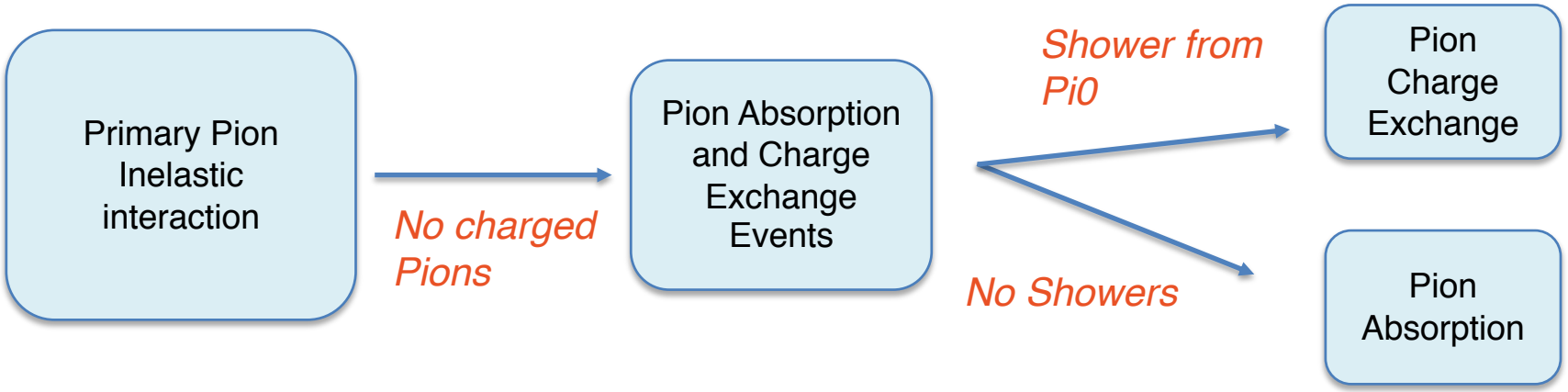
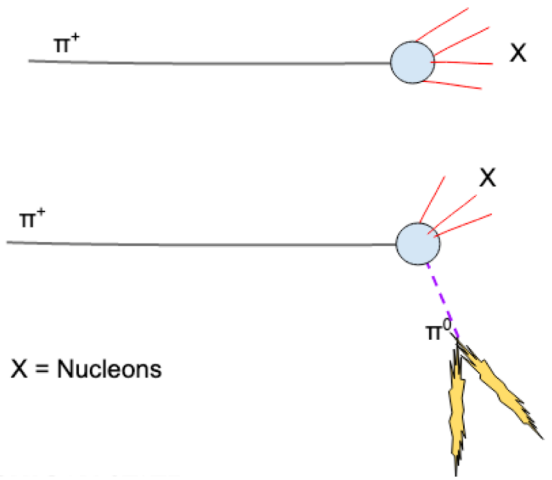
Looking at CNN shower Tag vs Pandora shower Tag

