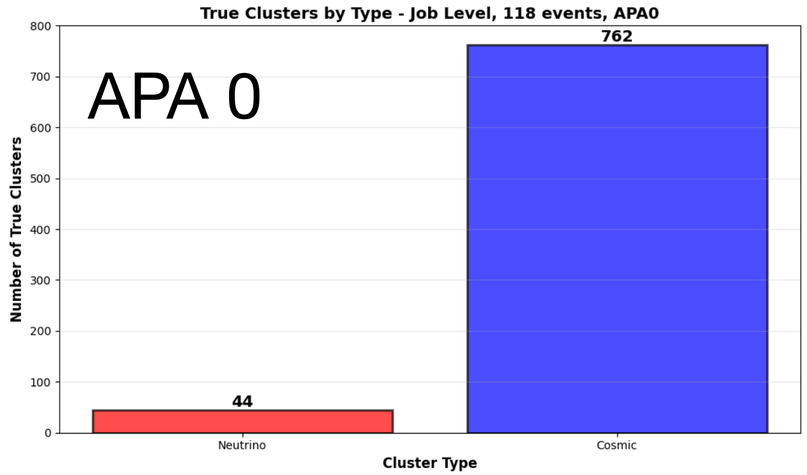
LOUISIANA STATE UNIVERSITY

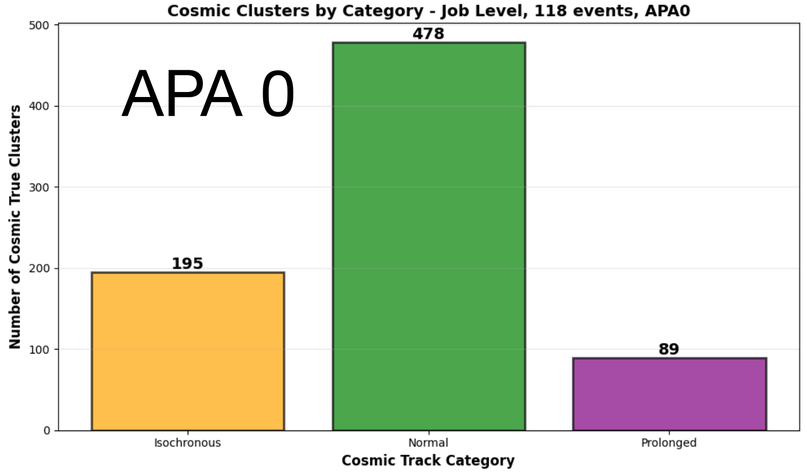
Status of WireCell Cluster Evaluation

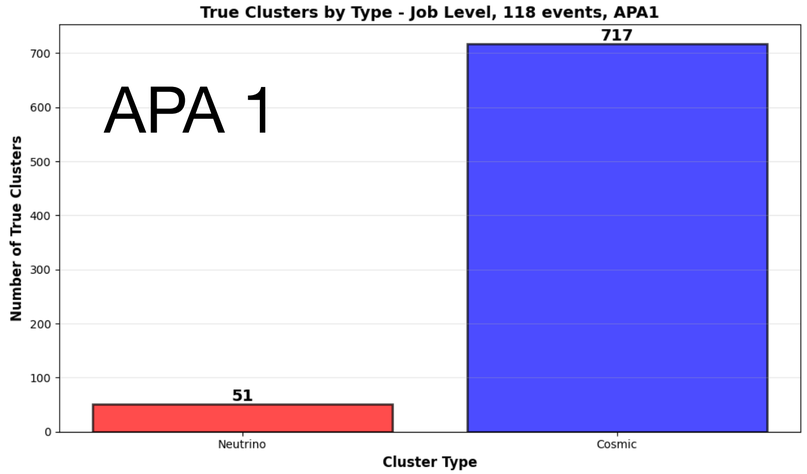
SBND Wirecell Meeting

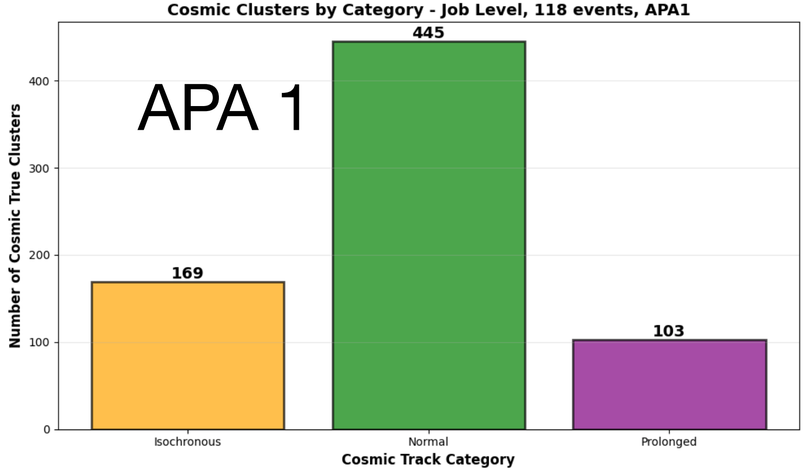
Prabhjot Singh, 16 July 2026

Sample for Evaluation







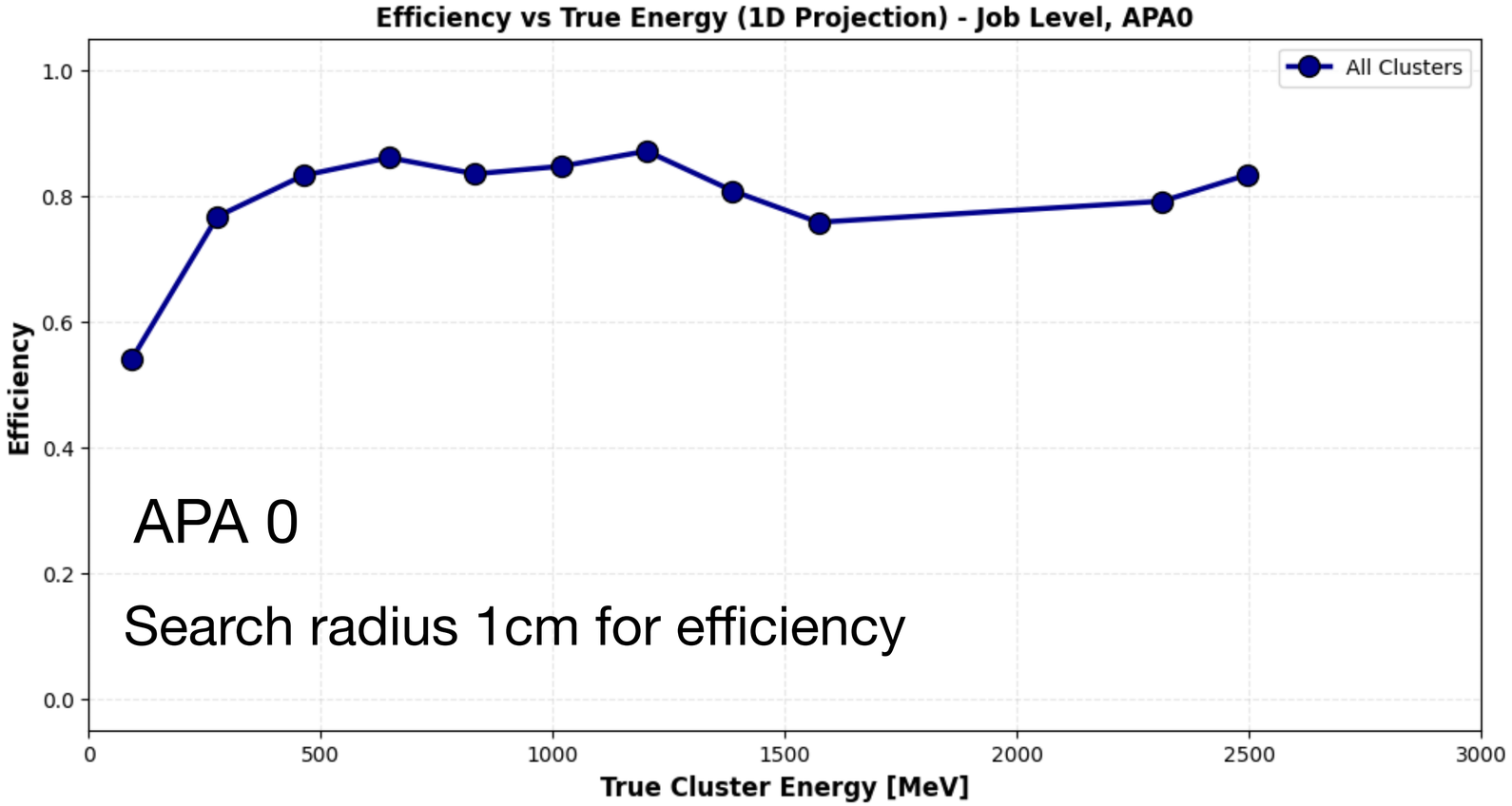


Number of Files	10
Events	118

- Though small but relatively good sample of neutrino and cosmic clusters for reconstruction efficiency and purity evaluations

- Isochronous cosmic clusters: $\theta_{xz} < 20^0$ (need to reduce to 15^0)
- Prolonged cosmic clusters: $\theta_{xz} > 70^0$ (need to increase to 70^0)
- Normal cosmic clusters: $\theta_{xz} : 20^0 - 70^0$ (need to change to $\theta_{xz} : 15^0 - 75^0$)

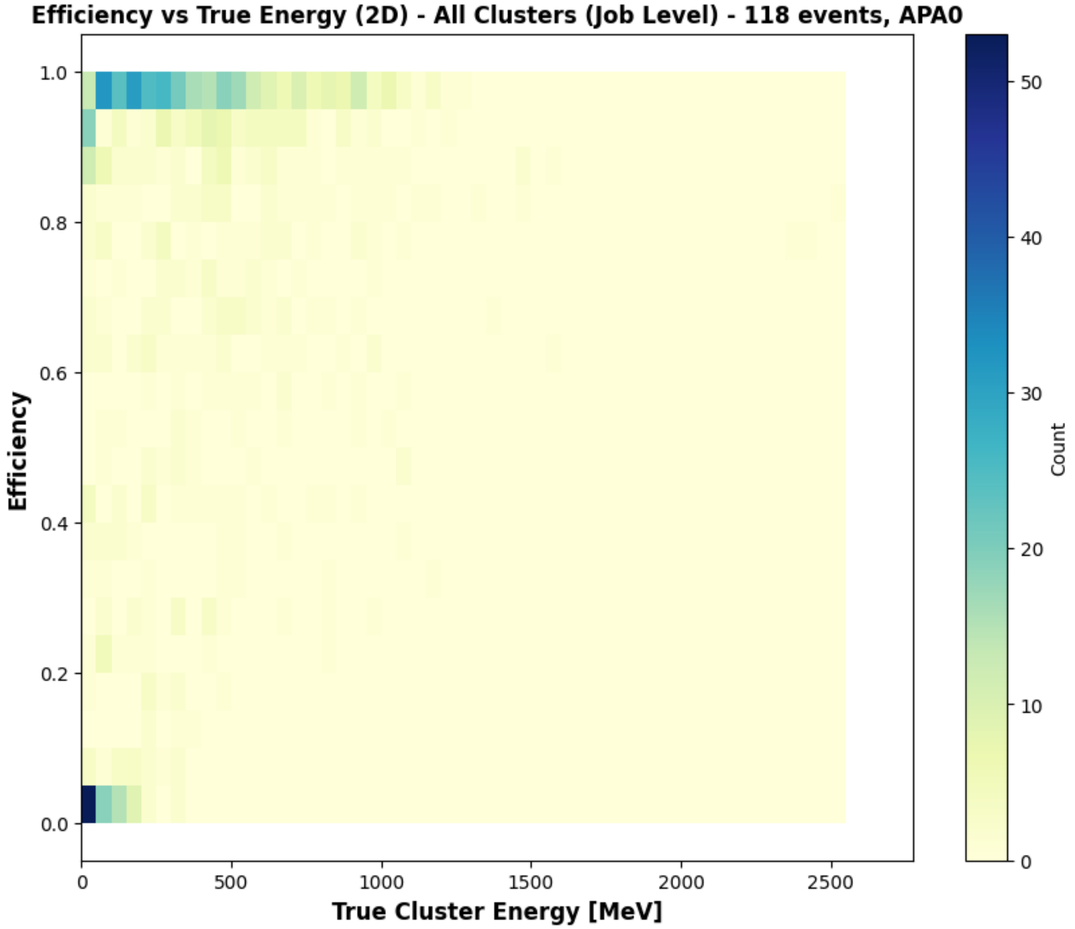
Efficiency as a function of True Cluster Energy



- This efficiency is estimated for all neutrinos and all cosmics
- As expected efficiency is flat and stable as a function of true energies of clusters
- Each point is an average efficiency in a given true energy bin
- Though the efficiency curve is flat and stable but it is not as high as we would have expected

Efficiency definition: fraction of true clusters matched to reco clusters

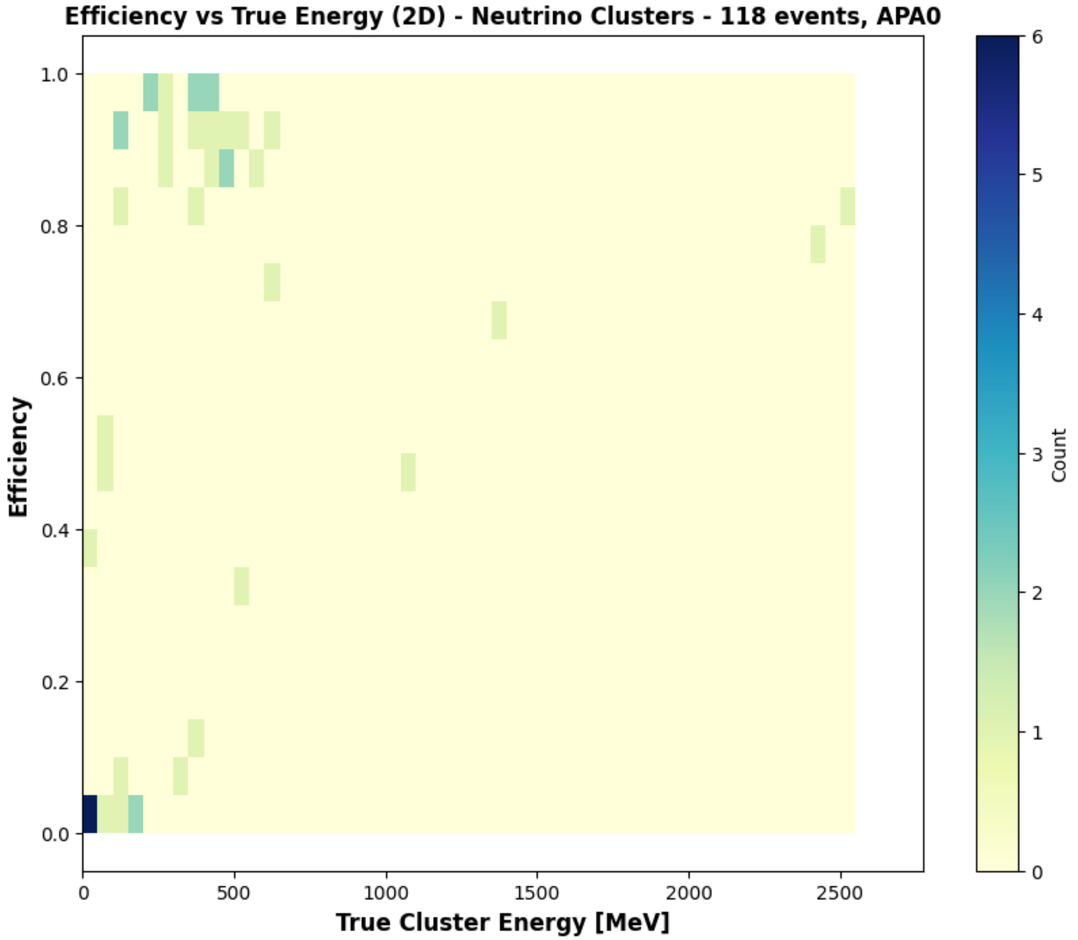
Efficiency as a function of True Cluster Energy



Efficiency definition: fraction of true clusters matched to reco clusters

- 1d efficiency plot is estimated from this 2d plot
- In each true energy bin we take the average efficiency
- As you can see a large fraction of true clusters are closer to 100% efficiency which we would expect
- There is also a large fraction of true clusters close to zero efficiency which we do not expect and it needs further investigation

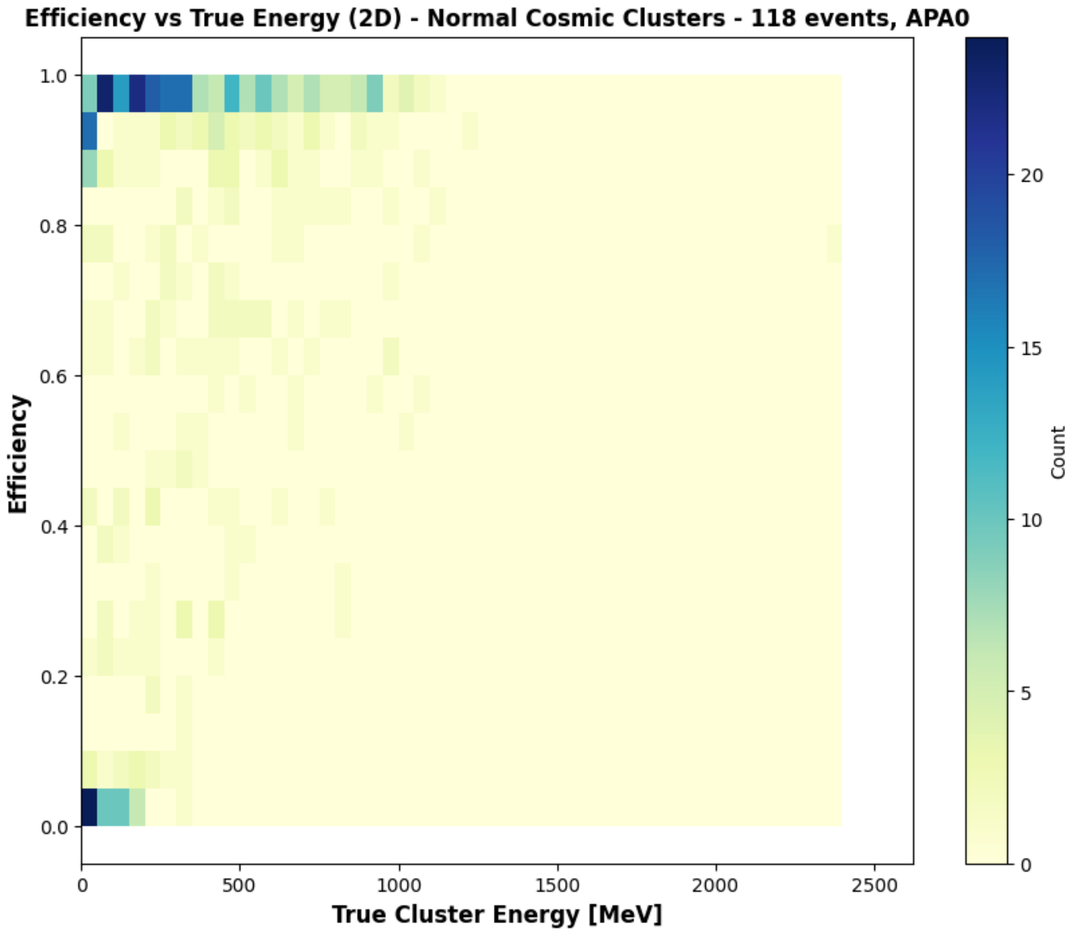
Efficiency as a function of True Cluster Energy



- Several neutrino clusters are at high efficiency regions but there are also some at very low efficiencies

Efficiency definition: fraction of true clusters matched to reco clusters

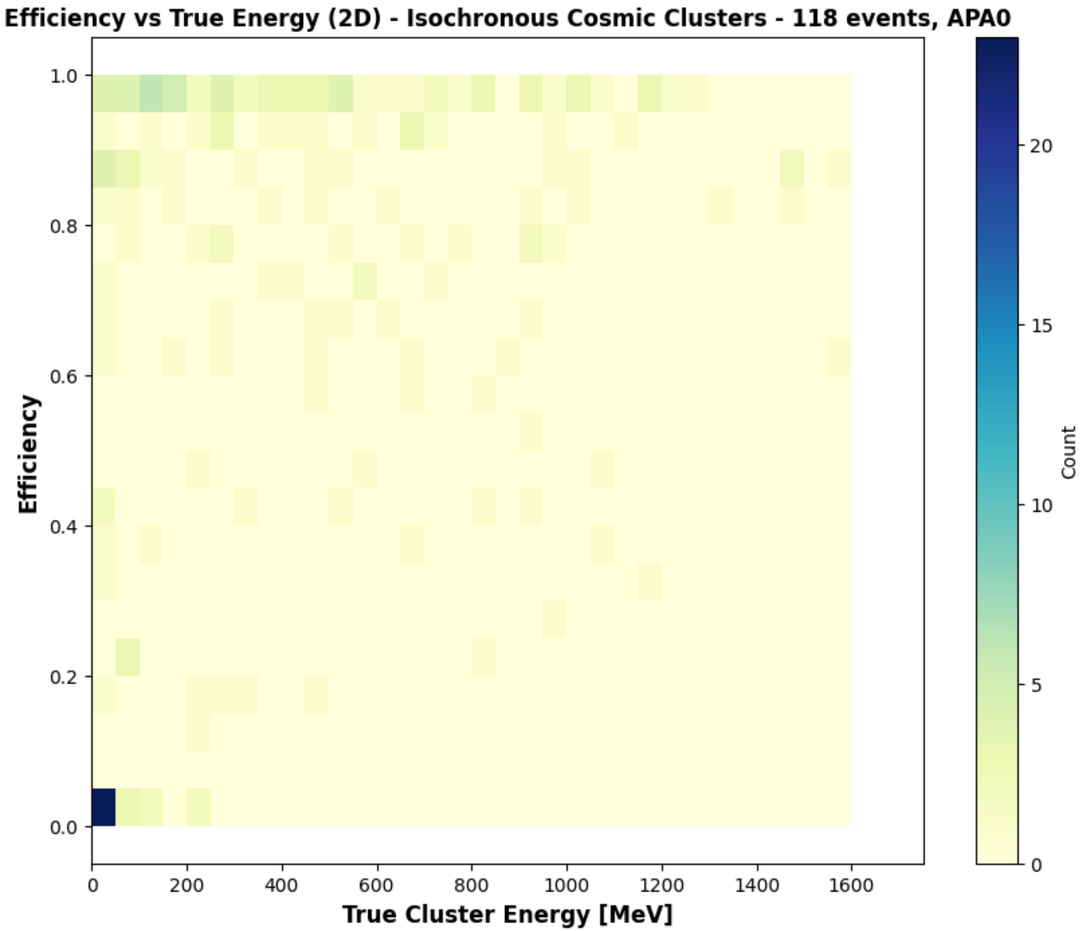
Efficiency as a function of True Cluster Energy



- Normal cosmic: non-isochronous and non-prolonged closures
- This plot shows some fraction of normal cosmic true clusters also are close to zero efficiency

