

(my) First look at SBND Wire-Cell PR

Matteo Vicenzi (mvicenzi@bnl.gov)

SBND Wire-Cell Meeting
Sep 17th, 2026



Production

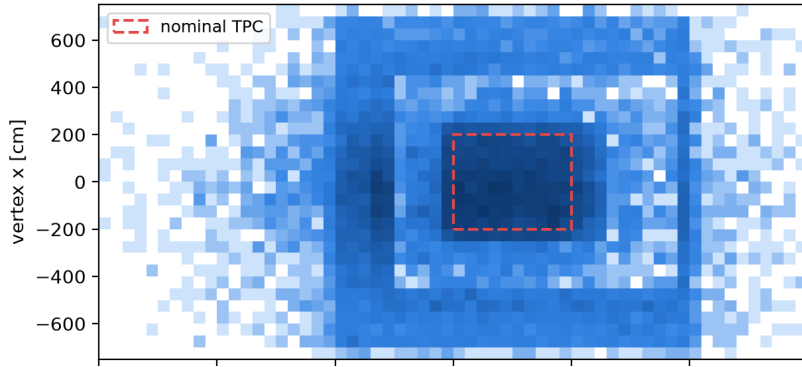
- Sample: **MC BNB CV** (BNB nu + cosmics, GENIE CV)
 - /exp/sbnd/data/users/yuhw/production-prep/r3-mc-cv-2026-09-09/run/
- MC truth from Bee .zip files (data/0/0-mc.json)
- Reco output from the ROOT files (tracking_pr*.root).
- Goals:
 - Familiarize myself with truth/reco data format
 - Evaluate preliminary PR performance (efficiency, purity, ..)

True vertex

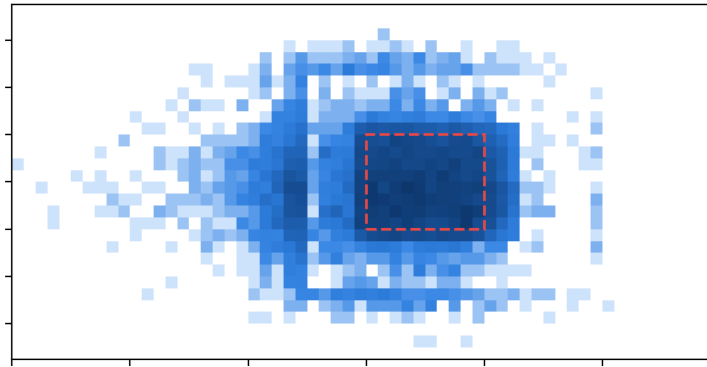
True signal
+ dirt background

True signal

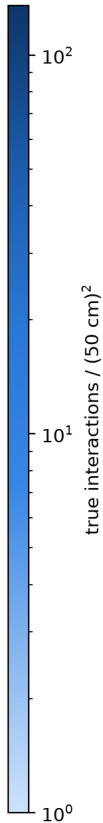
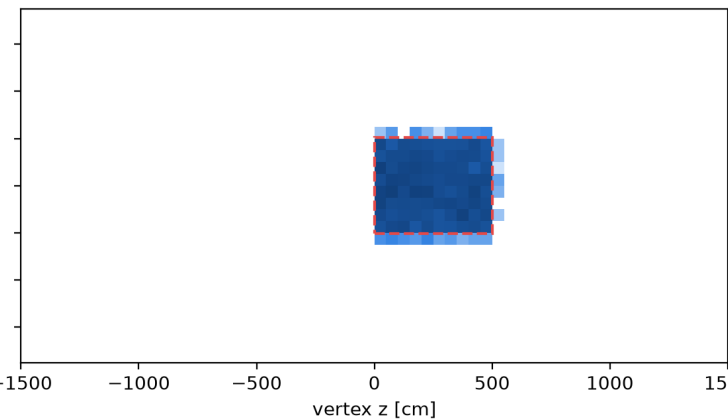
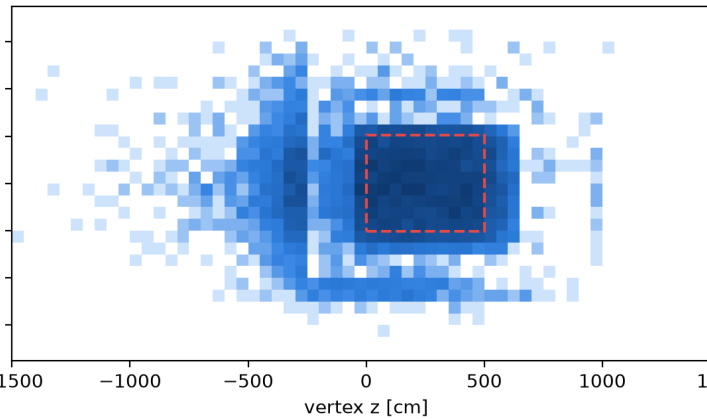
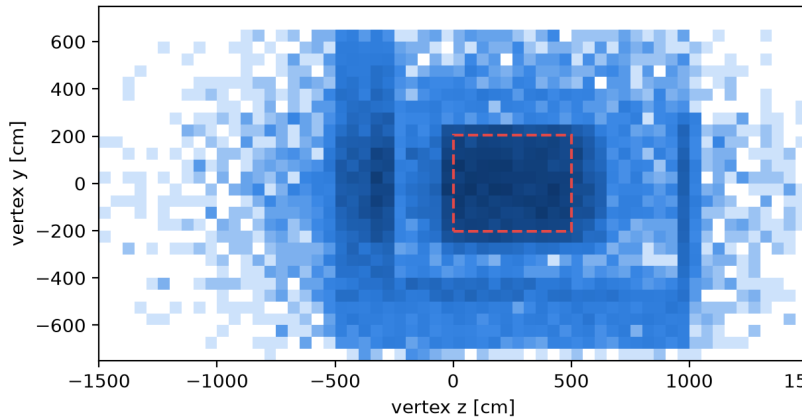
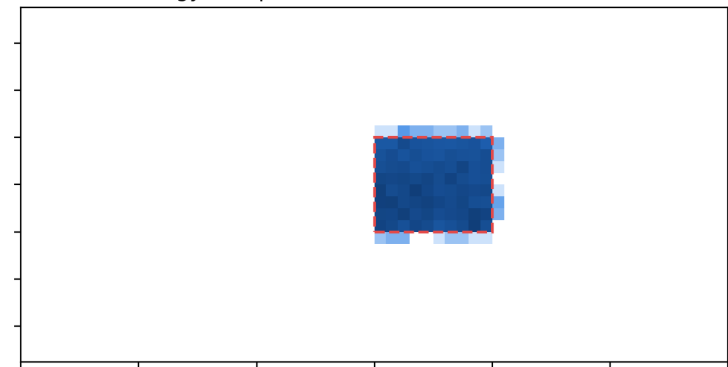
All true neutrino interactions



Visible energy (edep > 0) in active volume



Visible energy (edep > 0) & vertex in active volume (6367)



