Numu CC Xsec
Mock Data Challenge

Enrique Arrieta Díaz

Quantil Meeting

5/4/2017

INTRODUCTION

- This is an analysis done for the NOvA experiment with the purpose of testing a technique developed to measure a physical property called:
  - **Cross Section**
  - How often an interaction between particles occurs
  - Neutrinos with the materials of the NOvA detectors
  - The current measurements have large uncertainty ~20%
  - NOvA should present results with uncertainty ~10%
  - Useful to test current competing theoretical models describing the interactions

- Neutrinos are fundamental particles, i.e. without internal structure
  - Electric charge = 0
  - 10^{15} neutrinos go through your bodies each second, constantly
    - Most of them produced at the Sun, others produced in the atmosphere by cosmic rays
  - Only 2 in your lifetime will interact with your bodies
  - Extremely tough to catch, therefore, extremely tough to detect

- Brief introduction of the NOvA detectors

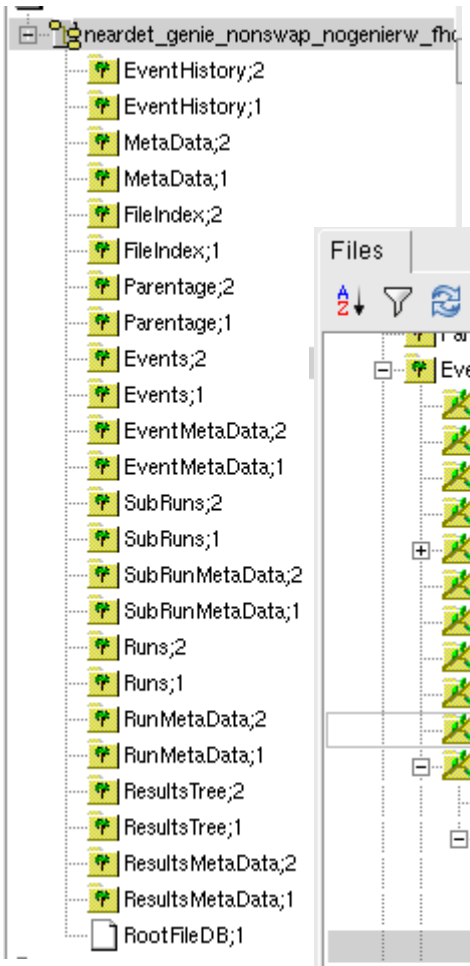
- Very little physics

- Focus on the analysis technique

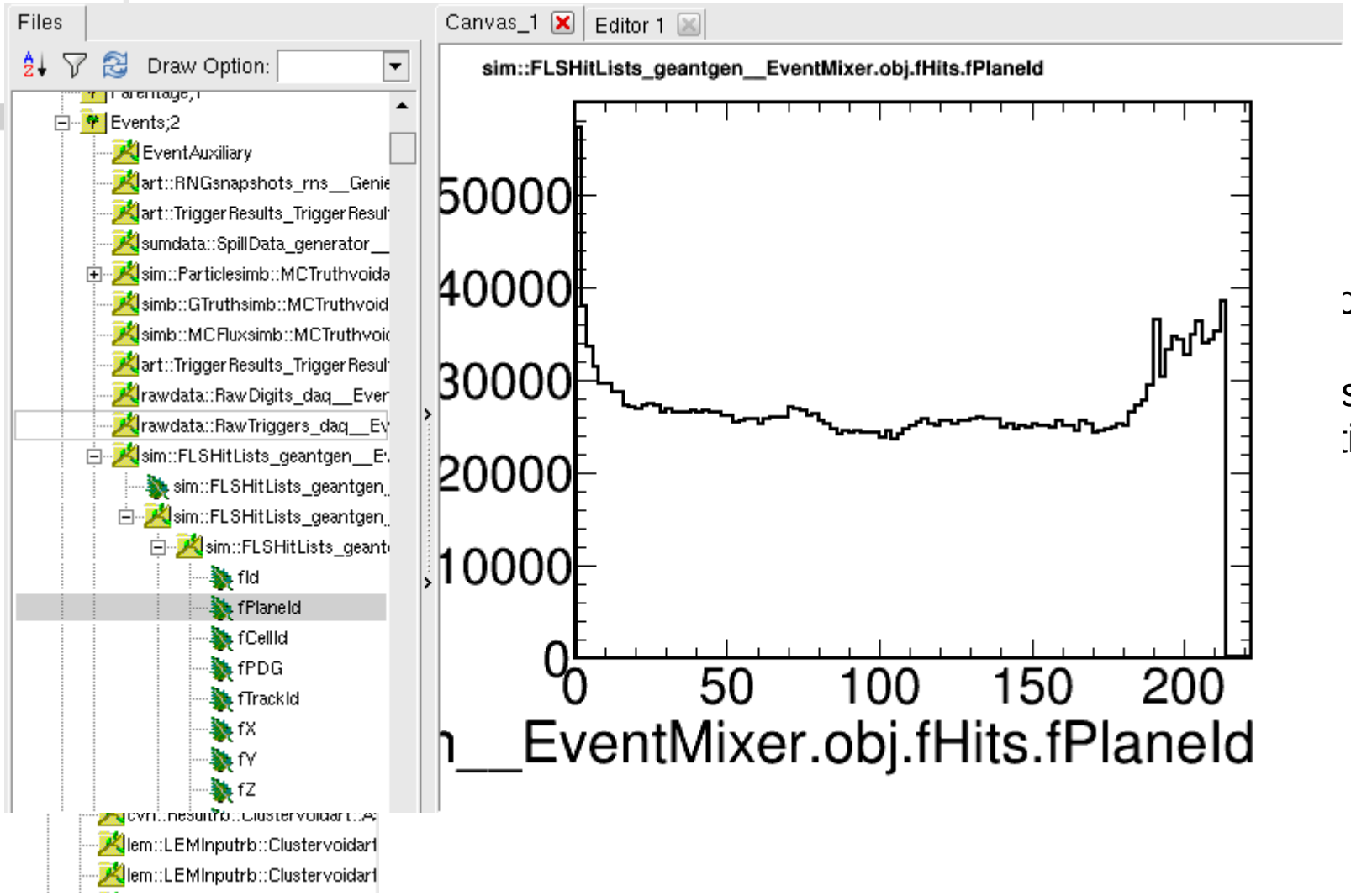
- NOvA simulates the entire scope of the experiment in order to:
  - Develop analysis algorithms
  - Develop particle and event identification algorithms
  - Calibration
  - Detectors performance and capabilities
  - Physics, how the interactions between neutrinos and matter occur
  - Simulators (Monte Carlo)
    - GEANT4: Simulates how the interactions between particles and matter propagate through
      matter (and vacuum)
    - GENIE: Simulates the physics of the interactions, i.e. what happens at the interaction point