Efficiency definition: fraction of true clusters matched to reco clusters

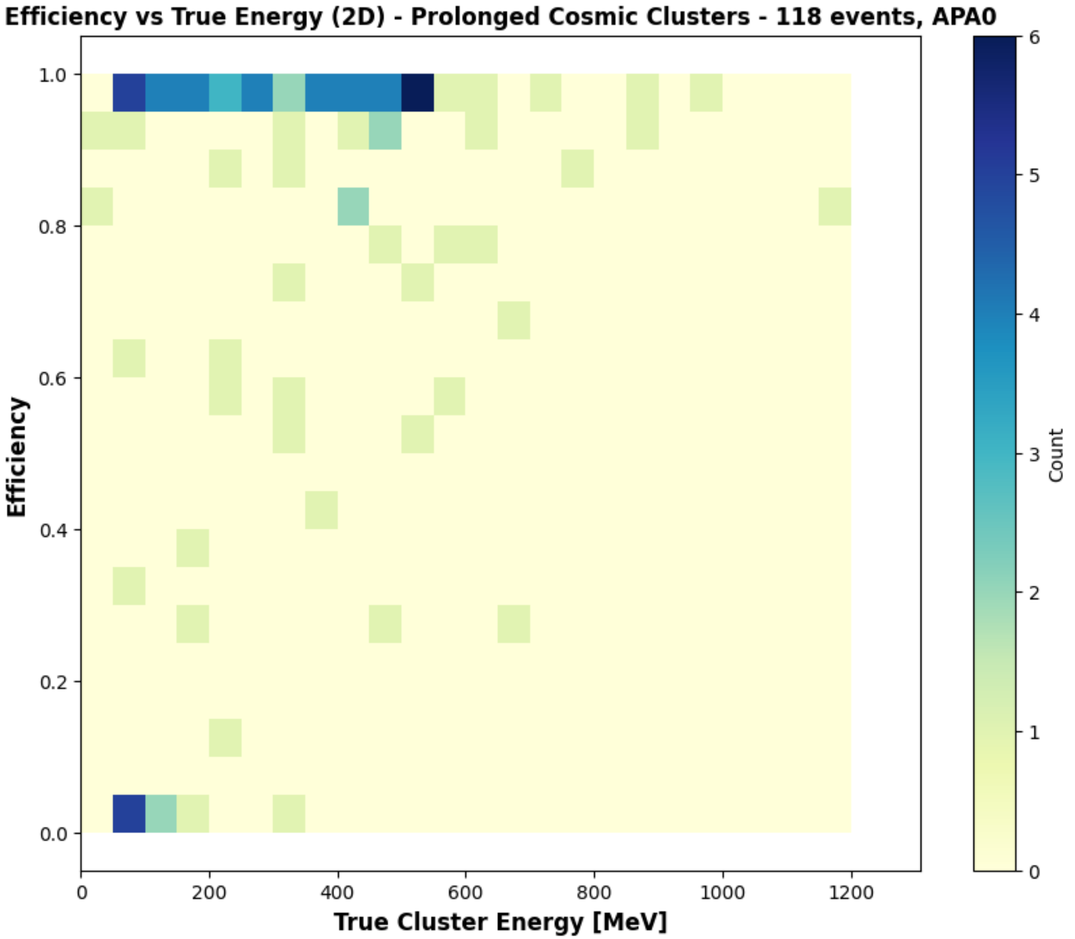
Efficiency as a function of True Cluster Energy



- Some isochronous true clusters are also at zero efficiency regions

Efficiency definition: fraction of true clusters matched to reco clusters

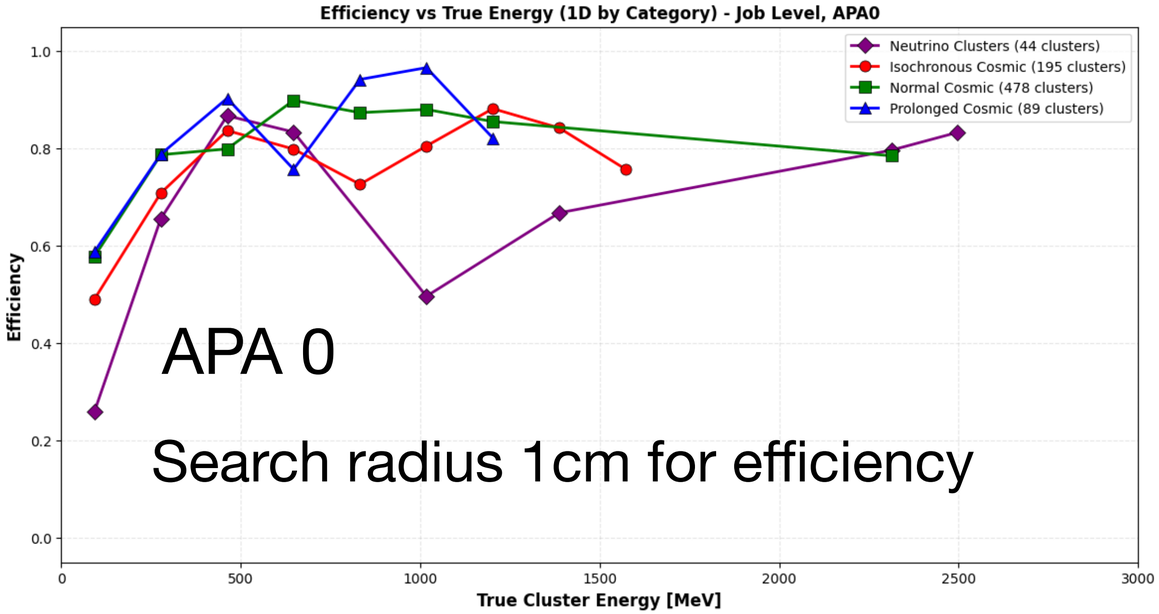
Efficiency as a function of True Cluster Energy

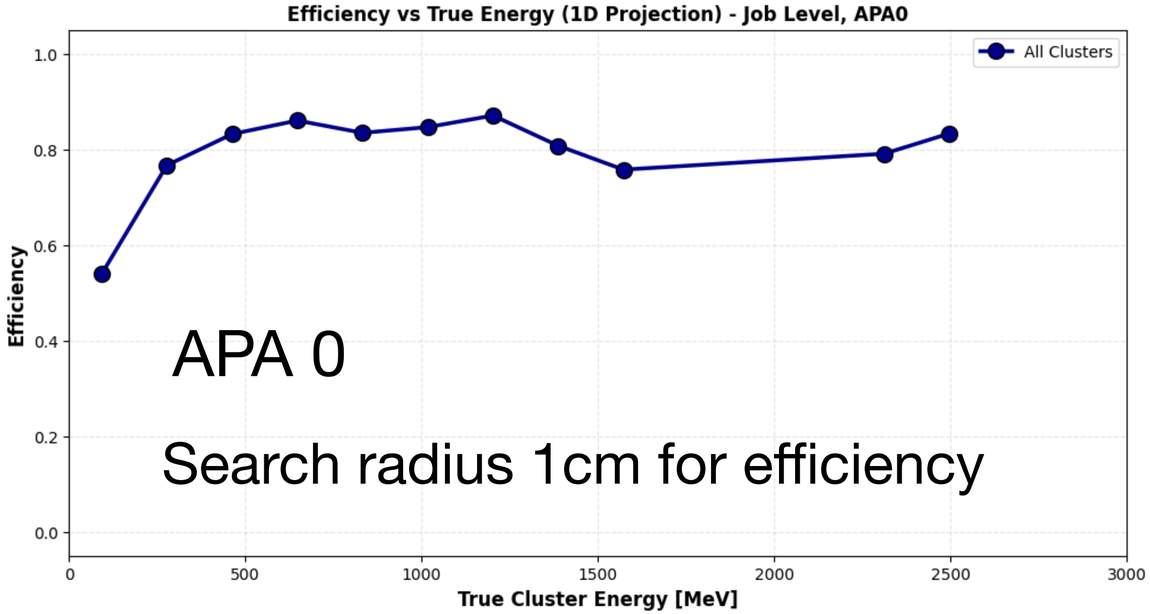


- A lot of prolonged cosmic clusters are at high efficiency regions and few are at low regions

Efficiency definition: fraction of true clusters matched to reco clusters

Efficiency as a function of True Cluster Energy





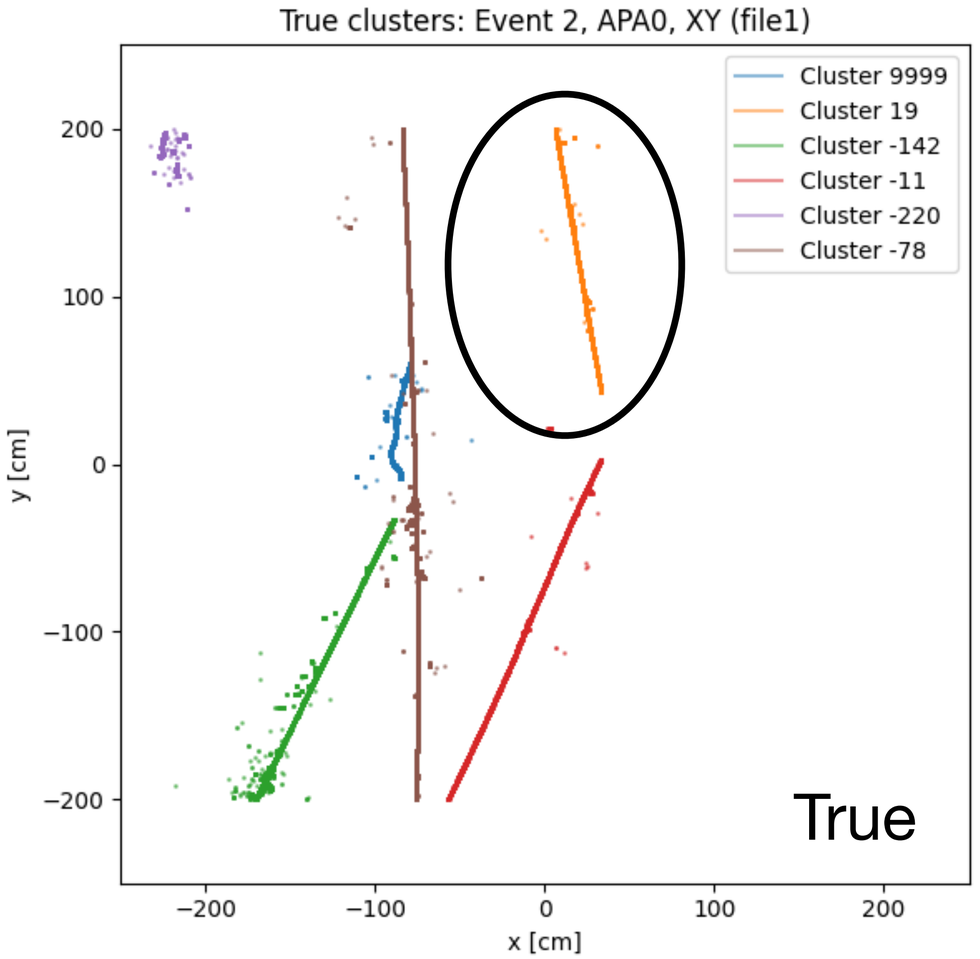
- These are summarized 1d efficiency plots for all categories: Neutrinos and Cosmics
- All of these need some further investigation to understand the low efficiency cases

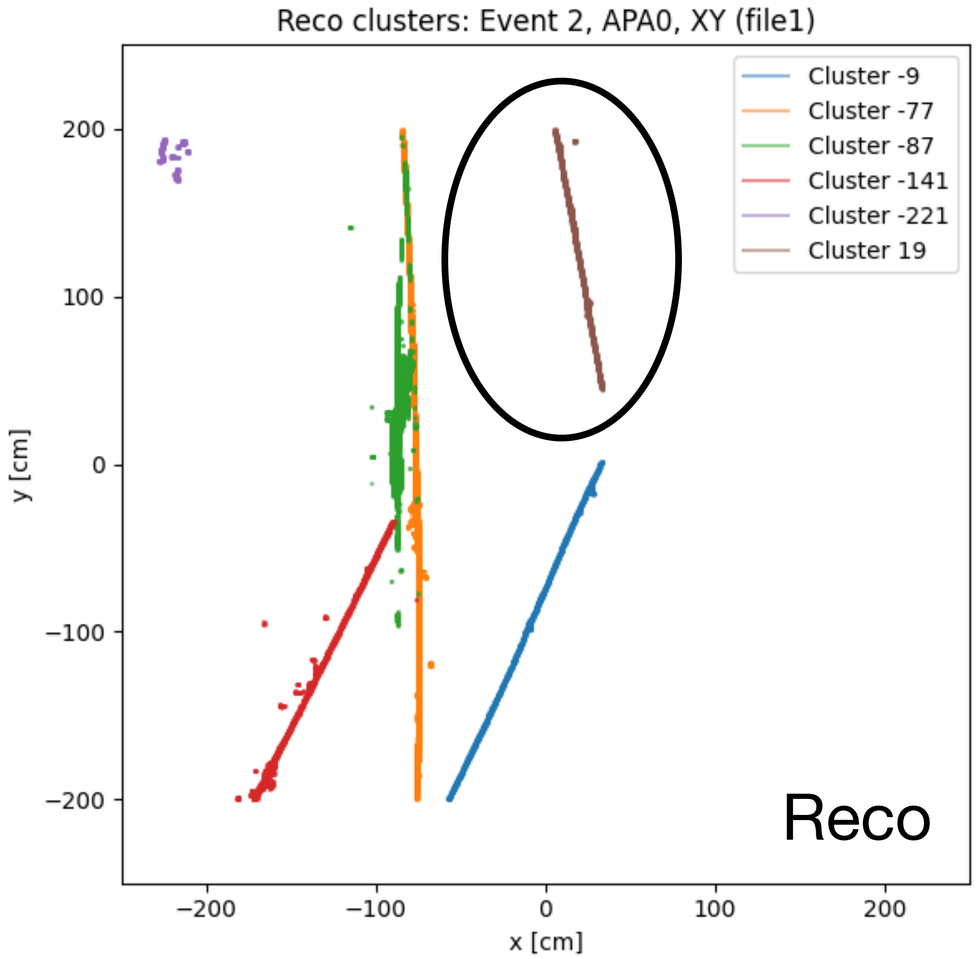
Reasons of low efficiencies

- A large fraction of low efficiency clusters is caused due to uncorrected space-charge effects
- Space-charge effects can cause reco clusters to be reconstructed slightly away from true clusters
- For efficiency estimates we include reco points within 1 cm radius from true clusters but if reco points are more than 1 cm away from true points then they are excluded from efficiency evaluation leading to low efficiencies
- There are also some cases where true clusters did not find any reco clusters which caused zero efficiency cases
- Some more investigation is needed to find may be other reasons for non-matching between true and reco points