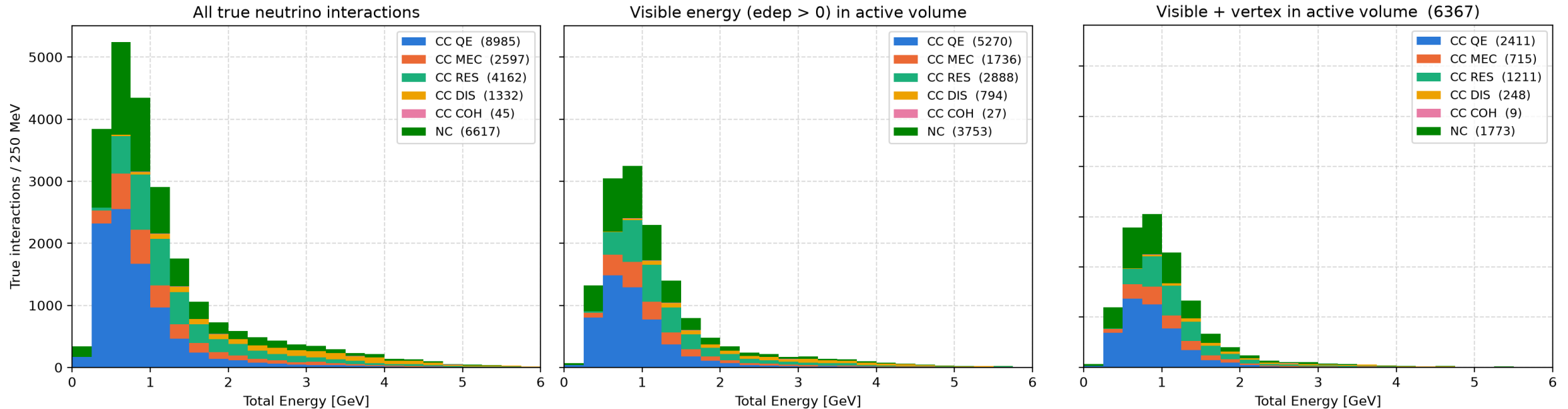
23738

14468

6367

Neutrino energy

Plots mix nue+numu
however, nue is negligible



numuCC 10612
numuNC 3726

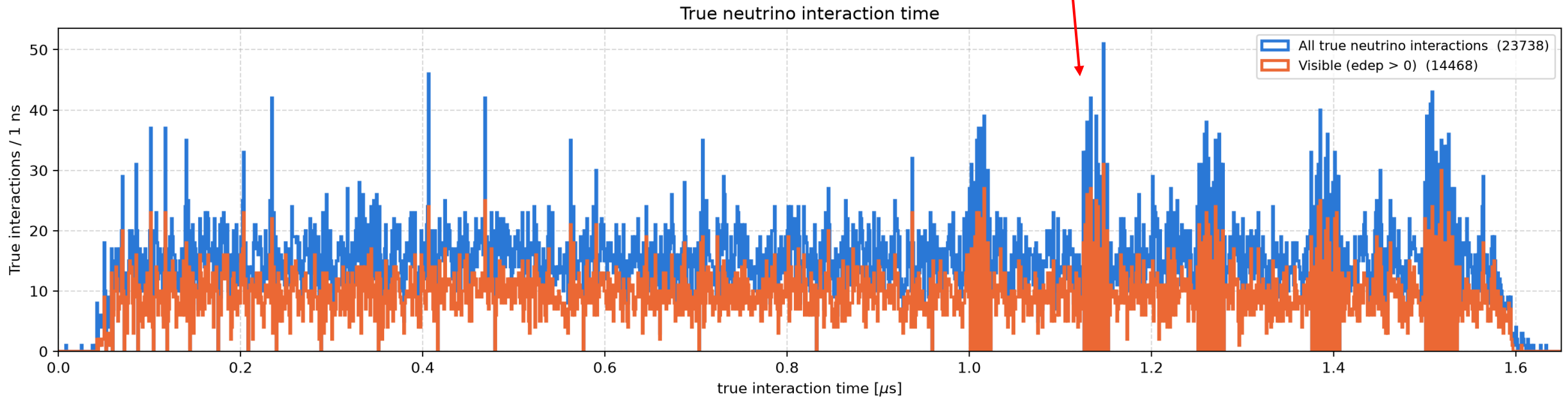
nueCC 103
nueNC 27

numuCC 4551
numuNC 1760

nueCC 43
nueNC 13

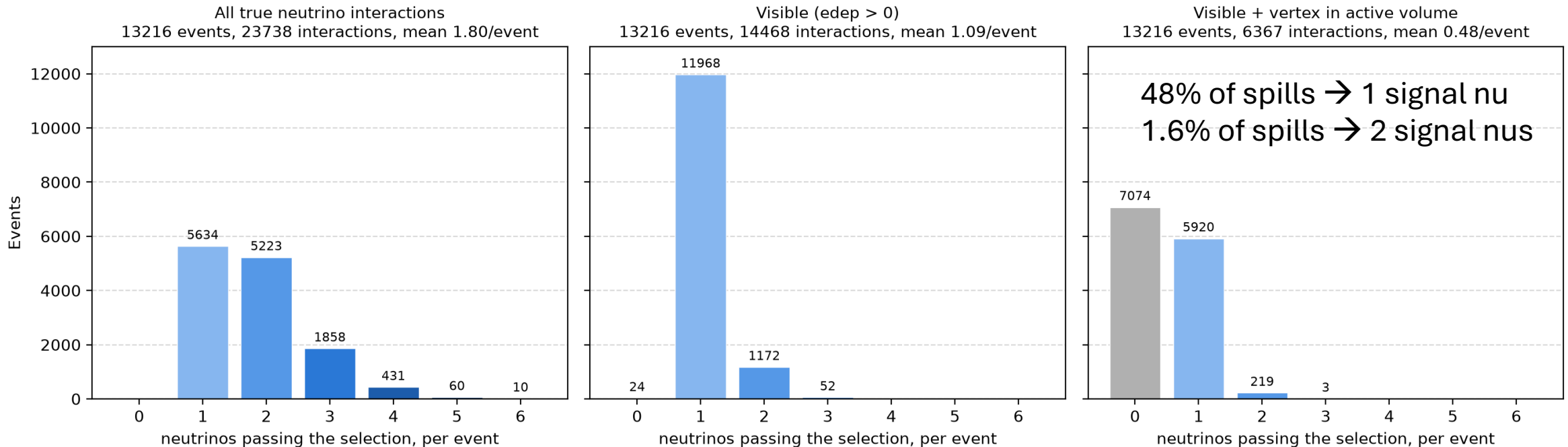
True neutrino time

- Flat in beam spill [0,1.6] μ s, except for ~5 peaks?
- (No 18ns bunch structure, but we don't really care)



Neutrinos per event

- How many neutrino interactions per MC event (= beam spill?)
 - MC uses a fixed POT amount per spill