- Simulation chain:
  - Entire process of propagation of neutrinos from source to detector
  - A neutrino interaction occurs inside the detectors
  - Every physical property of the interaction is known (True information)
  - All the reconstruction algorithms are applied to the Monte Carlo simulated data
    - Construction of hit clusters
    - Construction of tracks from hits
    - Particle identification
    - Energy estimation

- NOvA “cheats” by using the simulated data to know exactly what happened in an interaction in order to develop and test various analysis algorithms



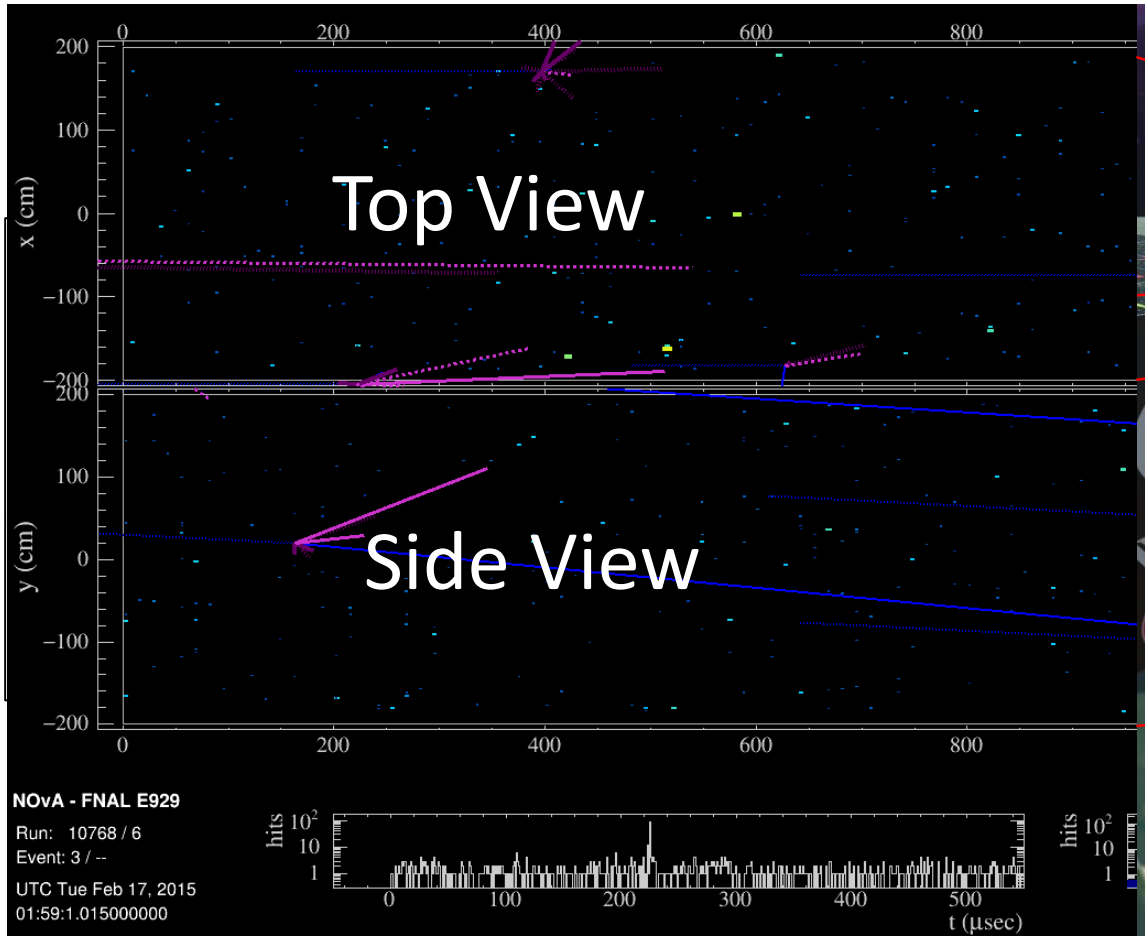
- NOvA produces large amounts of simulated and real world data
- Data is gathered in various types of files
  - Files have a tree structure that relates all interaction variables and quantities