Francesca Stocker

10.10.2019

Context: Pion Charge Exchange and Absorption Channel



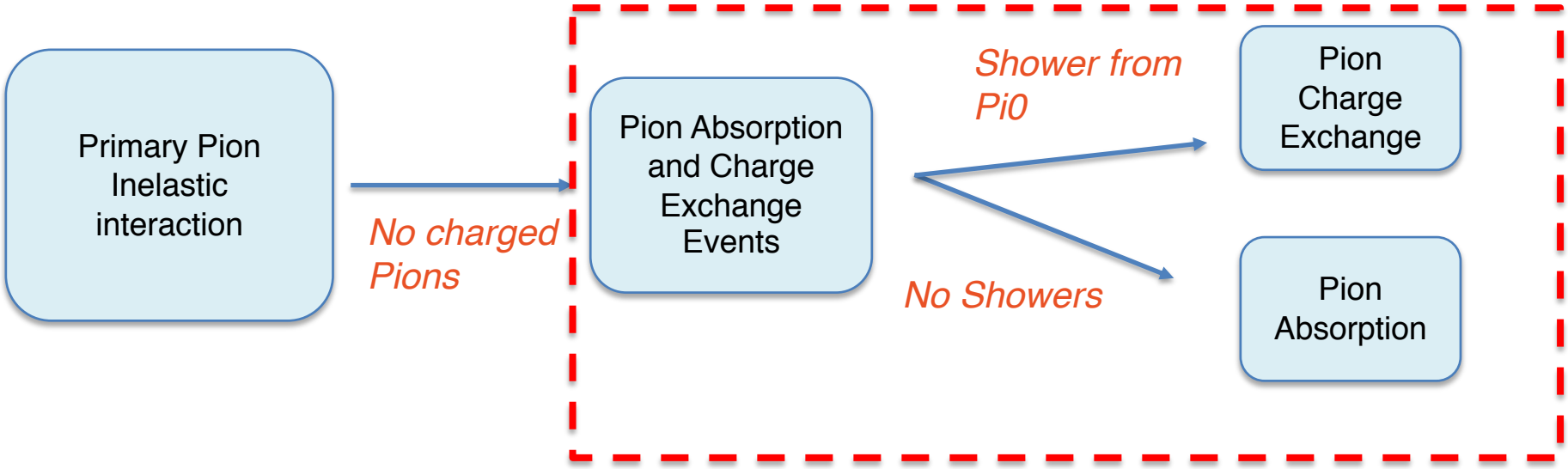


Absorption

Charge Exchange

Jake Calcutt:
https://indico.fnal.gov/event/21445/session/13/contribution/83/material/slides/0.pdf

Context: Pion Charge Exchange and Absorption Channel

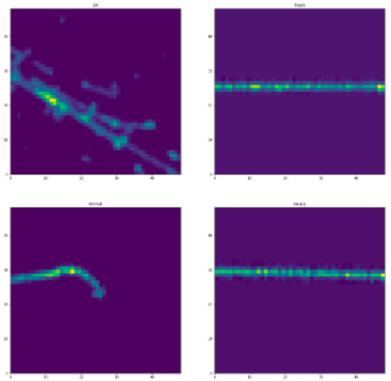


- The correct identification of showers is important to separate the Absorption from the Charge Exchange Channel
- Two options for showers:
  - Pandora Shower Tag
  - CNN track-like Score (Aidan Reynolds) https://indico.fnal.gov/event/20654/contribution/2/material/slides/0.pdf

Pandora Shower Tag / CNN

- Pandora Shower Tag: clustering, 2D then 3D tensor?
  - Haven't found any slides with efficiencies on this, maybe someone can point me to it?
- CNN: Initial goal use for calibration samples
  - Michel Electrons and delta ray removal for muon calibration
  - Takes 4 types of images for training EM, Track, Michel, Empty
  - Hit by hit track/shower separation
  - Trained on MCC11, SCE on, Fluid Flow on, All beam energies

Patch Type	EM	Track	Empty	Michel
Training	13,493,982	9,727,604	2,517,882	731,456
Validation	734,673	562,038	141,388	42,727
Test	764,659	518,805	139,987	39,674