Reco variables

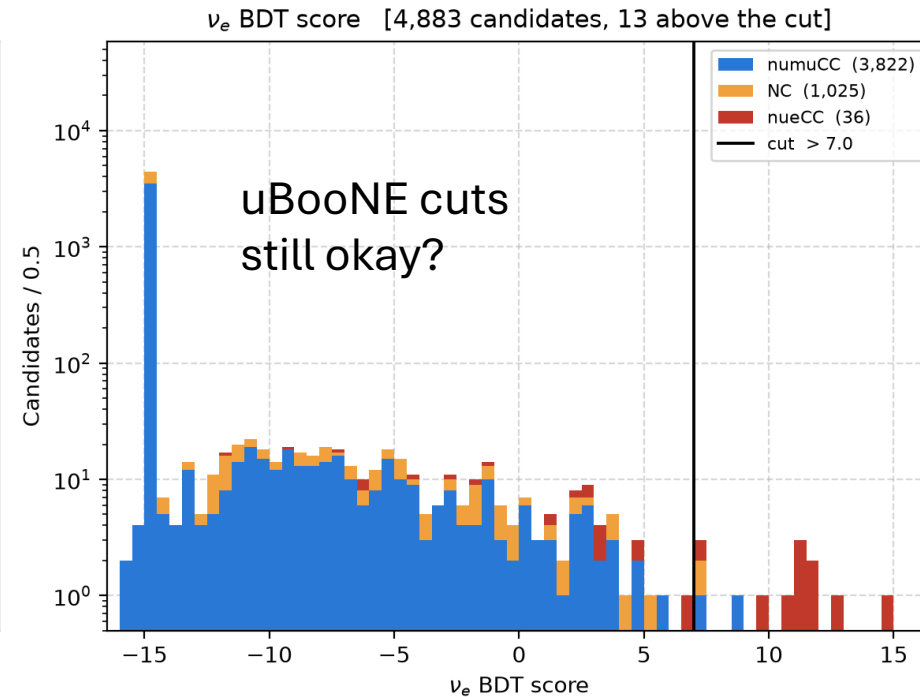
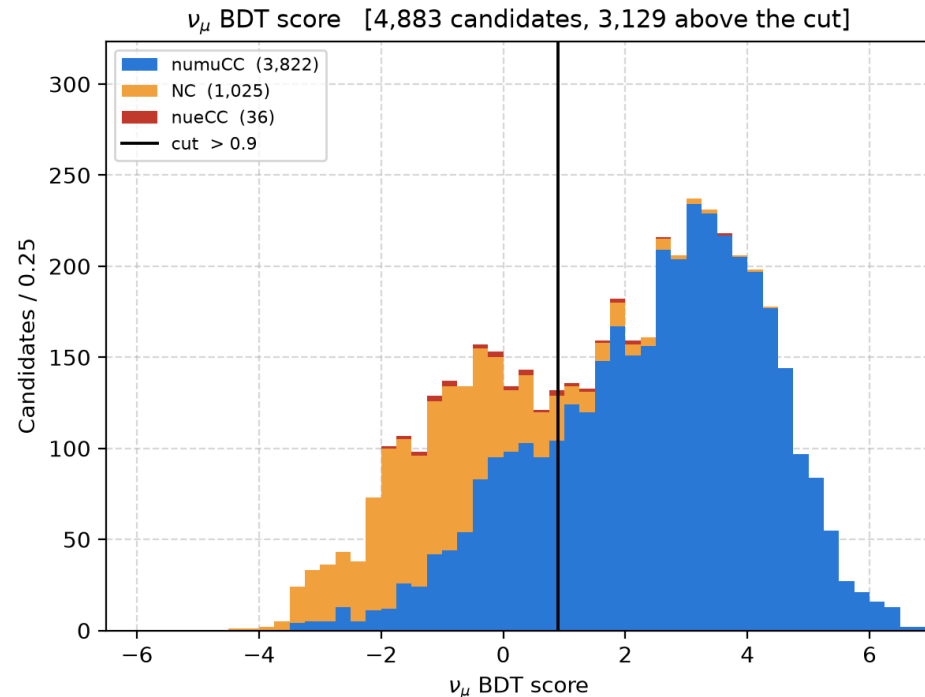
- I'm still wrapping my head around the data format.
- My current understanding:
 - If no Q-L bundles qualifies as candidate, nothing is saved.
 - (T_cluster contains list of bundles, but not much info)
 - Sometimes the *is_main* clusterID in a bundle changes, and gets reported differently across the trees. A bit confusing, but harmless (?).
 - **Beam gate cut is already included** in the candidate selection.
 - Candidate Q-L bundles seem **capped at one per drift volume** (E/W)
 - Is this by design?
 - What happens if there are two neutrinos in the side drift region?

Numu/nue BDT scores

- Given a candidate, reco class is assigned according to the score.
 - These are independent selections, not a classifier.
 - A candidate can pass both or neither

For these plots: using
unambiguous truth-reco matches
(1 signal nu, 1 reco candidate)

BDT scores by true class — one signal ν , one candidate, matched one-to-one



Definitions

Signal:

- true ν_μ CC (ν_e CC) interactions with $\text{edep} > 0$ and vertex in AV.

Background:

- cosmics + dirt ν 's ($\text{edep} > 0$, out of AV) + non-target signal ν 's ($\text{edep} > 0$, in AV).

$$\text{Efficiency} = \frac{\text{signal } \nu \text{ of class X, whose best-matched}^* \text{ candidate passes the score cut for X}}{\text{all signal } \nu \text{'s of class X}}$$

$$\text{Purity} = \frac{\text{selected candidates of class X whose best-matched}^* \text{ true } \nu \text{ is class X}}{\text{all selected candidates of class X}}$$

* both definitions require **matching** between true neutrinos and reco candidates

Matching strategy (I)

- Efficiency/purity definitions require having truth/reco matching
 - currently this needs to be **inferred** (not explicitly in the files?)
 - it can be non-trivial with >1 candidates/neutrinos in the same event.
- There are basically two questions:
 - Purity: Is this reco candidate correct? [“reco” side]
 - Efficiency: Is this true neutrino reconstructed? [“truth” side]
- I’m building these associations separately for **both modes**
 - Best-match based on 3D true/reco vertex proximity.
 - **No max distance cut-off, take the nearest.**
 - Perhaps some kind of hit overlap would be better.

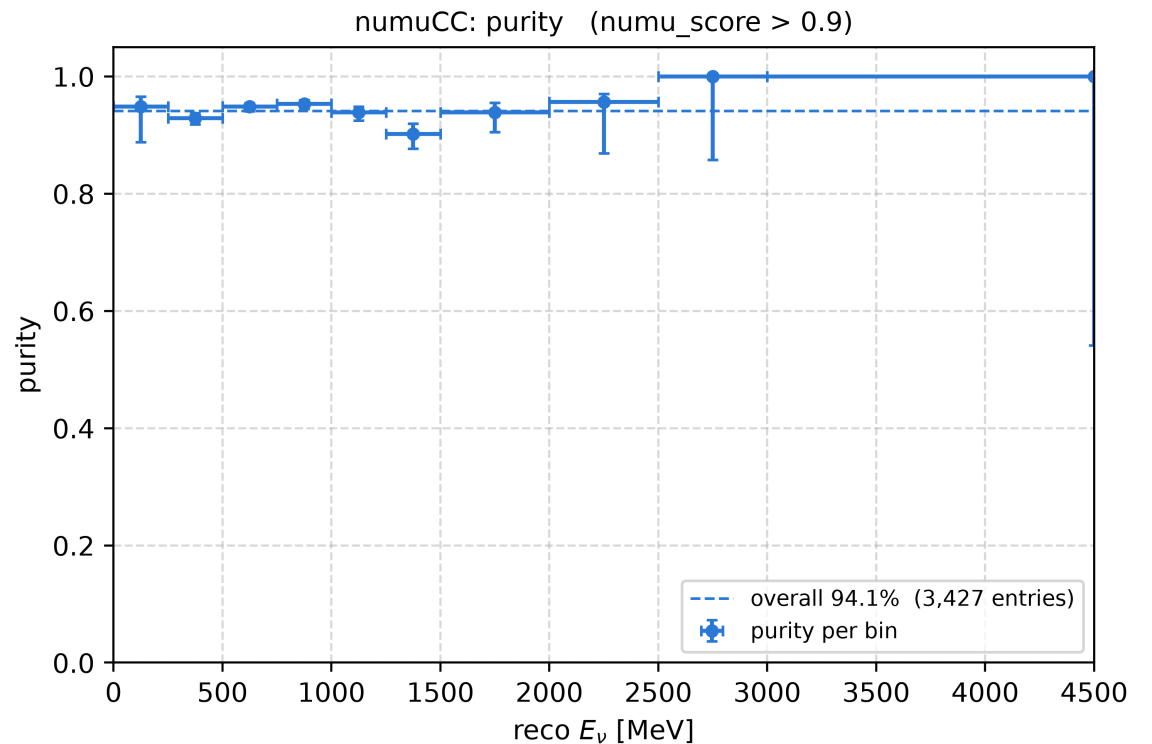
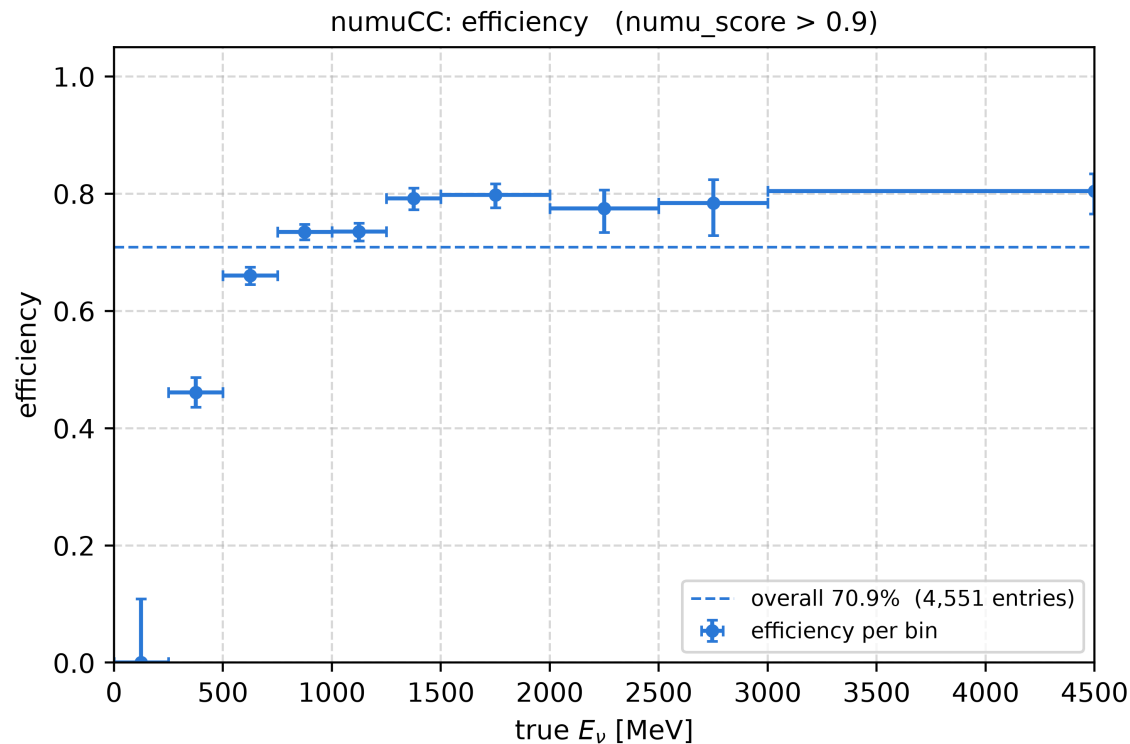
Matching strategy (II)