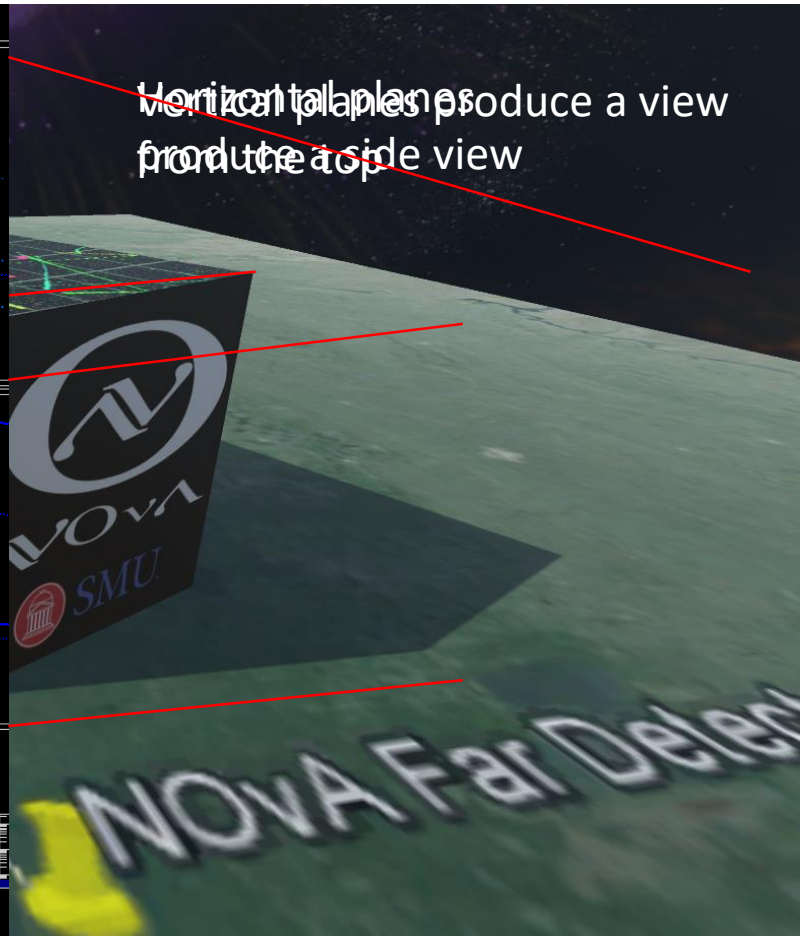


ctors

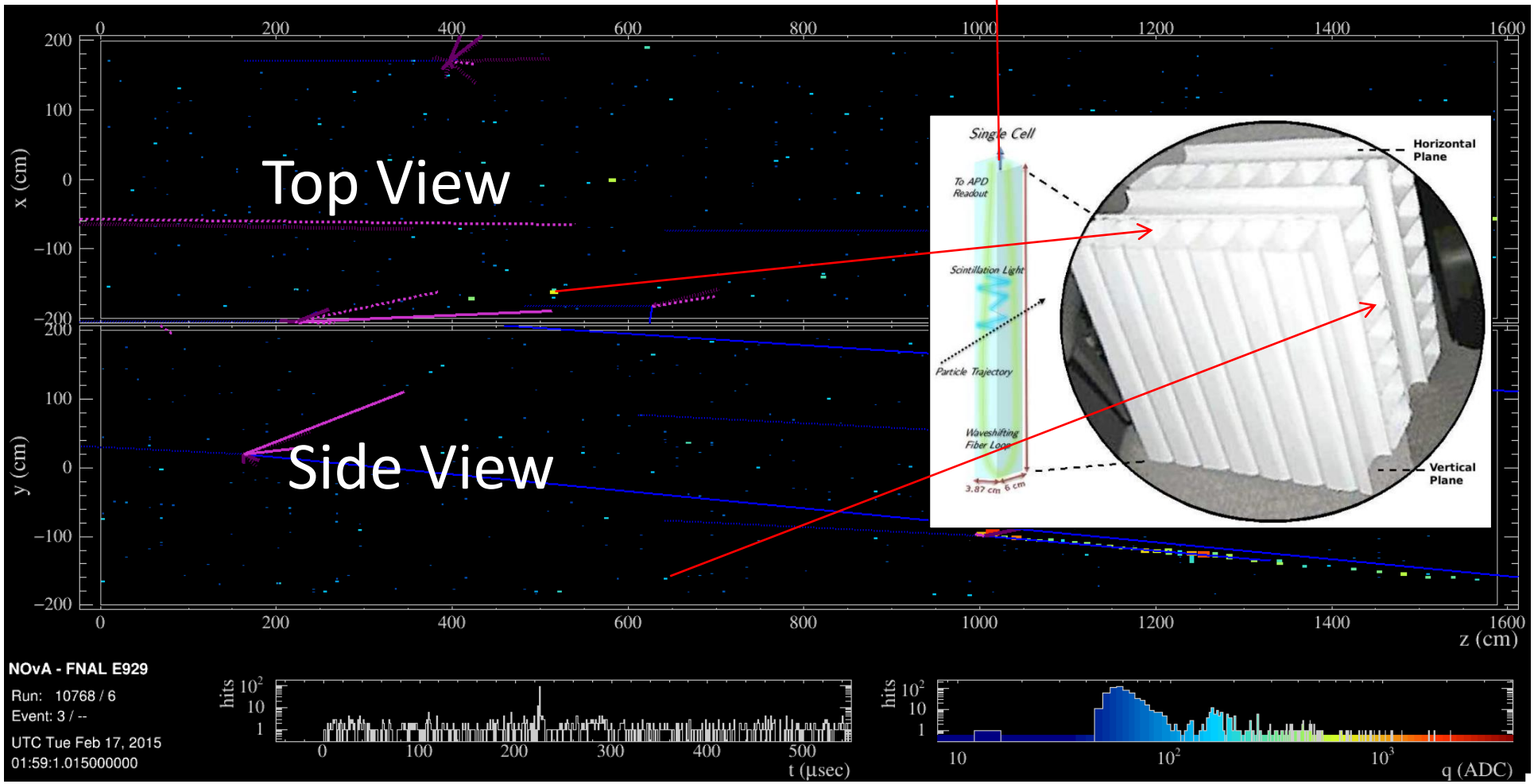
s

tion


Top View
Side View
x (cm)
y (cm)
NOvA - FNAL E929
Run: 10768 / 6
Event: 3 / --
UTC Tue Feb 17, 2015
01:59:1.015000000
hits
t (μsec)