Some cases of low efficiency: Event 2

Bee Link



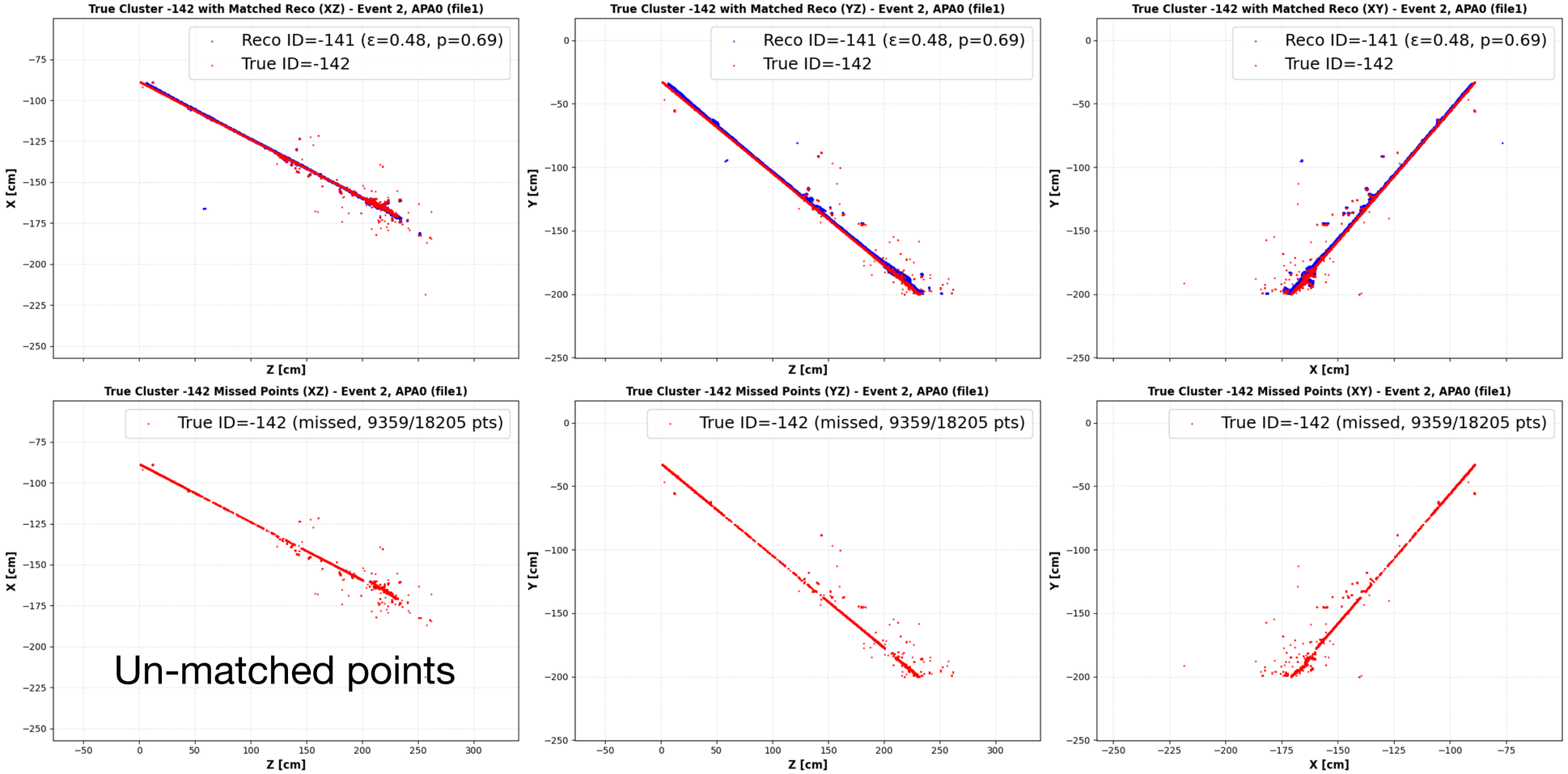


ε = 48%

Reason for low efficiency: Space-charge effect

Some cases of low efficiency: Event 2

Bee Link

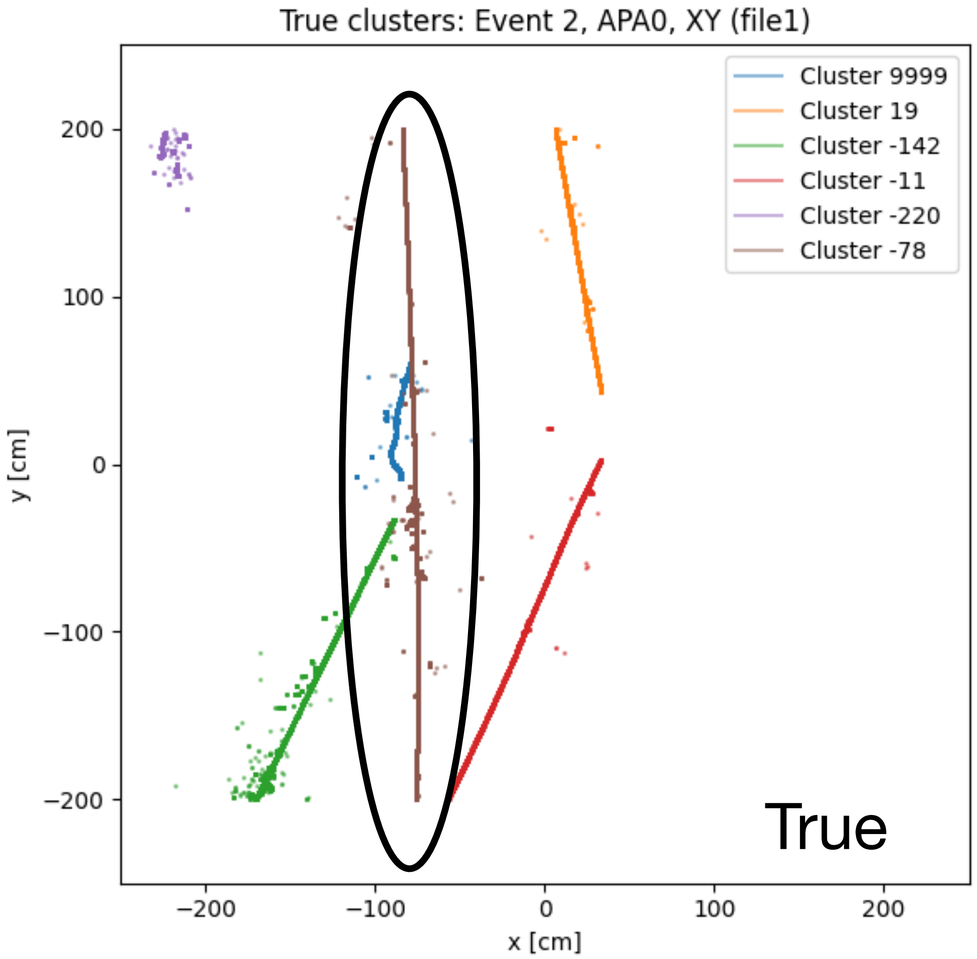


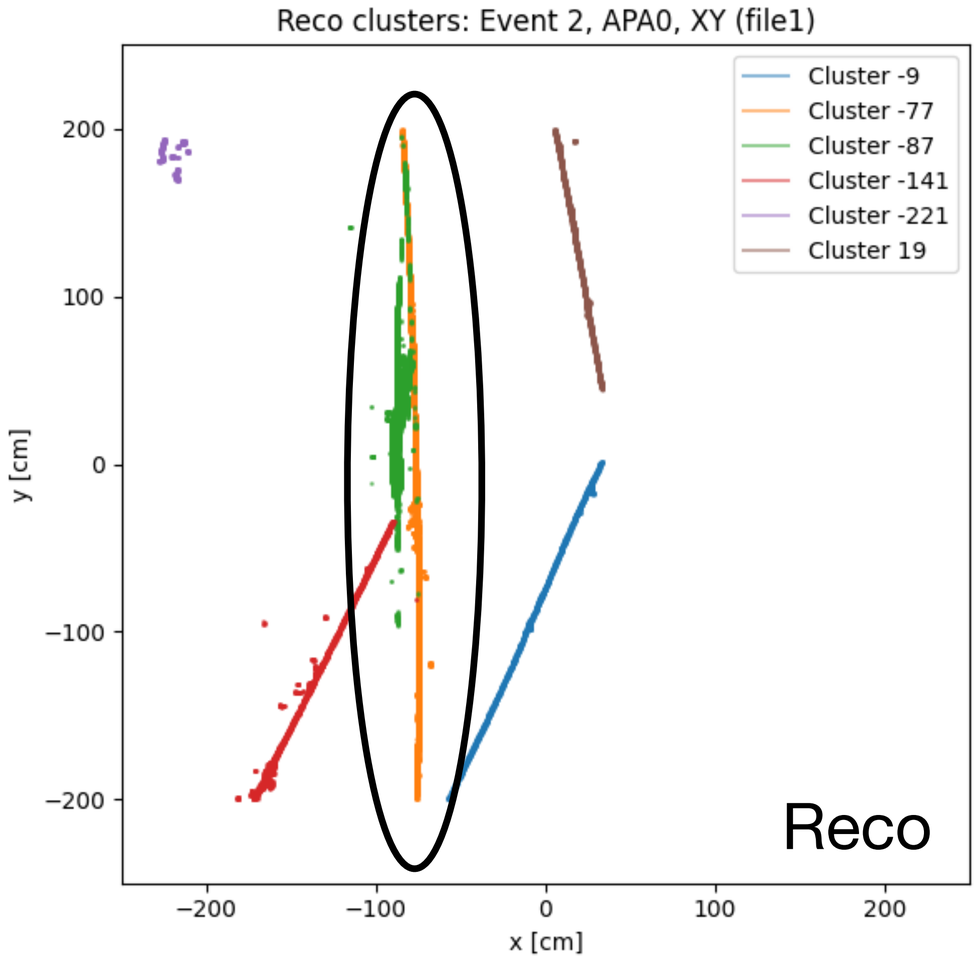
ε = 48%

Reason for low efficiency: Space-charge effect

Some cases of low efficiency: Event 2

Bee Link



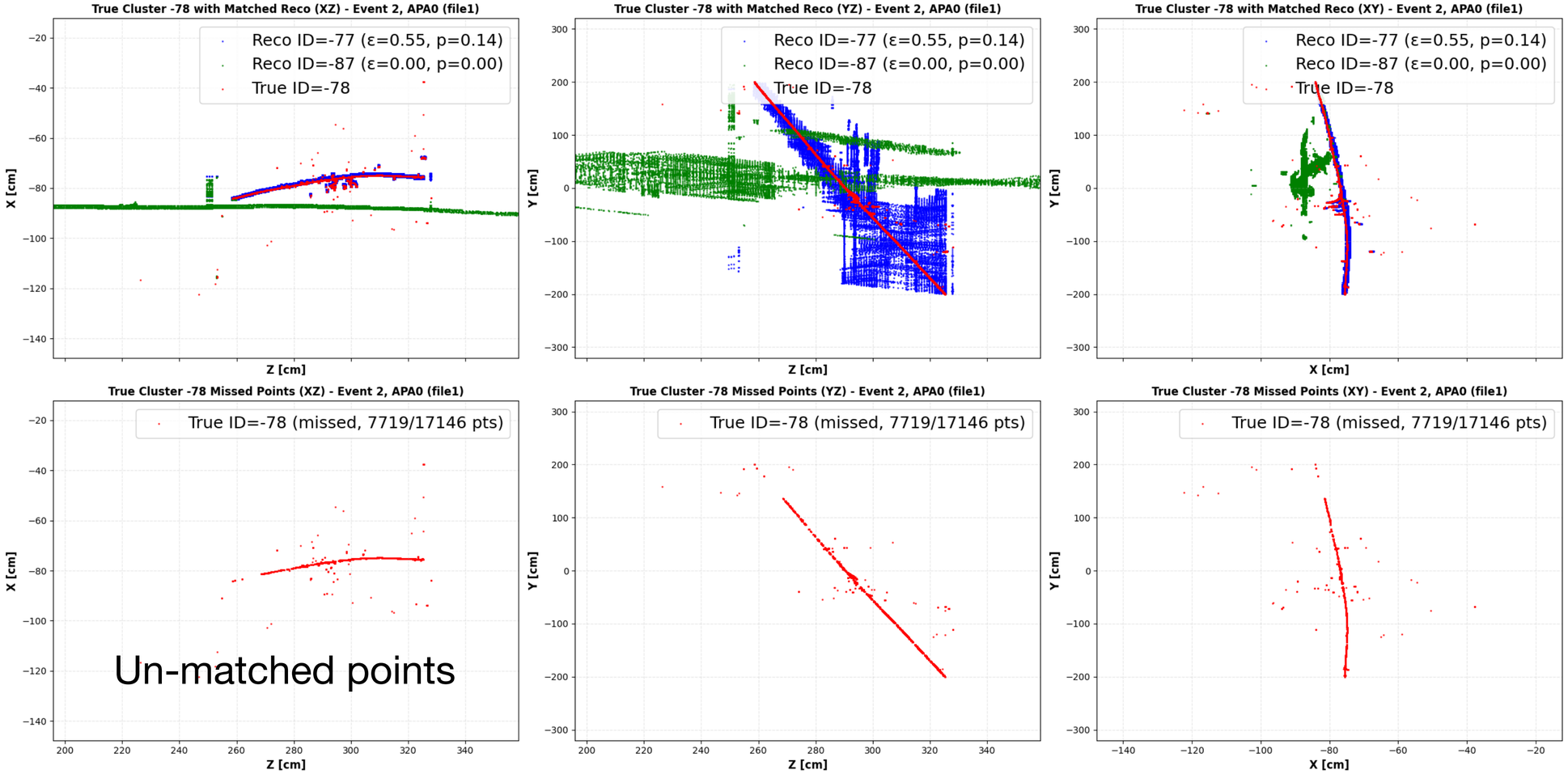


ε = 55%

Reason for low efficiency: Space-charge effect

Some cases of low efficiency: Event 2

Bee Link



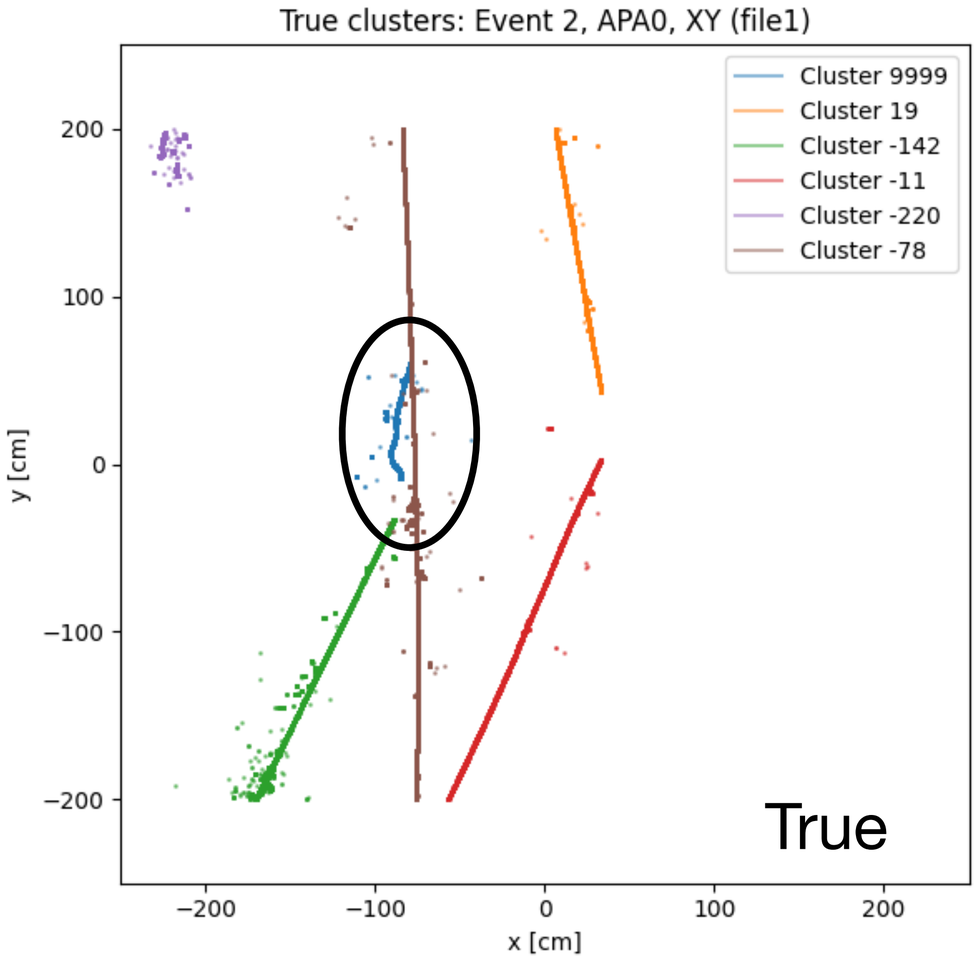
ε = 55%

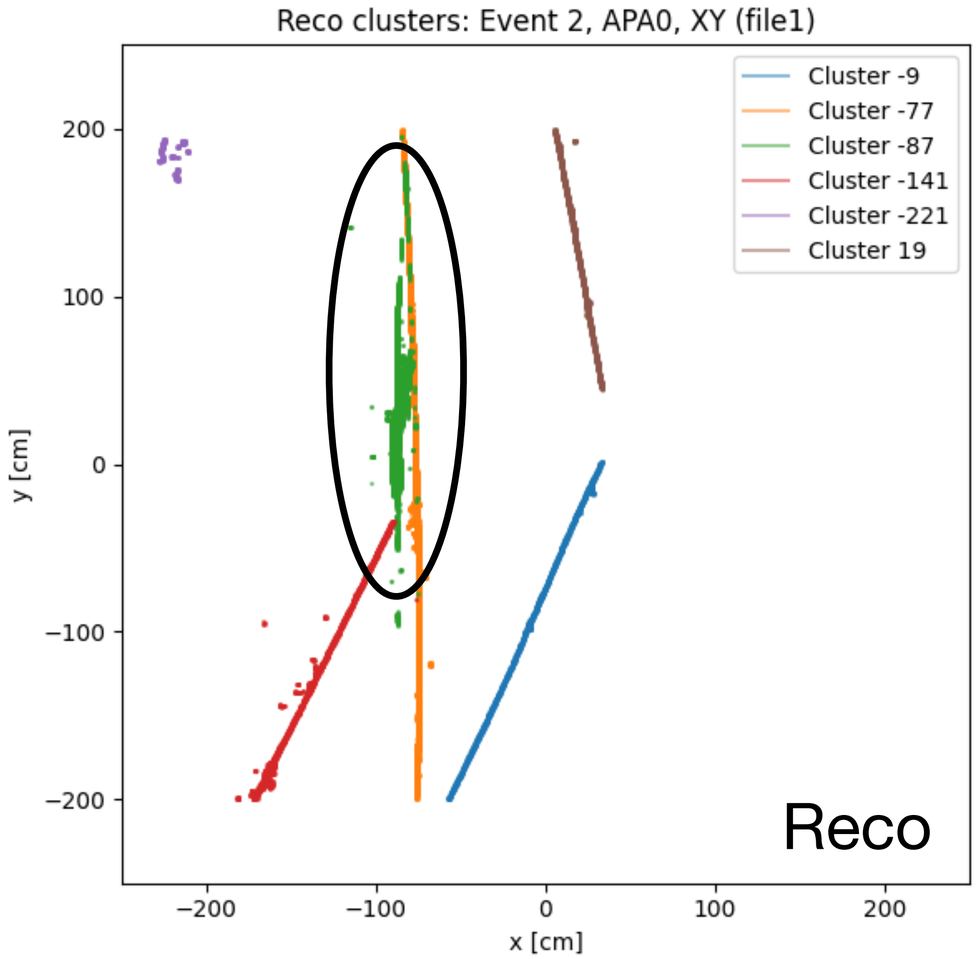
Reason for low efficiency: Space-charge effect

Some cases of low efficiency: Event 2

Bee Link

Neutrino Interaction





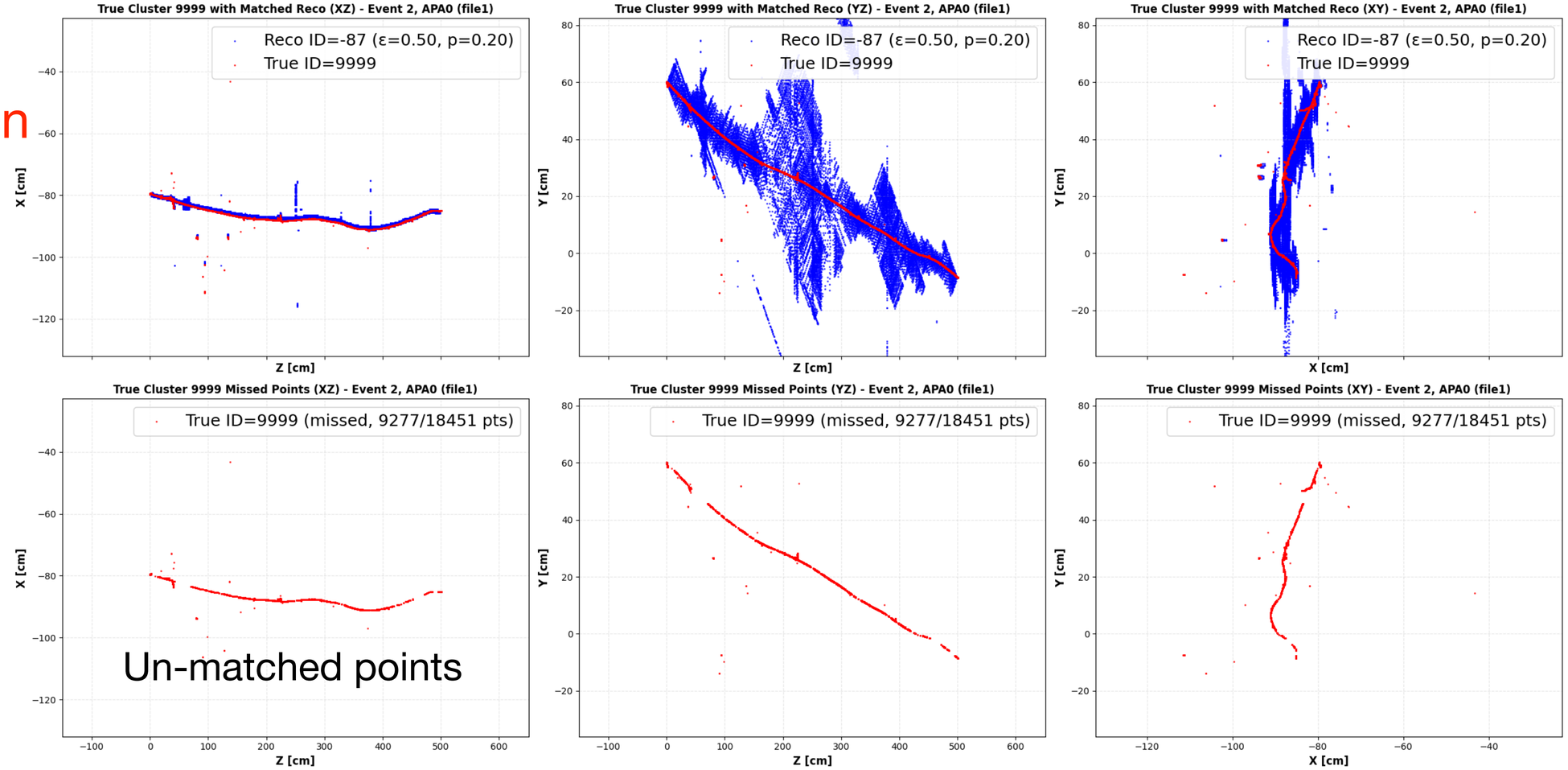
ε = 50%

Reason for low efficiency: Space-charge effect

Some cases of low efficiency: Event 2

Bee Link

Neutrino Interaction



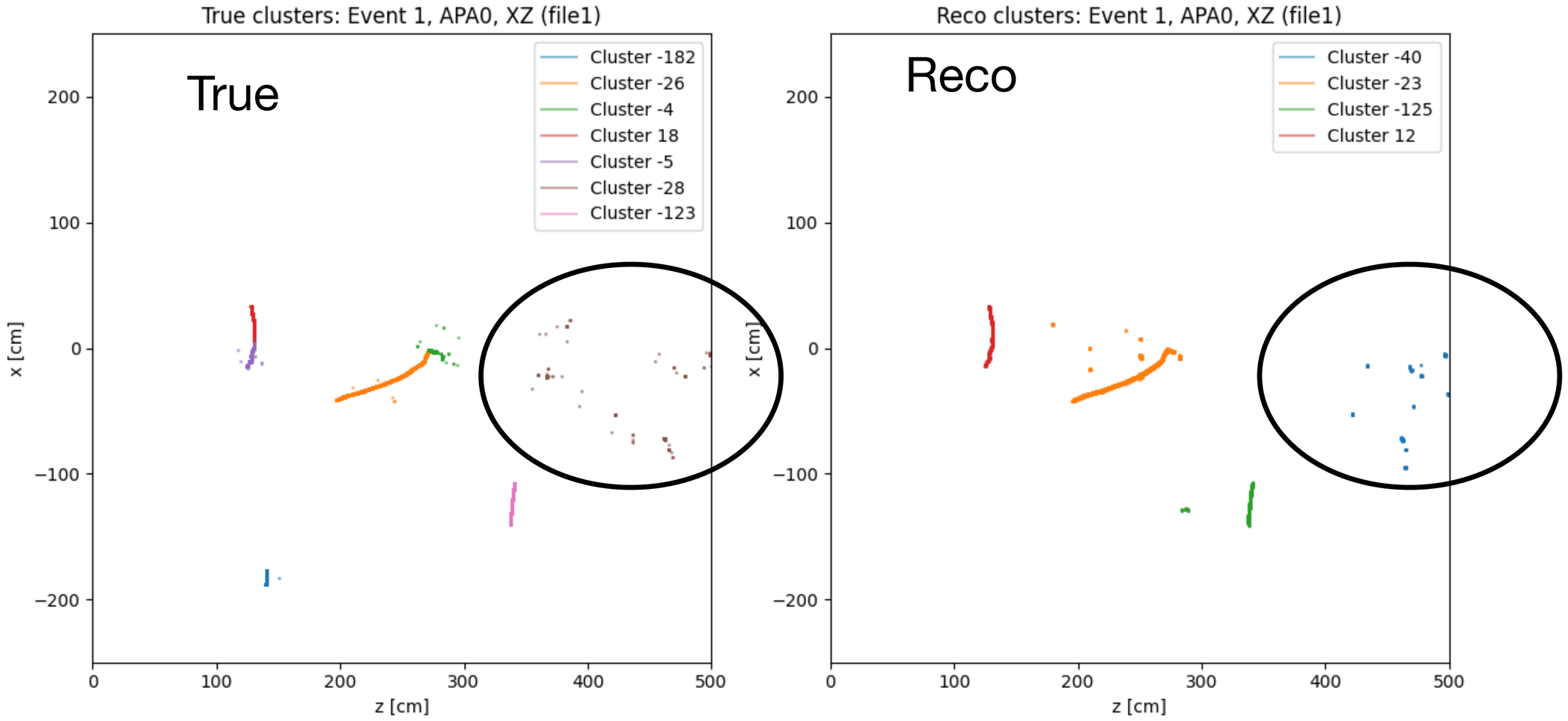
ε = 50%

Reason for low efficiency: Space-charge effect

Some cases of low efficiency: Event 1

Bee Link

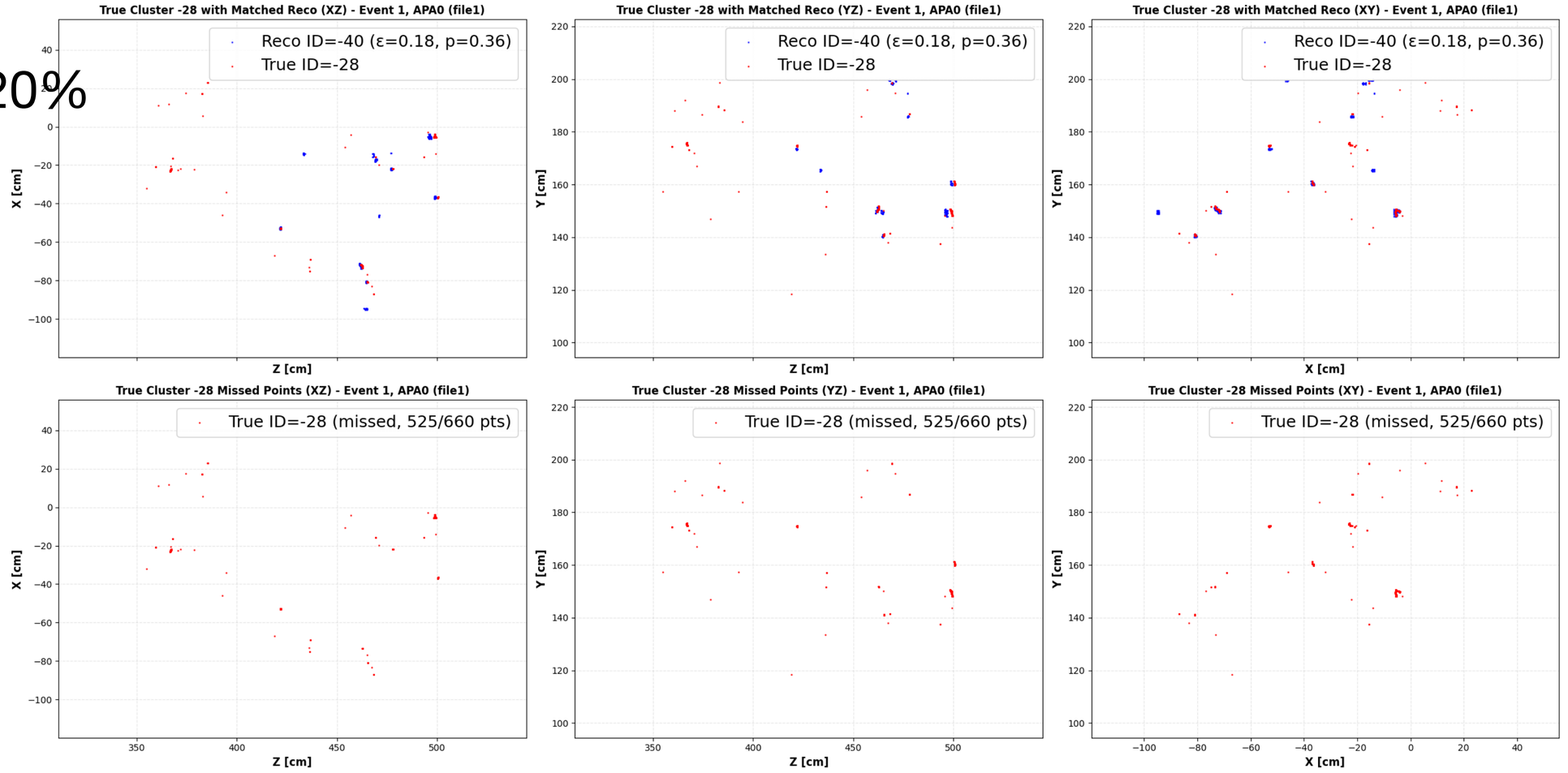
ε = 20%



Reason for low efficiency: Very few and scattered true points

Some cases of low efficiency: 1

ε = 20%



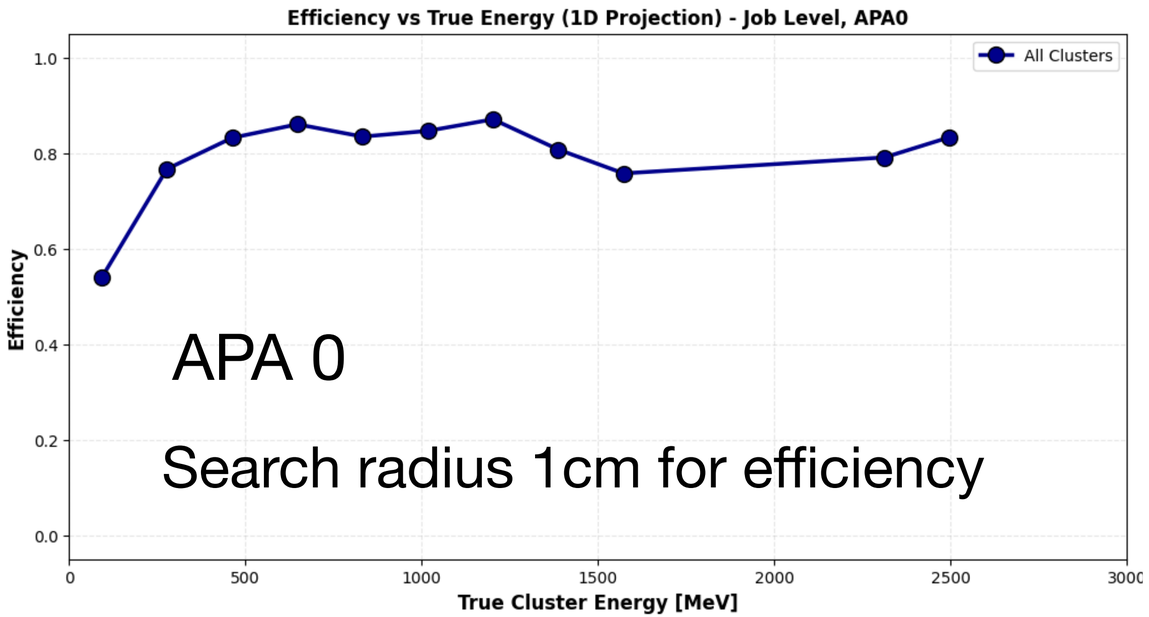
Reason for low efficiency: Very few and scattered true points

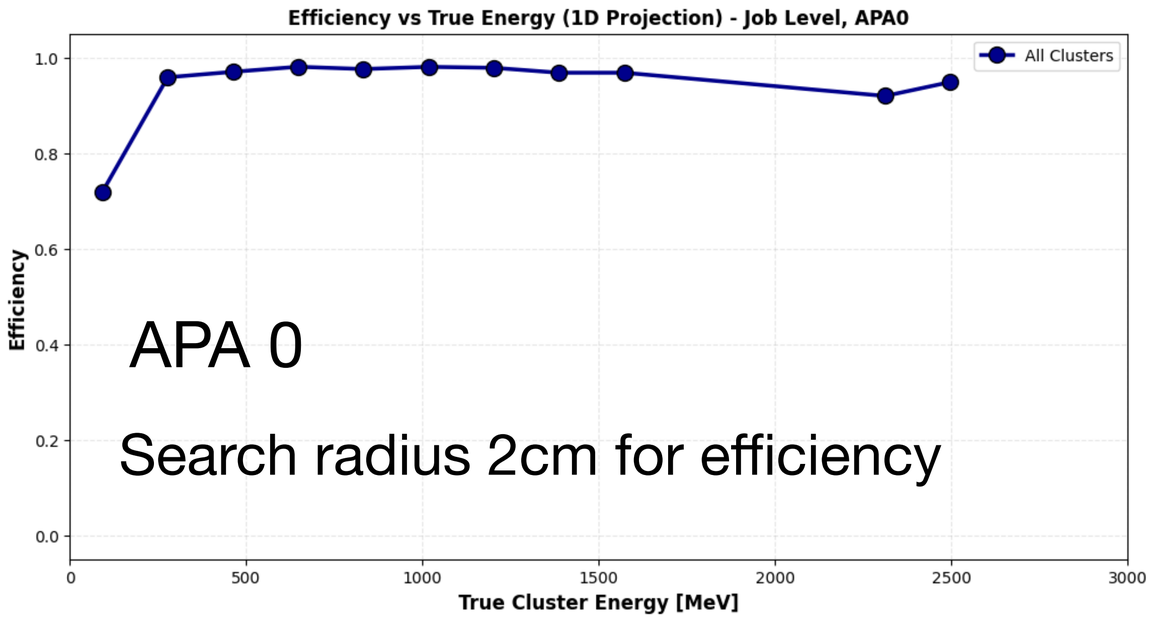
We need selections to identify and remove these kind of clusters