- Is this reco candidate correct? → **purity**
 - Take the true neutrino closest to it (vertex distance).
 - The candidate is correct only if that neutrino also points back to it and it belongs to the same class.
 - Two-way match, so one neutrino doesn't get matched to 2 bundles
 - If the event has no signal neutrino, it is not correct.
- Is this true neutrino reconstructed? → **efficiency**
 - Take the reco candidate closest to it.
 - The neutrino counts as reconstructed only if that candidate also points back to it and passes the nuscore selection.
 - Two-way match, so one candidate doesn't get matched to 2 ν 's
 - If the event has no reco candidates, it is a miss.

NumuCC

- Out of the box selection: candidate bundle + numu_score, no other cuts.
- Match with truth via smallest 3D distance (but no max distance)

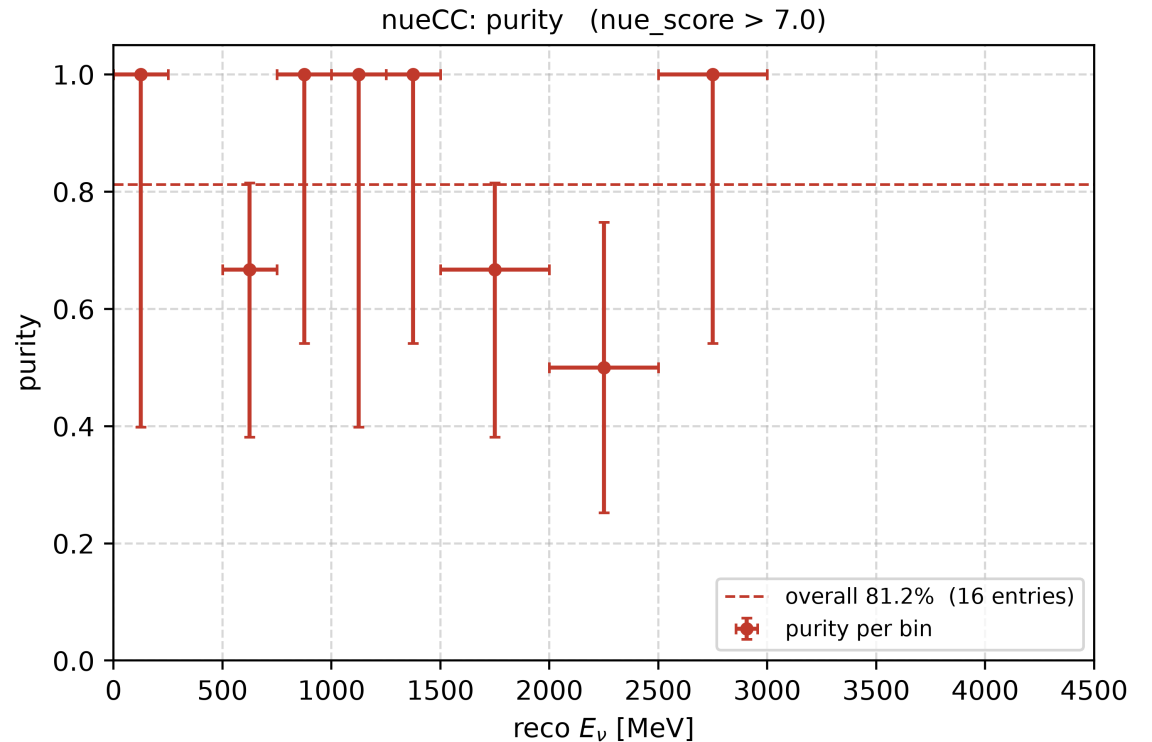
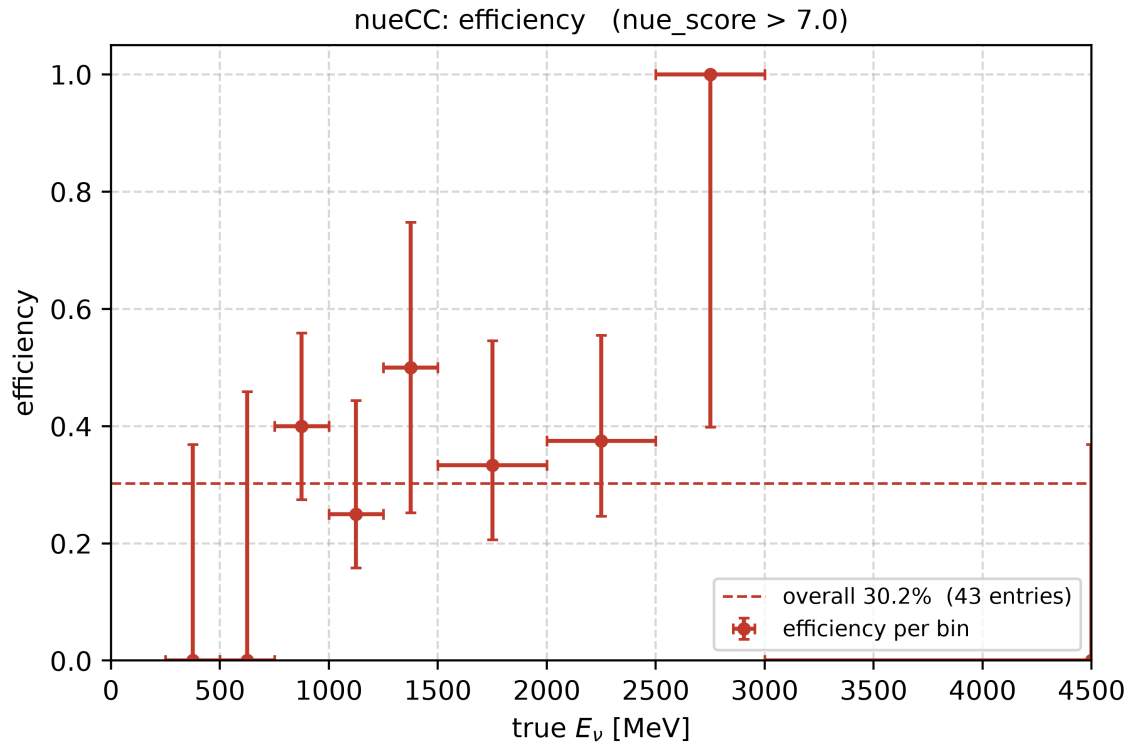
Efficiency (truth mode) and purity (reco mode); match = nearest signal ν