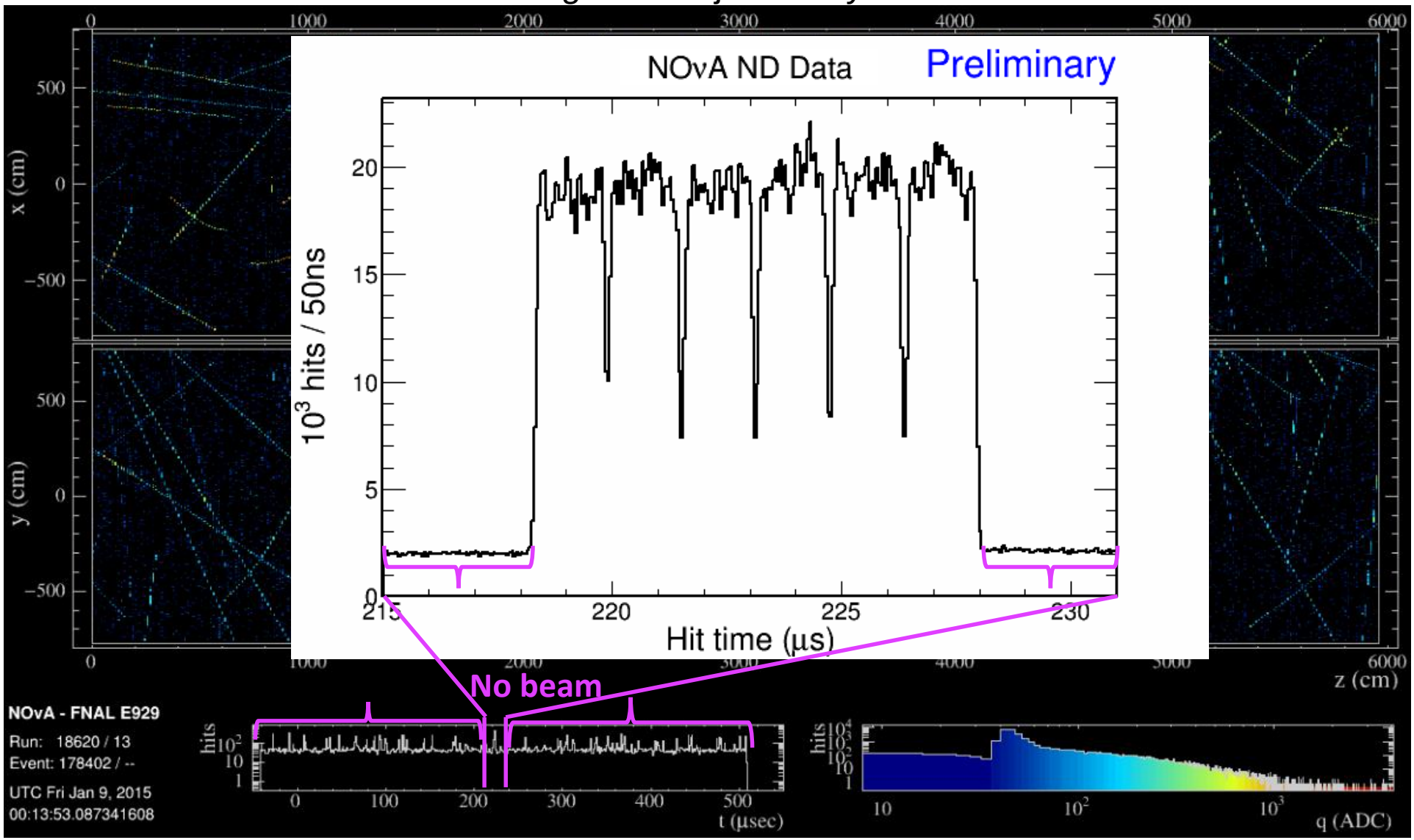

Horizontal planes produce a view from the top
Vertical planes produce a side view
SMU
NOvA Far Detec

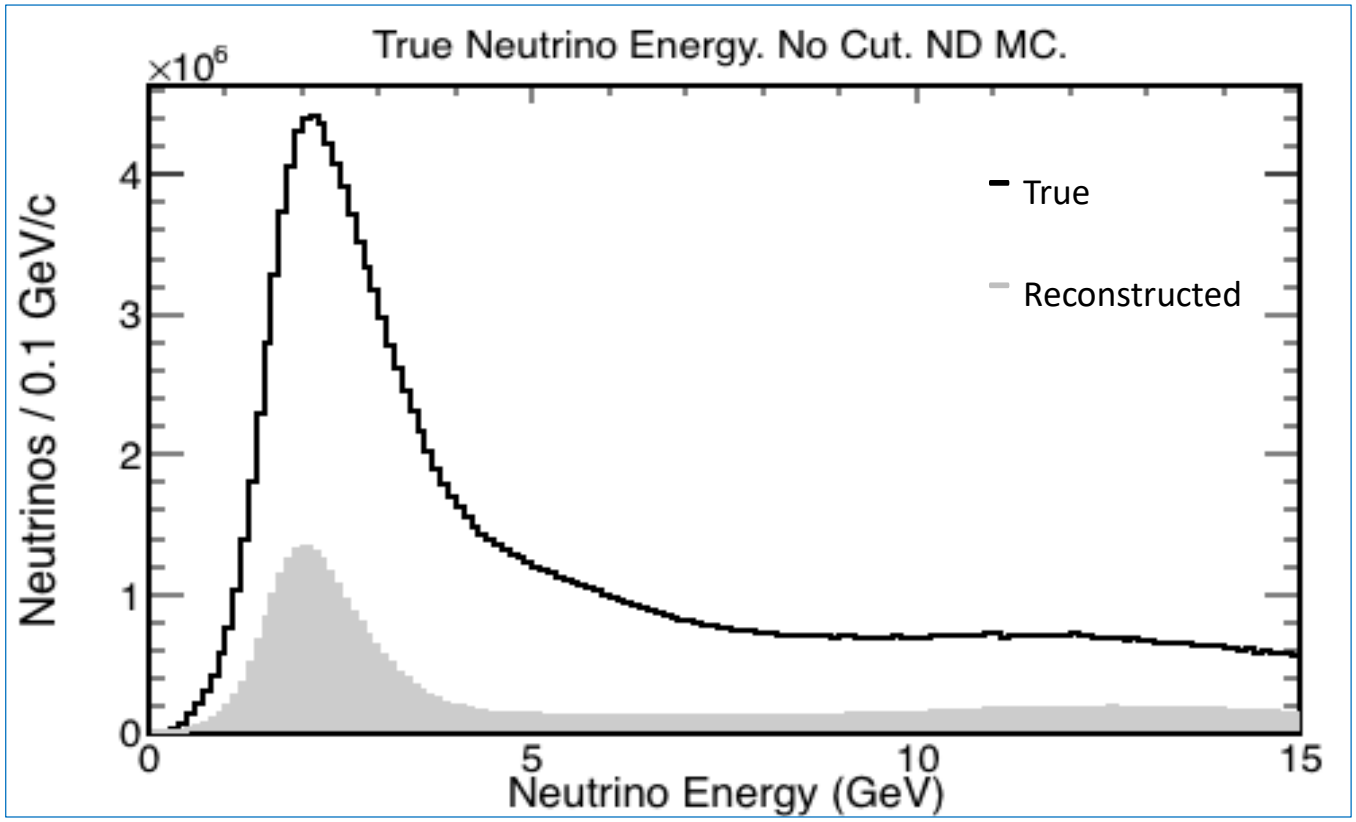
Each colored dot
represents a single cell



