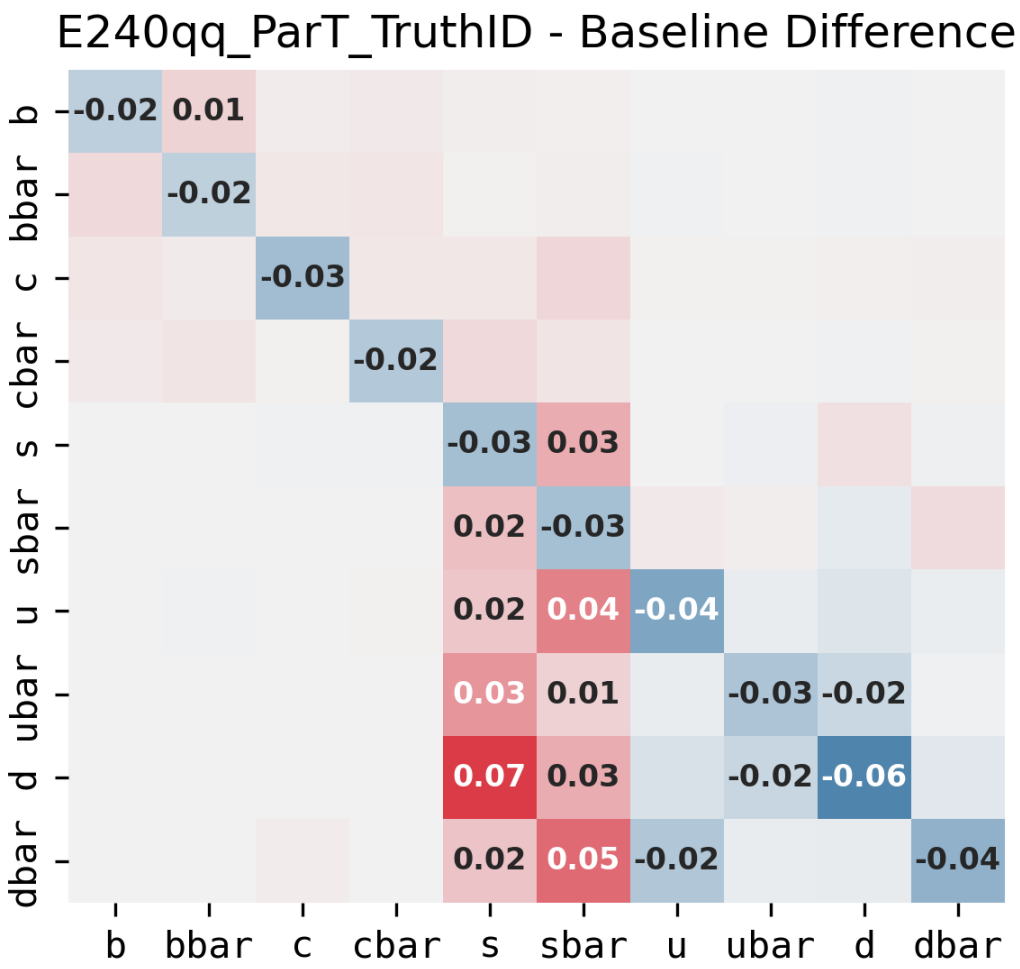
M10 in E91.2qq and E240qq

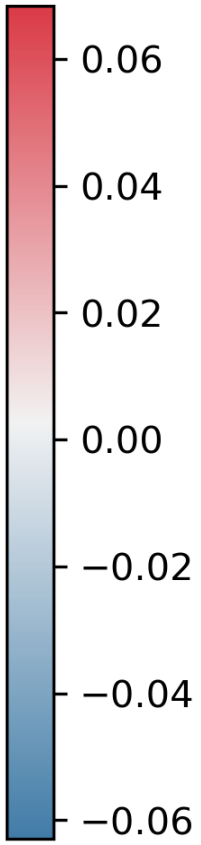




Difference in M10 Matrix





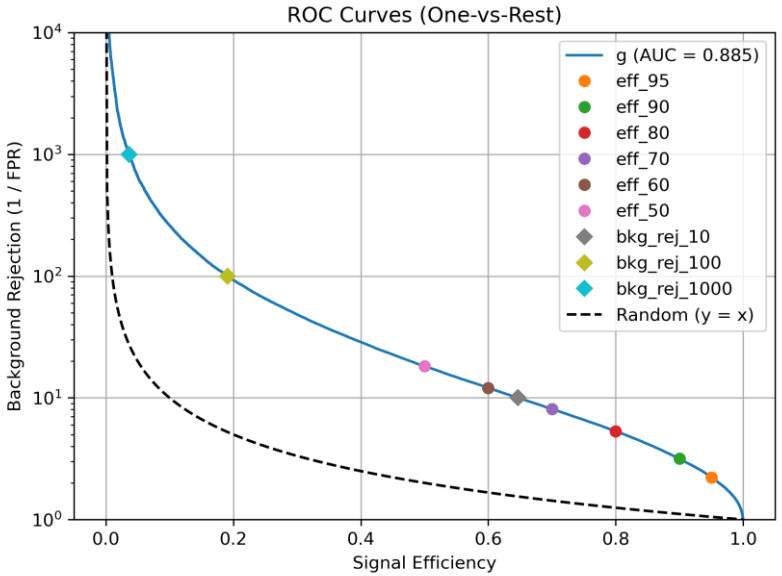


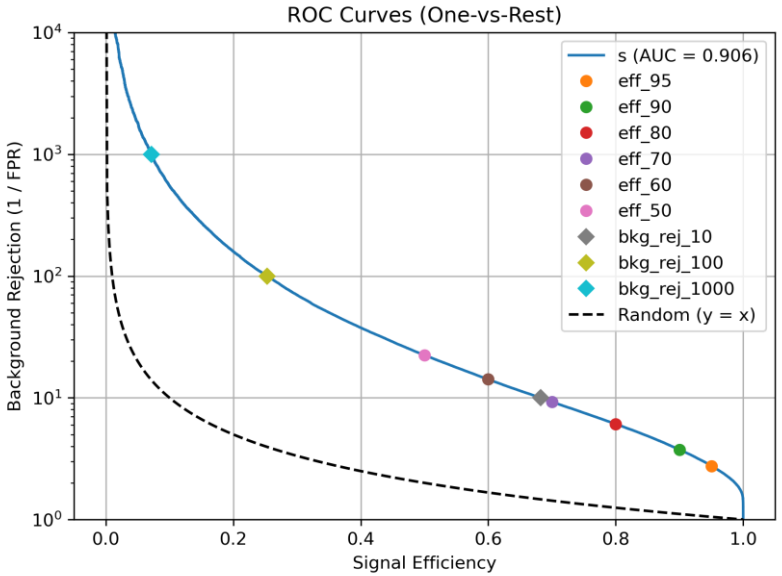
M11 -> M10 -> M2

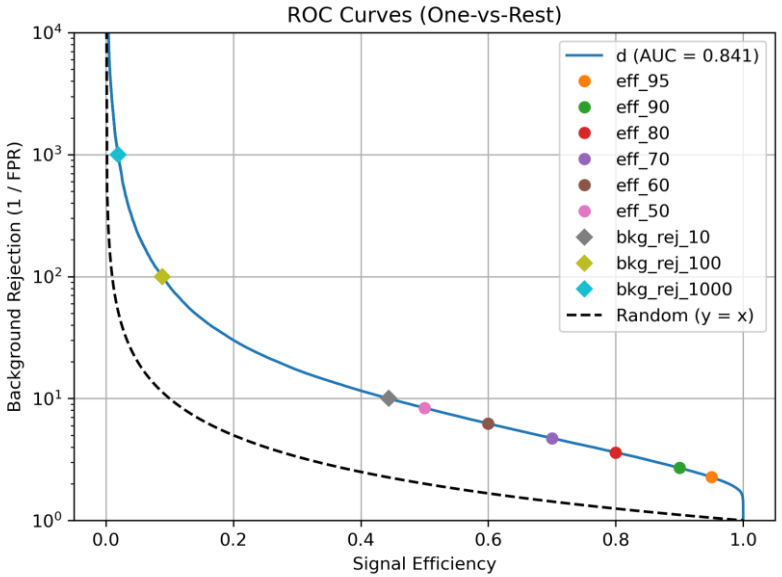
- To find the b tagging (b+bbar) WP on score

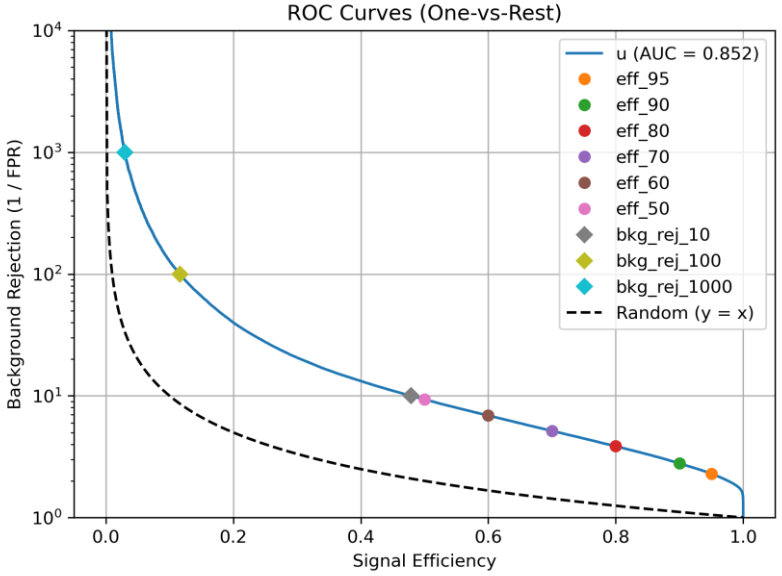
Working Point	Threshold	TPR	FPR	Bkg Rej
eff_95	0.378	0.95	0.009	106.6
eff_90	0.895	0.9	0.002	445
eff_80	0.99	0.8	0.001	1422.4
eff_70	0.996	0.7	0	2330.8
eff_60	0.997	0.6	0	3510.9
eff_50	0.998	0.5	0	4754.7
bkg_rej_10	0.015	0.987	0.1	10
bkg_rej_100	0.352	0.952	0.01	100
bkg_rej_1000	0.978	0.846	0.001	998.9

Roc curve - gsud









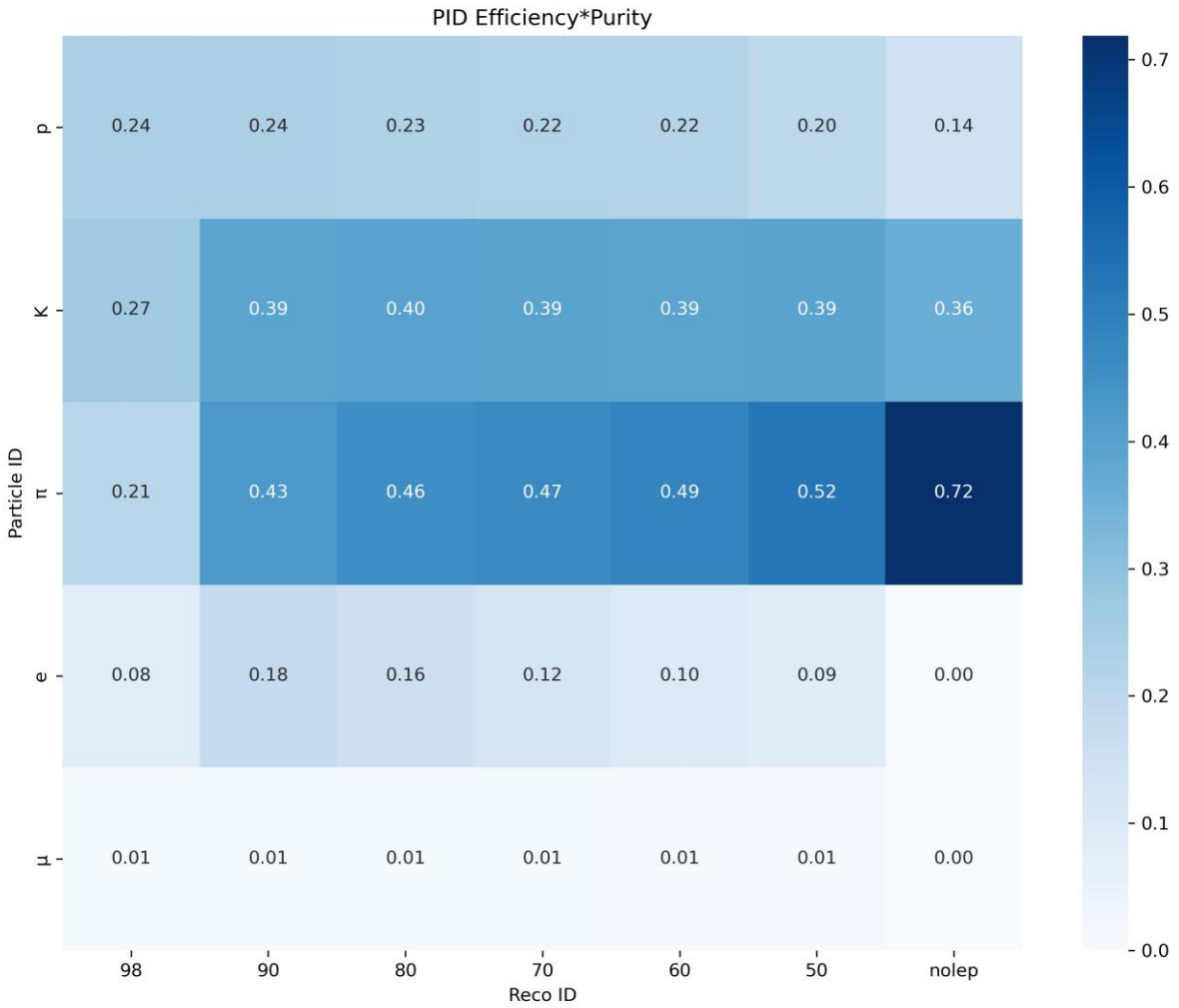
Temporary WP setting

This WP is under H decay hypothesis and assume equal ratio.
For different environment and ratios, WP shall be different.
B tagging can be ~0.3 for both eff and bkg rejection. (Also 0.5 is fine as for maximum likelihood)
For c\g\u\s\d. May use carefully.
Fpr b, Prefer Eff; For others, prefer purity?