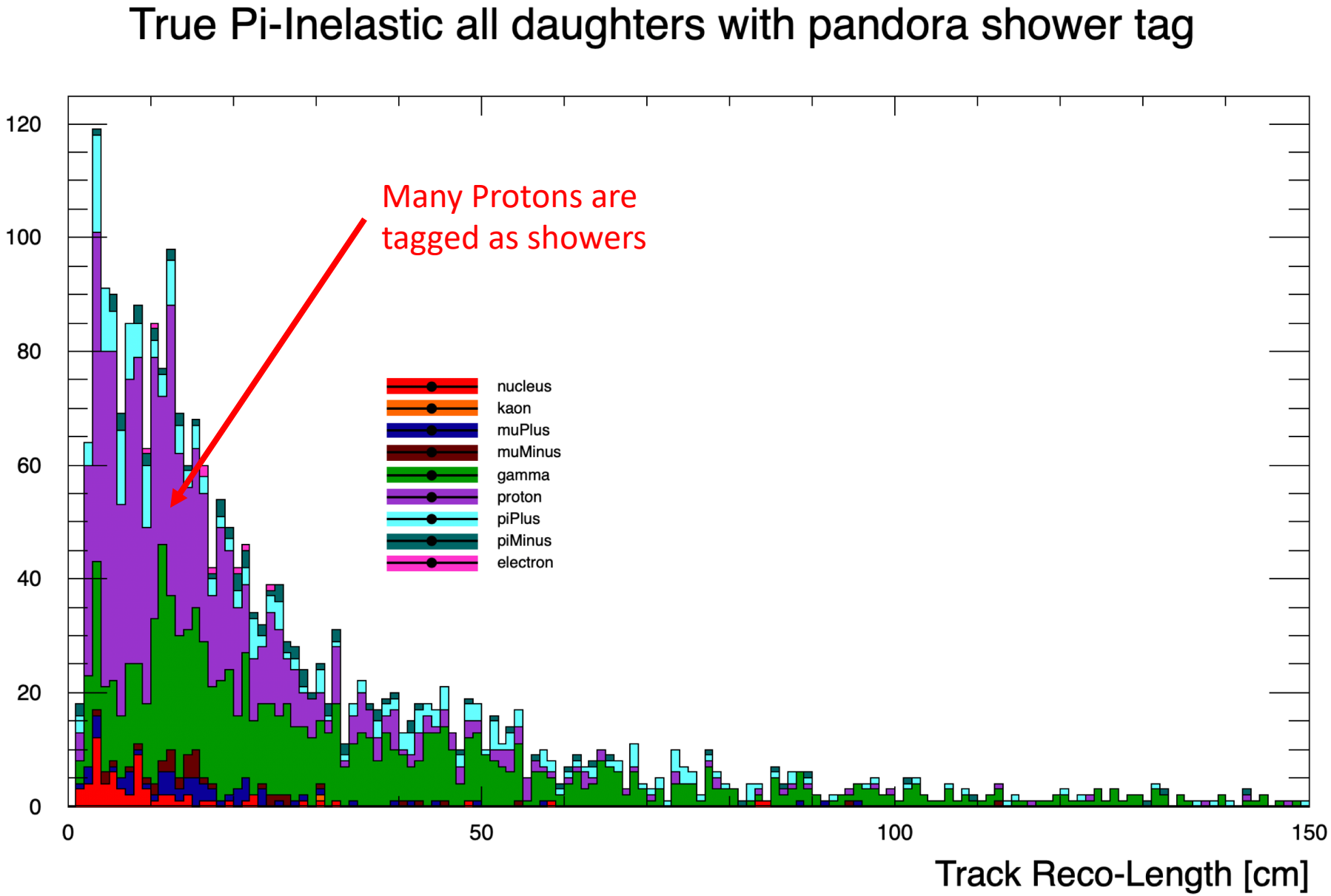


Starting Point

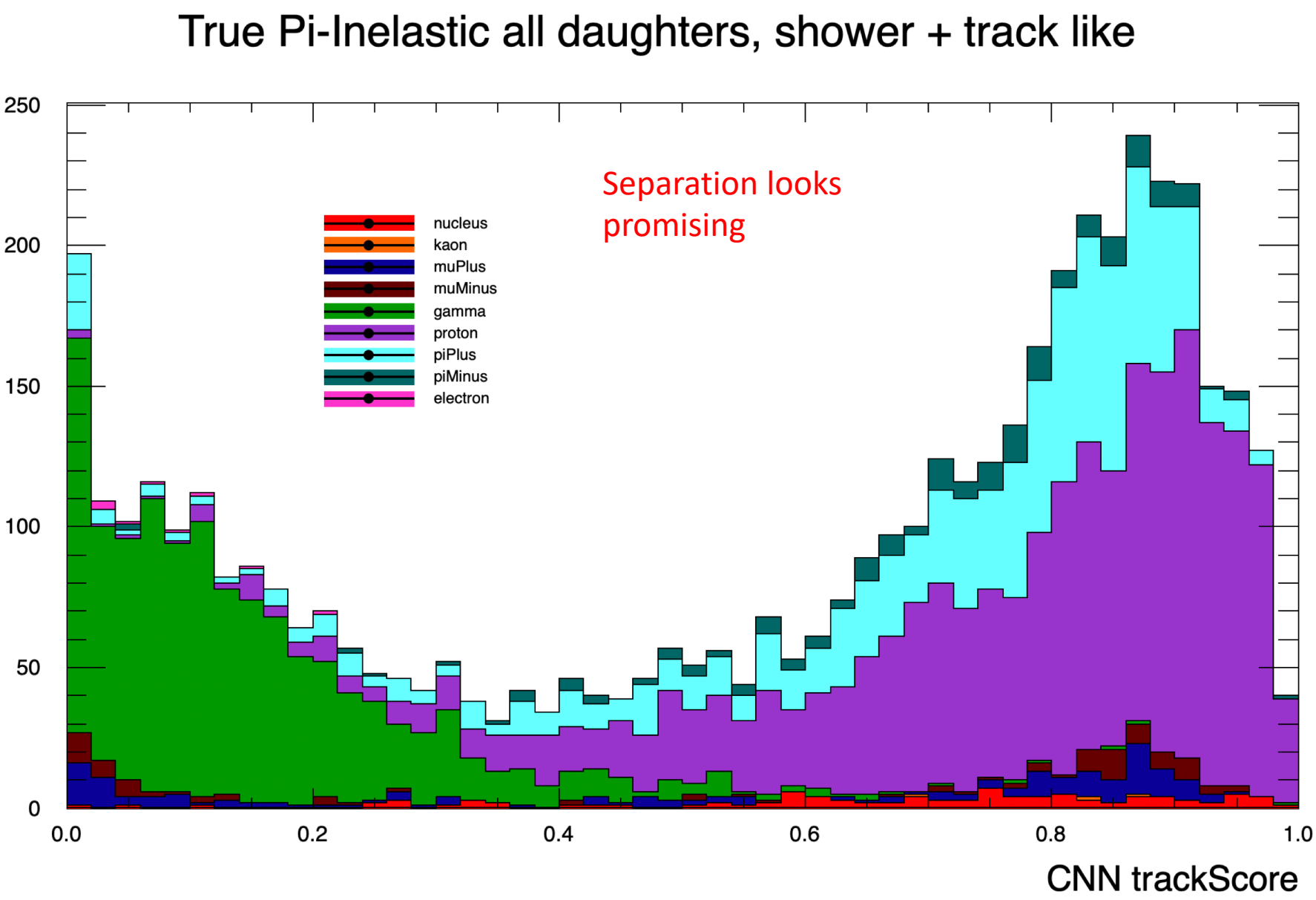
- Try to characterize how well the two work
- For CNN: average over the hits in a reconstructed object to get a track-like Score (0,1)
- I used MCC12 sample, 1GeV, sce-On and Jakes PionAnalyzer_MC module