Brief introduction to NOvA

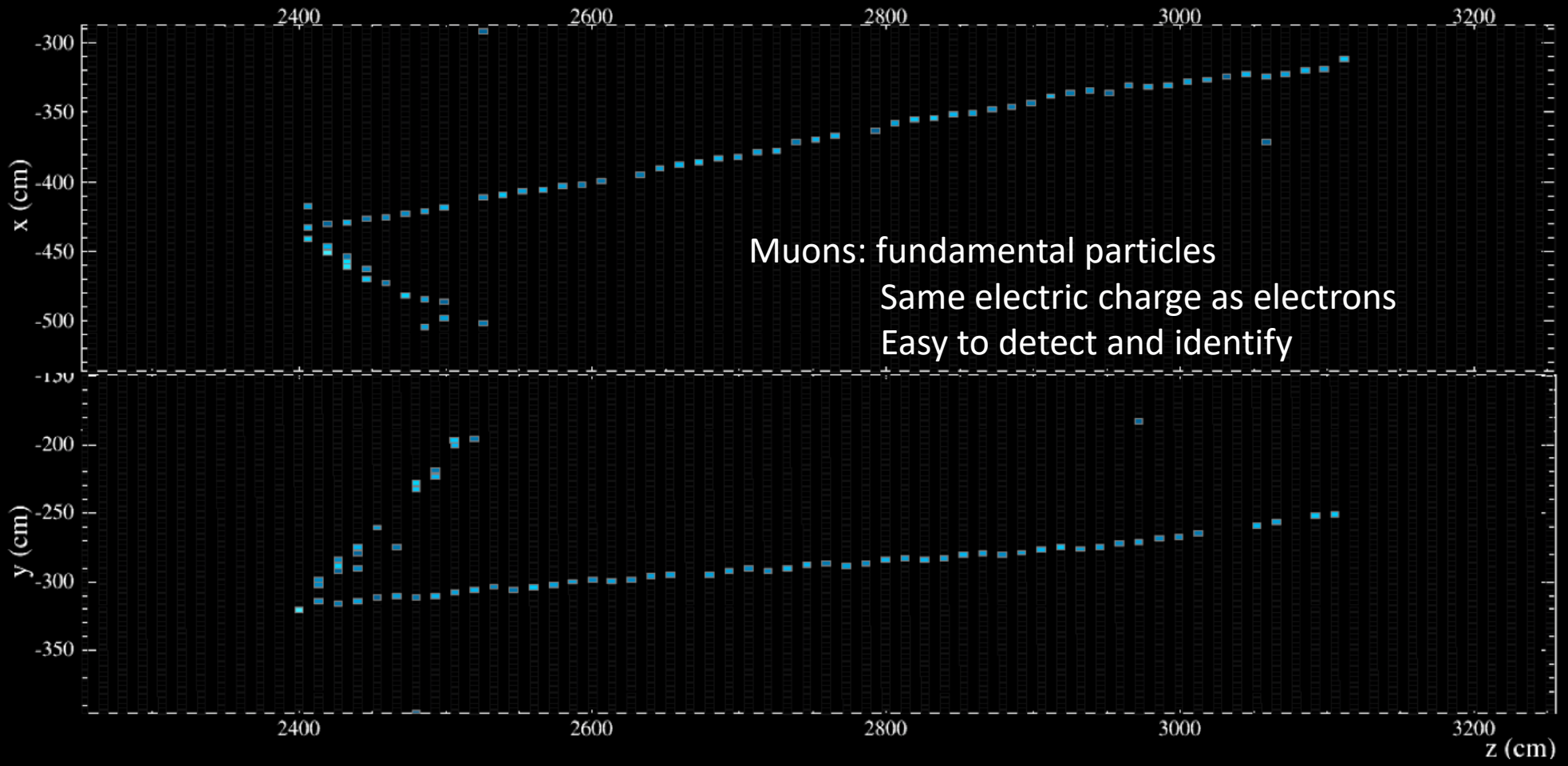
Cosmic background rejection by a factor of 10^7