Space-Charge Effects

- I saw several examples of space-charge effects which cause drop in efficiency
- We know space charge effects are maximally within 1cm
- So, if we loosen our 1cm efficiency radius to 2cm then it would cover the space-charge effects which is a good test of our assumption that space-charge effects are causing a big drop in efficiency
- Once we apply space-charge corrections in wirecell reconstruction then we should revert back our efficiency-radius to 1cm

Efficiency as a function of True Cluster Energy

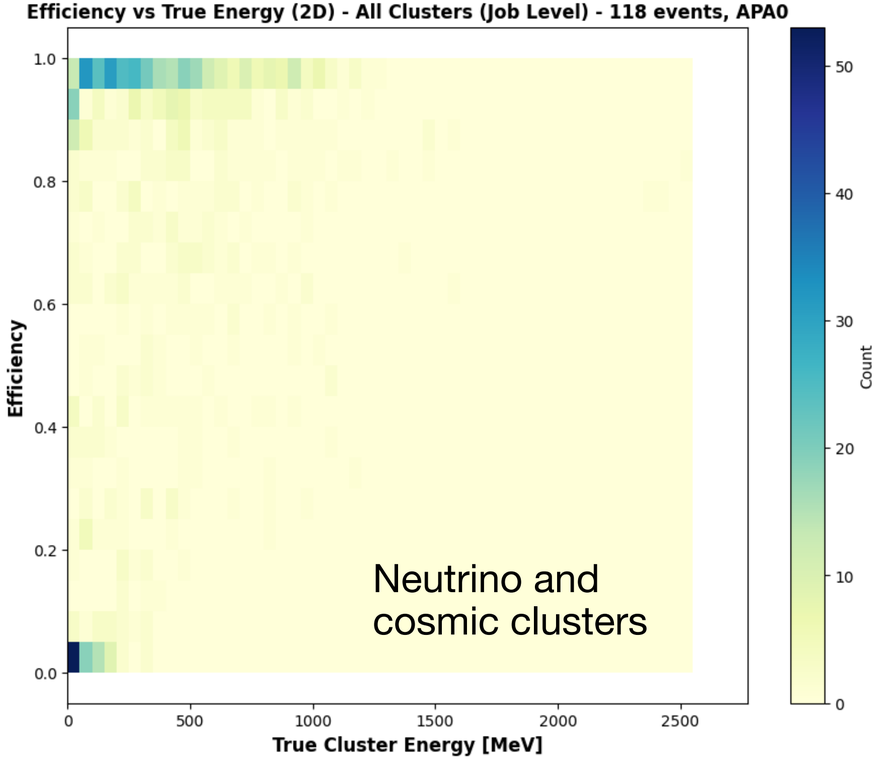




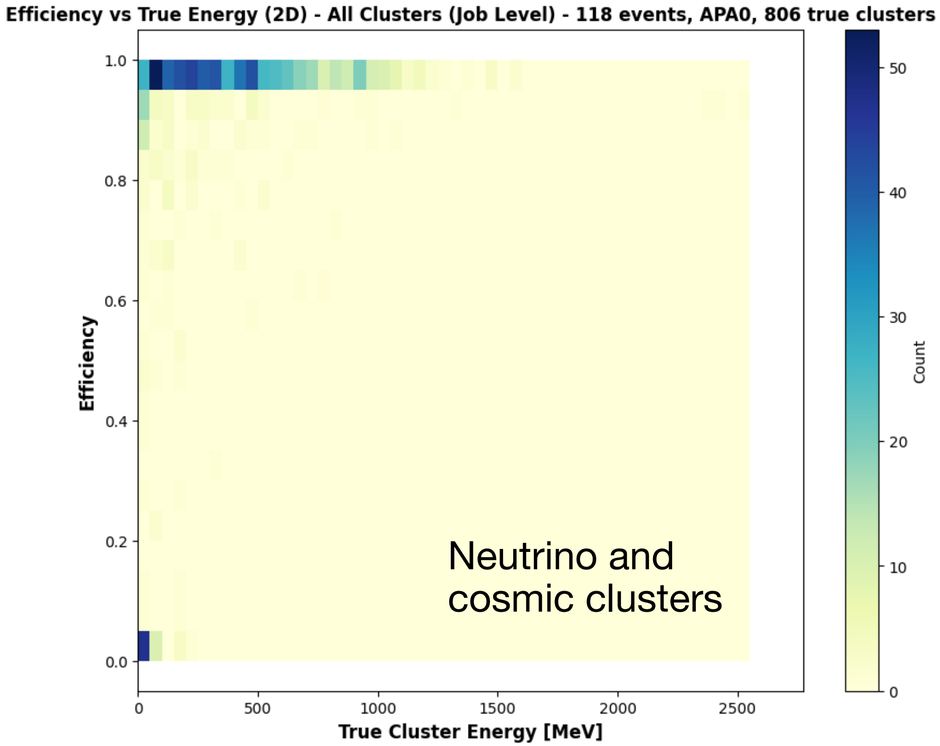
- As expected loosening efficiency-radius to 2cm from 1cm significantly improved efficiency
- This proves that the large drop in efficiency was due to space-charge effects

2D Efficiency as a function of True Cluster Energy

APA 0 Search radius 1cm for efficiency

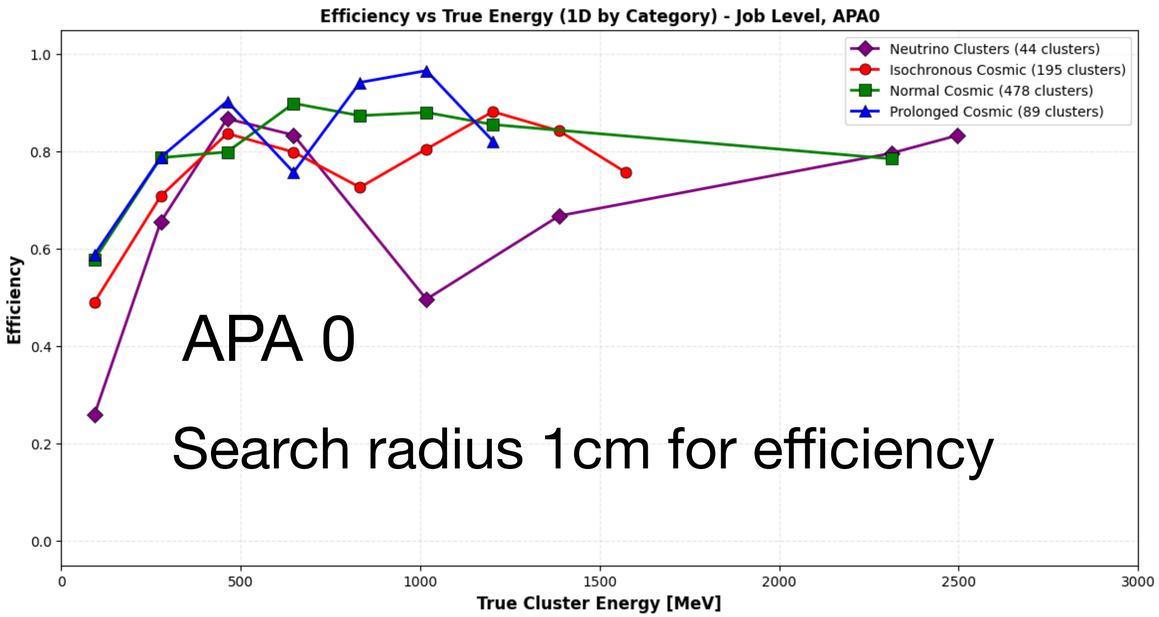


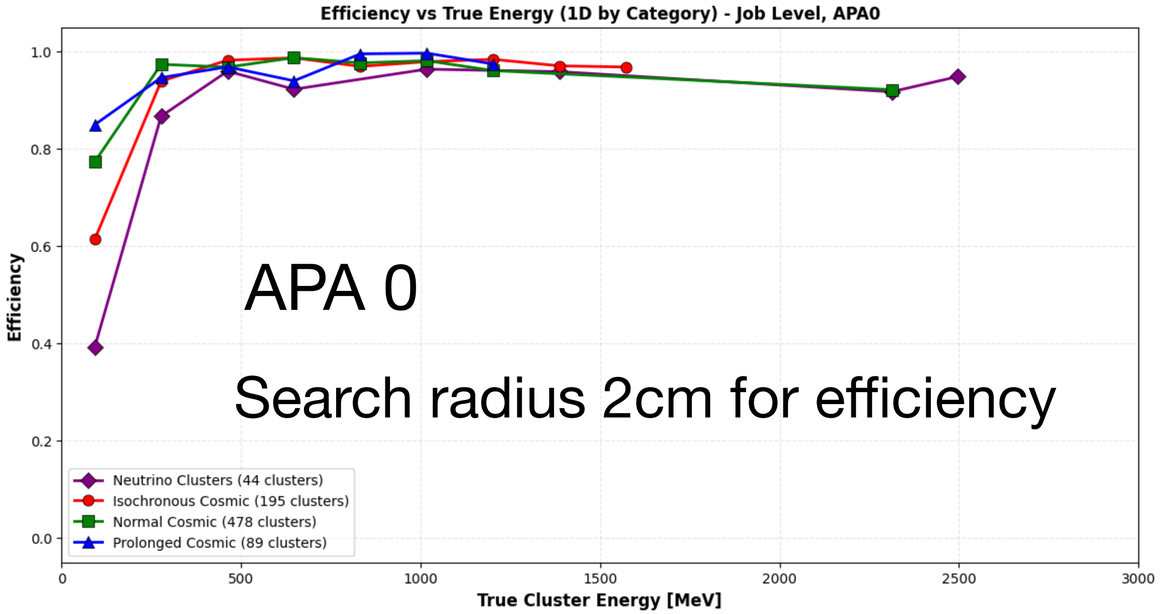
APA 0 Search radius 1cm for efficiency



- Significant improvement in efficiency
- More clusters closer to 100% efficiency which we would expect
- There are still some close to zero or very low efficiency

Efficiency as a function of True Cluster Energy

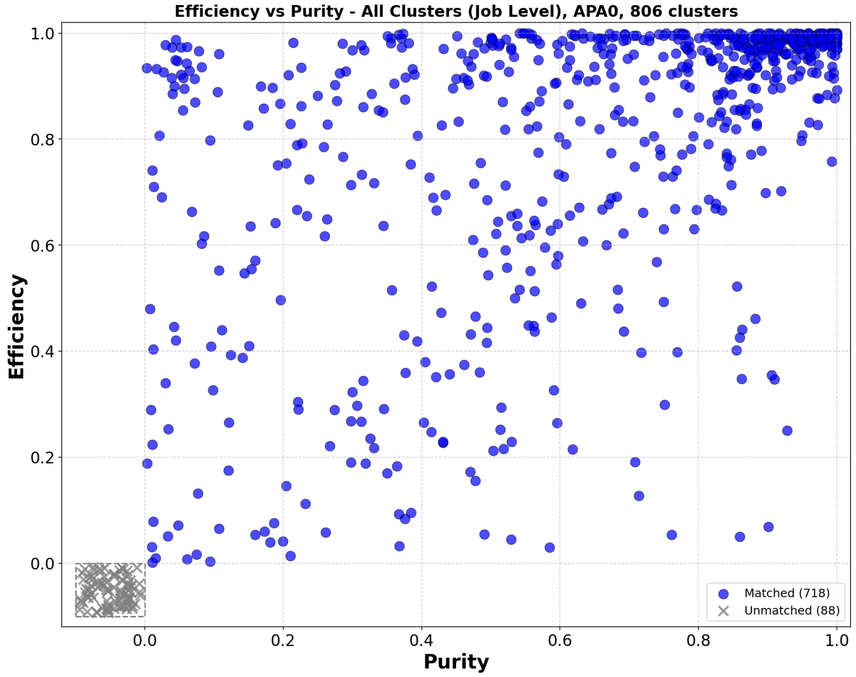




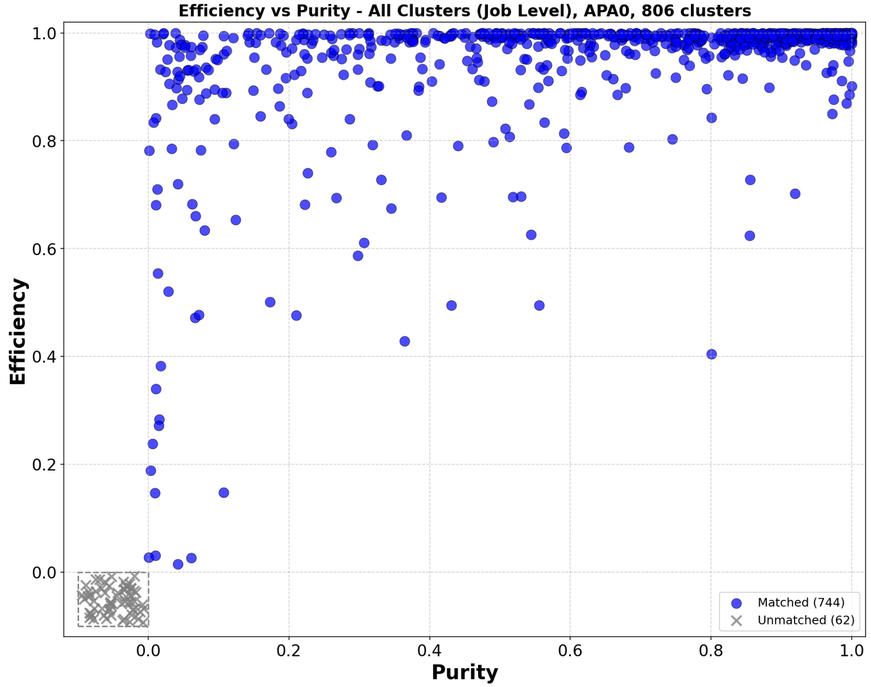
- Performance/efficiency of neutrinos, and all cosmic clusters also improve

Efficiency at Clustering

APA 0 1cm efficiency radius



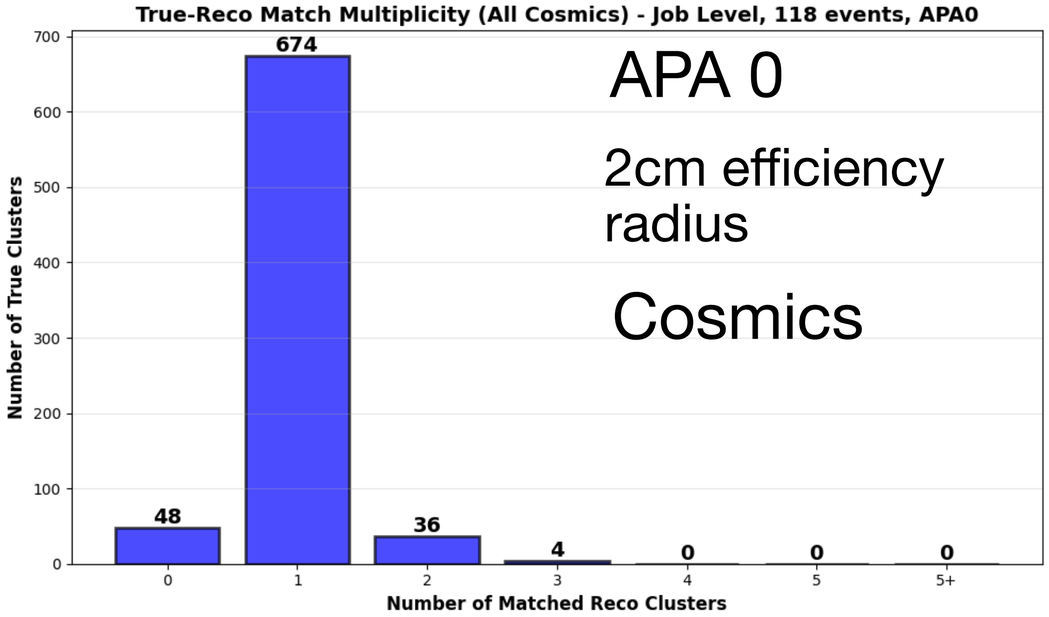
APA 0 2cm efficiency radius

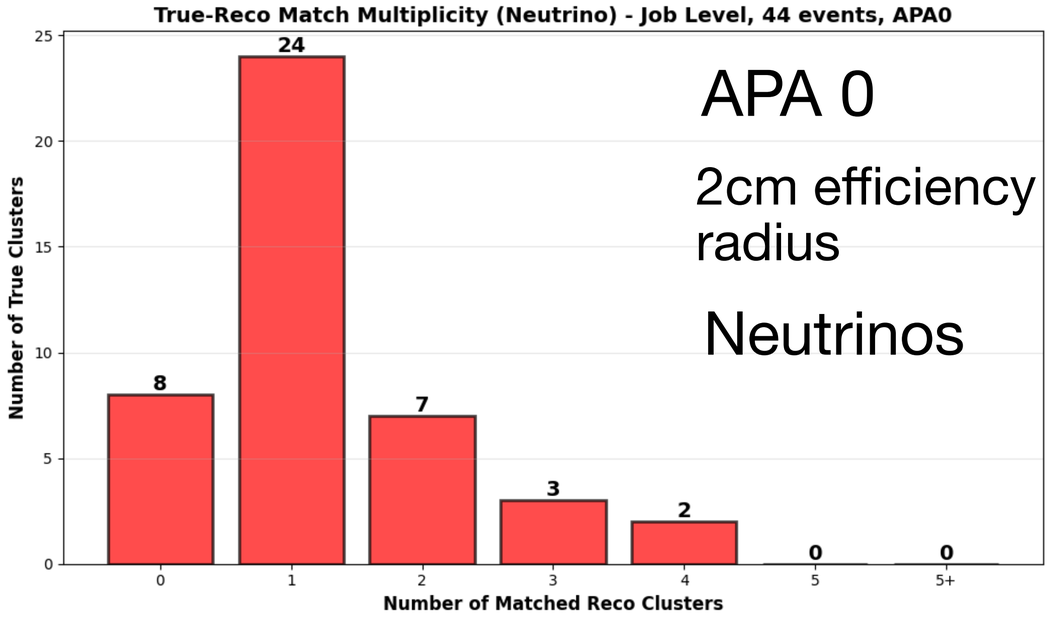


- To know the efficiency performance at clustering level, we pair one true cluster to one reco cluster
- If one true cluster matches with one reco cluster then all good but there could be cases that one true cluster matches with more than one true clusters
- In that case, we make true-reco pair by selecting the highest efficiency

- Each point in this plot is one true-reco pair
- As expected, after loosing search-radius to 2cm, to cover SCE, most of the pairs are concentrated at high efficiencies
- There is a fraction of true clusters unmatched to any reco cluster shown in bottom-left box. This needs investigation
- There are also clusters with low purity. This needs investigation also needs further investigation though many of those also come from isochronous clusters

True-Reco Matching





- These plots show how many true clusters did not find any or more than one reco clusters
- One-to-one match from true to reco is the best case scenario but we need to understand why some true clusters did not find any reco clusters
- Next I will show some examples, for neutrino interactions

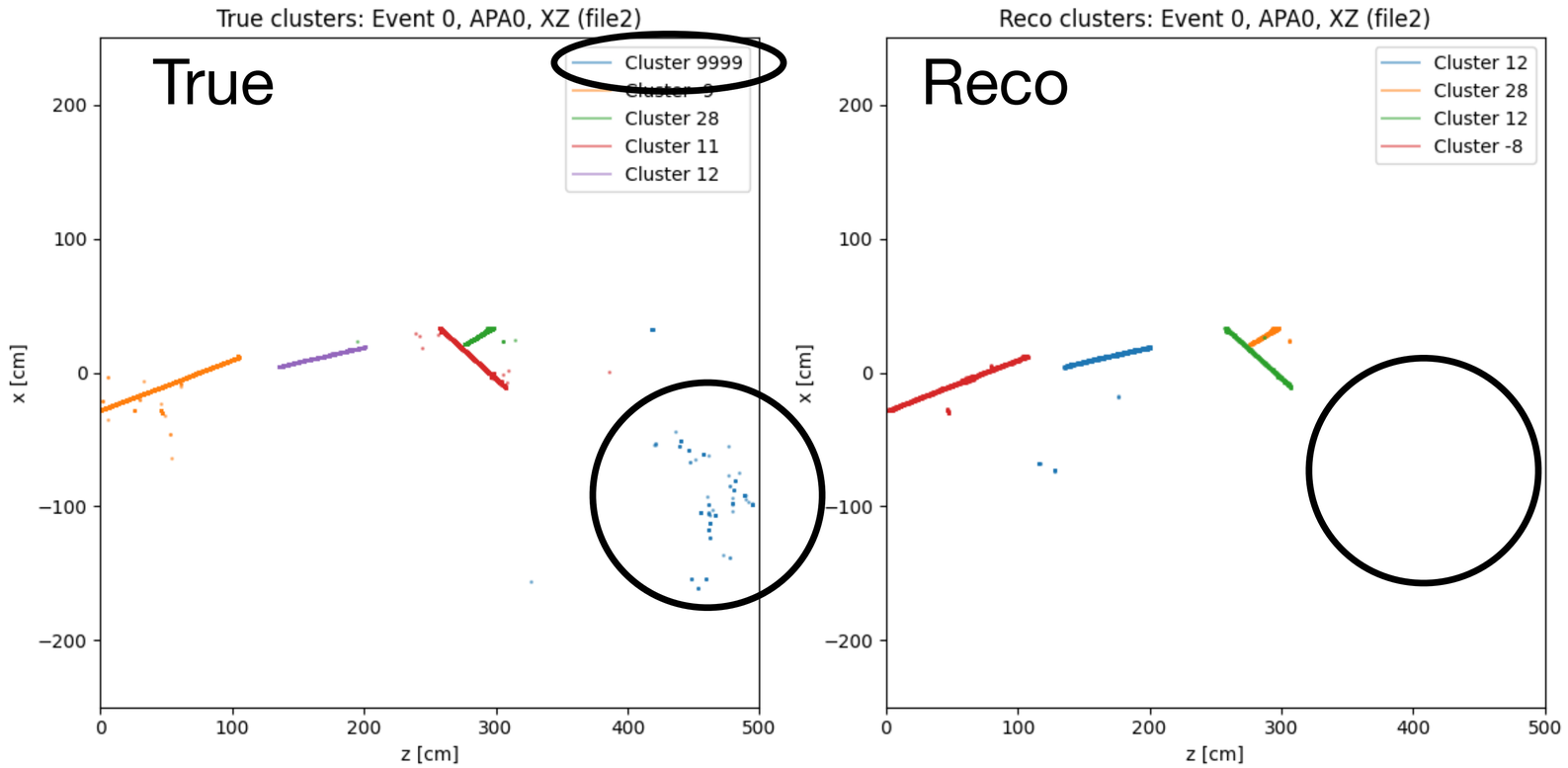
Examples of True Neutrinos not finding any reco clusters