	Threshold	Eff	Bkg Rej		Threshold	Eff	Bkg Rej
B	0.378	0.95	106.6		0.352	0.952	100
C	0.047	0.95	5.7		0.763	0.773	100
G	0.067	0.95	2.8		0.781	0.253	100
S	0.088	0.95	2.3		0.651	0.116	100
U	0.096	0.95	2.3		0.611	0.089	100
D	0.019	0.95	2.2		0.559	0.191	100

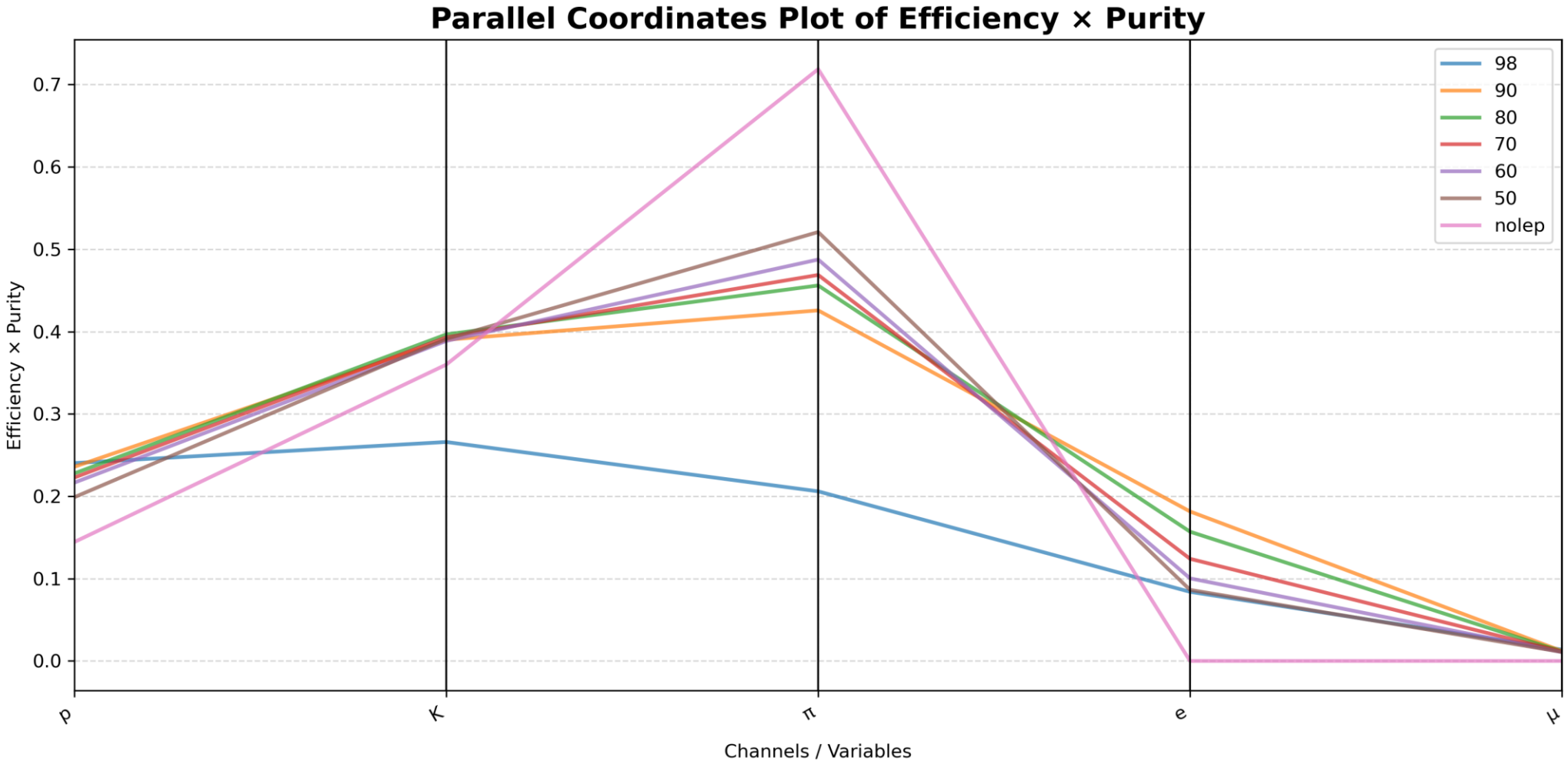
PID revisit in ZH: ubar

/hpcfs/cepc/higgsgpu/zhangkl/JOI_Fulltest/99_2_joi_E240_nnHuu.root

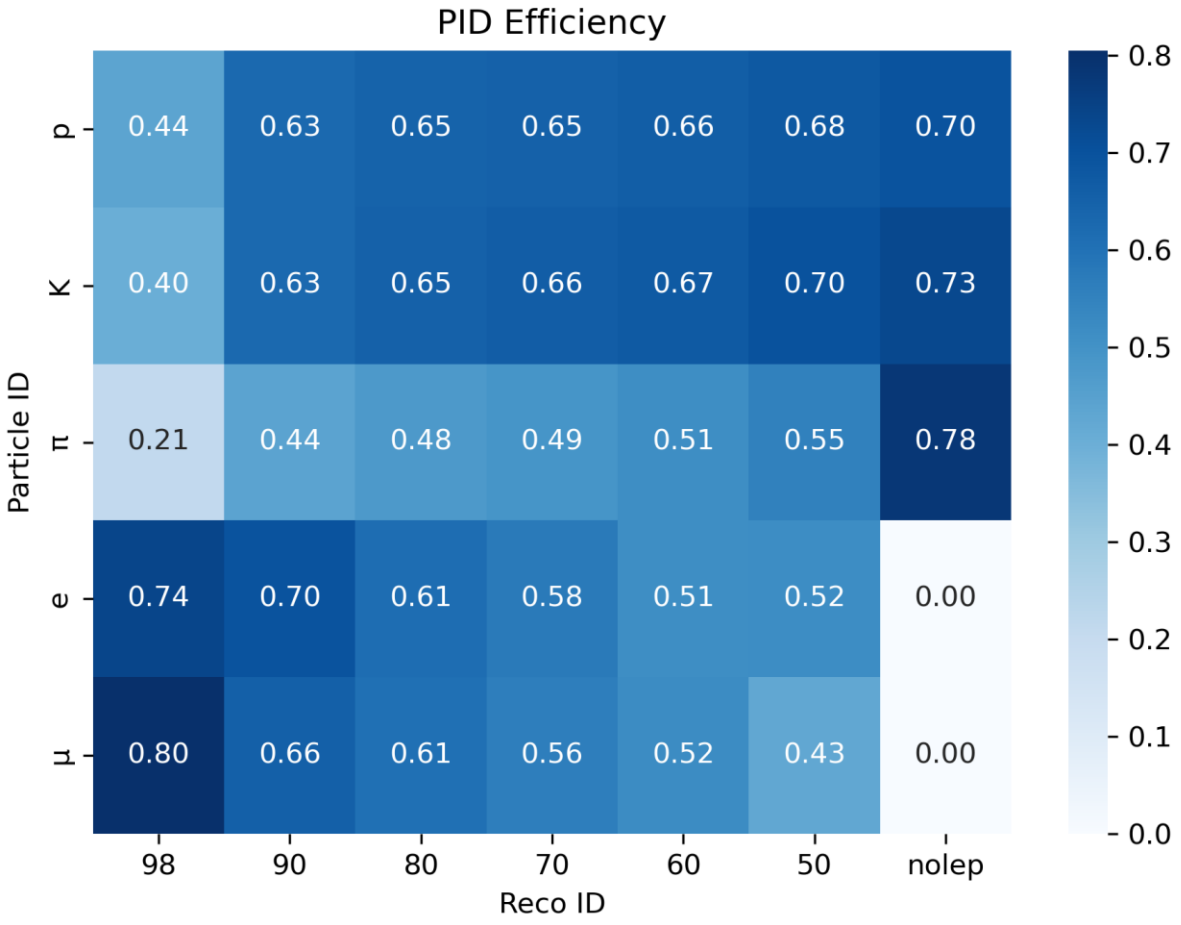


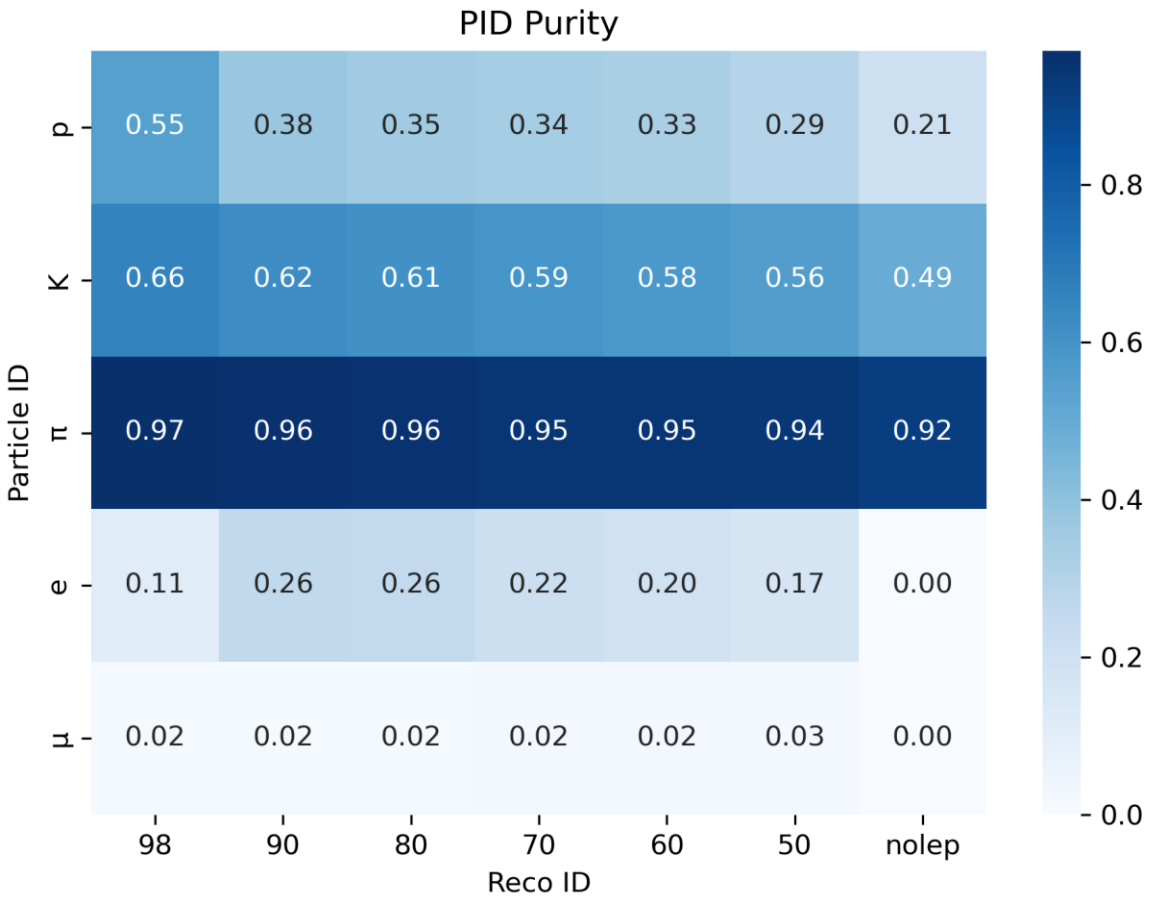
Total event:	9332
Total PFO:	293006
PFO per event:	31.4
Charged tracks:	106892
Neutral PFOs:	186114