NueCC

- Out of the box selection: candidate bundle + numu_score, no other cuts.
- Match with truth via smallest 3D distance (but no max distance)

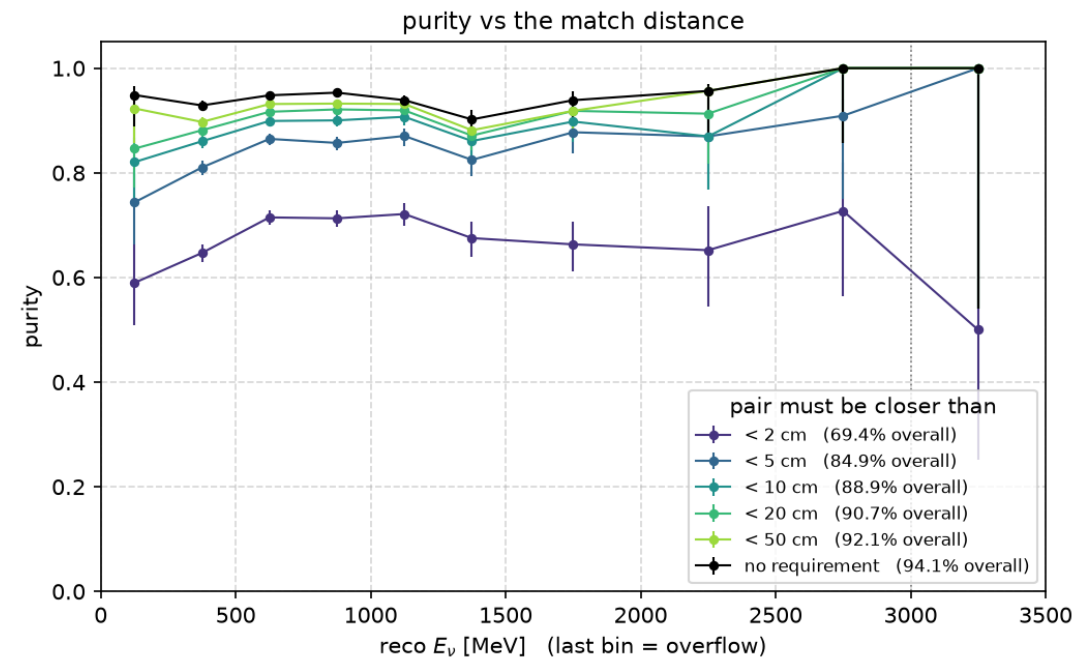
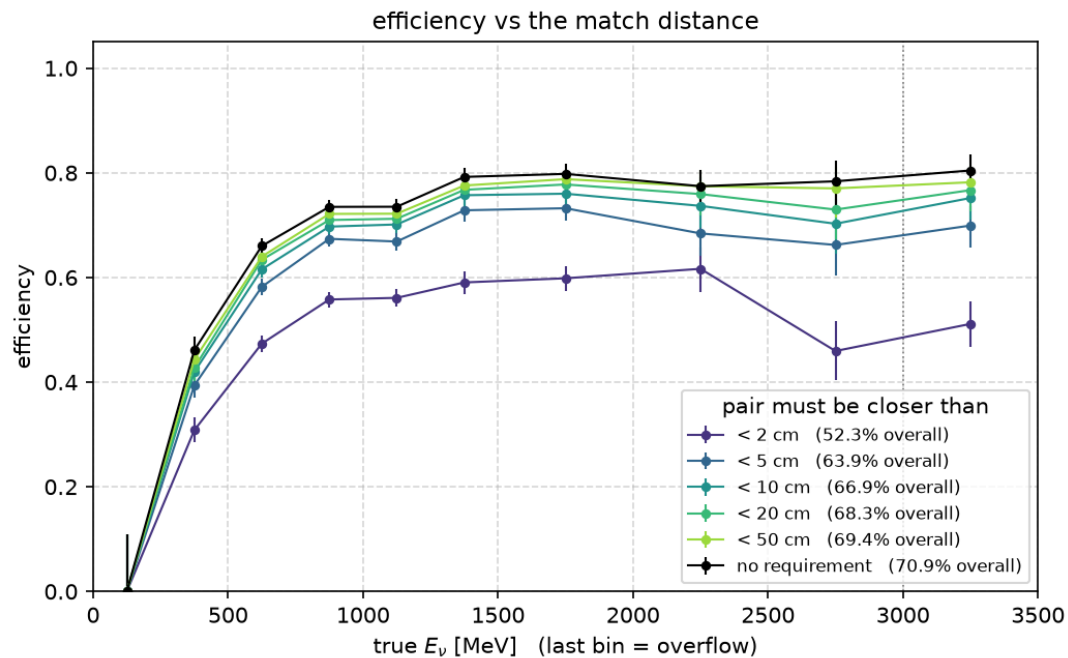
Efficiency (truth mode) and purity (reco mode); match = nearest signal ν



Aside: match quality

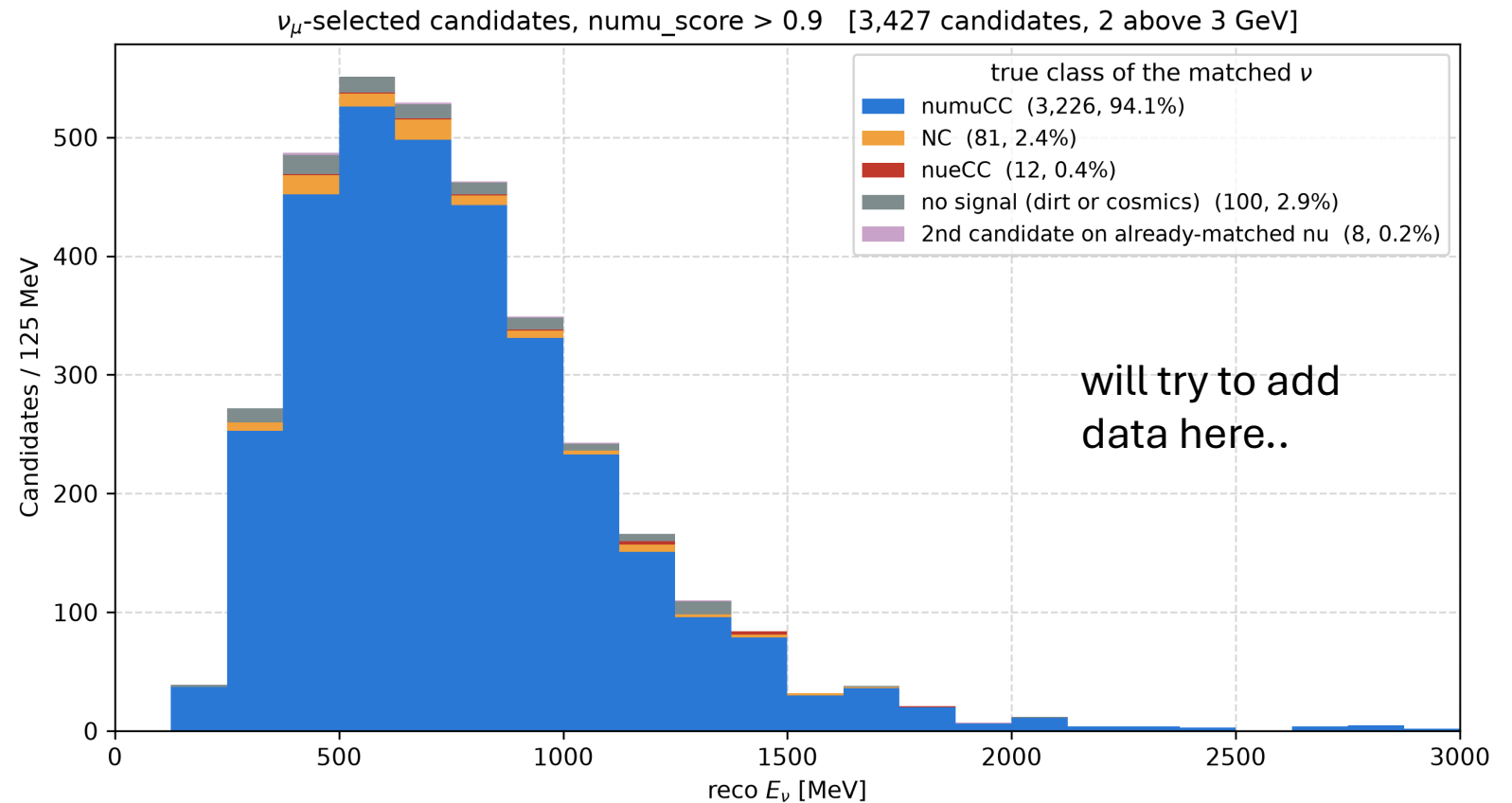
- If you define a max match distance, efficiency/purity drop (flat over energy) → not a selection cut, rather a definition.