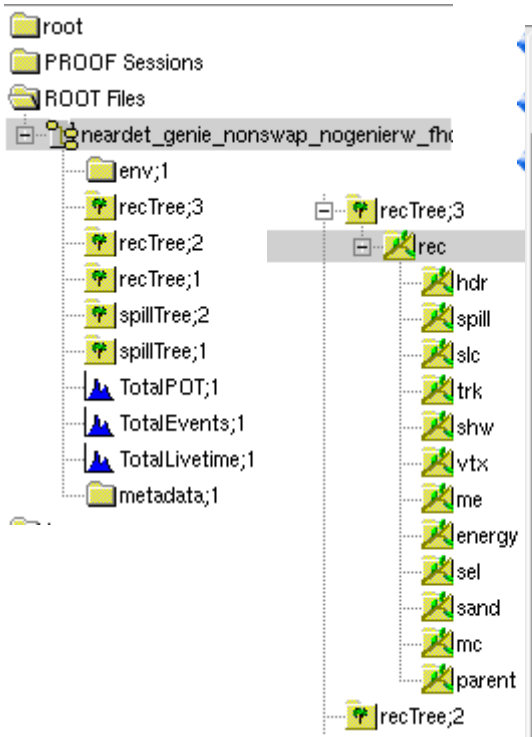


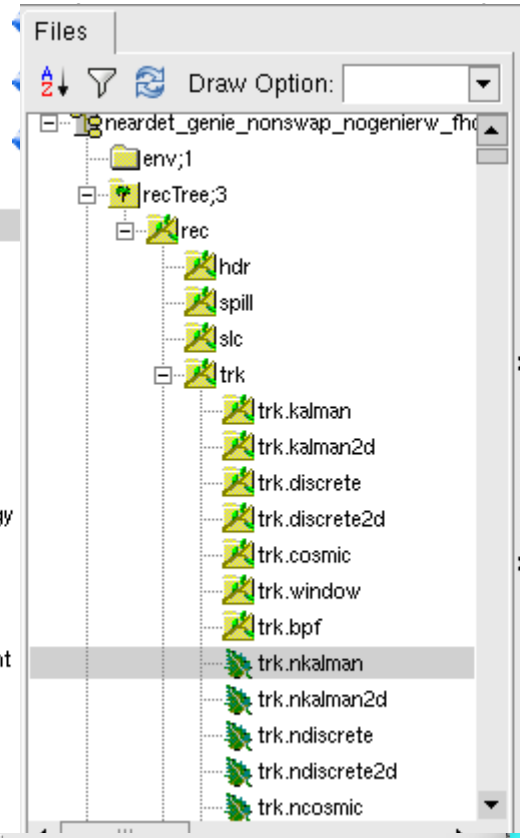


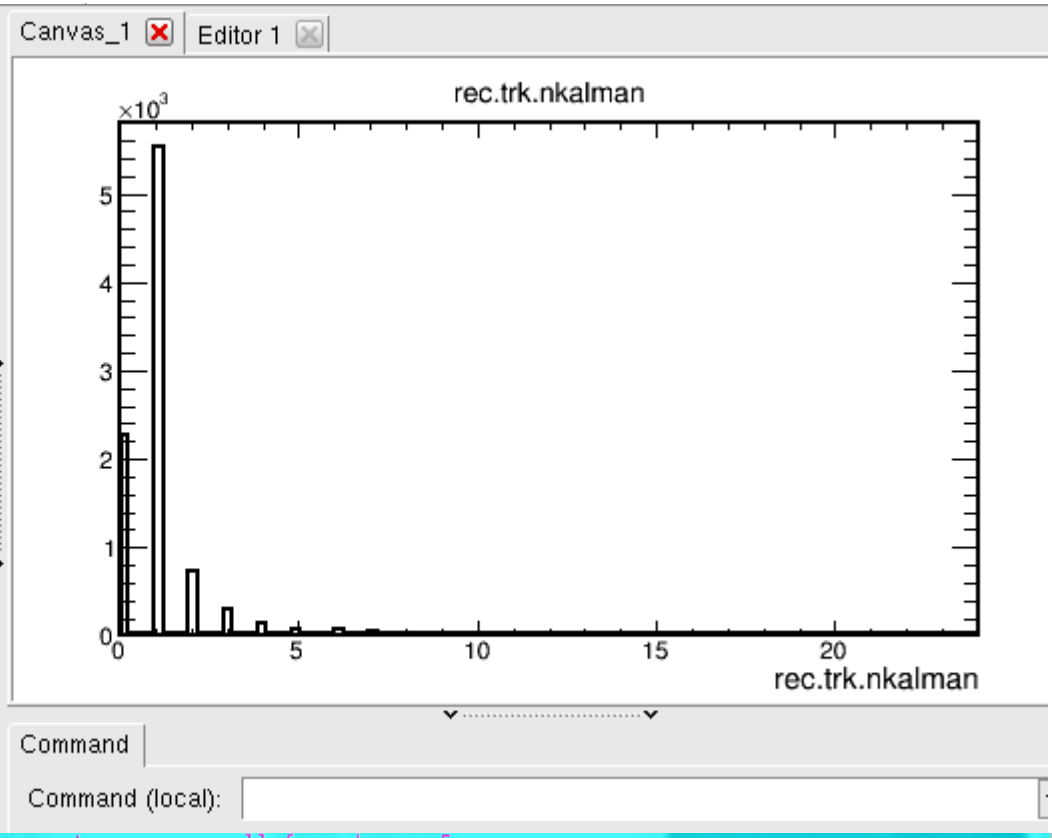
- Spectra from simulation
- True: All interactions occurring inside the detector
- Reconstructed: All interactions detected by NOvA's analysis algorithms