8 examples

Example of True Neutrinos not finding any reco clusters

File 2: Event 0

2cm efficiency radius
APA 0

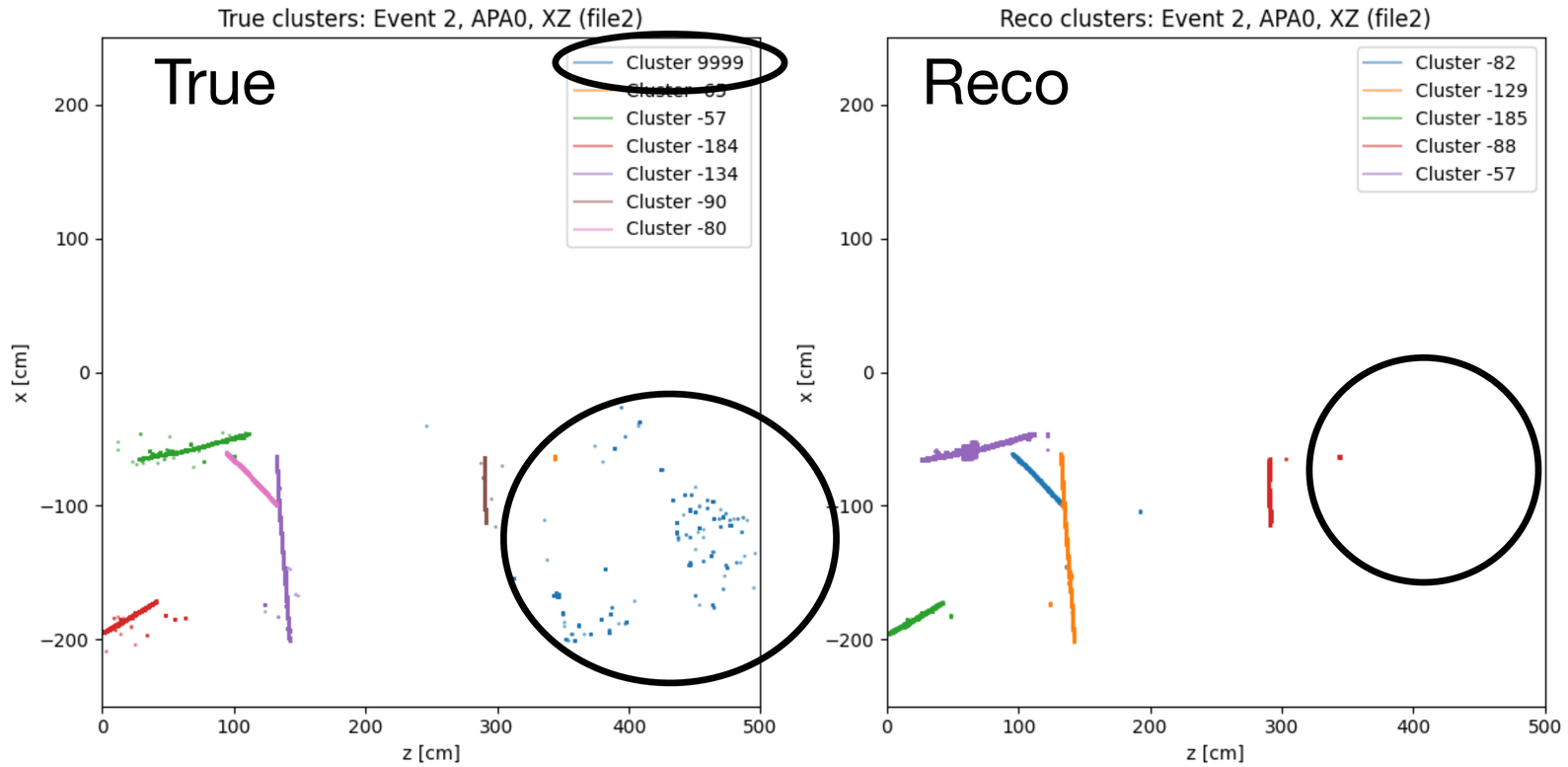


Bee Link

Example of True Neutrinos not finding any reco clusters

File 2: Event 2

2cm efficiency radius
APA 0

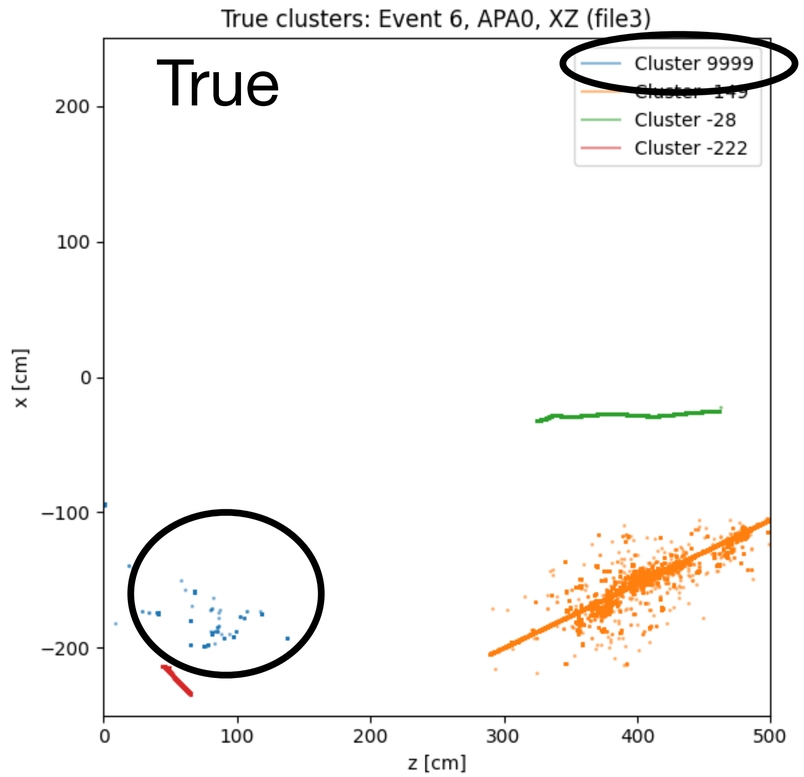


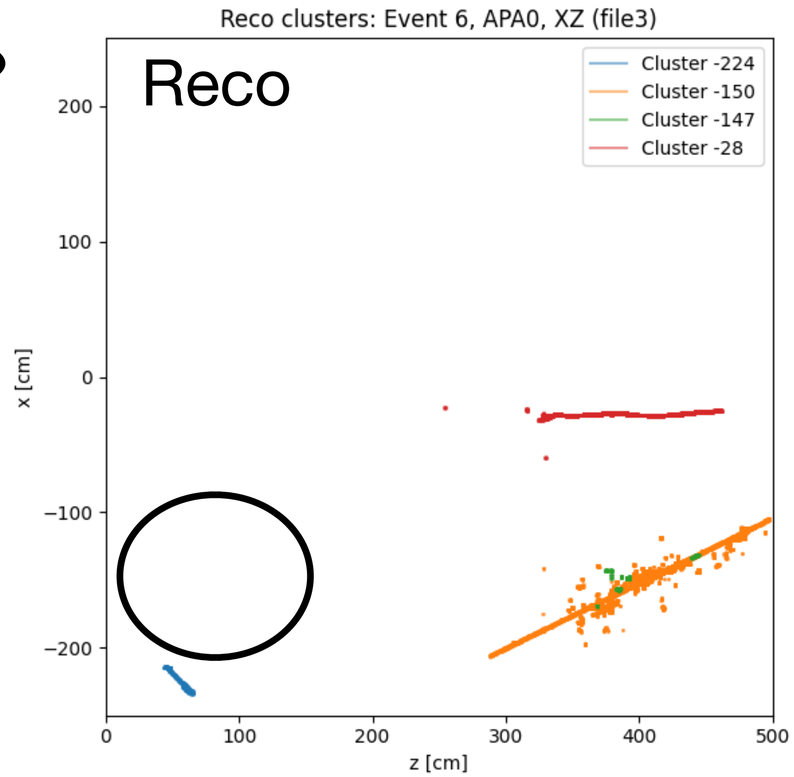
Bee Link

Example of True Neutrinos not finding any reco clusters

File 3: Event 6

2cm efficiency radius
APA 0





Bee Link

Other neutrino cases are also similar ones. I have them in backup slides

We need selections to identify and remove these kind of clusters

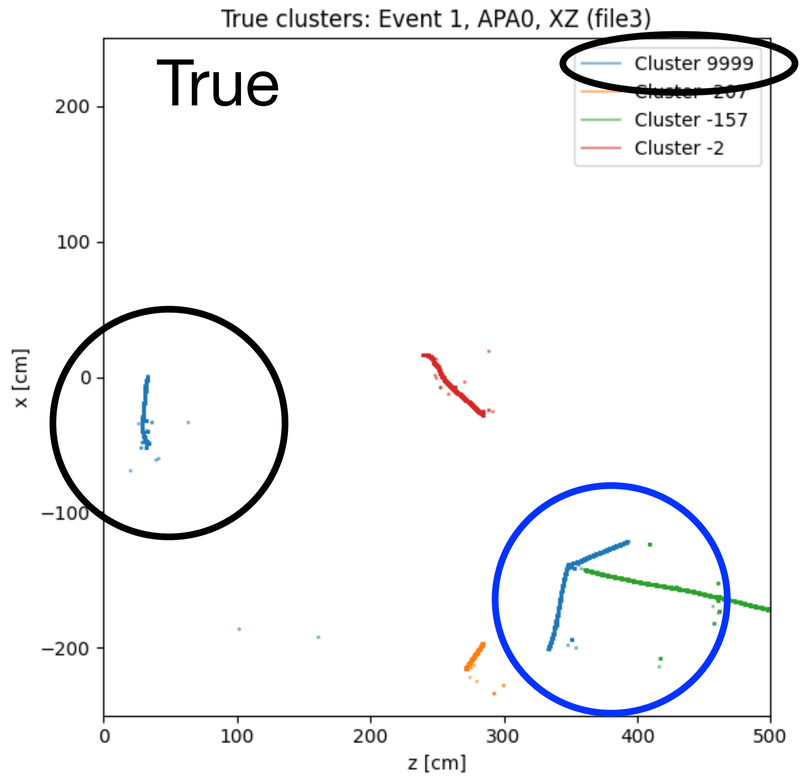
Examples of True Neutrinos finding 2 reco clusters

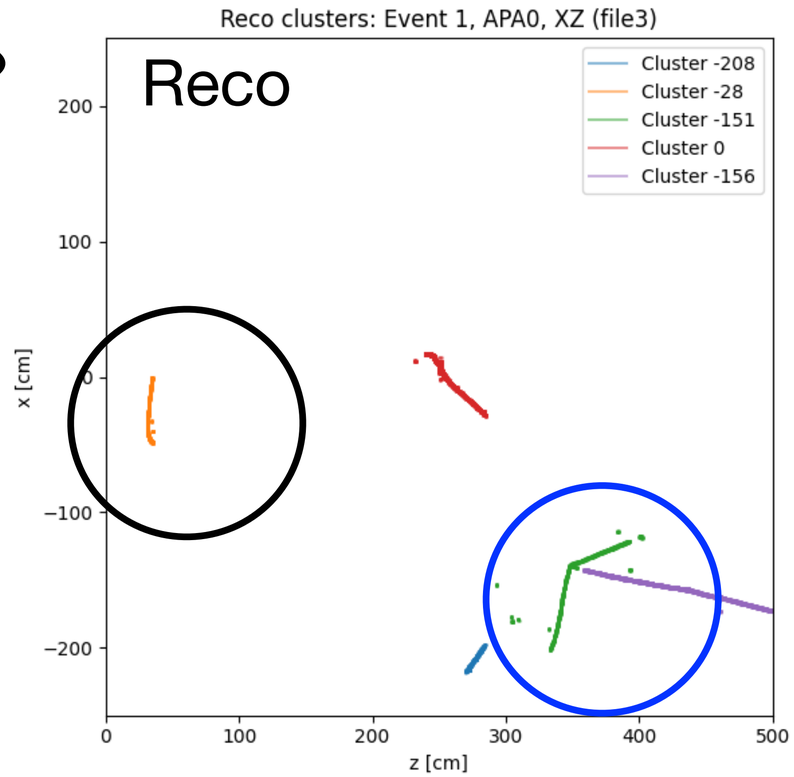
7 examples

Examples of True Neutrinos finding 2 reco clusters

File 3: Event 1

2cm efficiency radius
APA 0





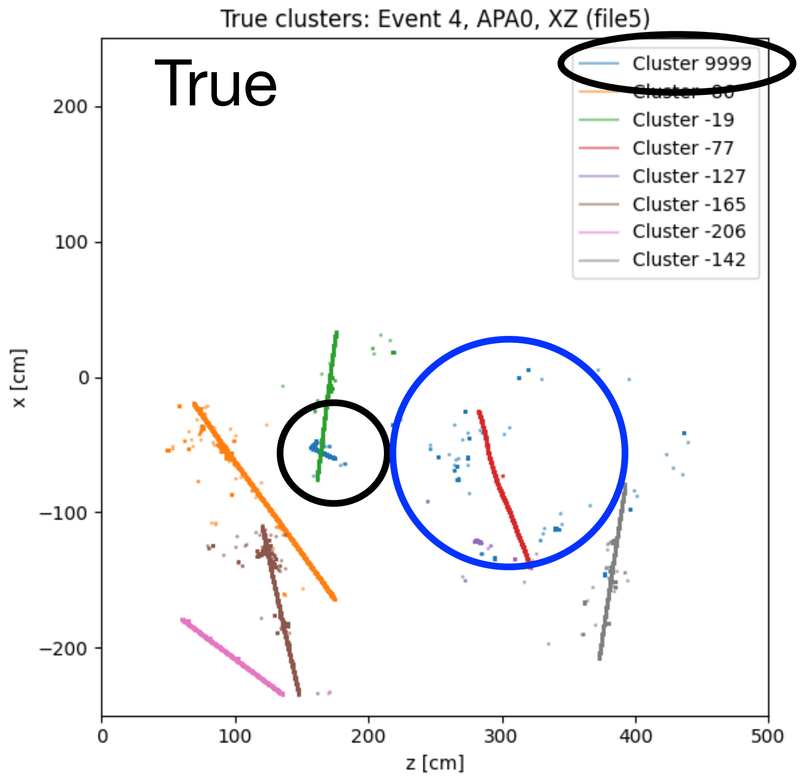
Bee Link

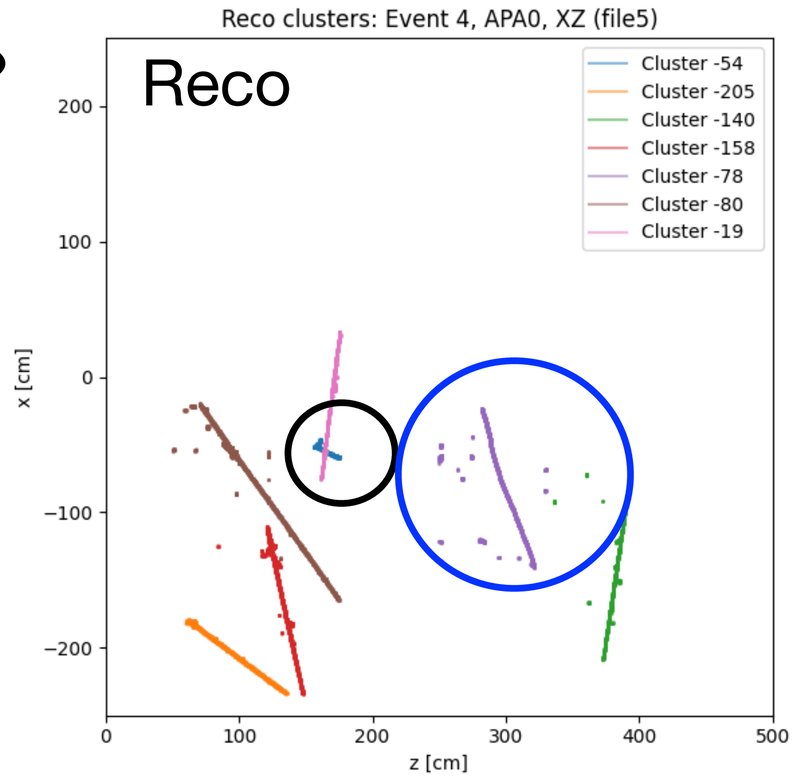
- Are there two neutrinos?
- Or is it one neutrino but separated by a distance?

Examples of True Neutrinos finding 2 reco clusters

File 5: Event 4

2cm efficiency radius
APA 0





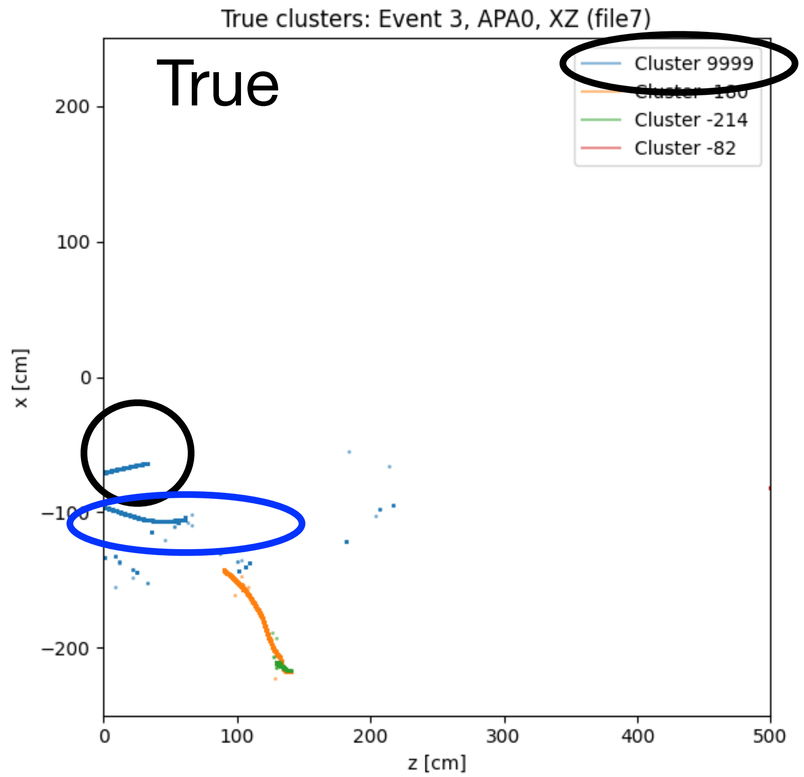
Bee Link

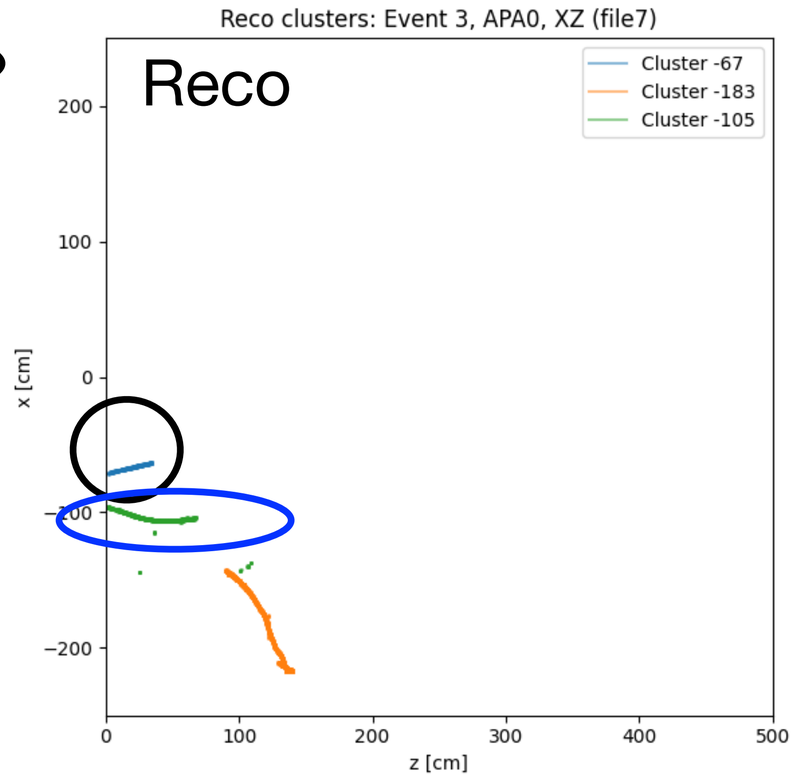
- Are there two neutrinos?
- Or is it one neutrino but separated by a distance?
- If two neutrinos, one is scattered

Examples of True Neutrinos finding 2 reco clusters

File 7: Event 3

2cm efficiency radius
APA 0





Bee Link

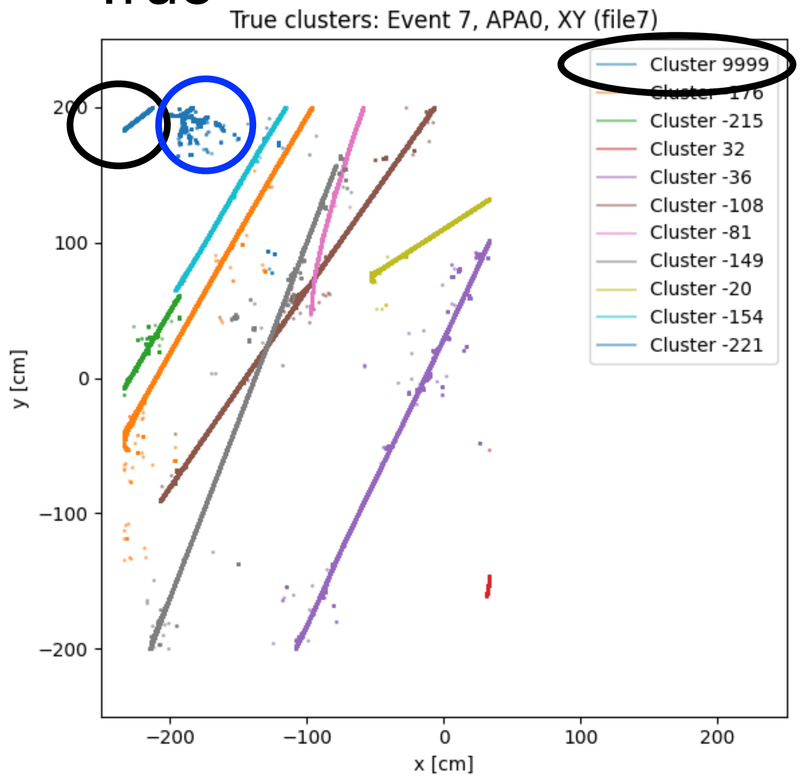
- Same: Are there two neutrinos?
- Or is it one neutrino but separated by a distance?

Examples of True Neutrinos finding 2 reco clusters

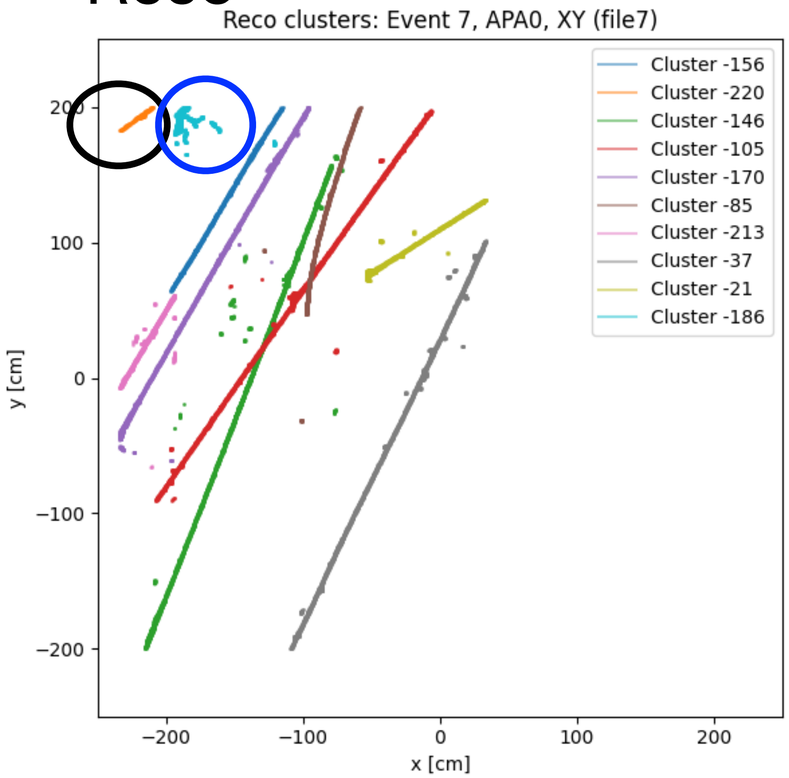
File 7: Event 7

2cm efficiency radius
APA 0

True



Reco



Bee Link

- Same: Are there two neutrinos?
- Or is it one neutrino but separated by a distance?