ν_μ selection (numu_score > 0.9)



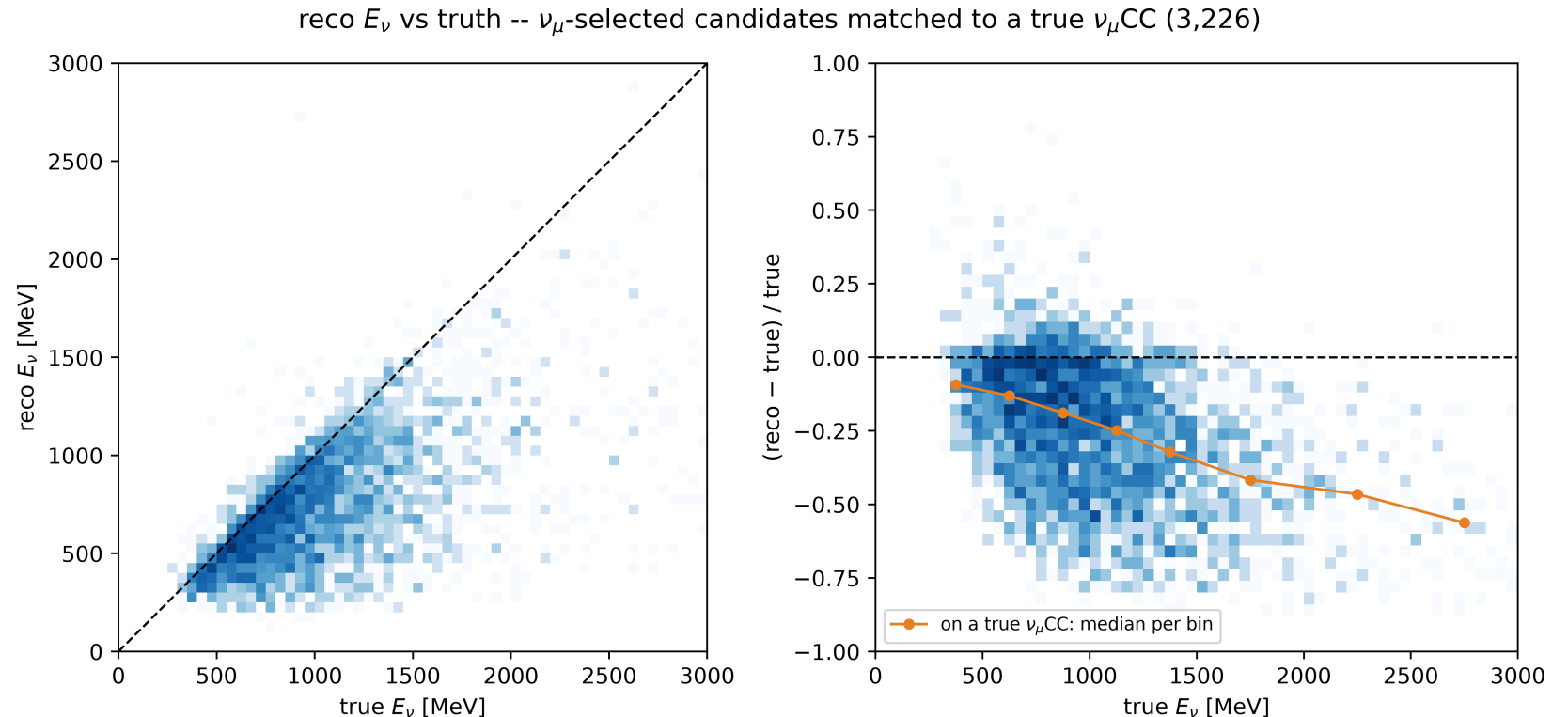
NumuCC – reco E_ν

- Same truth-matching used for purity.
 - can't separate dirt/cosmics currently.
 - similarly for 2nd candidate on already-matched neutrino.



NumuCC – reco E_ν

- Same truth-matching used for purity.
 - can't separate dirt/cosmics currently.
 - similarly for 2nd candidate on already-matched neutrino.
- **only matches to true ν_μ CC**



Summary & next steps

- Preliminary look at numu (nue) efficiency and purity.
- Reco candidates appear to be capped (one per drift length)
 - it might not be common, but in full dataset never >1 per drift.
 - some other reco weirdness in the back-up
- Currently inferring reco/truth matching with 3D distance
 - Is there a better way of doing this?
- Happy to hear any suggestion for direction to take:
 1. I still want to better understand what reco saves in case of multi-neutrinos (can two neutrinos be in the same bundle?)
 2. Try reco-Enu data/MC comparison (area normalized)

Back-up

Case 1: changing main clusterID

```
r719 s59 e6 8 trees: T_bad_ch, T_cluster, T_kine, T_proj, T_proj_data, T_rec_charge, T_tagger, Trun
74 clusters | 13 bundles (is_main & in_scope) | 1 of them in the beam window [0.2, 2.2) us

bundles (time-ordered):
  cluster_id flash_gid side n_assoc npoints length_cm t0_us flash_pe in_window
    6 1000014 W 0 12 1.8 -1353.534 540.5
    4 12 E 2 19 1.9 -1075.008 294.6
   10 1000018 W 0 8 1.3 -1036.047 70.4
    5 7 E 1 1156 59.5 -949.086 3096.7
    7 1000012 W 0 34 4.0 -865.998 818.2
   13 1000001 W 3 4069 450.9 -722.739 30107.2
   15 2 E 5 3924 424.7 -298.366 11701.1
   14 1000000 W 6 5423 523.0 -222.241 20225.9
    8 1000003 W 29 717 29.6 0.706 13021.5 yes
   16 1000004 W 6 1248 148.7 702.008 14977.4
   11 1000002 W 2 3022 234.9 726.477 28351.6
    1 1 E 4 10660 358.3 925.229 23849.0
    3 0 E 0 4793 289.2 940.731 41809.8
```

```
1 candidate(s) written:

candidate: flash GID 1000003 (W side) bundle mains [8] t0 0.706 us
T_kine.cluster_id 70 (is_associated, GID 1000003 = this candidate)
T_tagger.cluster_id 70 (is_associated, GID 1000003 = this candidate)
selected activity 8 (is_main, GID 1000003 = this candidate)
  <-- cluster_id reads 70 but the tagger selected 8: main-cluster swap
reco Enu 684.5 MeV neutrino_type 8 = NC
BDT scores numu -1.186 nue -8.086 cosmict +0.000
vertex (140.1, -170.4, 28.5) cm inside active volume

activities in this bundle:
  act_cluster_id act_length_cm act_is_selected act_is_demoted act_tgm act_stm act_fc act_evaluated
    8 29.598938 1 0 0 0 1 1

T_rec_charge (no nu_index -- one tree for the whole event): 378 space points under cluster_id [8]
73 flagged vertex, 157 shower, 73 with no origin cluster
from flash GID 1000003 = nu_index 0, 29 origin cluster(s):
  8:100 25:3 26:2 27:4 28:2 29:2 30:2 32:3 33:6 34:2 35:6 36:2
  37:2 38:3 39:3 40:2 41:4 42:5 43:2 44:3 45:7 46:10 47:6 48:2
  70:56 71:26 72:9 73:26 74:5

truth: 3 interaction(s), 1 signal, 0 background
  int_idx flavor mode curr etot_MeV edep_MeV t_us vtx_x vtx_y vtx_z in_av signal background d_nu0 dt_nu0
    1 numu MEC NC 3752.5 0.0 0.809 389.2180 -502.606 868.9490 False 0 0 937.4 -0.103
    2 numu MEC CC 2405.4 0.0 1.190 -98.2635 -690.726 -717.0190 False 0 0 939.9 -0.484
    3 numu RES NC 2227.0 740.5 0.570 94.6336 -198.354 25.5034 True 1 0 53.5 0.136
```