Per channel eff*purity

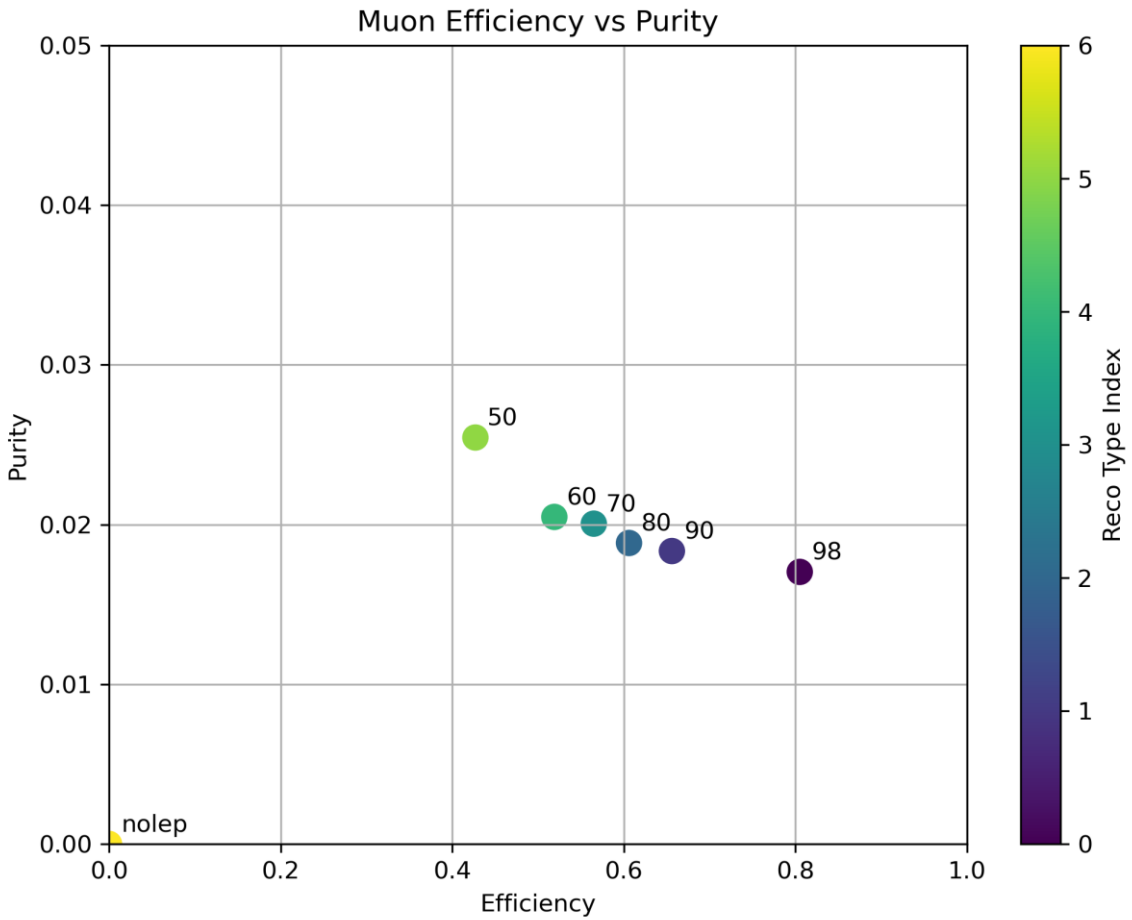


Eff and Purity

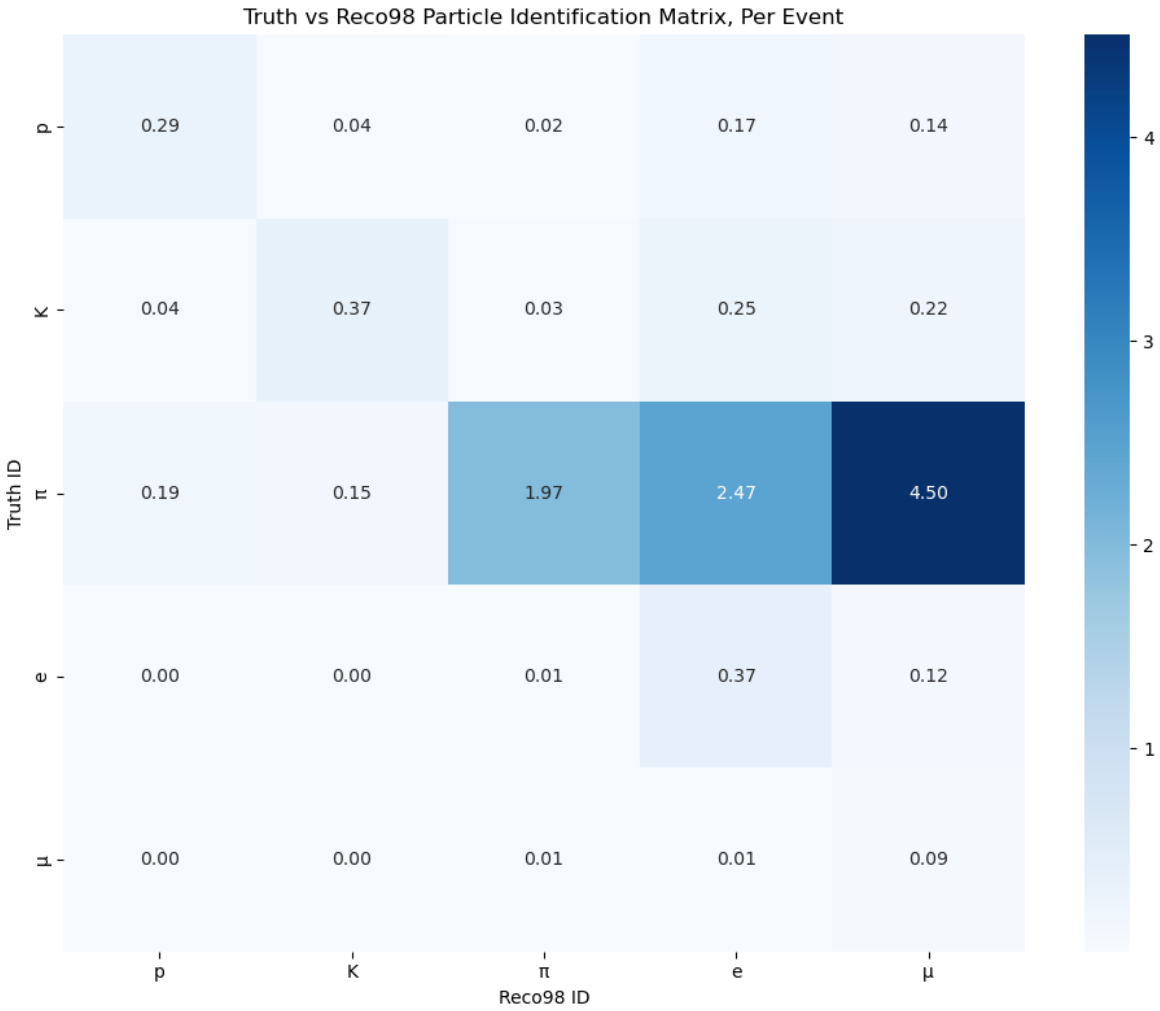


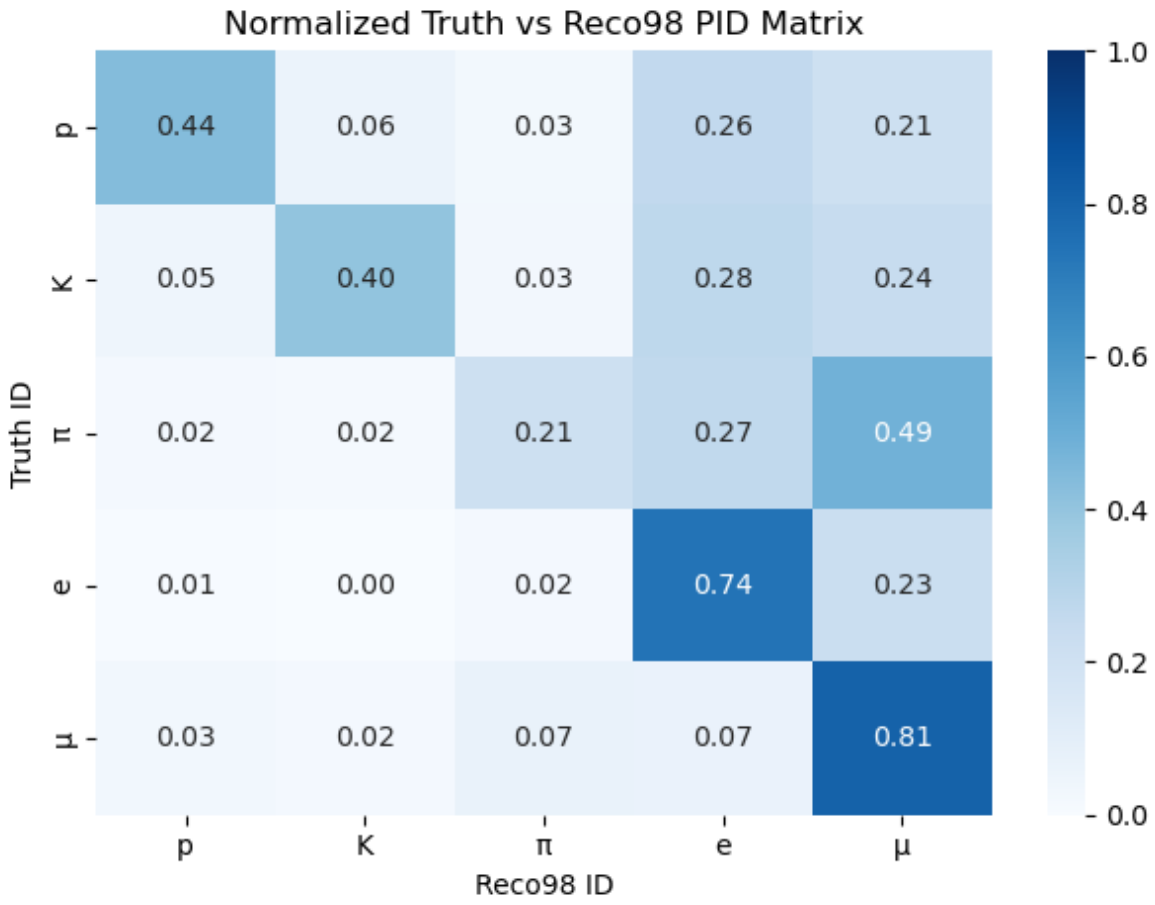


Muon WP

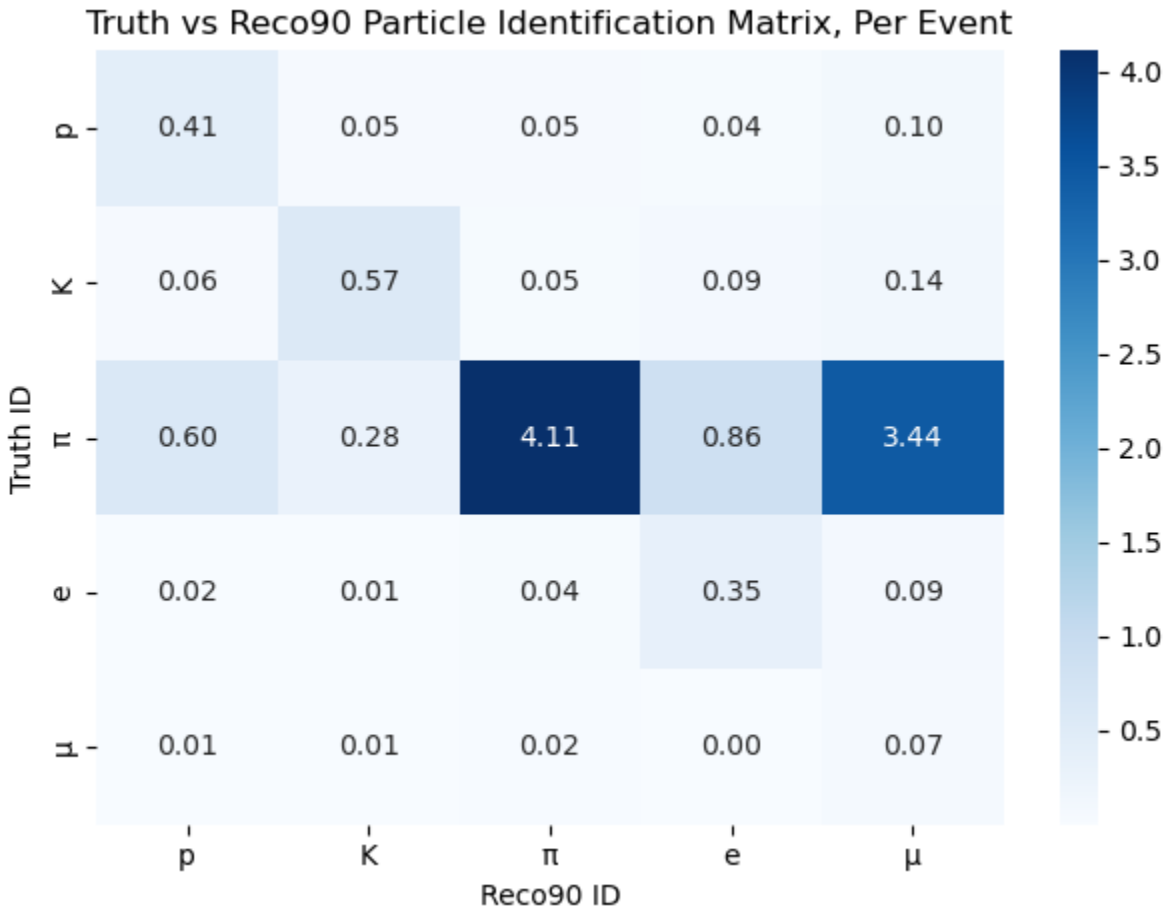

Muon Efficiency vs Purity
Purity
Efficiency
Reco Type Index
0.05
0.04
0.03
0.02
0.01
0.00
0.0
0.2
0.4
0.6
0.8
1.0
0
1
2
3
4
5
6
50
60
70
80
90
98
nolep

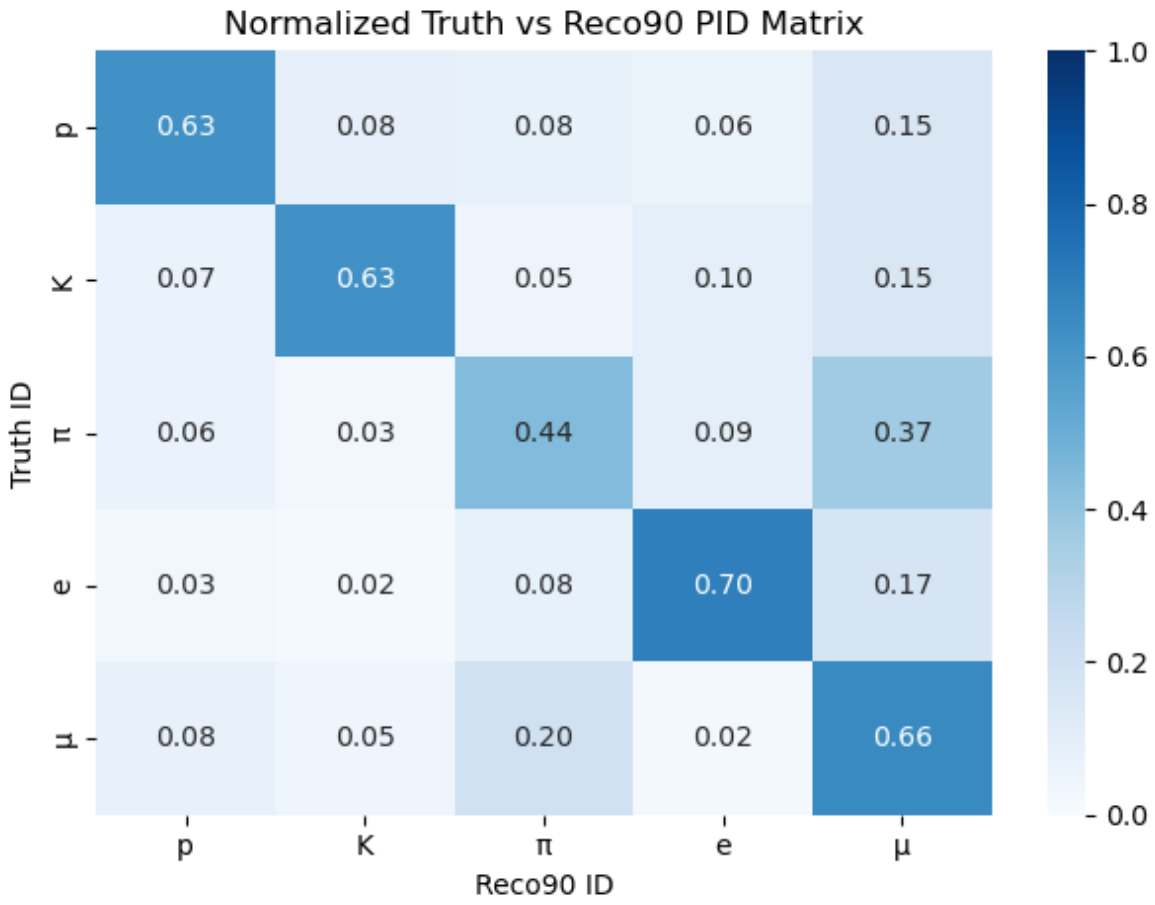
Reco98


Truth vs Reco98 Particle Identification Matrix, Per Event
Truth ID
Reco98 ID
p
K
π
e
μ
0.29 0.04 0.02 0.17 0.14
0.04 0.37 0.03 0.25 0.22
0.19 0.15 1.97 2.47 4.50
0.00 0.00 0.01 0.37 0.12
0.00 0.00 0.01 0.01 0.09
4
3
2
1

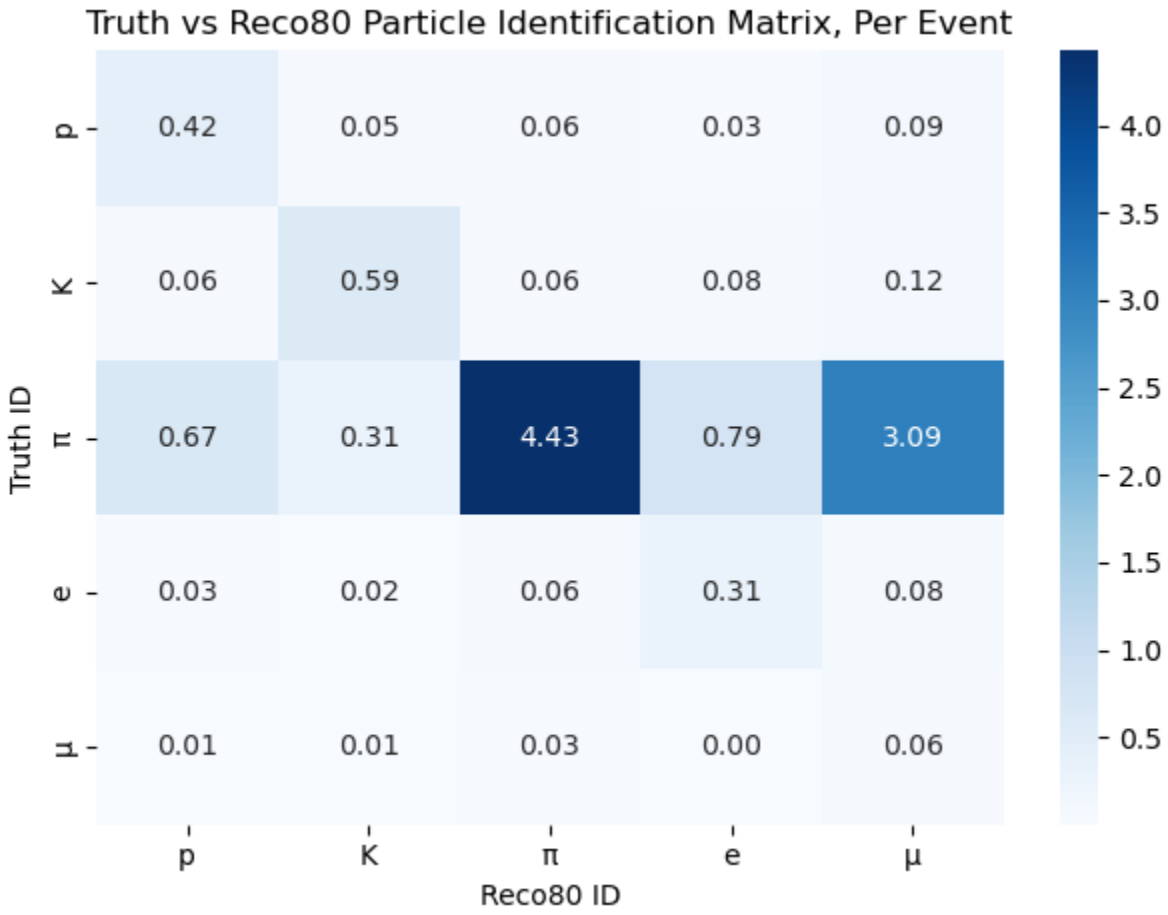

Normalized Truth vs Reco98 PID Matrix
Truth ID
Reco98 ID
p
K
π
e
μ
0.44 0.06 0.03 0.26 0.21
0.05 0.40 0.03 0.28 0.24
0.02 0.02 0.21 0.27 0.49
0.01 0.00 0.02 0.74 0.23
0.03 0.02 0.07 0.07 0.81
1.0
0.8
0.6
0.4
0.2
0.0