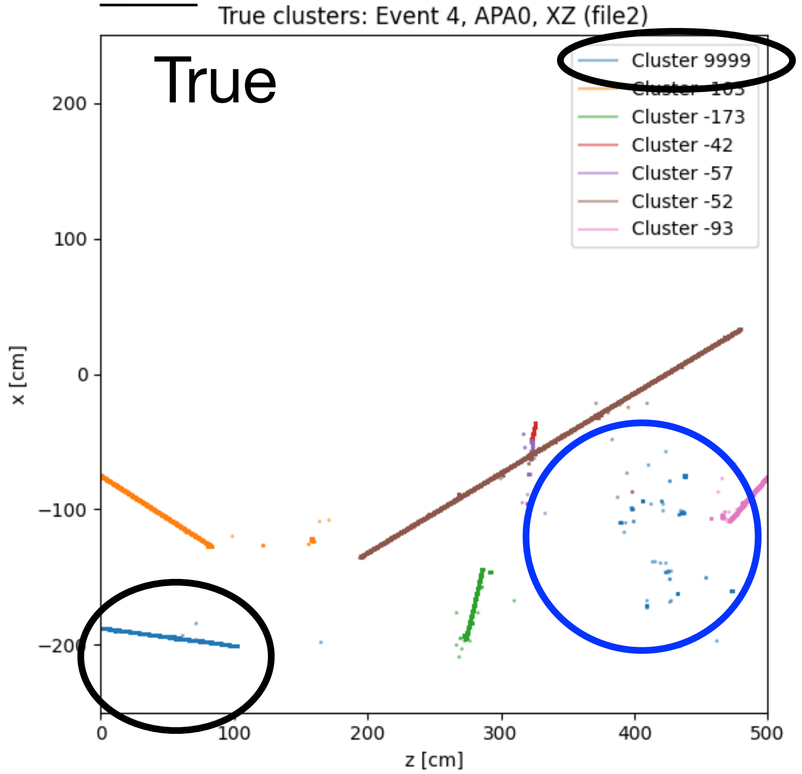
Examples of True Neutrinos finding 3 reco clusters

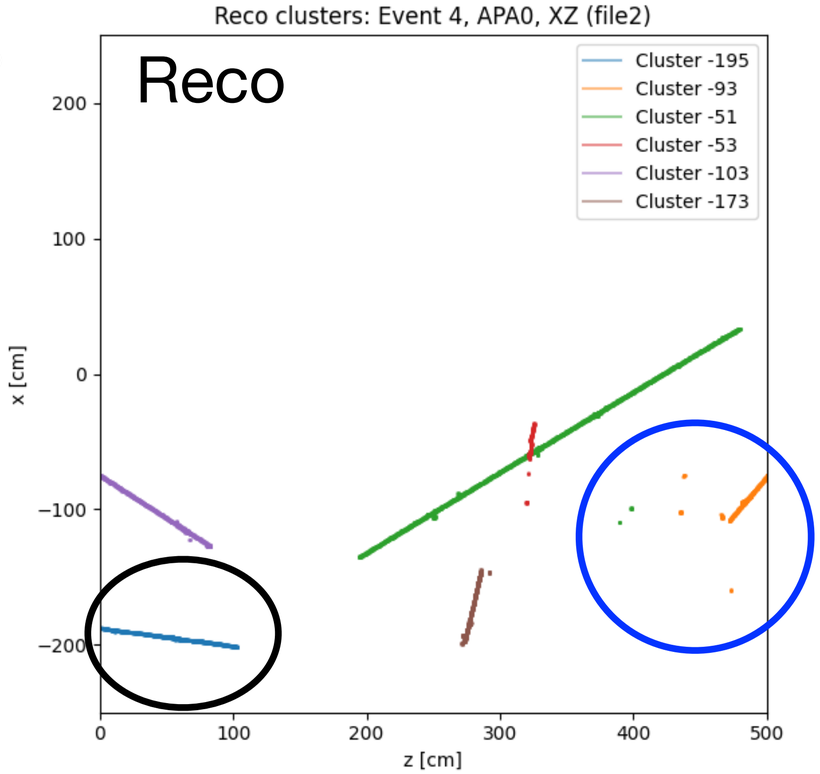
3 examples

Examples of True Neutrinos finding 3 reco clusters

File 2: Event 4

Bee Link

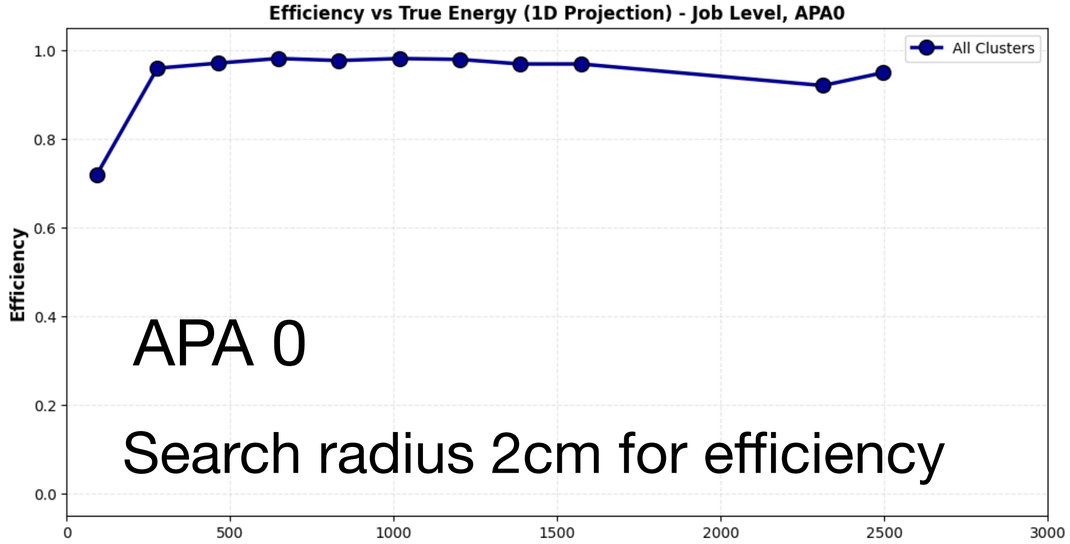


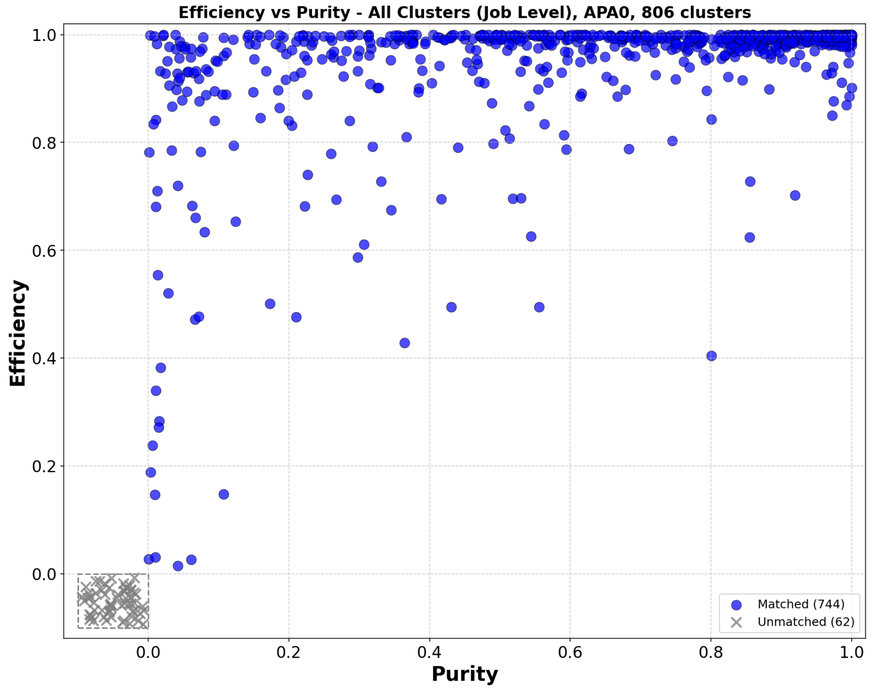


2 neutrinos and one of them overlapped tiny bit with another reco clusters

Conclusions

- Overall seems we are in a very good state
- There are some corner cases that need attention
- We are planning to present an update in the Simulation/ Reconstruction meeting soon to follow up our reconstruction work after the collaboration meeting





Backup

Revision: Current Selections for Cluster Evaluation

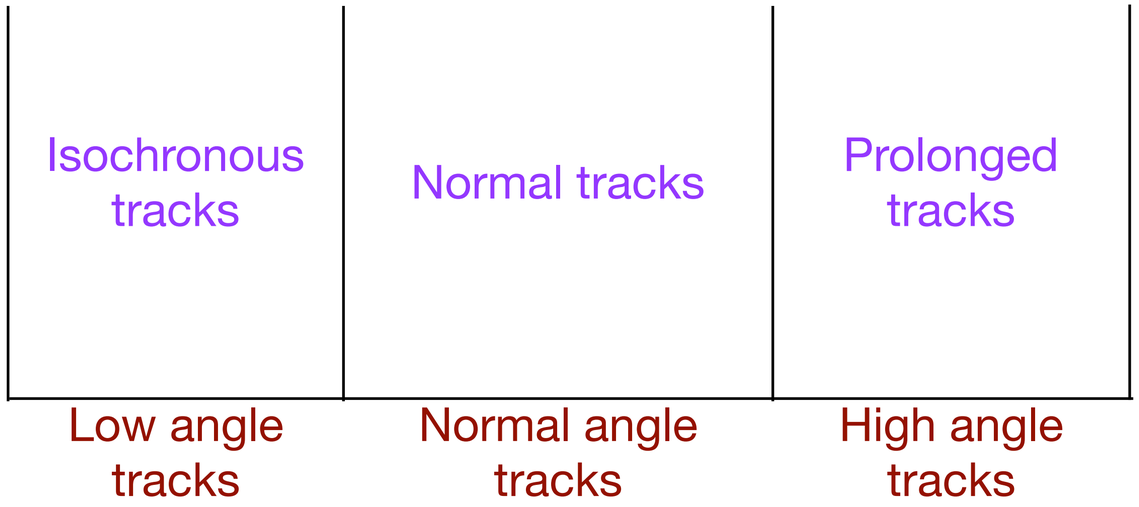
Selections	Accepted Values	Description	Applied? Yes or No
Energy Cut	E > 10 MeV	True clusters below 10 MeV are removed from evaluation	
Min true point cut	N > 200	Removes true and reco clusters with fewer than 200 points	
Wire readout sensitive yz plane cut	-200 cm < y < 200 cm 0.15 cm < y < 500.85 cm	Removes true and reco points output wirecell sensitive regions in y and z	
Time window cut	-205 μs < t < 1508.5 μs	Removes true points outside time window cut. Applied at CellTree level Reco points are already within time window cuts so we apply this to true clusters	
Dead Area cut		Reco points are NOT reconstructed within dead wire areas so we also need to remove true points from dead areas for correct reco-true comparisons	

Revision: Current Selections for Cluster Evaluation

Selections	Accepted Values	Description	Applied? Yes or No
Energy Cut	E > 10 MeV	True clusters below 10 MeV are removed from evaluation	All Applied
Min true point cut	N > 200	Removes true and reco clusters with fewer than 200 points	
Wire readout sensitive yz plane cut	-200 cm < y < 200 cm 0.15 cm < y < 500.85 cm	Removes true and reco points output wirecell sensitive regions in y and z	
Time window cut	-205 μs < t < 1509 μs	Removes true points outside time window cut. Applied at CellTree level Reco points are already within time window cuts so we apply this to true clusters	
Dead Area cut		Reco points are NOT reconstructed within dead wire areas so we also need to remove true points from dead areas for correct reco-true comparisons	

Revision: Plans for Isochronous and Prolonged Tracks

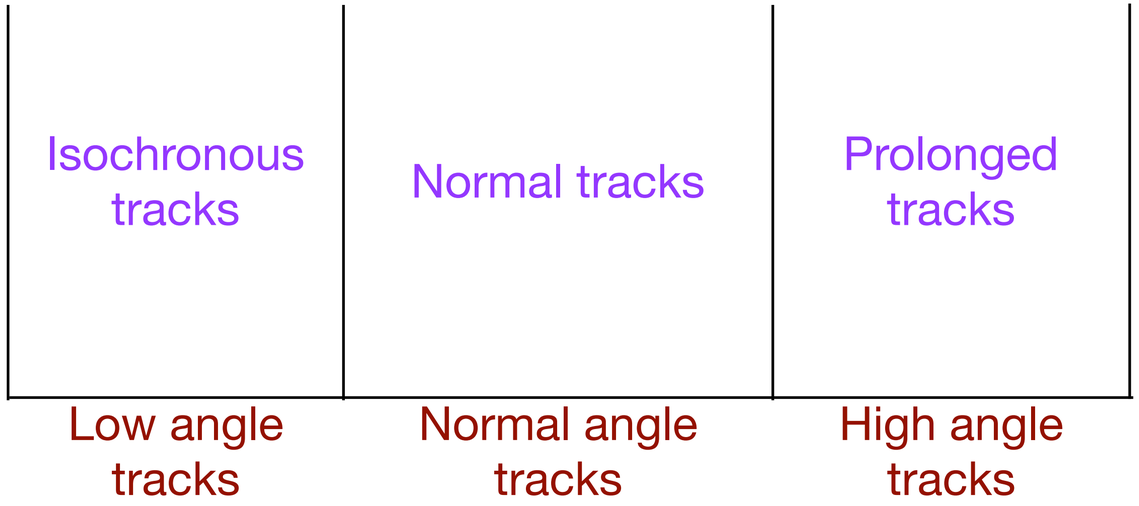
- Both isochronous and prolonged tracks affect the evaluation
- Isochronous tracks affect the purity due to their spread
- Prolonged tracks affect efficiency because sometime they show gaps irrespective of dead area
- For evaluation purposes, we need to separate isochronous, and prolonged tracks from rest of the tracks



The idea is to divide tracks into low, high and normal angle bins and then do evaluation

Revision: Plans for Isochronous and Prolonged Tracks

- Both isochronous and prolonged tracks affect the evaluation
- Isochronous tracks affect the purity due to their spread
- Prolonged tracks affect efficiency because sometime they show gaps irrespective of dead area
- For evaluation purposes, we need to separate isochronous, and prolonged tracks from rest of the tracks



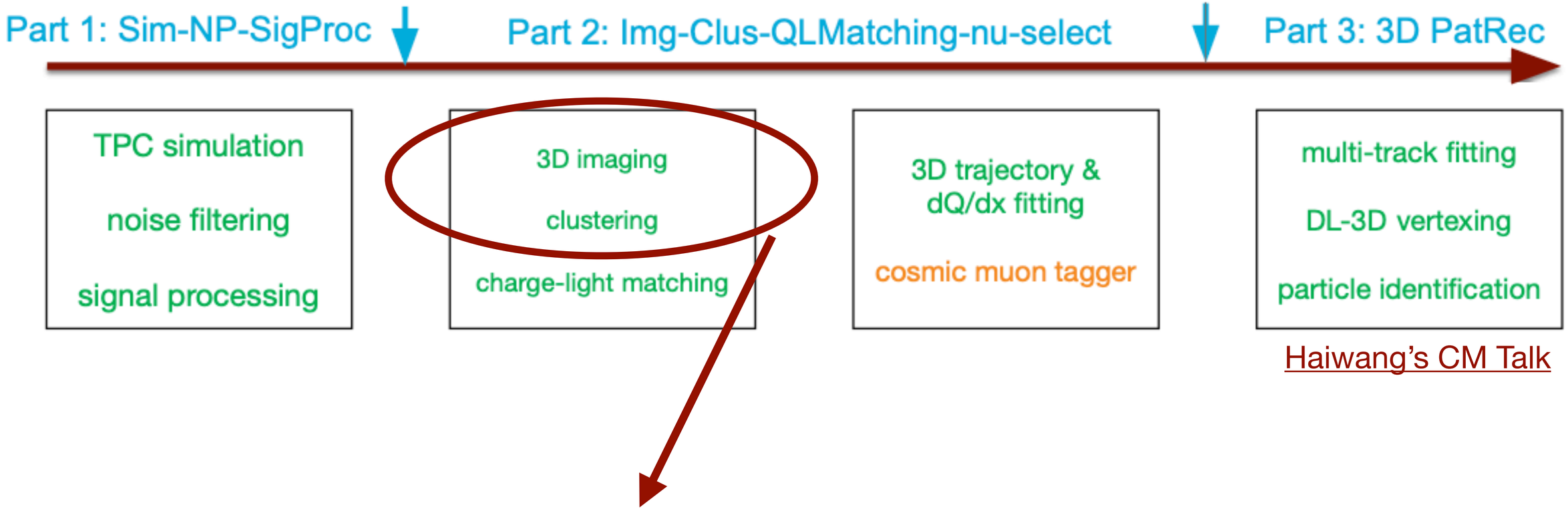
The idea is to divide tracks into low, high and normal angle bins and then do evaluation

All Applied

Efficiencies

- There are still some clusters that have low efficiencies so we will go through them to understand the reasons

Structure of Overall Progress



Haiwang's CM Talk

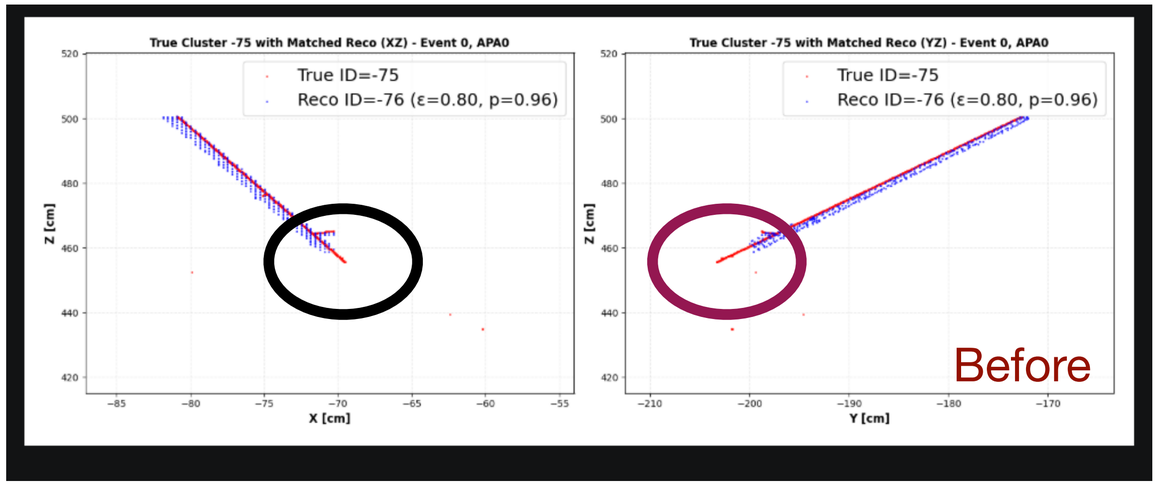
- Currently focussed on the imaging+clustering evaluation (scanning and numerical evaluation)
- Evaluation of charge-light matching after this

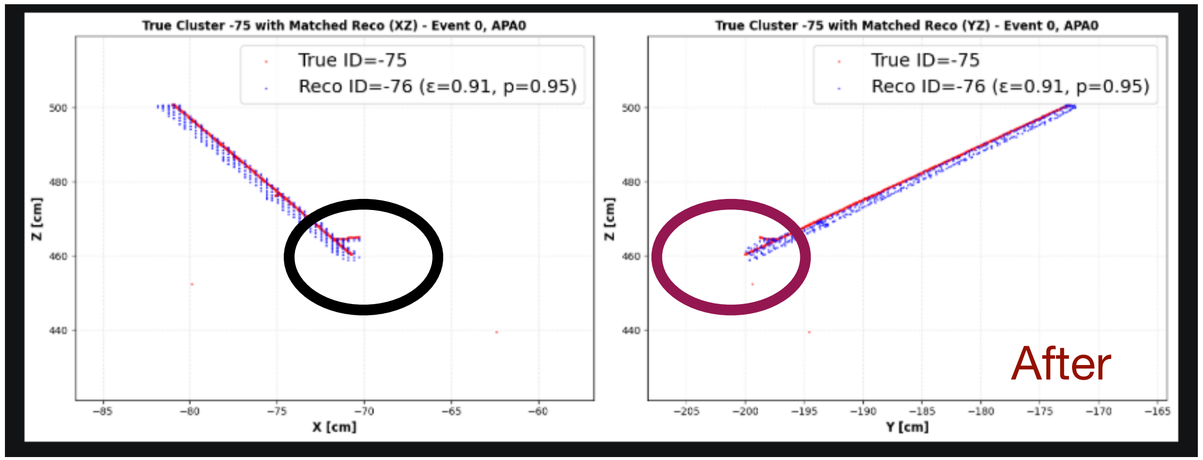
Revision: Current Selections for Cluster Evaluation

Selections	Accepted Values	Description	Applied? Yes or No
Energy Cut	E > 10 MeV	True clusters below 10 MeV are removed from evaluation	Yes
Min true point cut	N > 200	Removes true and reco clusters with fewer than 200 points	Yes
Wire readout sensitive yz plane cut	-200 cm < y < 200 cm 0.15 cm < y < 500.85 cm	Removes true and reco points output wirecell sensitive regions in y and z	Yes
Time window cut	-205 μs < t < 1509 μs	Removes true points outside time window cut. Applied at CellTree level Reco points are already within time window cuts so we apply this to true clusters	Yes
Dead Area cut		Reco points are NOT reconstructed within dead wire areas so we also need to remove true points from dead areas for reco-true comparisons	Yes

Wire readout sensitive yz plane cut

Ymin: -200 cm, Ymax: 200 cm, Zmin: 0.15 cm, Zmax: 500.85 cm (a bit smaller than detector active volume Y-Z)





	Efficiency	Purity
Before	**80%**	96%
After	**91%**	95%

Conclusion: efficiency increased by ~10%

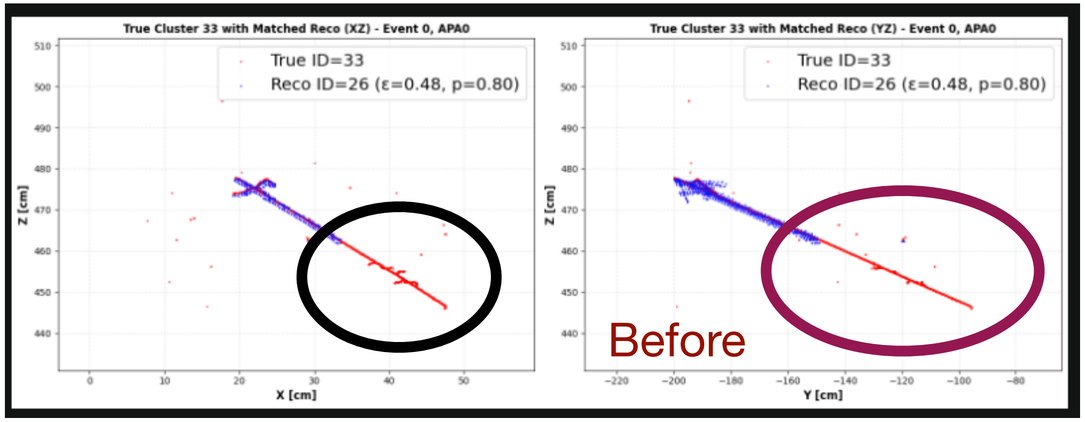
Caveat: This selection is only applied at evaluation level not at CellTree (True jobs) not at img-clustering (Reco jobs)

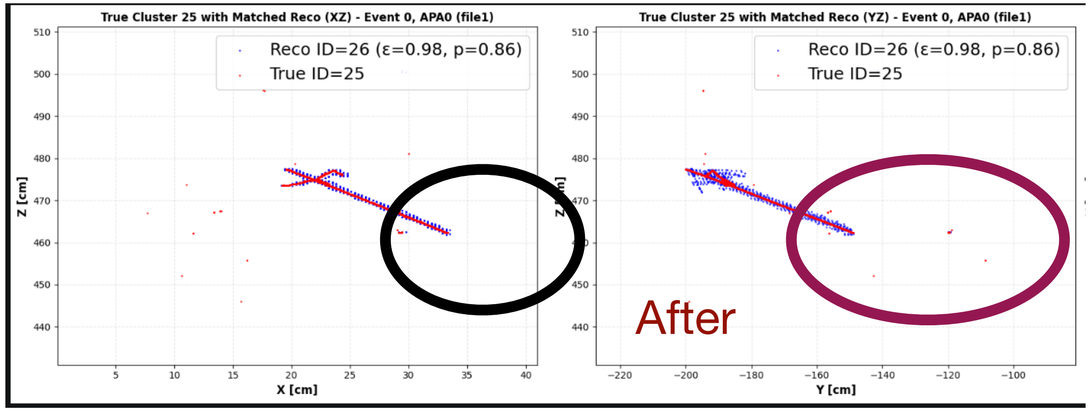
*More info in google-sheet for yz-plane cut

Time Window Cut

Removes true points outside time window cut

Reco points are already within time window cuts so we apply this to true clusters