1. **True primary beam pion Inelastic interaction**
    1. Who of the daughter particles was tagged by Pandora as Shower?
    2. What are the CNN track-like scores
2. **True ChEx + Absorption Events**
    1. Separate two channels by shower tagging with Pandora or CNN get efficiencies and purities

Pandora Shower Tag –who's who

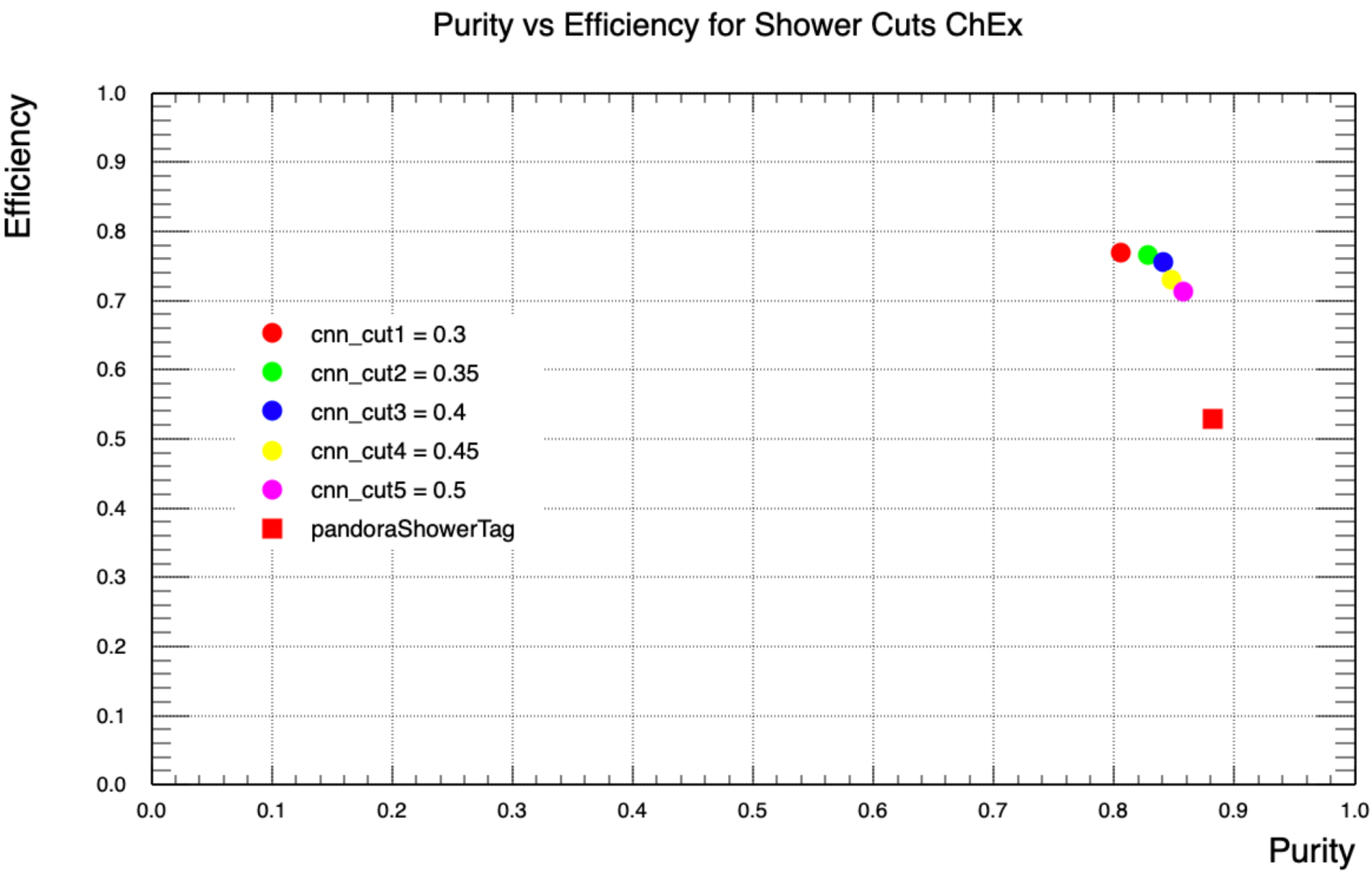