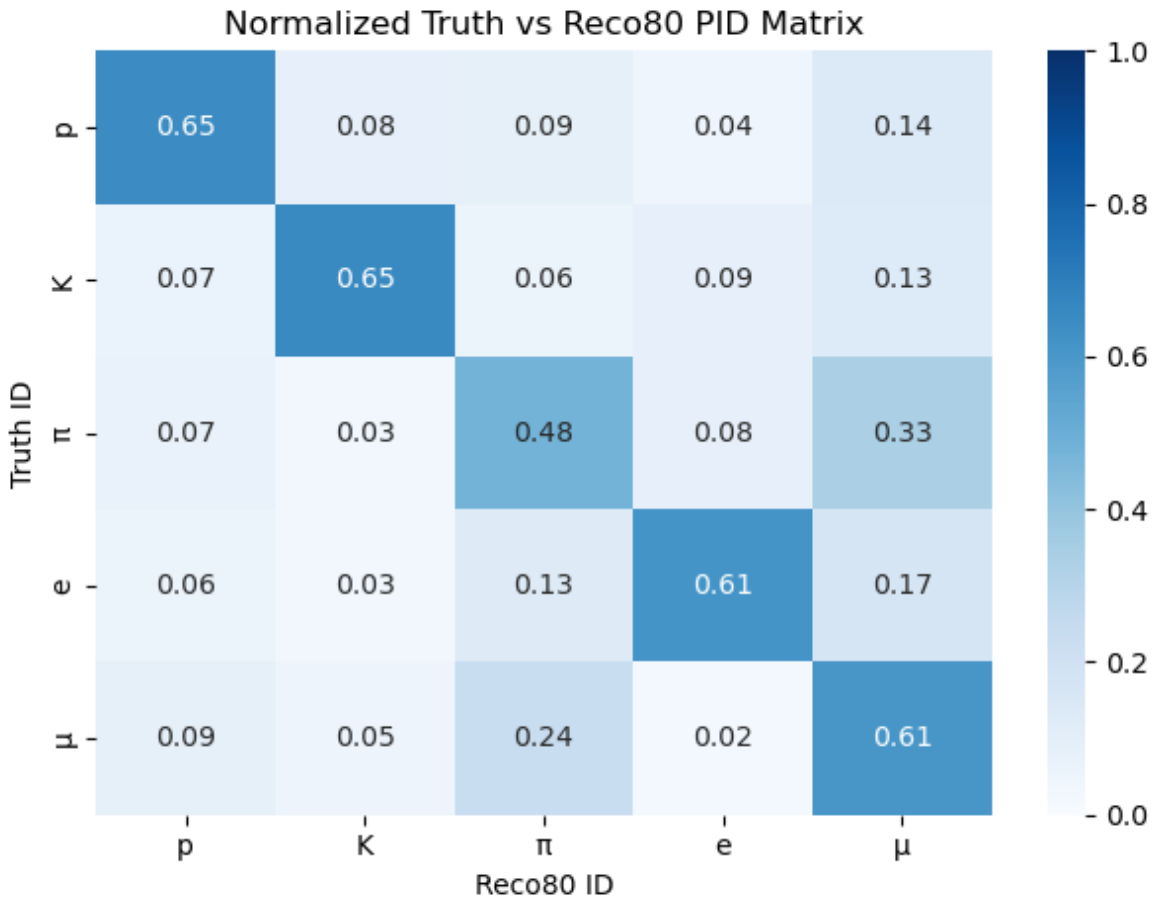
Reco90


Truth vs Reco90 Particle Identification Matrix, Per Event
Truth ID
Reco90 ID
p
K
π
e
μ
0.41 0.05 0.05 0.04 0.10
0.06 0.57 0.05 0.09 0.14
0.60 0.28 4.11 0.86 3.44
0.02 0.01 0.04 0.35 0.09
0.01 0.01 0.02 0.00 0.07
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5

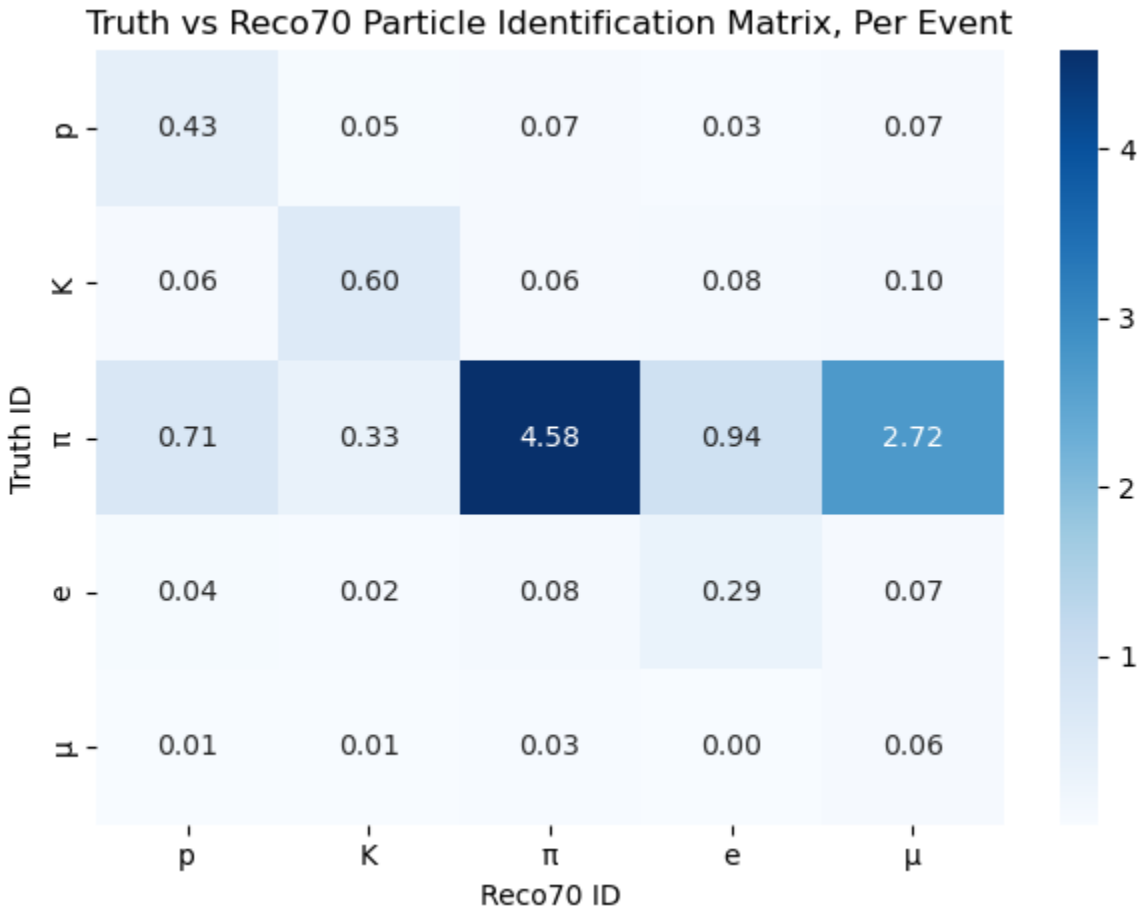

Normalized Truth vs Reco90 PID Matrix
Truth ID
Reco90 ID
p
K
π
e
μ
0.63 0.08 0.08 0.06 0.15
0.07 0.63 0.05 0.10 0.15
0.06 0.03 0.44 0.09 0.37
0.03 0.02 0.08 0.70 0.17
0.08 0.05 0.20 0.02 0.66
1.0
0.8
0.6
0.4
0.2
0.0

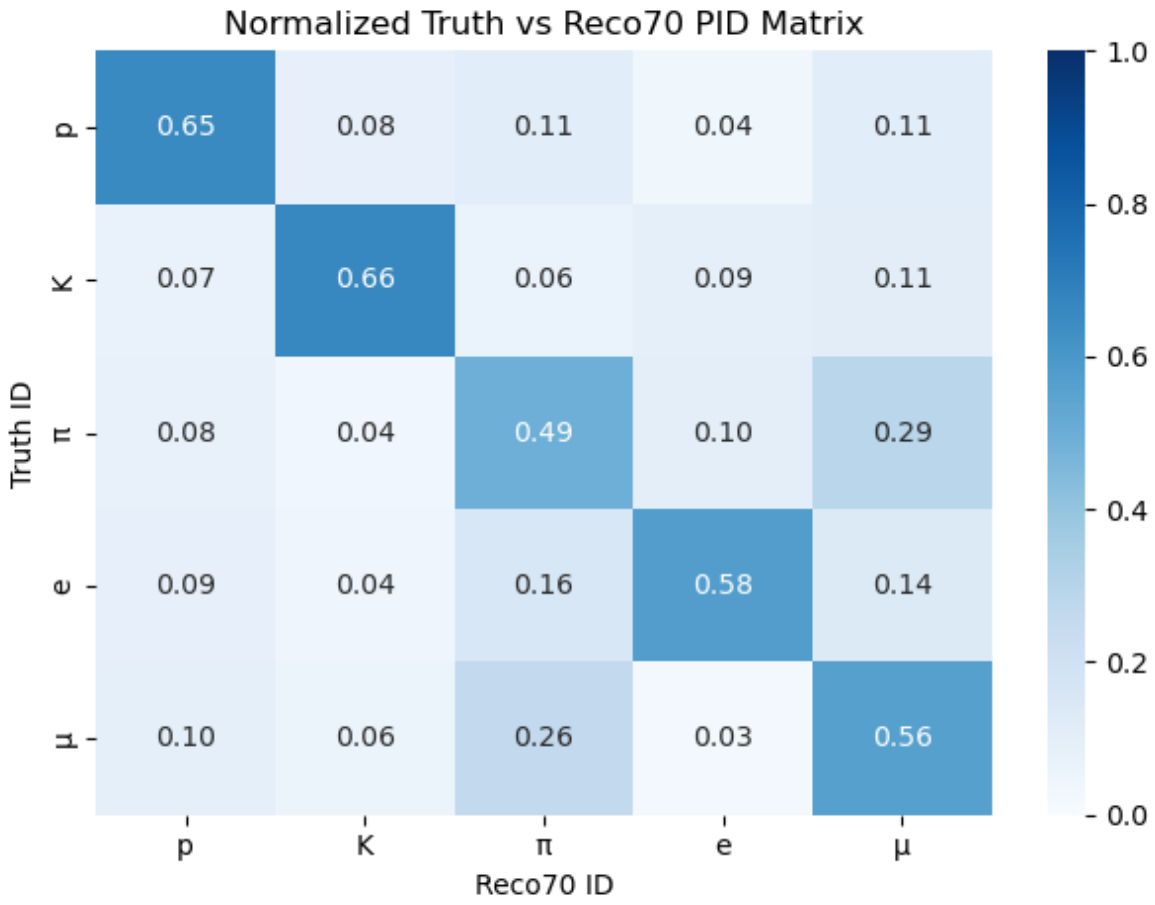
Reco80


Truth vs Reco80 Particle Identification Matrix, Per Event
Truth ID
Reco80 ID
p
K
π
e
μ
0.42 0.05 0.06 0.03 0.09
0.06 0.59 0.06 0.08 0.12
0.67 0.31 4.43 0.79 3.09
0.03 0.02 0.06 0.31 0.08
0.01 0.01 0.03 0.00 0.06
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5

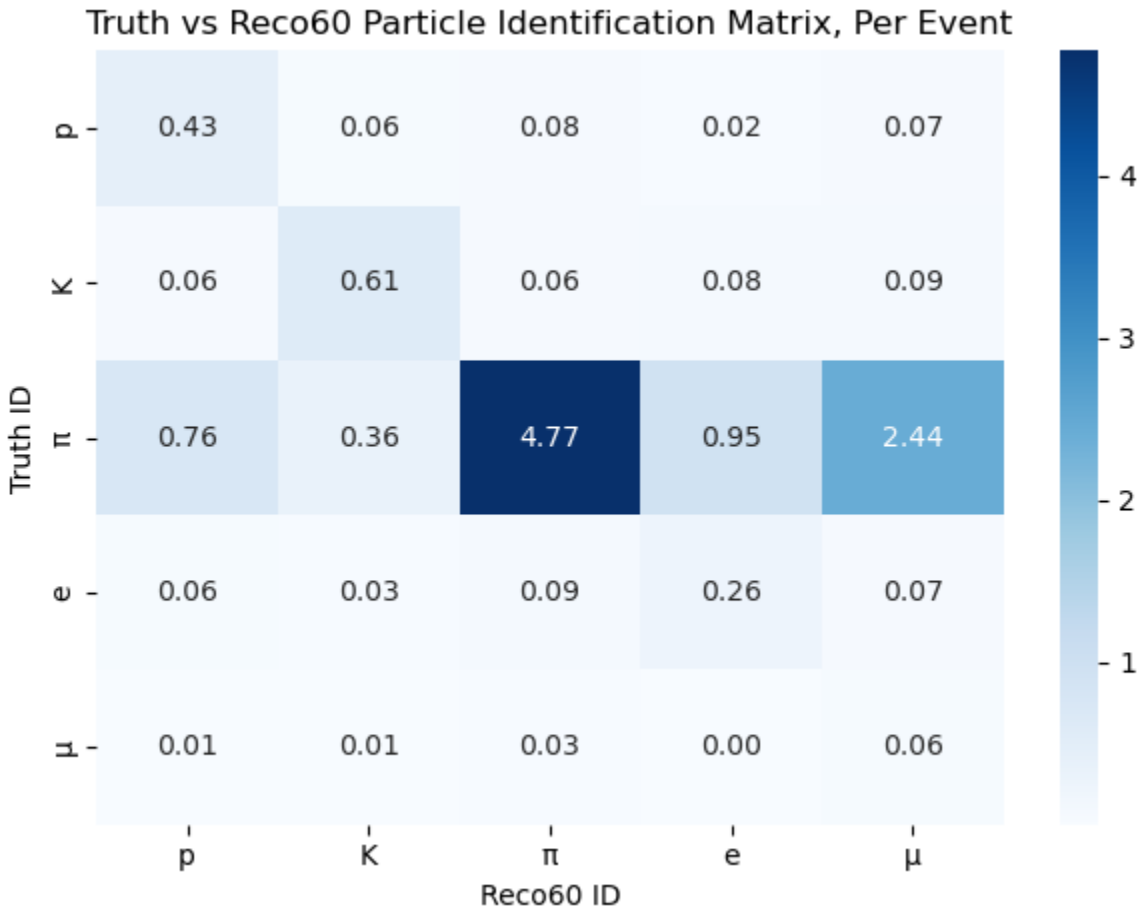

Normalized Truth vs Reco80 PID Matrix
Truth ID
Reco80 ID
p
K
π
e
μ
0.65 0.08 0.09 0.04 0.14
0.07 0.65 0.06 0.09 0.13
0.07 0.03 0.48 0.08 0.33
0.06 0.03 0.13 0.61 0.17
0.09 0.05 0.24 0.02 0.61
1.0
0.8
0.6
0.4
0.2
0.0