	Efficiency	Purity
Before	48%	80%
After	98%	86%

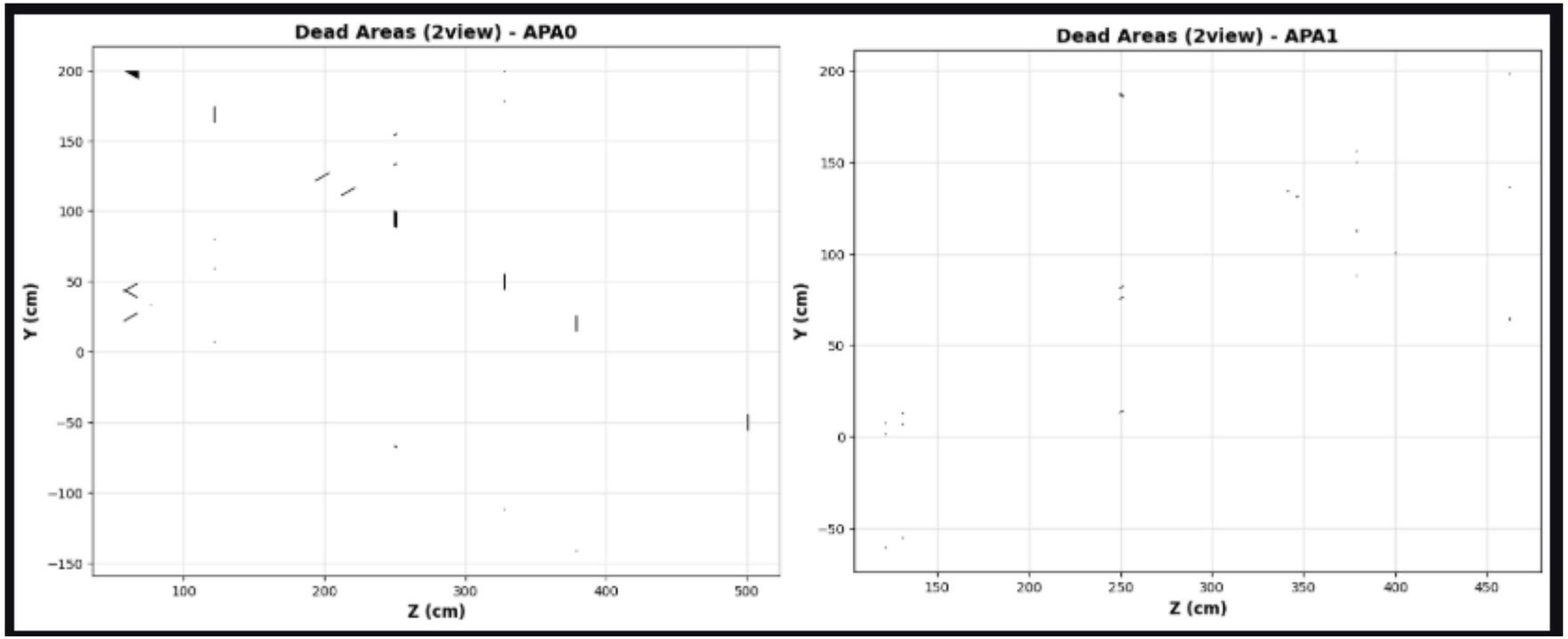
Conclusion: **efficiency highly improved.** Improvement depends on how much true was outside time window cut

Caveat: Fhicl level change for CellTree (true) job

*More info in google-sheet for time window cut

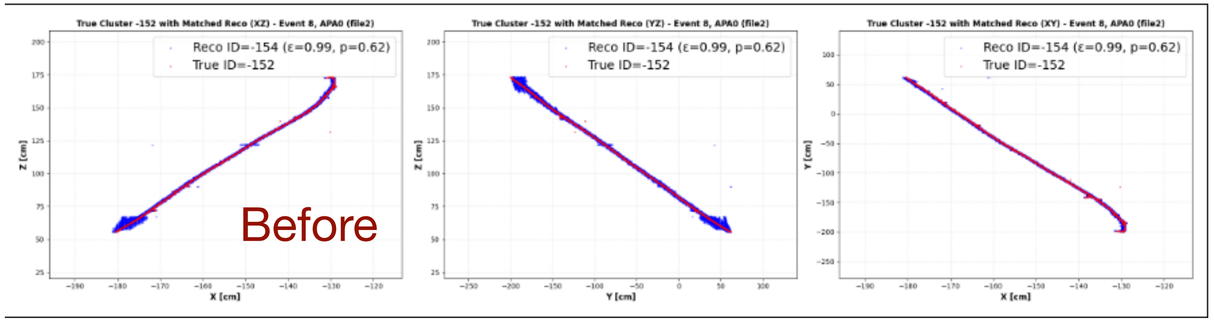
Dead Area Cut

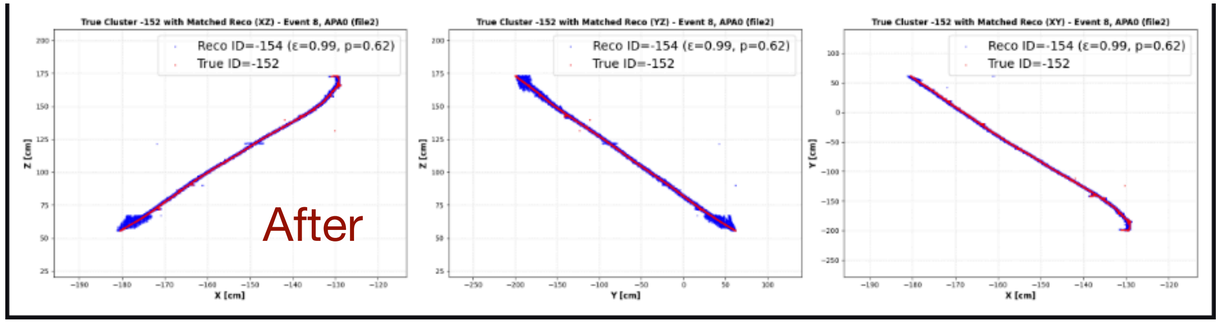
Reco points are NOT reconstructed within dead wire areas so we also need to remove true points from dead areas for reco-true comparisons



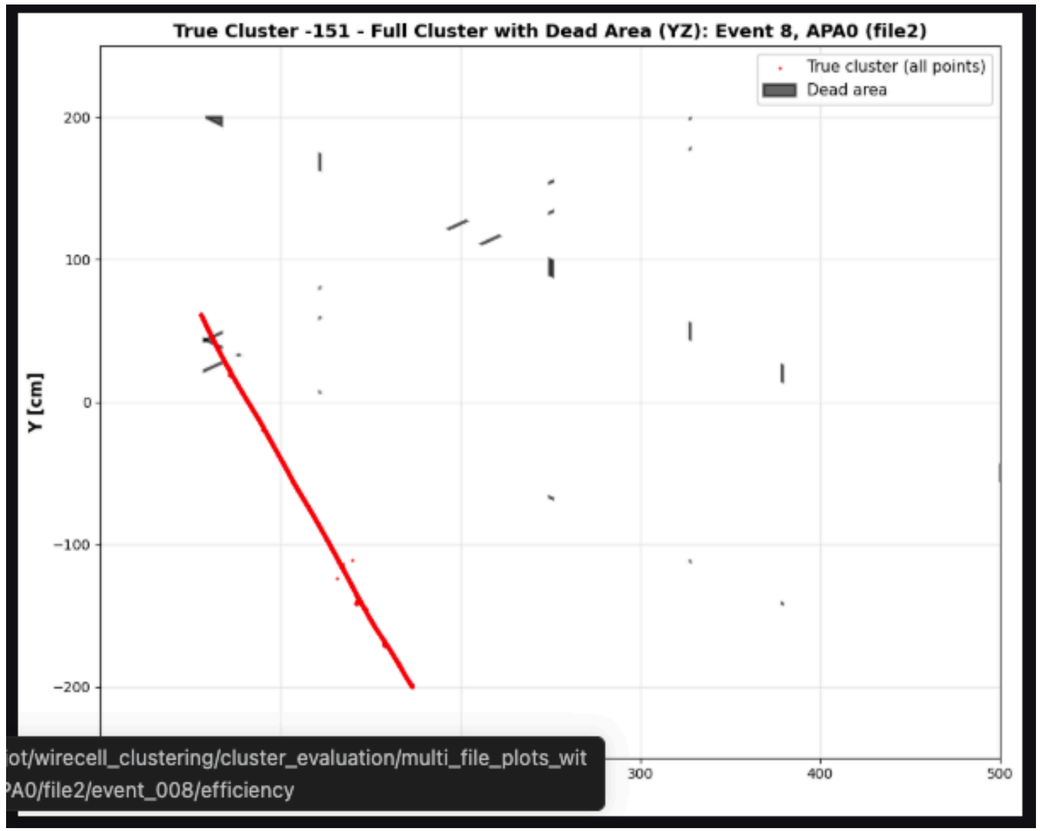
For 2view active, the dead area regions are very small, so the improvement on efficiency would be small but it is important to exclude true points in dead area for evaluation purpose

Dead Area Cut





	Efficiency	Purity
Before	99%	62%
After	99%	62%

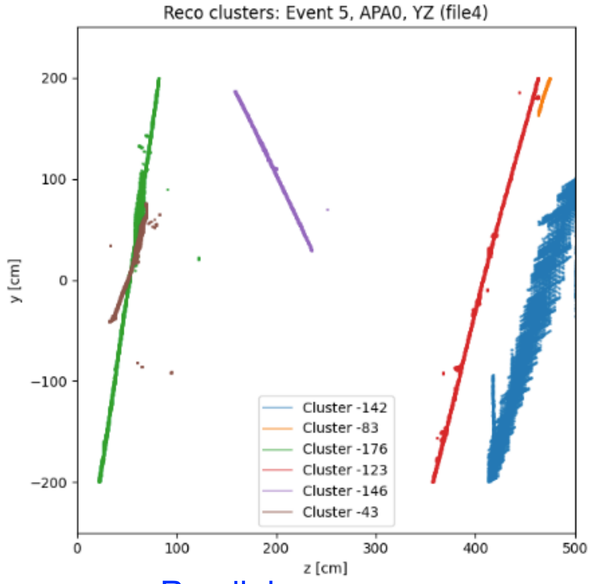


Conclusion: as expected not much effect on efficiency

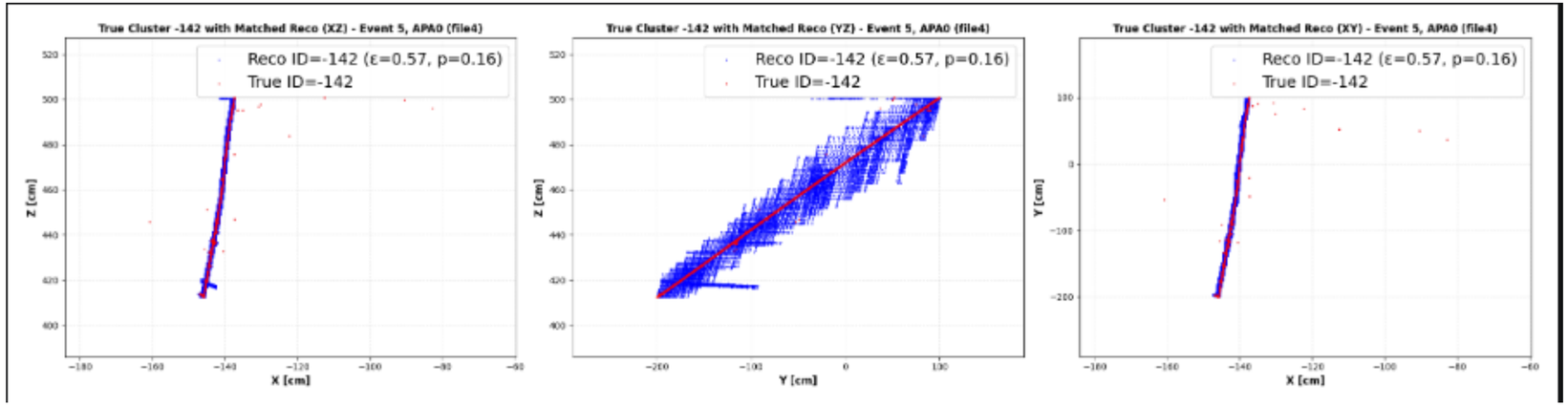
Caveat: Dead area cut doe not affect 2view active cases but would be very effective for old 3 view cases

Plans for Isochronous and Prolonged Tracks

- Both isochronous and prolonged tracks affect the evaluation
- Isochronous tracks affect the purity due to their spread



Bee link



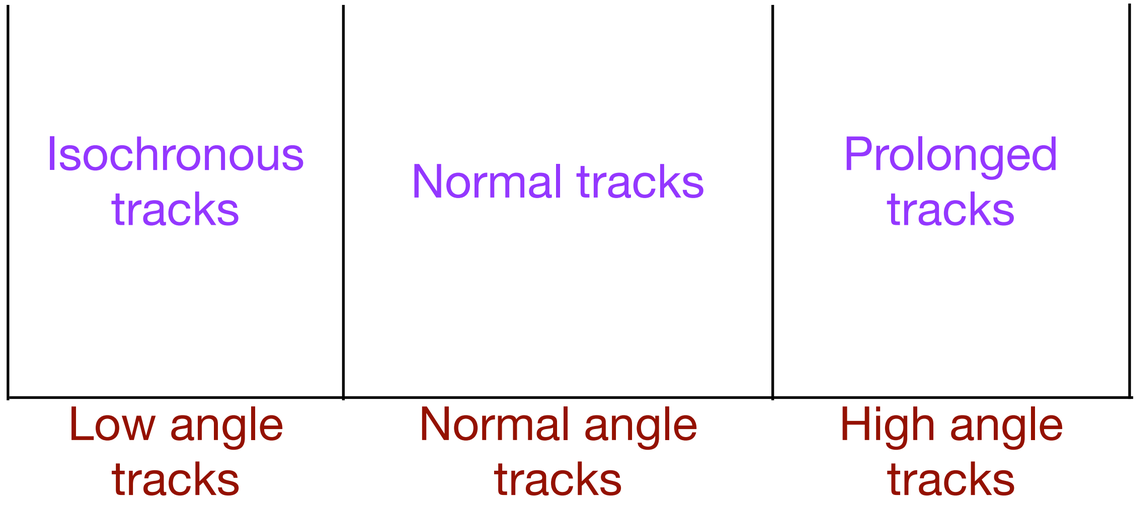
Track Type	Efficiency	Purity
Isochronous	57%	16%

Plans for Isochronous and Prolonged Tracks

- Both isochronous and prolonged tracks affect the evaluation
- Isochronous tracks affect the purity due to their spread
- Prolonged tracks affect efficiency because sometime they show gaps irrespective of dead area
- For evaluation purposes, we need to separate isochronous, and prolonged tracks from rest of the tracks

Plans for Isochronous and Prolonged Tracks

- Both isochronous and prolonged tracks affect the evaluation
- Isochronous tracks affect the purity due to their spread
- Prolonged tracks affect efficiency because sometime they show gaps irrespective of dead area
- For evaluation purposes, we need to separate isochronous, and prolonged tracks from rest of the tracks

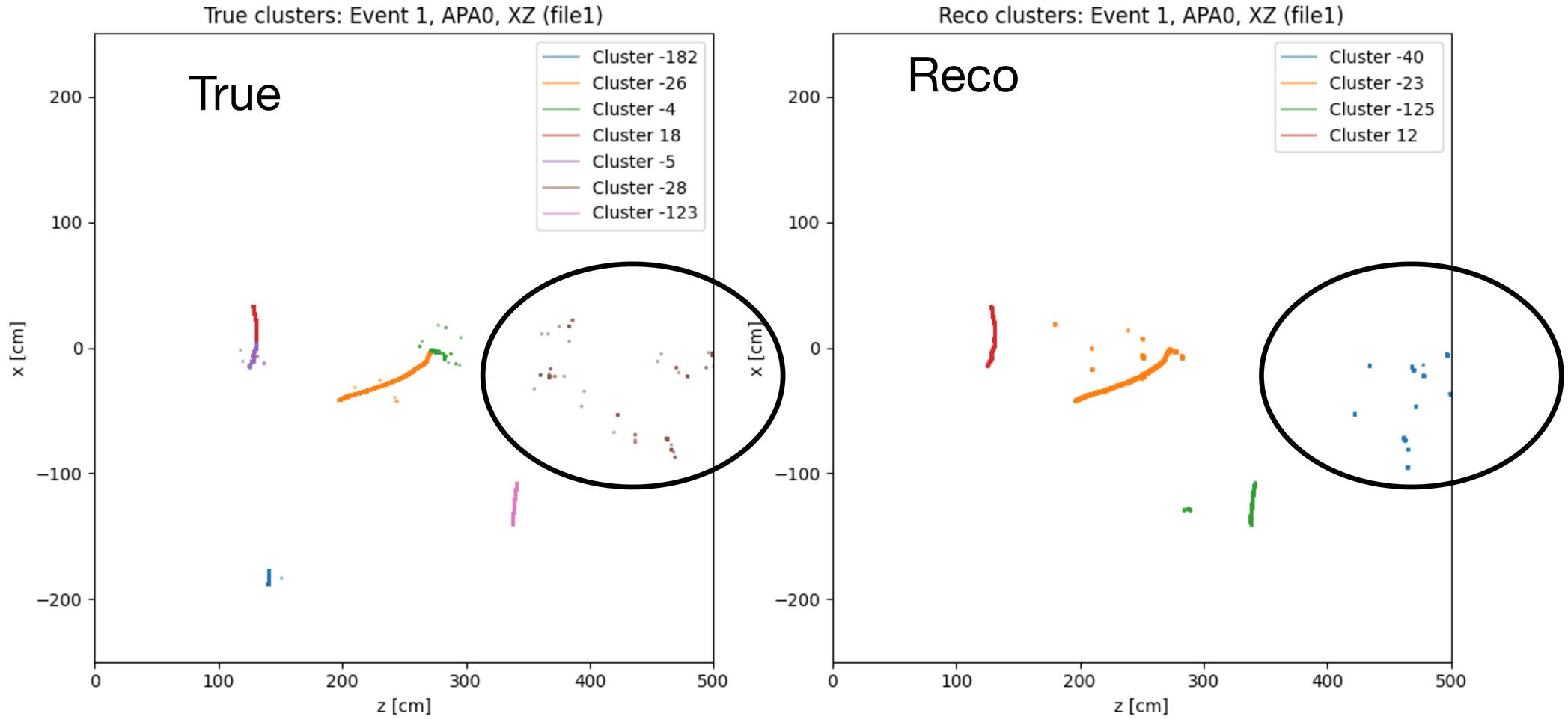


The idea is to divide tracks into low, high and normal angle bins and then do evaluation

Initial Plans for Generic Neutrino Selections

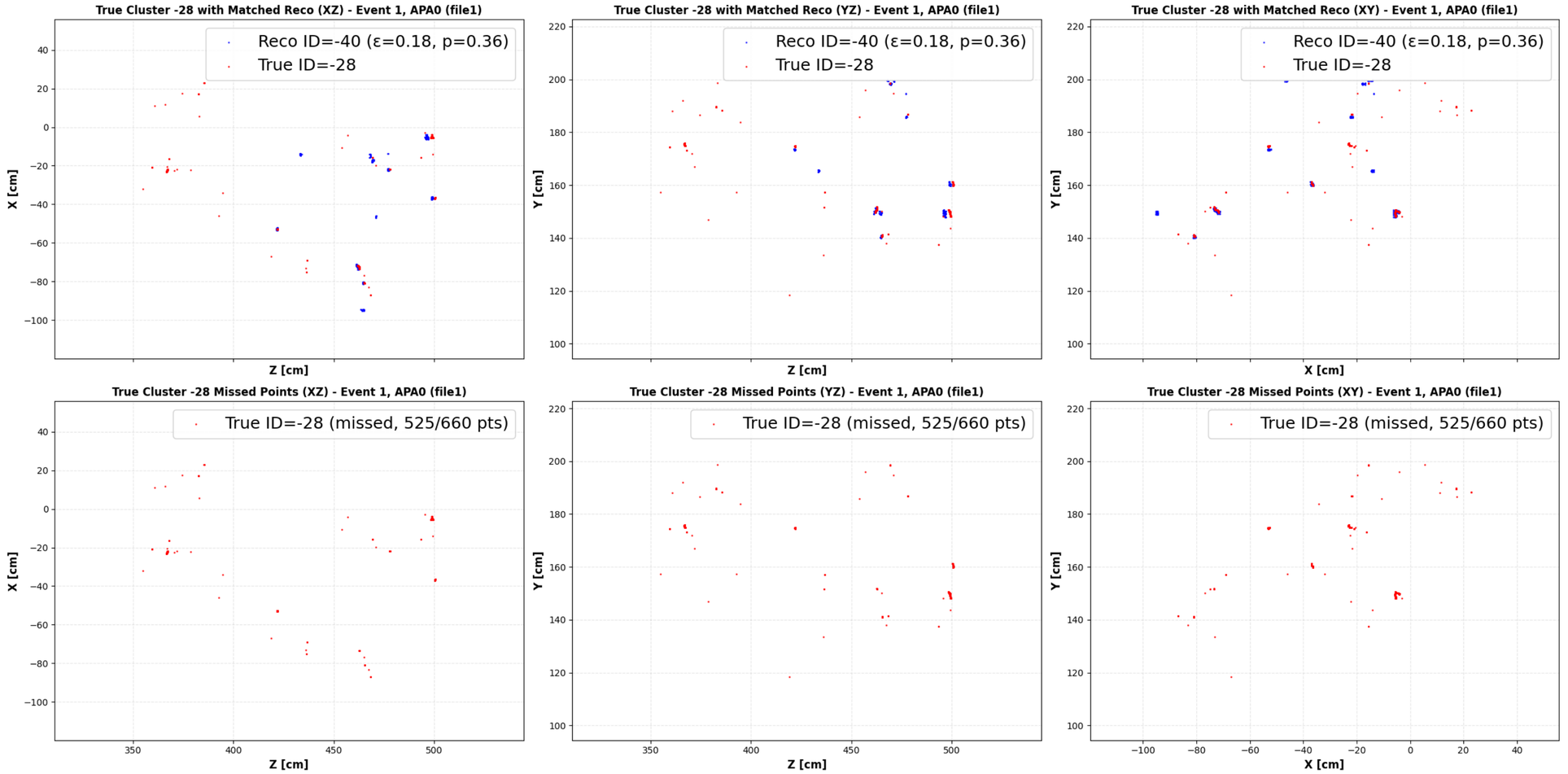
- Currently reading uBoone's Wirecell based neutrino selection paper at https://arxiv.org/pdf/2011.01375
- There are several ideas some of them is to check the Data and MC Distributions of
  - Reco clusters X, Y, Z
  - Time distributions to check clusters are within the time window
  - POT correction
  - Distributions as a function of reco charge information
- In next few meetings, I will show/discuss more on this topic