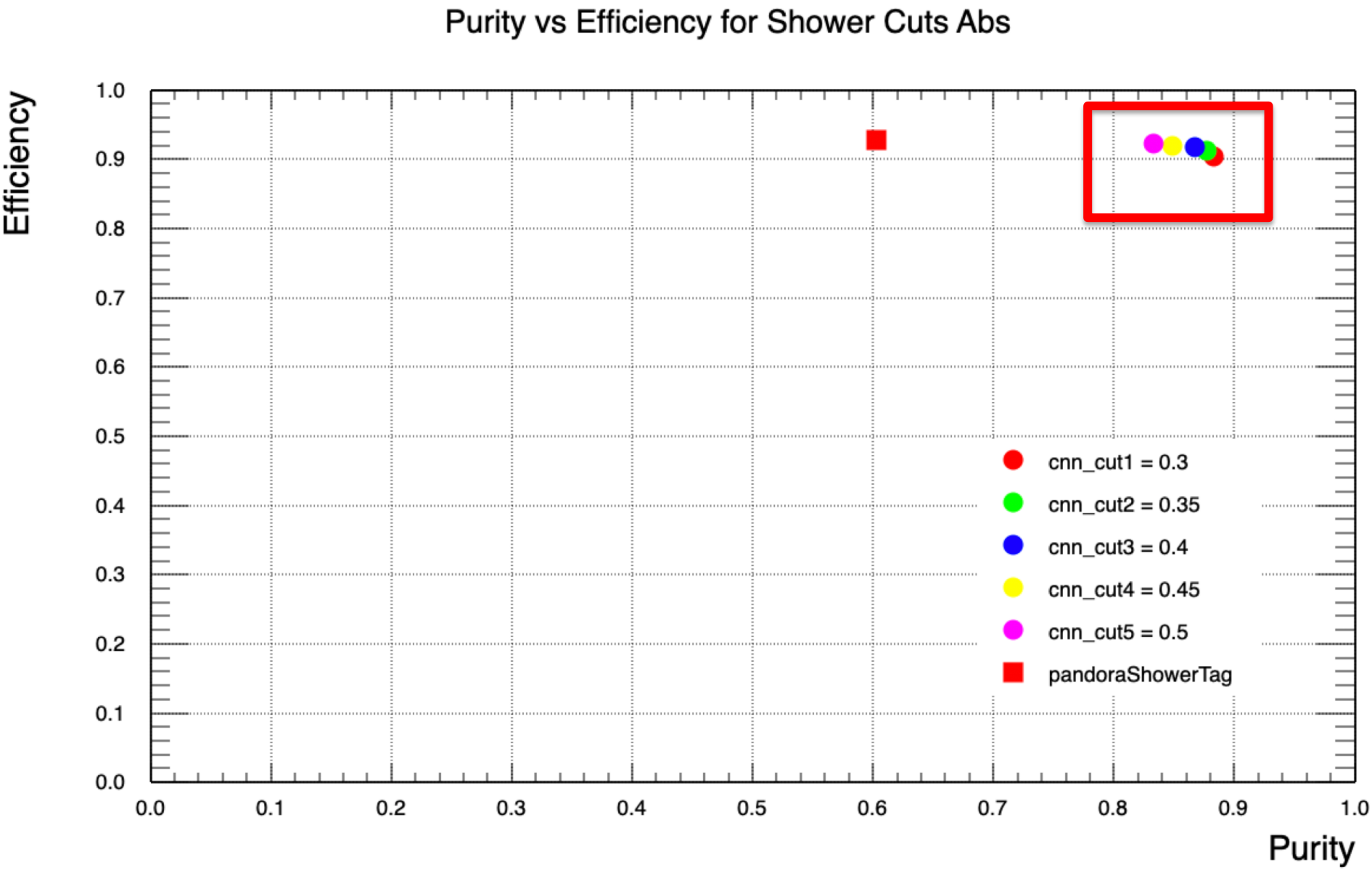
CNN track – score who's who



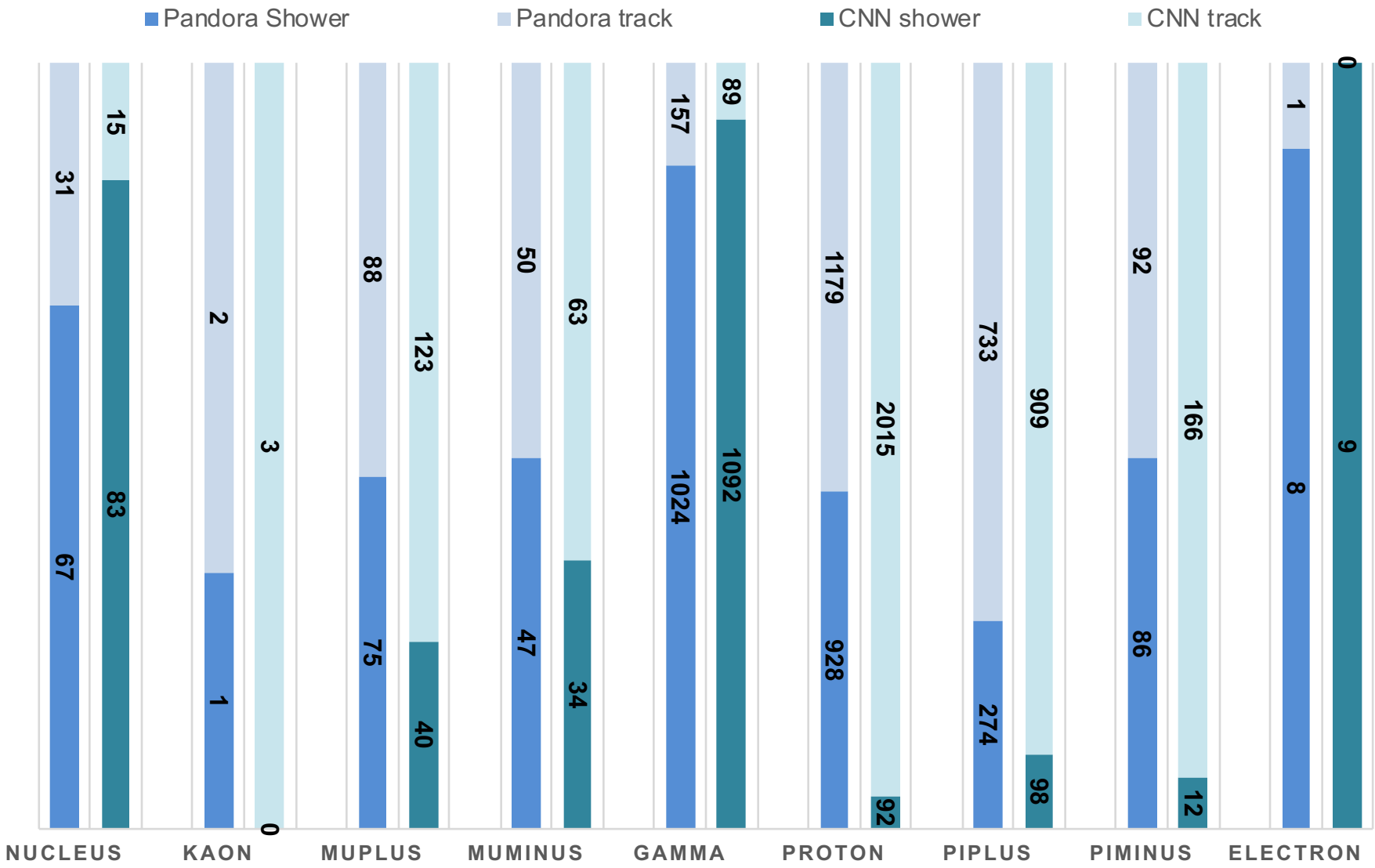
- From True ChEx + Abs Process use Shower Tag or CNN cut to separate channels
- $\text{Efficiency}_{\text{Abs}} = \frac{Match\ to\ True\ _{Abs}}{True\ _{Abs}}$ $\text{Purity}_{\text{Abs}} = \frac{Match\ to\ True\ _{Abs}}{Found\ _{Abs}}$