Case 2: empty/fake candidate bundle

```
r717 s32 e8 7 trees: T_bad_ch, T_cluster, T_kine, T_proj, T_rec_charge, T_tagger, Trun
79 clusters | 18 bundles (is_main & in_scope) | 1 of them in the beam window [0.2, 2.2) us
!! CHECK FAILED: no is_main & in_scope cluster carries flash GID 4
!! CHECK FAILED: empty placeholder -- Enu 0, nu_type 0, vertex at the origin, but flash GID 4
!! CHECK FAILED: T_rec_charge present but EMPTY: no saved 3D points

bundles (time-ordered):
  cluster_id  flash_gid side  n_assoc  npoints  length_cm  t0_us  flash_pe  in_window
    20      1000021    W      1        10       1.7 -1310.729   95.9
    11      1000002    W      0         1       0.5 -1206.350  17656.0
    16      1000001    W      0       364      38.2 -1146.088  19991.6
    19      1000007    W      5      1218      81.3  -876.011   6713.1
     8         19    E      1        12       1.3  -766.782   179.3
     9         15    E      2         10       1.9  -367.852   247.6
    22      1000004    W      2      4999     414.3  -264.555  13592.9
     7         6    E      1      1417     113.4  -253.268  3350.3
    12      1000012    W      0         7       1.5   -75.469   207.5
    17      1000005    W      7      2460     241.7   -20.148  11156.4
    23      1000006    W      1     12120     494.5    1.619  11963.6      yes
     6         1    E     26     8472     495.9   25.117  47249.4
    21      1000013    W      1         8       1.5   96.830   200.8
     5         3    E      4     3552     266.4   382.573  12868.9
    18      1000008    W      2     1734     187.8   388.311   8655.3
     4         2    E      3     1957     211.8   692.484  24607.9
     2        17    E      0         14       1.5   780.501    98.2
    15      1000000    W      0        767     83.0  1162.231  25357.7
```

```
1 candidate(s) written:

candidate: flash GID 4 (E side) bundle mains [] t0 nan us <-- EMPTY PLACEHOLDER
  T_kine.cluster_id      1 (is_main, GID 4 = this candidate)
  T_tagger.cluster_id    1 (is_main, GID 4 = this candidate)
  selected activity      1 (is_main, GID 4 = this candidate)
  reco Enu 0.0 MeV  neutrino_type 0 = none(0)
  BDT scores  numu -1.942  nue -15.000  cosmict +0.000  <-- placeholder default
  vertex (0.0, 0.0, 0.0) cm  inside active volume

activities in this bundle:
  act_cluster_id  act_length_cm  act_is_selected  act_is_demoted  act_tgm  act_stm  act_fc  act_evaluated
                1      0.798573          1              0          0          0          0          1

T_rec_charge: EMPTY -- the tree is there but holds no 3D points

truth: 2 interaction(s), 0 signal, 2 background
  int_idx  flavor  mode  curr  etot_MeV  edep_MeV  t_us  vtx_x  vtx_y  vtx_z  in_av  signal  background  d_nu0  dt_nu0
    1      numu  RES   CC    2116.3      2.4  1.137  491.5140  482.842  515.027  False    0          1  860.2   NaN
    2      numu  RES   CC    2277.9     1149.1  1.467  -87.3433  -126.560  -285.643  False    0          1  324.4   NaN
```

Case 3: T_rec_charge with clusterID -1

```
r474 s72 e31 8 trees: T_bad_ch, T_cluster, T_kine, T_proj, T_proj_data, T_rec_charge, T_tagger, Trun
67 clusters | 16 bundles (is_main & in_scope) | 2 of them in the beam window [0.2, 2.2) us
!! CHECK FAILED: T_rec_charge cluster_ids [-1, 9] are not the selected activities [9, 29]

bundles (time-ordered):
  cluster_id flash_gid side n_assoc npoints length_cm t0_us flash_pe in_window
      3         0     E      0        9       1.3 -1138.919 29858.7
      5        15     E      0       15       1.7 -411.160   368.5
      6         6     E     11     3581     401.1 -170.314 12993.8
     10         9     E      1       15       2.3 -113.451  1108.1
     23         3     E      5     4250     436.4 -66.328  28742.0
      9         2     E      1     2167     203.3   0.427  21930.5      yes
     25    1000011     W      8     3515     234.9   1.137   7529.5      yes
     24    1000007     W      4     4666     423.9  33.339  15227.6
     21         5     E      4     4126     414.3  68.480  15439.1
     16    1000006     W      2     4824     211.4 159.517  15707.3
      4        17     E      0        16       1.7  361.525   132.2
     13    1000001     W      0     2344     247.0 383.940  22382.9
     22    1000003     W      1     1623     180.2 472.142  19380.2
     17    1000008     W      1     4555     352.1 486.810  17047.9
     15    1000004     W      0     1457     168.5 835.220  29720.7
     19    1000000     W      4     3919     326.1 1015.671 35393.6
```

```
2 candidate(s) written:

nu_index 0: flash GID 2 (E side) bundle mains [9] t0 0.427 us
T_kine.cluster_id      9 (is_main, GID 2 = this candidate)
T_tagger.cluster_id    9 (is_main, GID 2 = this candidate)
selected activity      9 (is_main, GID 2 = this candidate)
reco Enu 840.8 MeV      neutrino_type 4 = numuCC
BDT scores numu +4.107 nue -15.000 cosmict +0.000 <-- nue at its -15 default: that BDT never ran on this candidate
vertex (-192.2, 99.1, 318.2) cm inside active volume

nu_index 1: flash GID 1000011 (W side) bundle mains [25] t0 1.137 us
T_kine.cluster_id      29 (is_associated, GID 1000011 = this candidate)
T_tagger.cluster_id    29 (is_associated, GID 1000011 = this candidate)
selected activity      29 (is_associated, GID 1000011 = this candidate)
reco Enu 810.6 MeV      neutrino_type 4 = numuCC
BDT scores numu +4.250 nue -15.000 cosmict +0.000 <-- nue at its -15 default: that BDT never ran on this candidate
vertex (53.3, 153.6, 427.5) cm inside active volume

T_rec_charge (no nu_index -- one tree for the whole event): 725 space points under cluster_id [-1, 9]
25 flagged vertex, 0 shower, 25 with no origin cluster
from flash GID 2 = nu_index 0, 2 origin cluster(s):
  9:382 41:2
from flash GID 1000011 = nu_index 1, 8 origin cluster(s):
  29:292 61:5 62:2 63:2 64:2 65:2 66:2 67:9

truth: 3 interaction(s), 2 signal, 0 background
int_idx flavor mode curr etot_MeV edep_MeV t_us vtx_x vtx_y vtx_z in_av signal background d_nu0 dt_nu0 d_nu1 dt_nu1
  1 numu RES CC 1018.5 535.3 1.001 53.4719 154.8820 428.019 True 1 0 274.8 -0.574 1.4 0.136
  2 numu MEC CC 1180.2 666.9 0.297 -191.9540 99.1389 318.318 True 1 0 0.3 0.130 273.9 0.840
  3 numu QE CC 203.1 0.0 1.193 -127.8340 -626.6950 994.712 False 0 0 994.3 -0.766 981.5 -0.056
```