Some cases of low efficiency: 1



ε = 20%

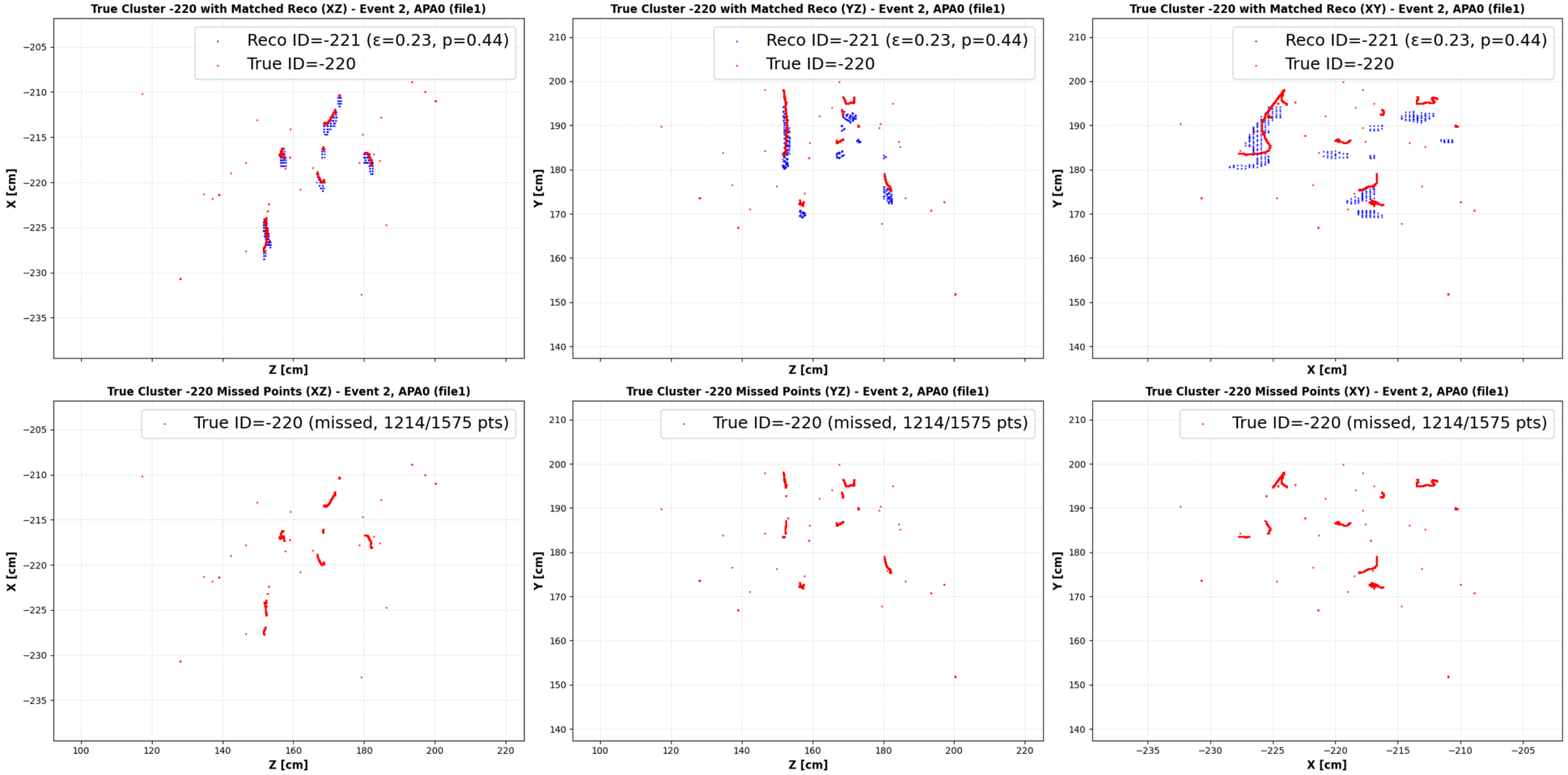
Some cases of low efficiency: 1



ε = 20%

Reason for low efficiency: Very few and scattered true points

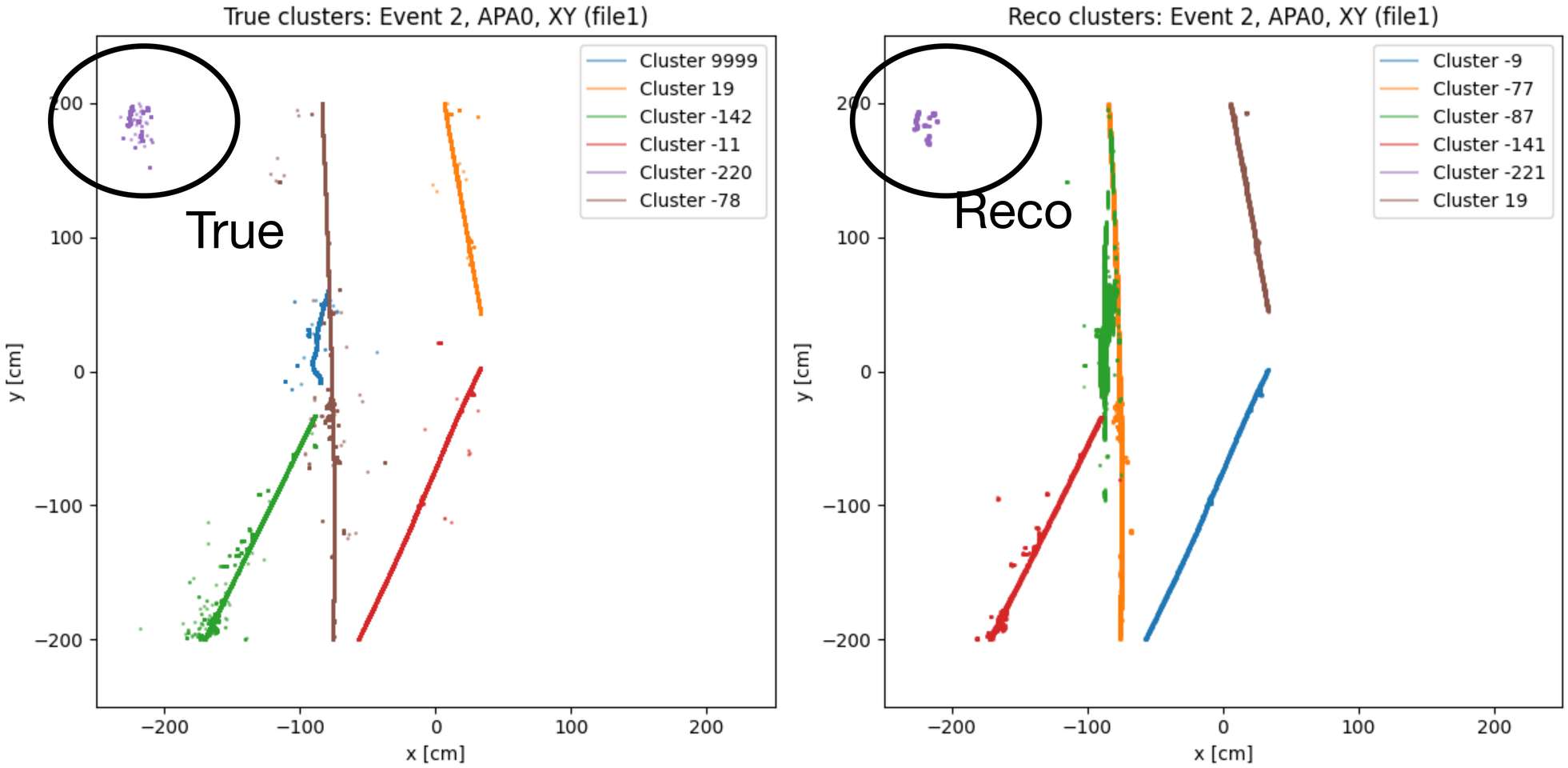
Some cases of low efficiency: 2



ε = 23%

Reason for low efficiency: Very few and scattered true points

Some cases of low efficiency: 2



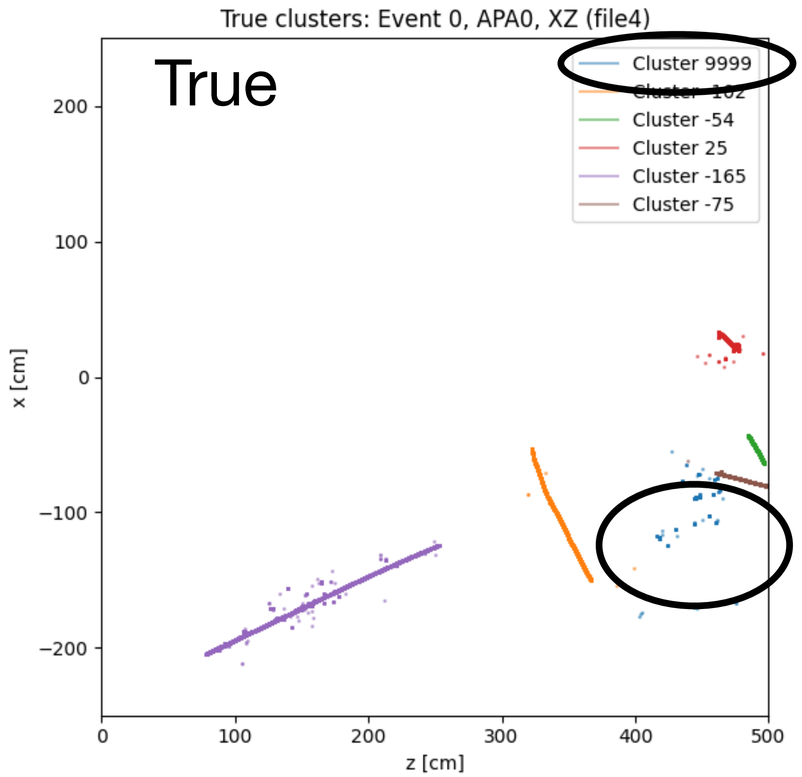
ε = 23%

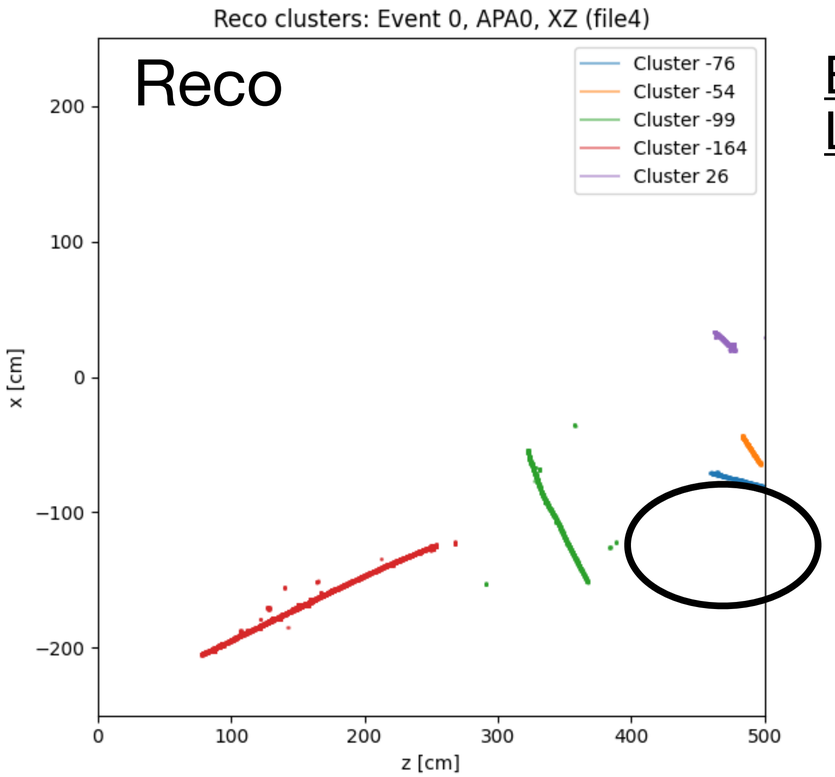
Example of True Neutrinos not finding any reco clusters

2cm efficiency radius

APA 0

File 4: Event 0



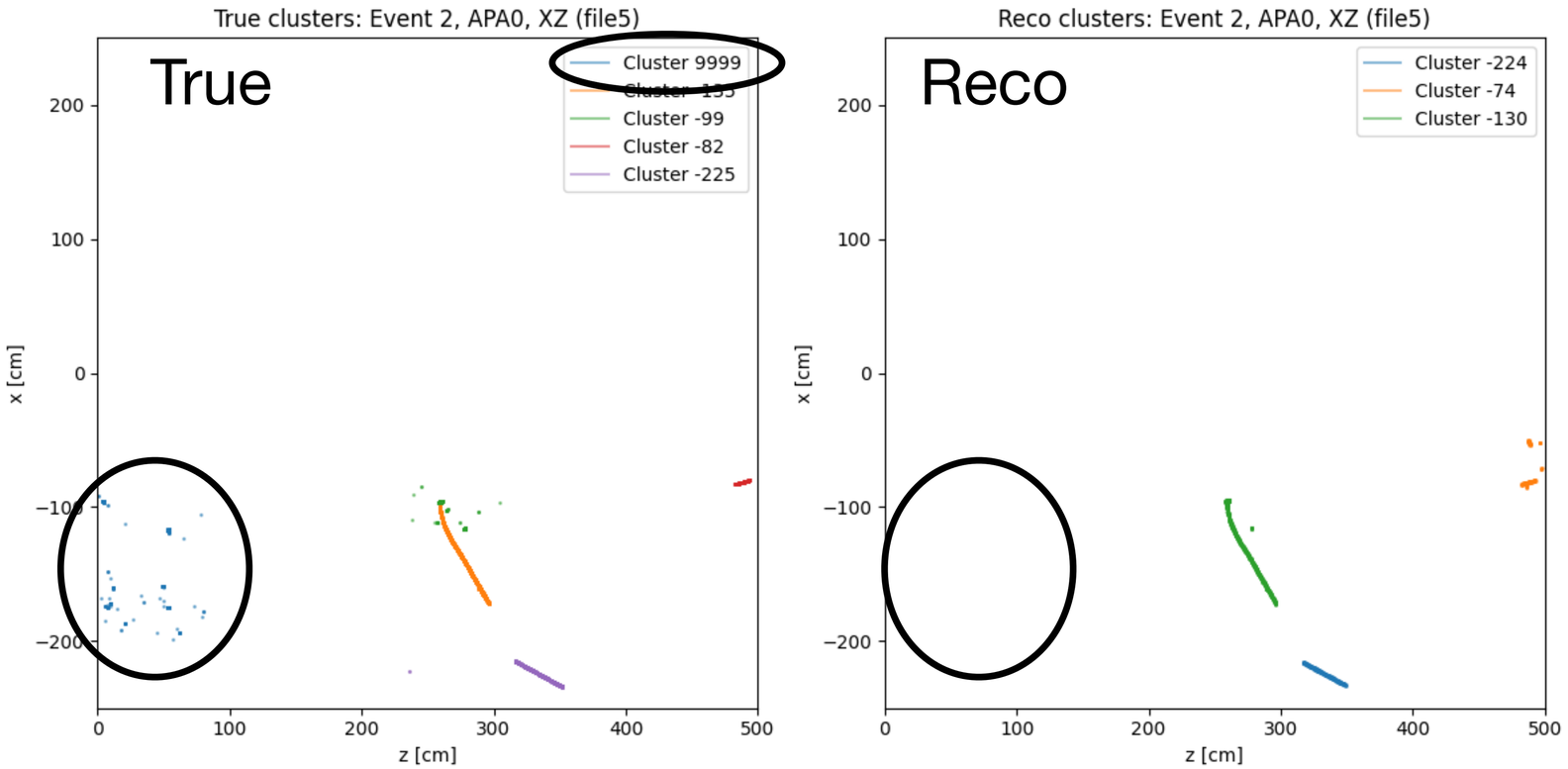


Bee Link

Example of True Neutrinos not finding any reco clusters

File 5: Event 2

2cm efficiency radius
APA 0



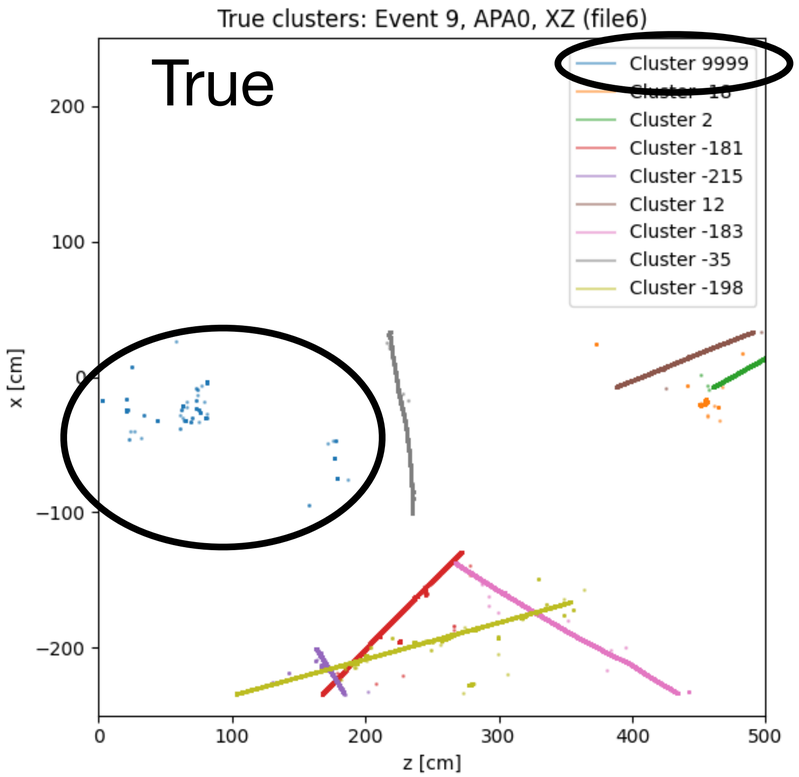
Bee Link

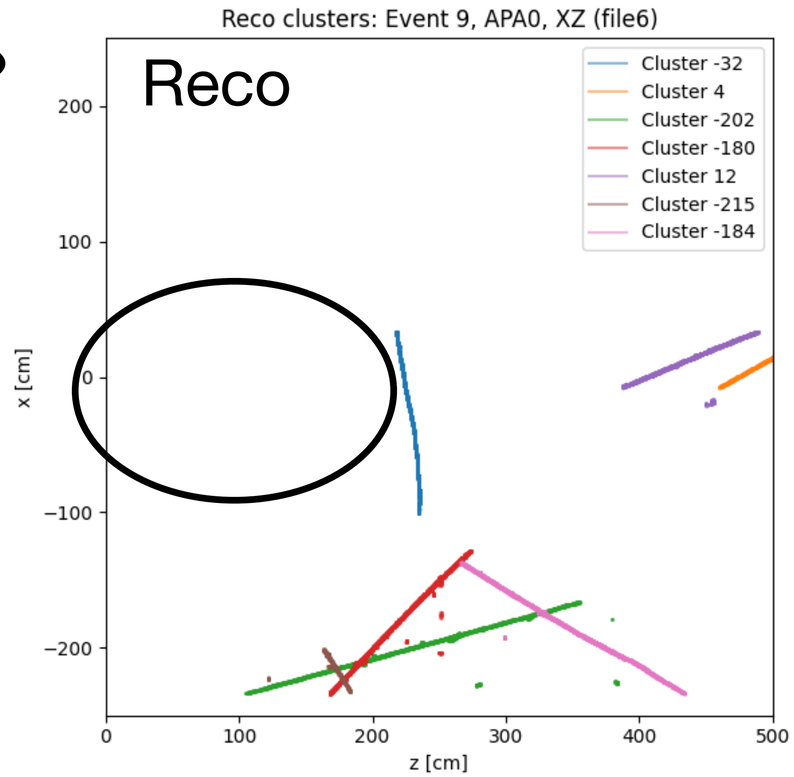
Example of True Neutrinos not finding any reco clusters

File 6: Event 9

2cm efficiency radius

APA 0





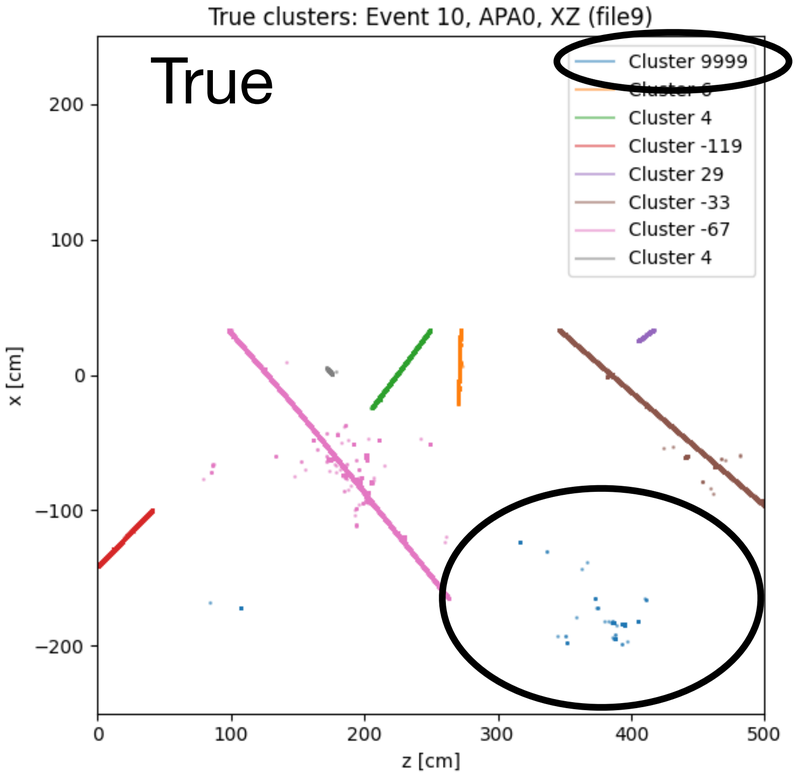
Bee Link

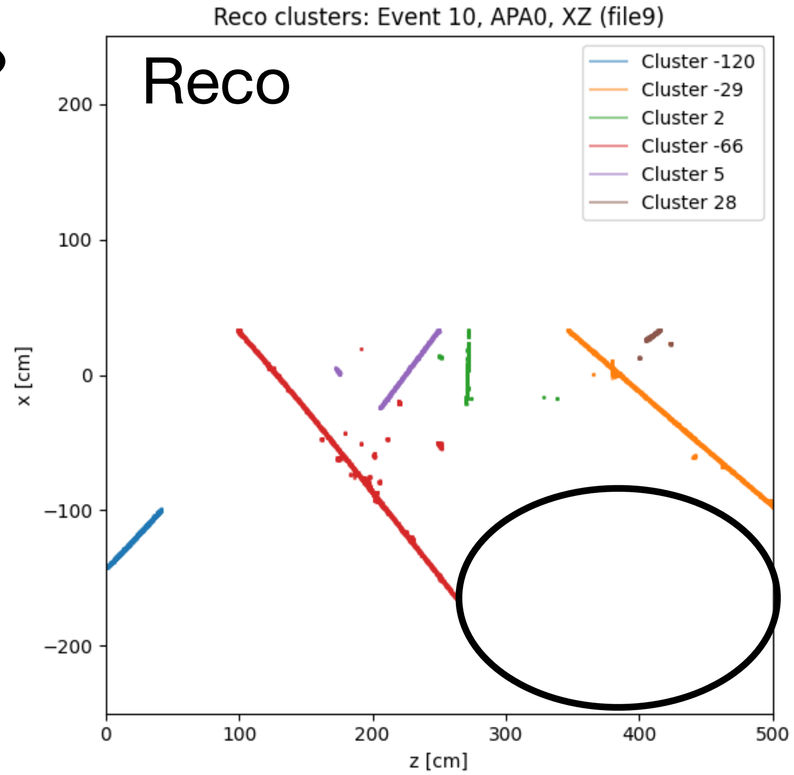
Example of True Neutrinos not finding any reco clusters

File 9: Event 10

2cm efficiency radius

APA 0





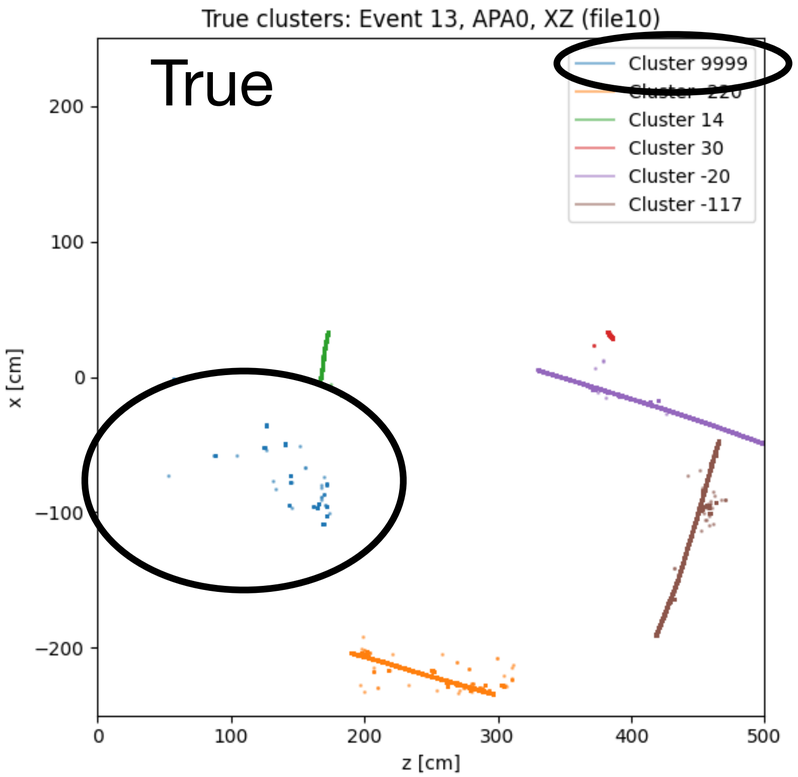
Bee Link

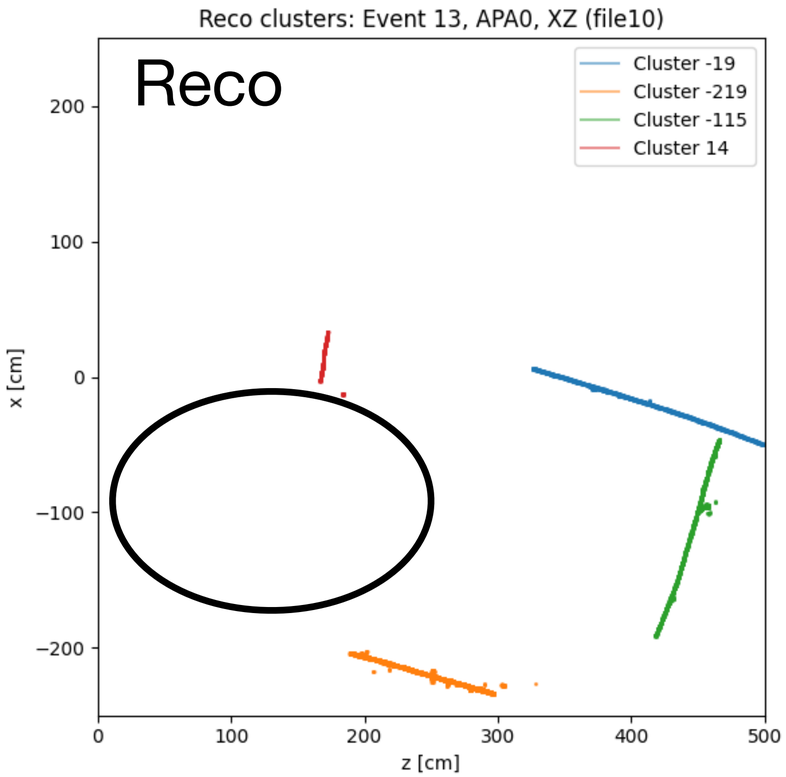
Example of True Neutrinos not finding any reco clusters

File 10: Event 13

2cm efficiency radius

APA 0





Bee Link