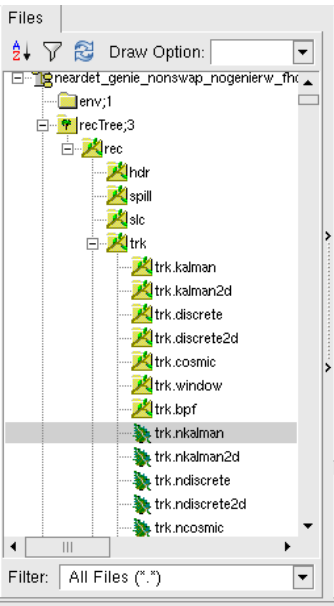
NOvA ν_μ Charged-current candidate

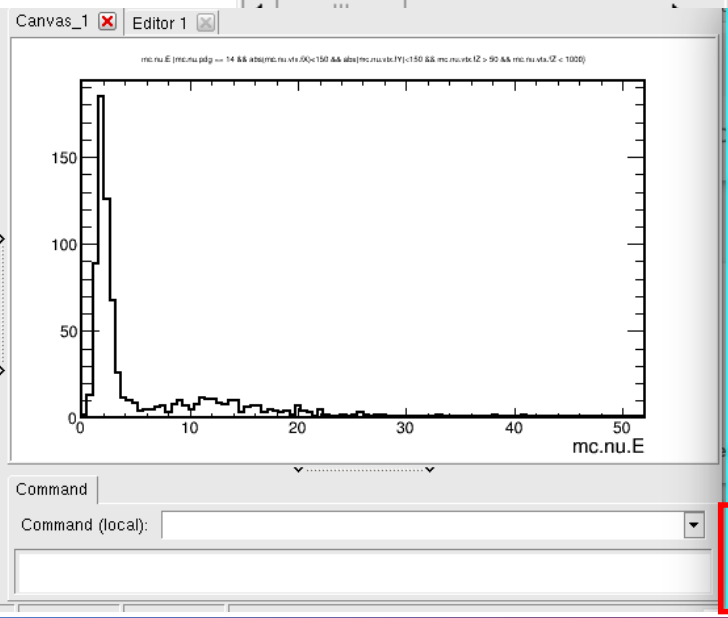






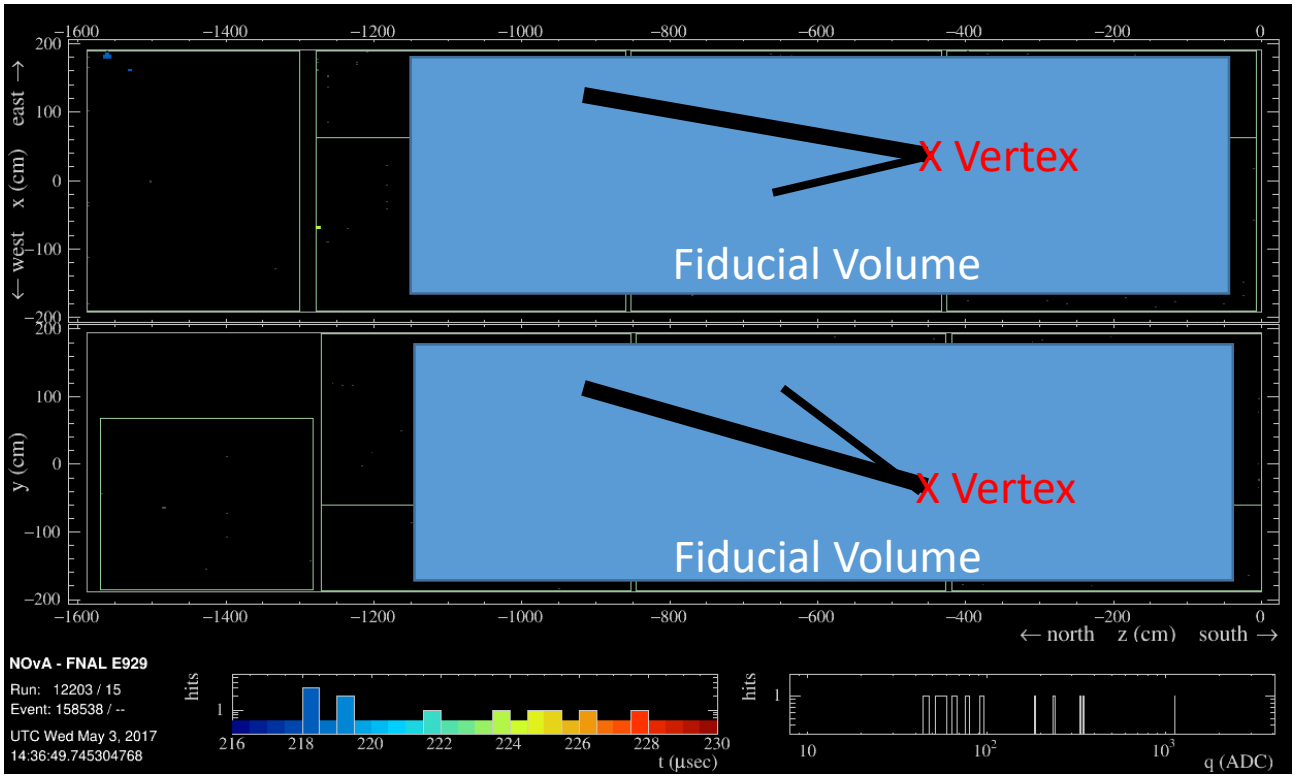







```
parent.mrccpar.ncellsfromedge = -5
parent.mrccpar.firstplane = -5
parent.mrccpar.lastplane = -5
parent.mrccpar.remid = -5.000000
parent.mrccpar.muonstart.fUniqueID = 0
parent.mrccpar.muonstart.fBits = 33554432
parent.mrccpar.muonstart.fX = -5
parent.mrccpar.muonstart.fY = -5
parent.mrccpar.muonstart.fZ = -5
parent.mrccpar.muonstop.fUniqueID = 0
parent.mrccpar.muonstop.fBits = 33554432
parent.mrccpar.muonstop.fX = -5
parent.mrccpar.muonstop.fY = -5
parent.mrccpar.muonstop.fZ = -5
parent.mrccpar.muonyattrans = -5.000000
parent.mrccpar.muonfwdcell = -5
parent.mrccpar.muonbkcell = -5
parent.mrccpar.numuE = -5.000000
parent.mrccpar.muE = -5.000000
parent.mrccpar.hadEinmucat = -5.000000
root [3] recTree->Draw("mc.nu.E","mc.nu.pdg == 14 && abs(mc.nu.vtx.fX)<150 && abs(mc.nu.vtx.fY)<150 && mc.nu.vtx.fZ > 50 && mc.nu.vtx.fZ < 1000")
(Long64_t)770
root [4]
```

- Fiducial Volume, for vertex containment:
  - -176 cm < X < 177 cm
  - -172 cm < Y < 179 cm
  - 25 cm < Z < 1150 cm