Two candidates here (one E, one W), but it happens for single candidates also
 → most common bad occurrence

Case 4: Two candidates, one empty/fake

```
r720 s4 e18 8 trees: T_bad_ch, T_cluster, T_kine, T_proj, T_proj_data, T_rec_charge, T_tagger, Trun
81 clusters | 11 bundles (is_main & in_scope) | 1 of them in the beam window [0.2, 2.2) us
!! CHECK FAILED: nu_index 1: no is_main & in_scope cluster carries flash GID 1
!! CHECK FAILED: nu_index 1: empty placeholder -- Enu 0, nu_type 0, vertex at the origin, but flash GID 1
!! CHECK FAILED: T_rec_charge cluster_id is -1 for every point
```

bundles (time-ordered):

cluster_id	flash_gid	side	n_assoc	npoints	length_cm	t0_us	flash_pe	in_window
17	1000002	W	0	5984	431.1	-1068.989	24384.2	
18	2	E	2	4262	410.0	-808.956	20637.5	
3	8	E	0	64	5.3	-641.182	2394.8	
1	16	E	0	24	3.3	-526.702	210.4	
2	11	E	0	202	25.2	-278.900	1066.6	
19	1000005	W	38	2265	207.6	0.482	7919.6	yes
11	1000014	W	0	8	1.6	49.875	233.1	
9	0	E	18	17814	501.4	269.207	29921.3	
7	3	E	4	3028	296.2	574.234	23631.2	
5	10	E	0	131	12.1	803.766	1820.5	
16	1000003	W	0	952	91.0	1091.791	14806.1	

2 candidate(s) written:

```
nu_index 0: flash GID 1000005 (W side) bundle mains [19] t0 0.482 us
T_kine.cluster_id 20 (is_associated, GID 1000005 = this candidate)
T_tagger.cluster_id 20 (is_associated, GID 1000005 = this candidate)
selected activity 20 (is_associated, GID 1000005 = this candidate)
reco Enu 529.3 MeV neutrino_type 10 = cosmic+NC
BDT scores numu -2.896 nue -10.718 cosmict +0.000
vertex (-179.2, 106.3, 0.8) cm inside active volume
```

```
nu_index 1: flash GID 1 (E side) bundle mains [] t0 nan us <-- EMPTY PLACEHOLDER
T_kine.cluster_id 8 (is_main, GID 1 = this candidate)
T_tagger.cluster_id 8 (is_main, GID 1 = this candidate)
selected activity 8 (is_main, GID 1 = this candidate)
reco Enu 0.0 MeV neutrino_type 0 = none(0)
BDT scores numu -1.942 nue -15.000 cosmict +0.000 <-- placeholder default
vertex (0.0, 0.0, 0.0) cm inside active volume
```

T_rec_charge (no nu_index -- one tree for the whole event): 322 space points under cluster_id [-1]

86 flagged vertex, 103 shower, 86 with no origin cluster

from flash GID 1000005 = nu_index 0, 38 origin cluster(s):

```
20:108 45:2 46:3 47:5 48:9 49:2 50:6 51:10 52:2 53:2 54:2 55:2
56:2 57:2 58:3 59:2 60:2 61:5 62:2 63:2 64:2 65:3 66:2 67:3
68:7 69:2 70:2 71:5 72:8 73:2 74:2 75:2 76:2 77:9 78:2 79:2
80:2 81:6
```

truth: 1 interaction(s), 0 signal, 1 background

int_idx	flavor	mode	curr	etot_MeV	edep_MeV	t_us	vtx_x	vtx_y	vtx_z	in_av	signal	background	d_nu0	dt_nu0	d_nu1	dt_nu1
1	numu	DIS	NC	6514.4	995.4	0.334	-123.755	196.445	-280.374	False	0	1	300.4	0.148	364.0	NaN

Case 5: Same nu seen on both sides

```
r472 s36 e40 8 trees: T_bad_ch, T_cluster, T_kine, T_proj, T_proj_data, T_rec_charge, T_tagger, Trun
48 clusters | 17 bundles (is_main & in_scope) | 2 of them in the beam window [0.2, 2.2) us
```

bundles (time-ordered):

cluster_id	flash_gid	side	n_assoc	npoints	length_cm	t0_us	flash_pe	in_window
22	1000009	W	2	1660	155.4	-1205.954	5138.5	
21	1000008	W	1	2016	139.2	-820.297	5315.7	
12	1000016	W	0	15	2.1	-621.137	671.7	
15	1000013	W	0	125	8.9	-567.688	951.7	
1	10	E	0	7	1.5	-493.620	72.2	
11	5	E	1	1122	104.2	1.577	11426.2	yes
23	1000006	W	2	4755	391.4	1.583	16718.8	yes
10	4	E	5	7118	452.4	69.368	23100.0	
18	1000005	W	7	3974	243.5	245.577	13513.3	
13	1000014	W	0	19	2.9	295.316	2043.3	
16	1000021	W	0	45	6.0	309.603	273.2	
9	6	E	1	1147	101.7	374.098	9522.1	
7	3	E	3	9896	412.4	425.143	21264.1	
3	2	E	0	1340	142.2	722.031	17403.1	
14	1000003	W	0	2107	194.0	774.205	24264.1	
17	1000002	W	0	705	86.6	978.463	20171.1	
6	0	E	3	3041	139.0	1435.903	42614.4	

2 candidate(s) written:

```
nu_index 0: flash GID 1000006 (W side) bundle mains [23] t0 1.583 us
T_kine.cluster_id      23 (is_main, GID 1000006 = this candidate)
T_tagger.cluster_id    23 (is_main, GID 1000006 = this candidate)
selected activity      23 (is_main, GID 1000006 = this candidate)
reco Enu 1313.7 MeV  neutrino_type 4 = numuCC
BDT scores  numu +3.433  nue -15.000  cosmict +0.000  <-- nue at its -15 default: that BDT never ran on this candidate
vertex (0.8, -140.7, 142.4) cm  inside active volume
```

```
nu_index 1: flash GID 5 (E side) bundle mains [11] t0 1.577 us
T_kine.cluster_id      11 (is_main, GID 5 = this candidate)
T_tagger.cluster_id    11 (is_main, GID 5 = this candidate)
selected activity      11 (is_main, GID 5 = this candidate)
reco Enu 849.4 MeV  neutrino_type 8 = NC
BDT scores  numu +1.123  nue -15.000  cosmict +0.000  <-- nue at its -15 default: that BDT never ran on this candidate
vertex (-10.4, -192.5, 10.1) cm  inside active volume
```

T_rec_charge (no nu_index -- one tree for the whole event): 1197 space points under cluster_id [11, 23]

20 flagged vertex, 92 shower, 20 with no origin cluster

from flash GID 5 = nu_index 1, 1 origin cluster(s):

11:205

from flash GID 1000006 = nu_index 0, 3 origin cluster(s):

23:683 47:2 48:287

truth: 1 interaction(s), 1 signal, 0 background

int_idx	flavor	mode	curr	etot-MeV	edep-MeV	t_us	vtx_x	vtx_y	vtx_z	in_av	signal	background	d_nu0	dt_nu0	d_nu1	dt_nu1
1	numu	RES	CC	2766.6	1765.1	1.443	-10.4684	-193.253	9.18667	True	1	0	143.7	0.14	1.2	0.134