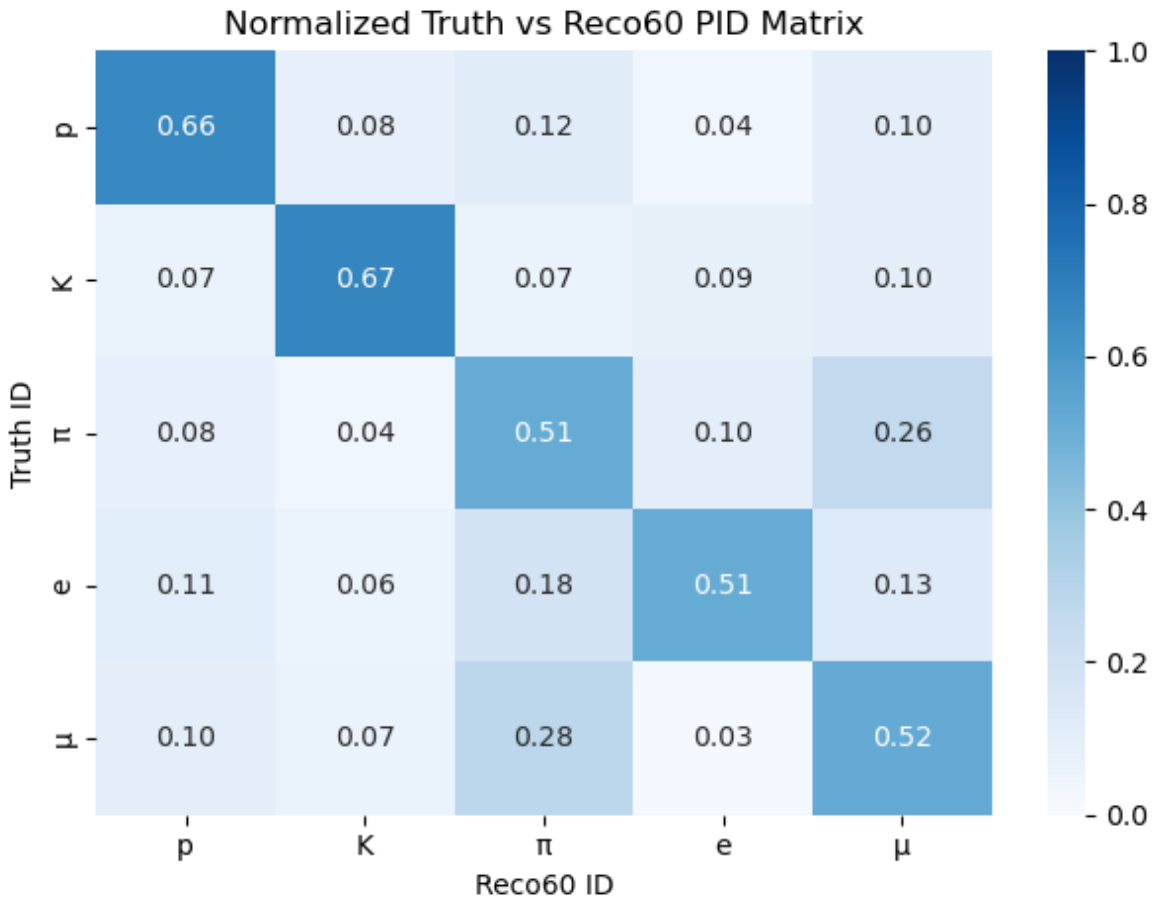
Reco70


Truth vs Reco70 Particle Identification Matrix, Per Event
Truth ID
Reco70 ID
p
K
π
e
μ
0.43 0.05 0.07 0.03 0.07
0.06 0.60 0.06 0.08 0.10
0.71 0.33 4.58 0.94 2.72
0.04 0.02 0.08 0.29 0.07
0.01 0.01 0.03 0.00 0.06
4
3
2
1

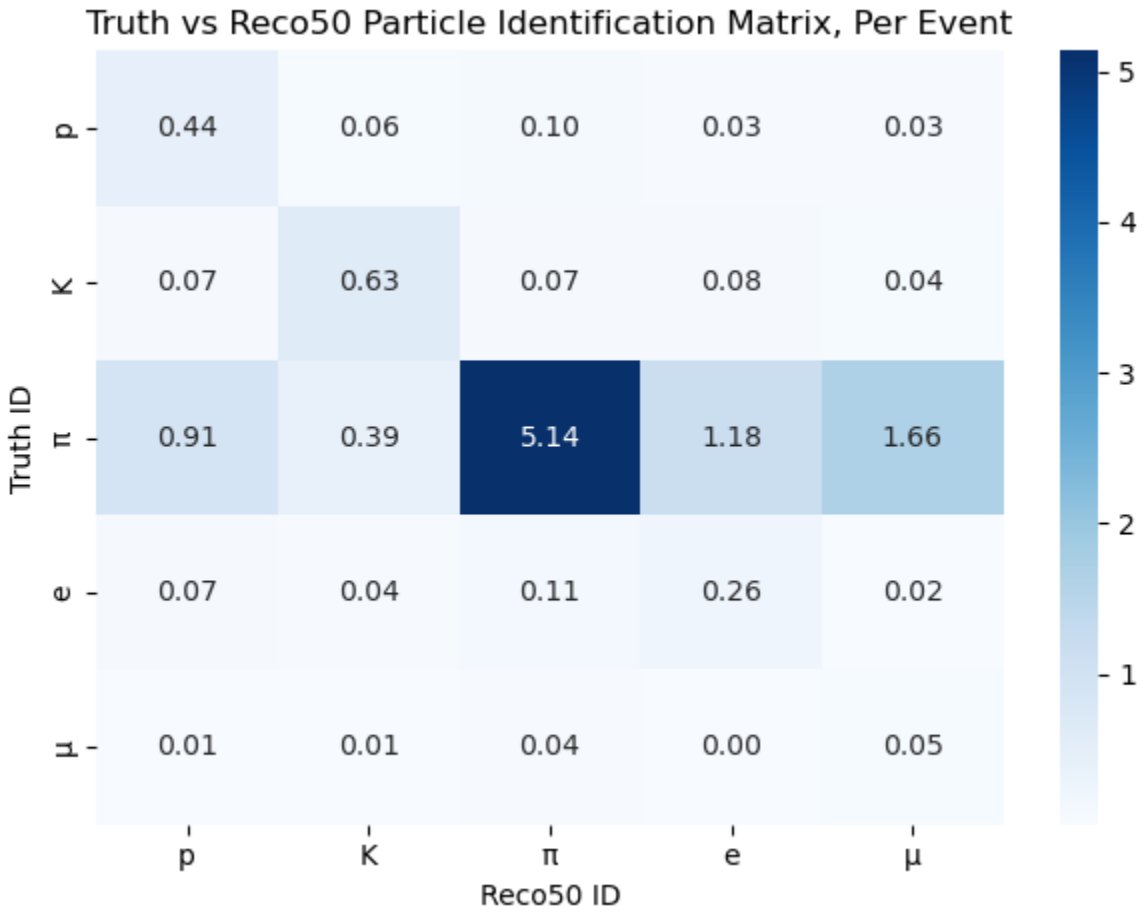

Normalized Truth vs Reco70 PID Matrix
Truth ID
Reco70 ID
p
K
π
e
μ
0.65 0.08 0.11 0.04 0.11
0.07 0.66 0.06 0.09 0.11
0.08 0.04 0.49 0.10 0.29
0.09 0.04 0.16 0.58 0.14
0.10 0.06 0.26 0.03 0.56
1.0
0.8
0.6
0.4
0.2
0.0

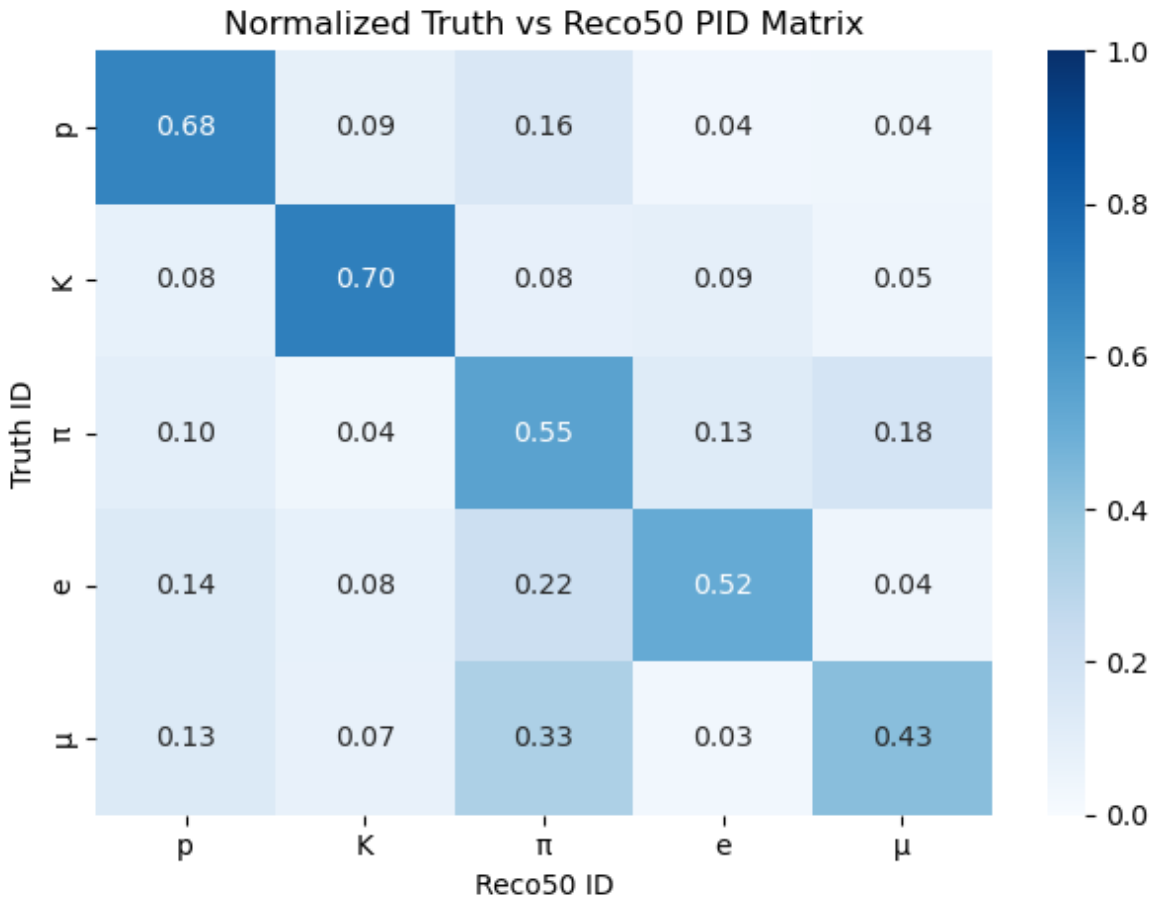
Reco60


Truth vs Reco60 Particle Identification Matrix, Per Event
Truth ID
Reco60 ID
p
K
π
e
μ
0.43 0.06 0.08 0.02 0.07
0.06 0.61 0.06 0.08 0.09
0.76 0.36 4.77 0.95 2.44
0.06 0.03 0.09 0.26 0.07
0.01 0.01 0.03 0.00 0.06
4
3
2
1

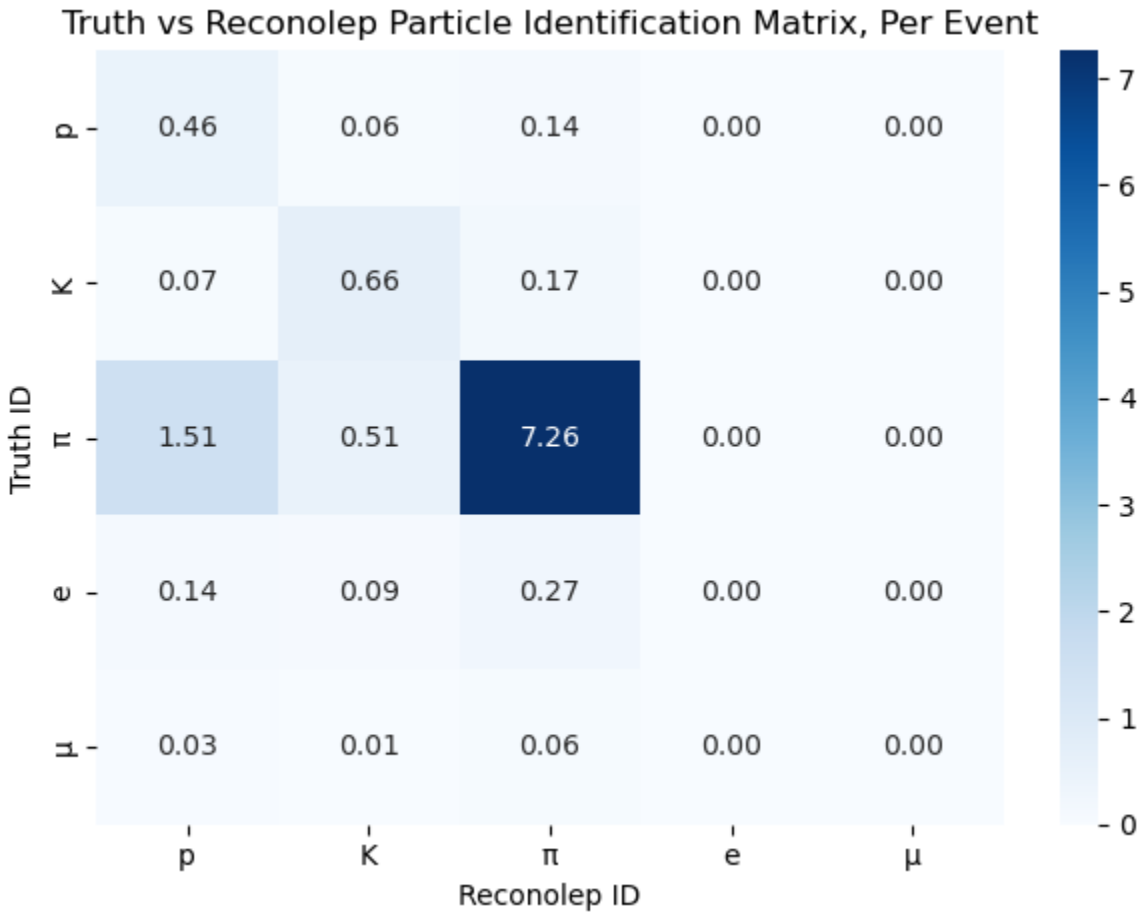

Normalized Truth vs Reco60 PID Matrix
Truth ID
Reco60 ID
p
K
π
e
μ
0.66 0.08 0.12 0.04 0.10
0.07 0.67 0.07 0.09 0.10
0.08 0.04 0.51 0.10 0.26
0.11 0.06 0.18 0.51 0.13
0.10 0.07 0.28 0.03 0.52
1.0
0.8
0.6
0.4
0.2
0.0