				True ChEx	True Abs					
MC truth: ChEx + Abs	**1245**			**407**	**838**					
	Found ChEx	**Found Abs**		**Match to True ChEx**	**Match to True Abs**		**Efficiency ChEx**	**Purity ChEx**	**Efficiency Abs**	**Purity Abs**
Pand Shower Tag	679	544		359	505		**0.88**	**0.53**	**0.60**	**0.93**
CNN cut1 = 0.3	426	819		328	740		**0.81**	**0.77**	**0.88**	**0.90**
CNN cut2 = 0.35	440	805		337	735		**0.83**	**0.77**	**0.88**	**0.91**
CNN cut3 = 0.4	453	792		342	727		**0.84**	**0.75**	**0.87**	**0.92**
CNN cut4 = 0.45	472	773		345	711		**0.85**	**0.73**	**0.85**	**0.92**
CNN cut5 = 0.5	489	756		349	698		**0.86**	**0.71**	**0.83**	**0.92**

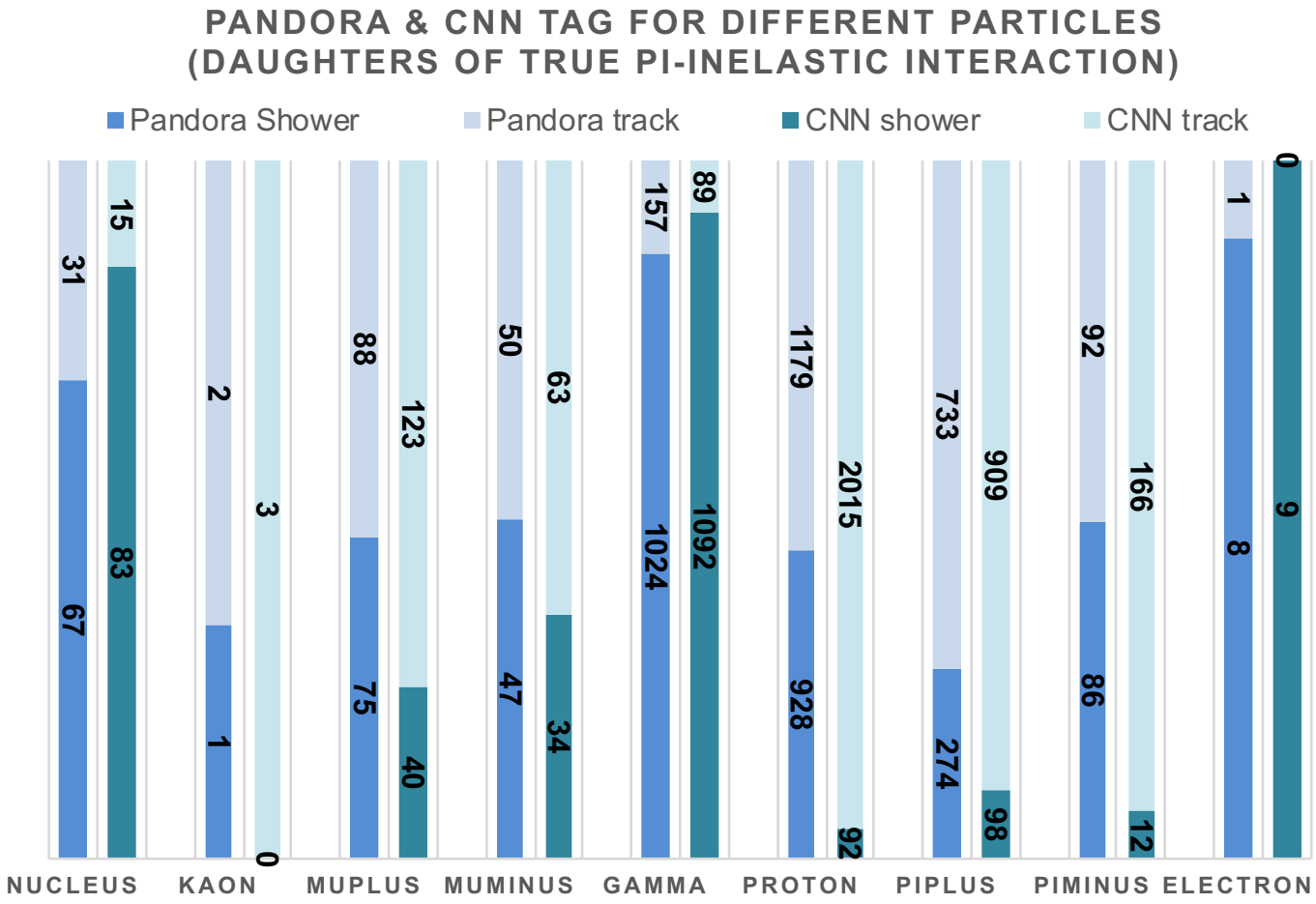

Purity vs Efficiency for Shower Cuts ChEx
Efficiency
Purity
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
cnn_cut1 = 0.3
cnn_cut2 = 0.35
cnn_cut3 = 0.4
cnn_cut4 = 0.45
cnn_cut5 = 0.5
pandoraShowerTag


Purity vs Efficiency for Shower Cuts Abs
Efficiency
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
Purity
cnn_cut1 = 0.3
cnn_cut2 = 0.35
cnn_cut3 = 0.4
cnn_cut4 = 0.45
cnn_cut5 = 0.5
pandoraShowerTag

PANDORA & CNN TAG FOR DIFFERENT PARTICLES (DAUGHTERS OF TRUE PI-INELASTIC INTERACTION)



- Pandora Tag works well for true Photons
- Pandora Tag doesn't work well for protons (~50% tagged as showers)
- CNN seems to do a better job,
  - still 8% of the photons not shower tagged though
- piPlus and Muons?



Conclusions

- CNN looks more promising
  - 3D graph-CNN done by Saul Monsalve and Leigh Whitehead (worth to look into those values? See how easy it is to get results from it.
  - https://indico.cern.ch/event/781262/contributions/3380328/attachments/1823851/2984124/ep-nu-sam-04-04-19.pdf
- Why do Protons get CNN/shower Tag? How could this be worked around? Chi2? Cuts?