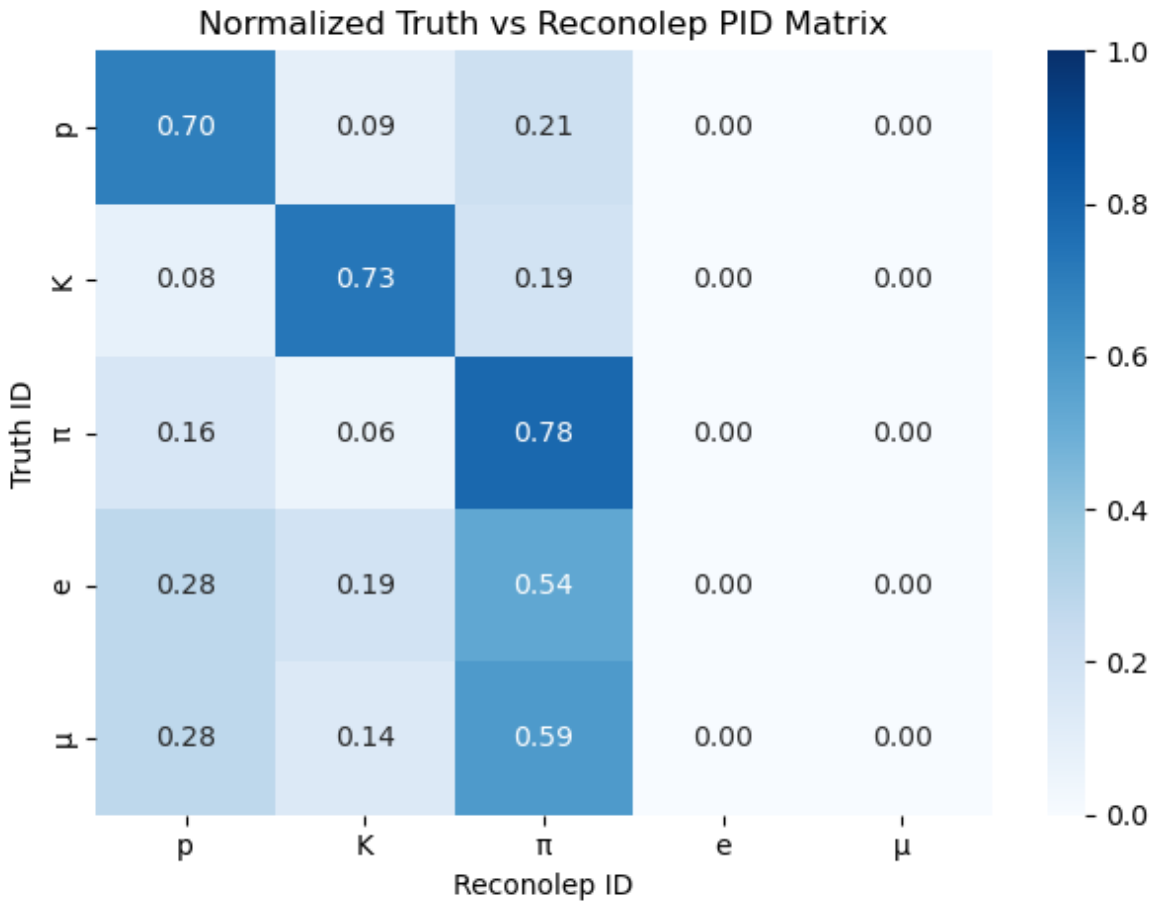
Reco50


Truth vs Reco50 Particle Identification Matrix, Per Event
Truth ID
Reco50 ID
p
K
π
e
μ
0.44 0.06 0.10 0.03 0.03
0.07 0.63 0.07 0.08 0.04
0.91 0.39 5.14 1.18 1.66
0.07 0.04 0.11 0.26 0.02
0.01 0.01 0.04 0.00 0.05
5
4
3
2
1


Normalized Truth vs Reco50 PID Matrix
Truth ID
Reco50 ID
p
K
π
e
μ
0.68 0.09 0.16 0.04 0.04
0.08 0.70 0.08 0.09 0.05
0.10 0.04 0.55 0.13 0.18
0.14 0.08 0.22 0.52 0.04
0.13 0.07 0.33 0.03 0.43
1.0
0.8
0.6
0.4
0.2
0.0

Reco Nolep


Truth vs Reconolep Particle Identification Matrix, Per Event
Truth ID
Reconolep ID
p
K
π
e
μ
0.46 0.06 0.14 0.00 0.00
0.07 0.66 0.17 0.00 0.00
1.51 0.51 7.26 0.00 0.00
0.14 0.09 0.27 0.00 0.00
0.03 0.01 0.06 0.00 0.00
0
1
2
3
4
5
6
7

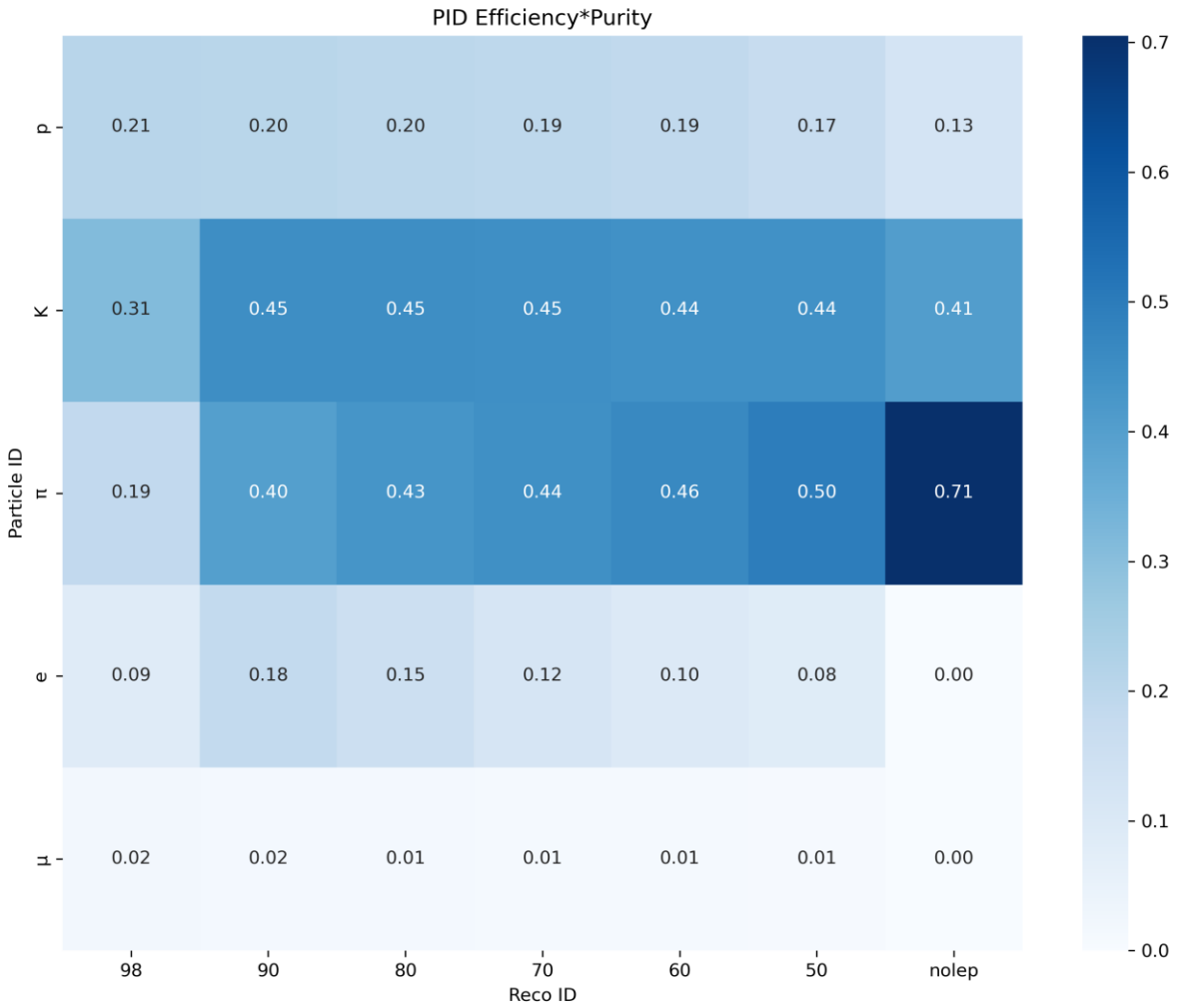

Normalized Truth vs Reconolep PID Matrix
Truth ID
Reconolep ID
p
K
π
e
μ
0.70 0.09 0.21 0.00 0.00
0.08 0.73 0.19 0.00 0.00
0.16 0.06 0.78 0.00 0.00
0.28 0.19 0.54 0.00 0.00
0.28 0.14 0.59 0.00 0.00
0.0
0.2
0.4
0.6
0.8
1.0

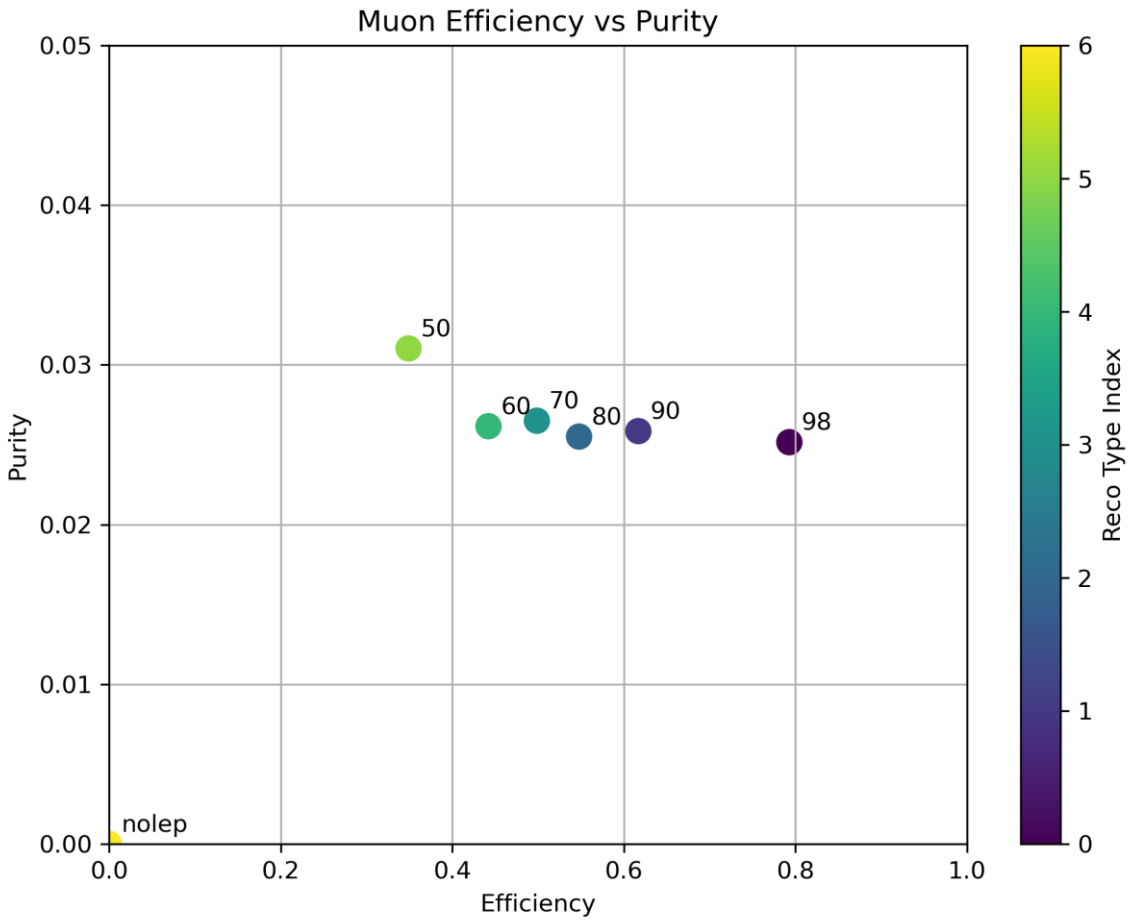
On ZH jet:

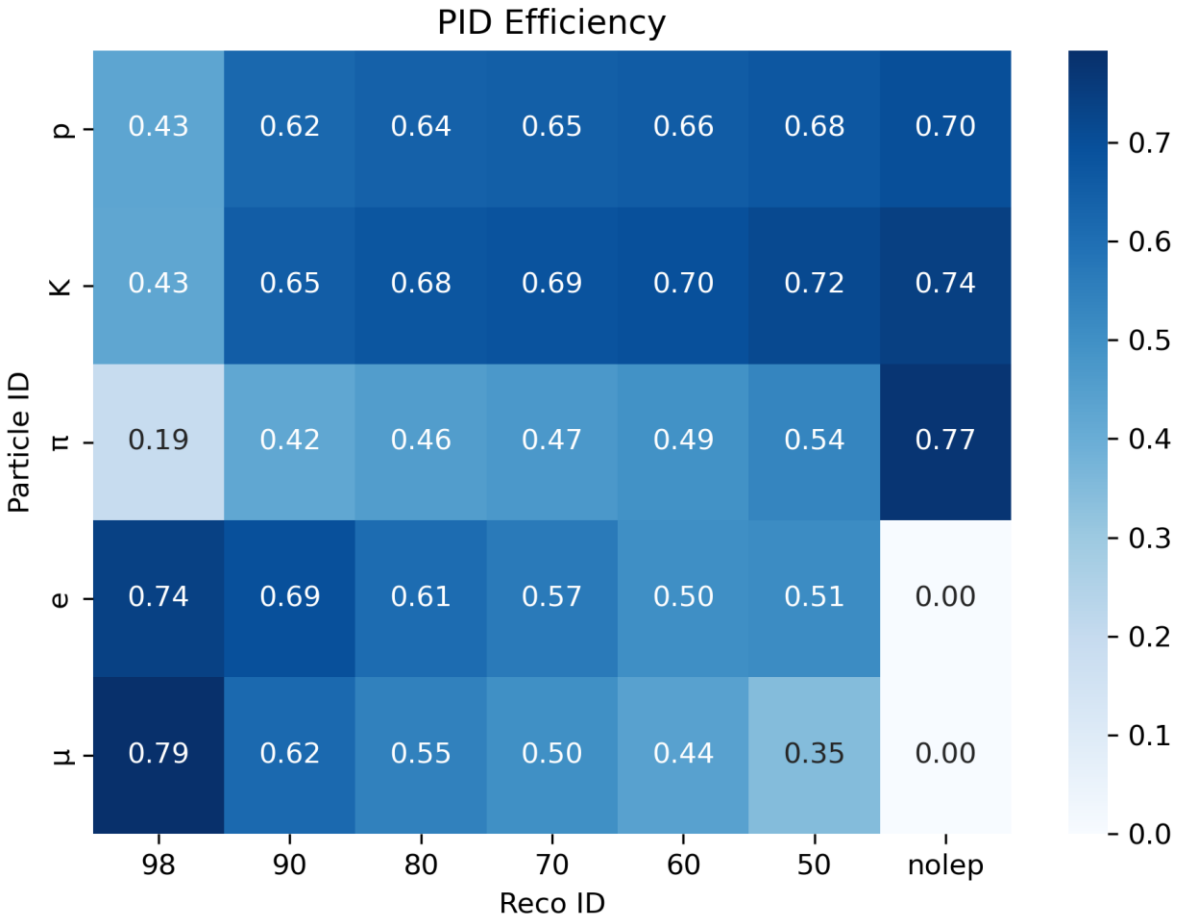
/hpcfs/cepc/higgsgpu/zhangkl/JOI_Fulltest/99*.root