- Go with CNN for shower tagging? What work/direction should be chosen to improve this?
  - Aidan: CNN was not trained with many protons
  - Train with more protons?
  - CNN was trained with MCC11 change and train with MCC12? Benefit? Workload?
  - Or work with the above mentioned 3D CNN? Easy/Quick?

OutLook

- See if there are discriminative shower properties
- Look at some failure event displays (photons/protons)

				True ChEx	True Abs					
MC truth: ChEx + Abs	**1245**			**407**	**838**					
	Found ChEx	**Found Abs**		**Match to True ChEx**	**Match to True Abs**		**Efficiency ChEx**	**Purity ChEx**	**Efficiency Abs**	**Purity Abs**
Pand Shower Tag	679	544		359	505		**0.88**	**0.53**	**0.60**	**0.93**
CNN cut1 = 0.3	426	819		328	740		**0.81**	**0.77**	**0.88**	**0.90**
CNN cut2 = 0.35	440	805		337	735		**0.83**	**0.77**	**0.88**	**0.91**
CNN cut3 = 0.4	453	792		342	727		**0.84**	**0.75**	**0.87**	**0.92**
CNN cut4 = 0.45	472	773		345	711		**0.85**	**0.73**	**0.85**	**0.92**
CNN cut5 = 0.5	489	756		349	698		**0.86**	**0.71**	**0.83**	**0.92**
Combined Cuts Pandora -> CNN										
cut1	399	544		319	505		**0.78**	**0.80**	**0.60**	**0.93**
cut2	411	544		327	505		**0.80**	**0.80**	**0.60**	**0.93**
cut3	424	544		332	505		**0.82**	**0.78**	**0.60**	**0.93**
cut4	437	544		335	505		**0.82**	**0.77**	**0.60**	**0.93**
cut5	450	544		338	505		**0.83**	**0.75**	**0.60**	**0.93**
Combined Cuts CNN --> Pandora										
cut1	404	522		322	489		**0.79**	**0.80**	**0.58**	**0.94**
cut2	416	520		330	488		**0.81**	**0.79**	**0.58**	**0.94**
cut3	429	520		335	488		**0.82**	**0.78**	**0.58**	**0.94**
cut4	444	516		335	488		**0.82**	**0.75**	**0.58**	**0.95**
cut5	460	515		341	484		**0.84**	**0.74**	**0.58**	**0.94**