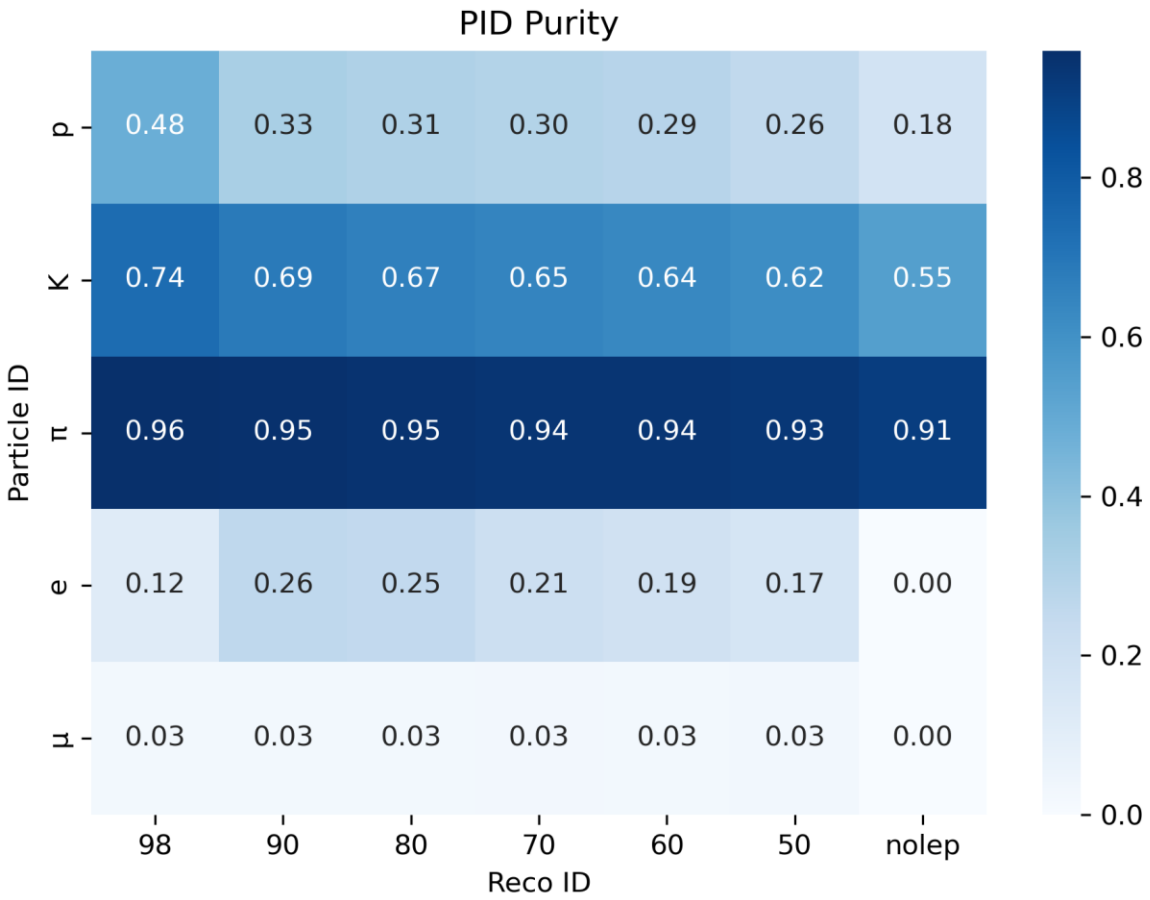
120000 jets

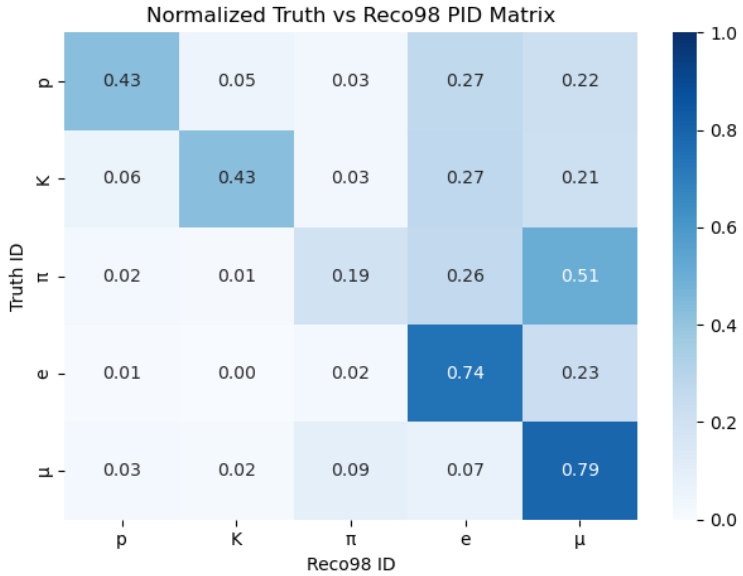
b/c/g/u/d/s

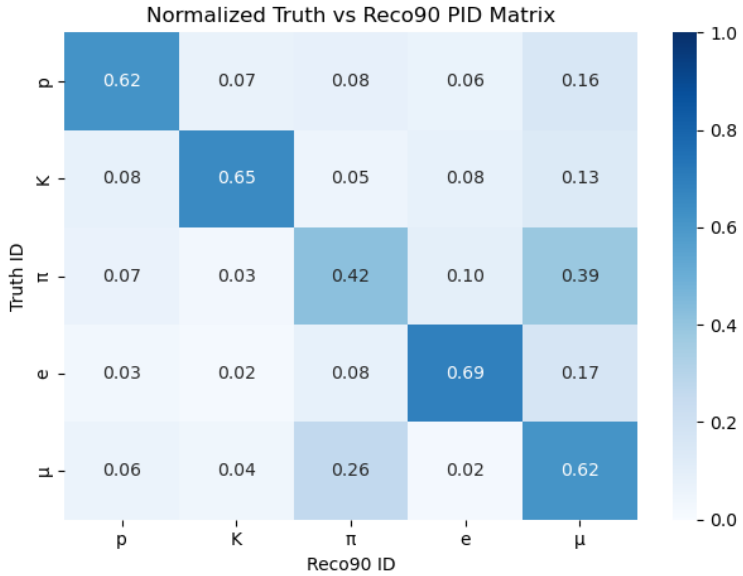

PID Efficiency*Purity
Particle ID
p
K
π
e
μ
0.21 0.20 0.20 0.19 0.19 0.17 0.13
0.31 0.45 0.45 0.45 0.44 0.44 0.41
0.19 0.40 0.43 0.44 0.46 0.50 0.71
0.09 0.18 0.15 0.12 0.10 0.08 0.00
0.02 0.02 0.01 0.01 0.01 0.01 0.00
98 90 80 70 60 50 nolep
Reco ID
0.7 0.6 0.5 0.4 0.3 0.2 0.1 0.0

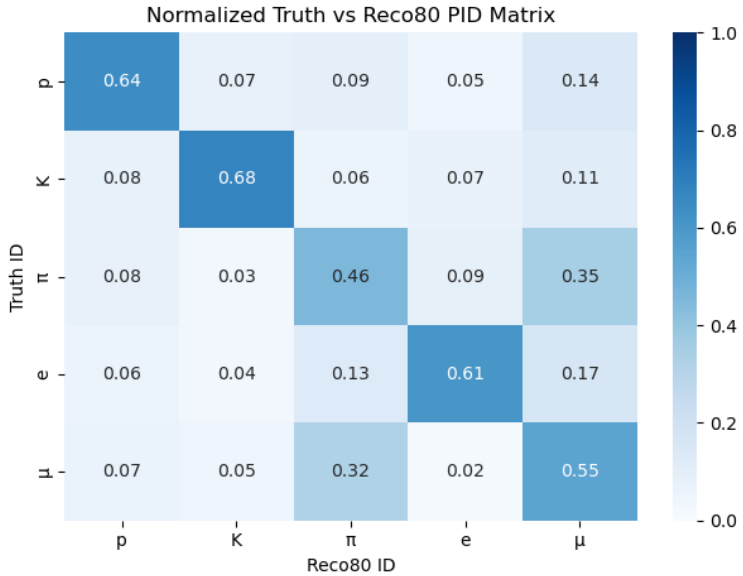

Muon Efficiency vs Purity
Purity
Efficiency
Reco Type Index
0.05
0.04
0.03
0.02
0.01
0.00
0.0
0.2
0.4
0.6
0.8
1.0
0
1
2
3
4
5
6
50
60
70
80
90
98
nolep

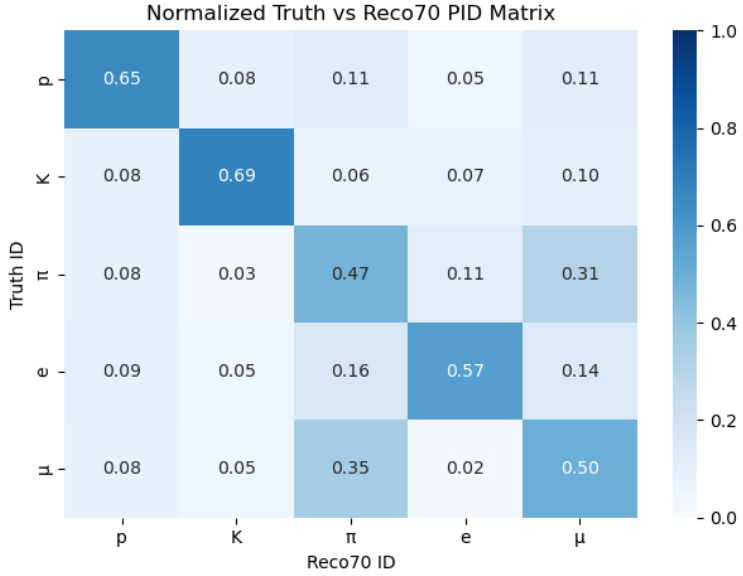

PID Efficiency
Particle ID
p
K
π
e
μ
0.43 0.62 0.64 0.65 0.66 0.68 0.70
0.43 0.65 0.68 0.69 0.70 0.72 0.74
0.19 0.42 0.46 0.47 0.49 0.54 0.77
0.74 0.69 0.61 0.57 0.50 0.51 0.00
0.79 0.62 0.55 0.50 0.44 0.35 0.00
98 90 80 70 60 50 nolep
Reco ID
0.7 0.6 0.5 0.4 0.3 0.2 0.1 0.0

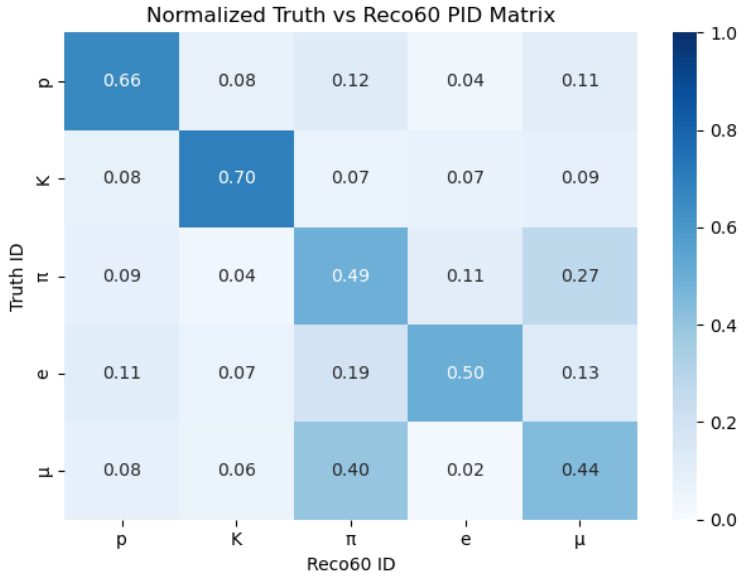

PID Purity
Particle ID
p
K
π
e
μ
0.48 0.33 0.31 0.30 0.29 0.26 0.18
0.74 0.69 0.67 0.65 0.64 0.62 0.55
0.96 0.95 0.95 0.94 0.94 0.93 0.91
0.12 0.26 0.25 0.21 0.19 0.17 0.00
0.03 0.03 0.03 0.03 0.03 0.03 0.00
98 90 80 70 60 50 nolep
Reco ID
0.8 0.6 0.4 0.2 0.0

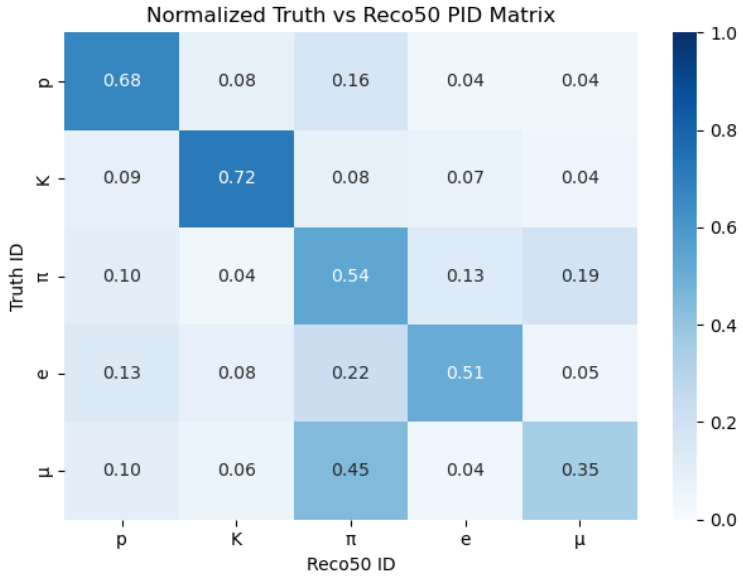

Normalized Truth vs Reco98 PID Matrix
Truth ID
Reco98 ID
p K π e μ
0.43 0.05 0.03 0.27 0.22
0.06 0.43 0.03 0.27 0.21
0.02 0.01 0.19 0.26 0.51
0.01 0.00 0.02 0.74 0.23
0.03 0.02 0.09 0.07 0.79
1.0 0.8 0.6 0.4 0.2 0.0

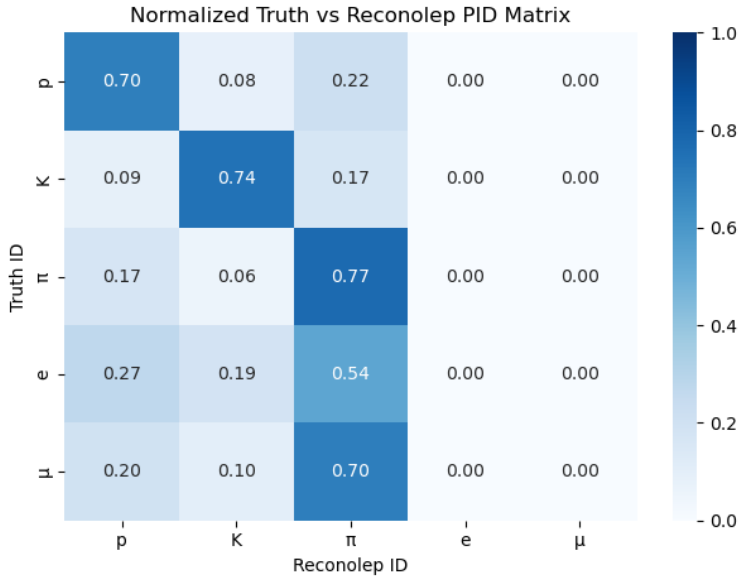

Normalized Truth vs Reco90 PID Matrix
Truth ID
Reco90 ID
p K π e μ
0.62 0.07 0.08 0.06 0.16
0.08 0.65 0.05 0.08 0.13
0.07 0.03 0.42 0.10 0.39
0.03 0.02 0.08 0.69 0.17
0.06 0.04 0.26 0.02 0.62
1.0 0.8 0.6 0.4 0.2 0.0


Normalized Truth vs Reco80 PID Matrix
Truth ID
Reco80 ID
p K π e μ
0.64 0.07 0.09 0.05 0.14
0.08 0.68 0.06 0.07 0.11
0.08 0.03 0.46 0.09 0.35
0.06 0.04 0.13 0.61 0.17
0.07 0.05 0.32 0.02 0.55
1.0 0.8 0.6 0.4 0.2 0.0


Normalized Truth vs Reco70 PID Matrix
Truth ID
Reco70 ID
p K π e μ
0.65 0.08 0.11 0.05 0.11
0.08 0.69 0.06 0.07 0.10
0.08 0.03 0.47 0.11 0.31
0.09 0.05 0.16 0.57 0.14
0.08 0.05 0.35 0.02 0.50
1.0 0.8 0.6 0.4 0.2 0.0


Normalized Truth vs Reco60 PID Matrix
Truth ID
Reco60 ID
p
K
π
e
μ
0.66 0.08 0.12 0.04 0.11
0.08 0.70 0.07 0.07 0.09
0.09 0.04 0.49 0.11 0.27
0.11 0.07 0.19 0.50 0.13
0.08 0.06 0.40 0.02 0.44
1.0
0.8
0.6
0.4
0.2
0.0


Normalized Truth vs Reco50 PID Matrix
Truth ID
Reco50 ID
p
K
π
e
μ
0.68 0.08 0.16 0.04 0.04
0.09 0.72 0.08 0.07 0.04
0.10 0.04 0.54 0.13 0.19
0.13 0.08 0.22 0.51 0.05
0.10 0.06 0.45 0.04 0.35
1.0
0.8
0.6
0.4
0.2
0.0


Normalized Truth vs Reconolep PID Matrix
Truth ID
Reconolep ID
p
K
π
e
μ
0.70 0.08 0.22 0.00 0.00
0.09 0.74 0.17 0.00 0.00
0.17 0.06 0.77 0.00 0.00
0.27 0.19 0.54 0.00 0.00
0.20 0.10 0.70 0.00 0.00
1.0
0.8
0.6
0.4
0.2
0.0

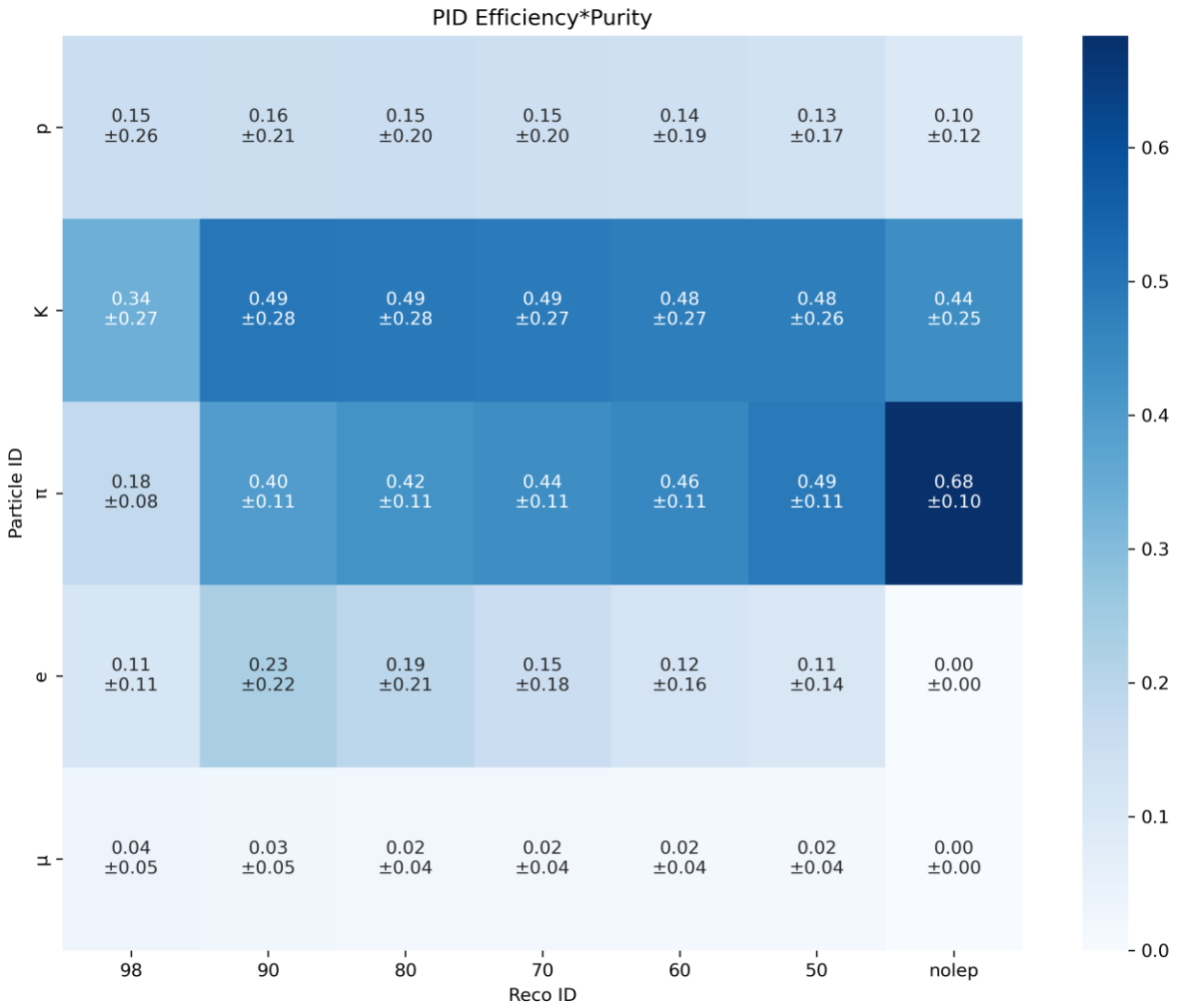
On bb jet:

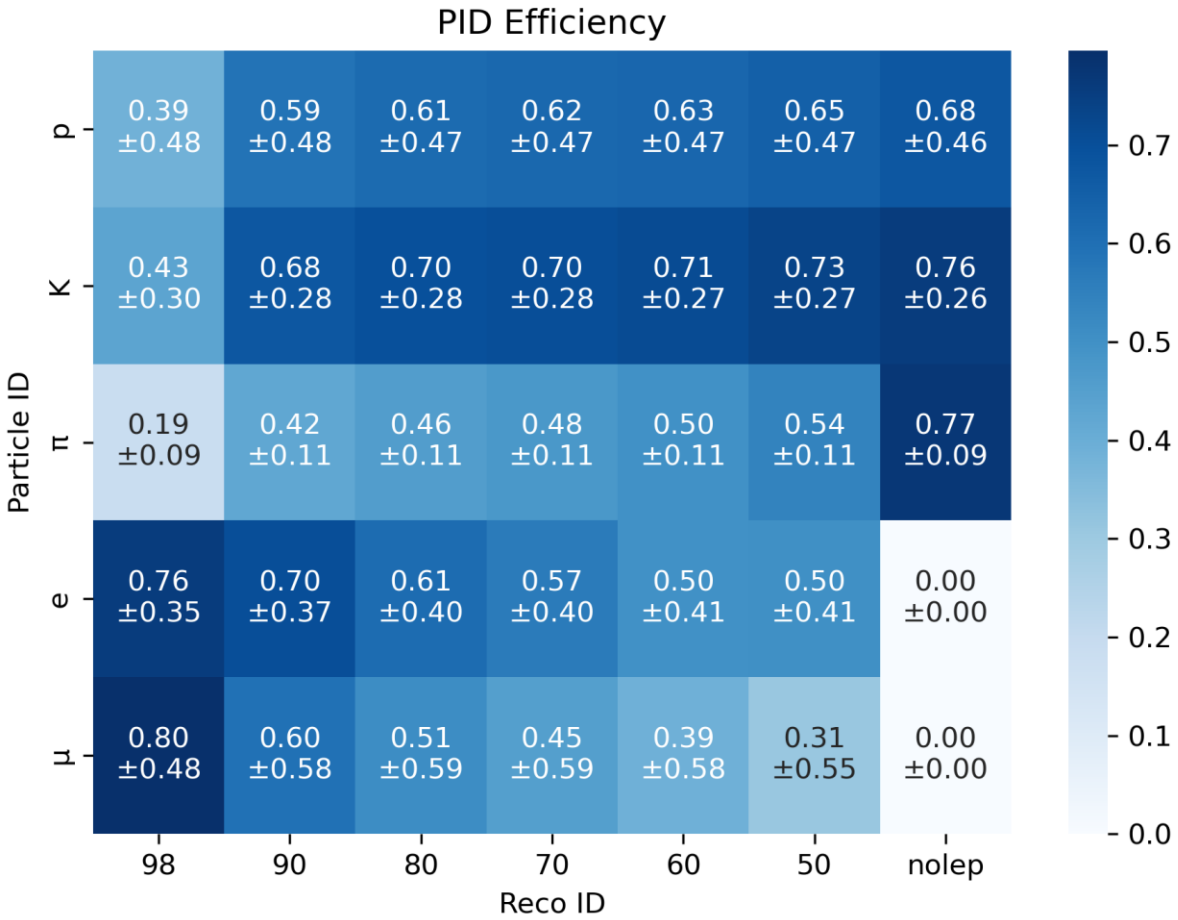
/hpcfs/cepc/higgsgpu/zhangkl/JOI_Fulltest/99*bb.root

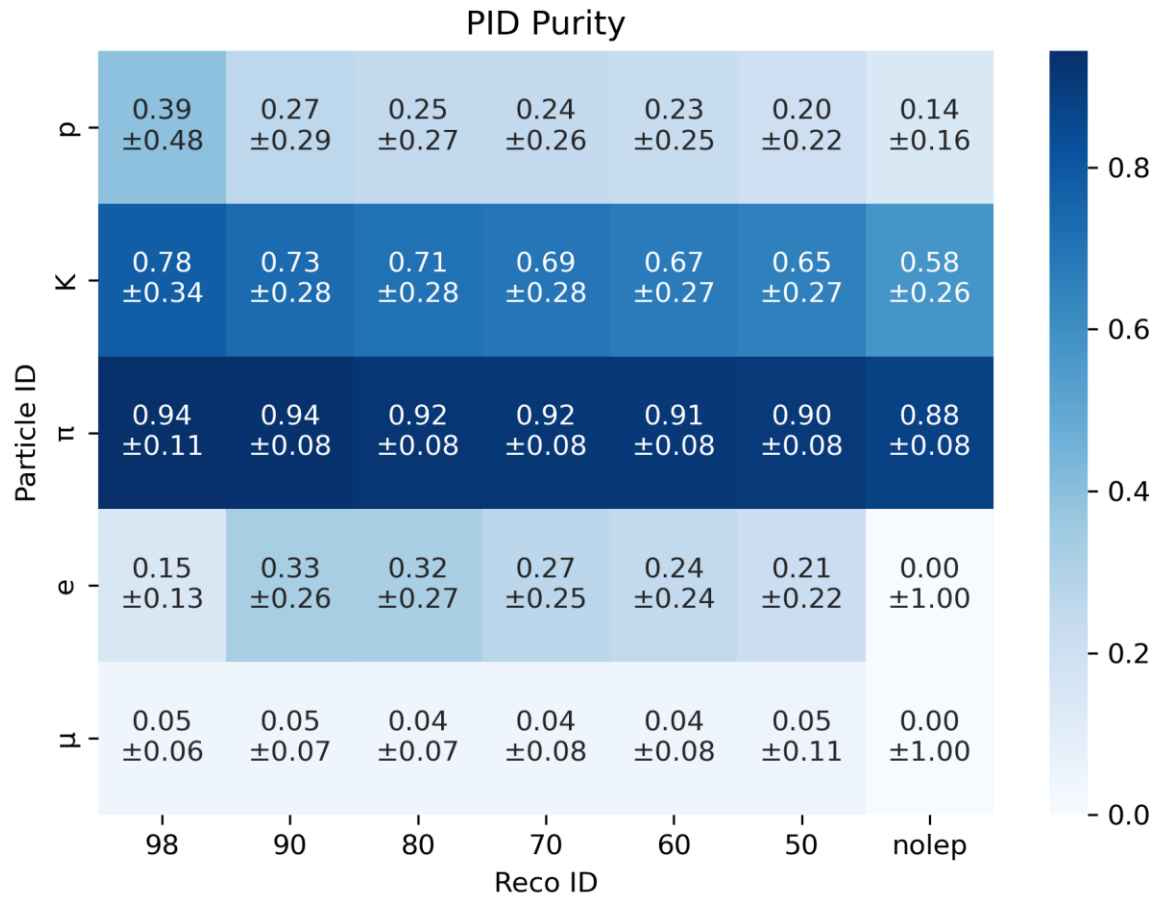
20000 jets

-> Dominant in Hqq environment

With uncertainty




PID Efficiency
Particle ID
p
K
π
e
μ
0.39 ±0.48
0.59 ±0.48
0.61 ±0.47
0.62 ±0.47
0.63 ±0.47
0.65 ±0.47
0.68 ±0.46
0.43 ±0.30
0.68 ±0.28
0.70 ±0.28
0.70 ±0.28
0.71 ±0.27
0.73 ±0.27
0.76 ±0.26
0.19 ±0.09
0.42 ±0.11
0.46 ±0.11
0.48 ±0.11
0.50 ±0.11
0.54 ±0.11
0.77 ±0.09
0.76 ±0.35
0.70 ±0.37
0.61 ±0.40
0.57 ±0.40
0.50 ±0.41
0.50 ±0.41
0.00 ±0.00
0.80 ±0.48
0.60 ±0.58
0.51 ±0.59
0.45 ±0.59
0.39 ±0.58
0.31 ±0.55
0.00 ±0.00
98
90
80
70
60
50
nolep
Reco ID
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0


PID Purity
Particle ID
p
K
π
e
μ
0.39 ±0.48
0.27 ±0.29
0.25 ±0.27
0.24 ±0.26
0.23 ±0.25
0.20 ±0.22
0.14 ±0.16
0.78 ±0.34
0.73 ±0.28
0.71 ±0.28
0.69 ±0.28
0.67 ±0.27
0.65 ±0.27
0.58 ±0.26
0.94 ±0.11
0.94 ±0.08
0.92 ±0.08
0.92 ±0.08
0.91 ±0.08
0.90 ±0.08
0.88 ±0.08
0.15 ±0.13
0.33 ±0.26
0.32 ±0.27
0.27 ±0.25
0.24 ±0.24
0.21 ±0.22
0.00 ±1.00
0.05 ±0.06
0.05 ±0.07
0.04 ±0.07
0.04 ±0.08
0.04 ±0.08
0.05 ±0.11
0.00 ±1.00
98
90
80
70
60
50
nolep
Reco ID
0.8
0.6
0.4
0.2
0.0