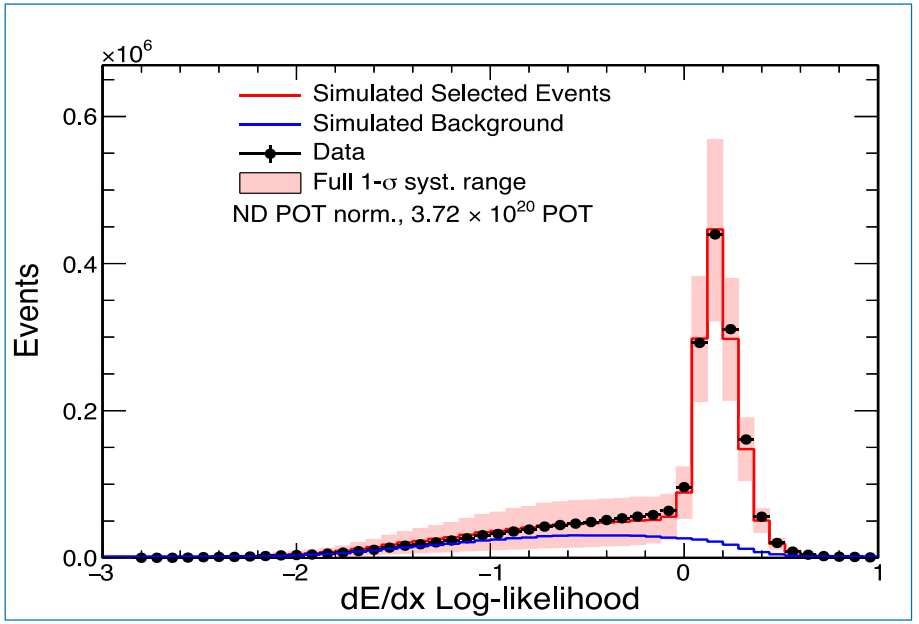


- Quality:
  - slc.nhit > 20
  - slc.ncontplanes > 4
  - sel.remid.pid > 0
  - energy.numusimp.trkccE > 0

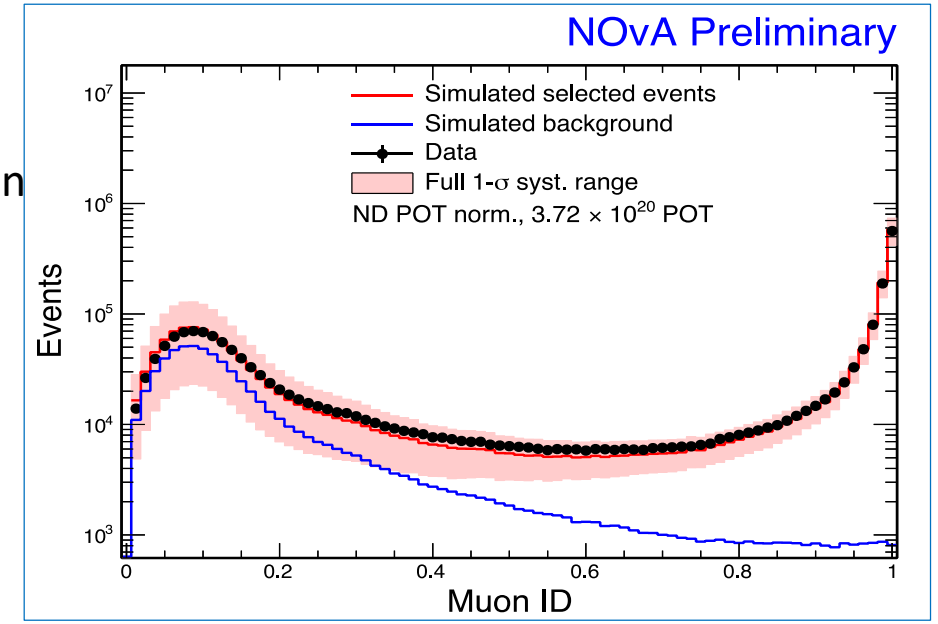
- Containment, CAFAna:
  - trk.nkalman > sel.remid.bestidx
  - slc.ncellsfromedge > 4
  - slc.firstplane > 4
  - slc.lastplane < 212
  - sel.contain.kalfwdcellnd > 4
  - sel.contain.kalbakcellnd > 8
  - trk.kalman[sr->.remid.bestidx].stop.fZ < 1275 || sel.contain.kalyposattrans < 55
  - energy.numusimp.ndhadcalcatE + energy.numusimp.ndhadcaltranE) < 0.03

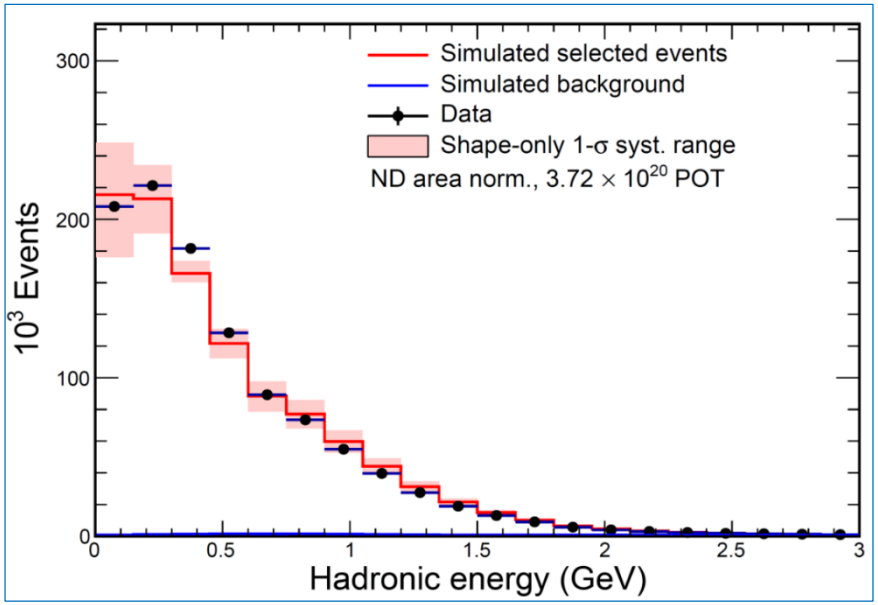
- RemID Score cut applied

- To identify muons, NOvA uses a set of variables
  - The main one is muon energy deposition per unit length: dE/dx
  - Muons track length
  - Number of consecutive planes with hits
  - A few other

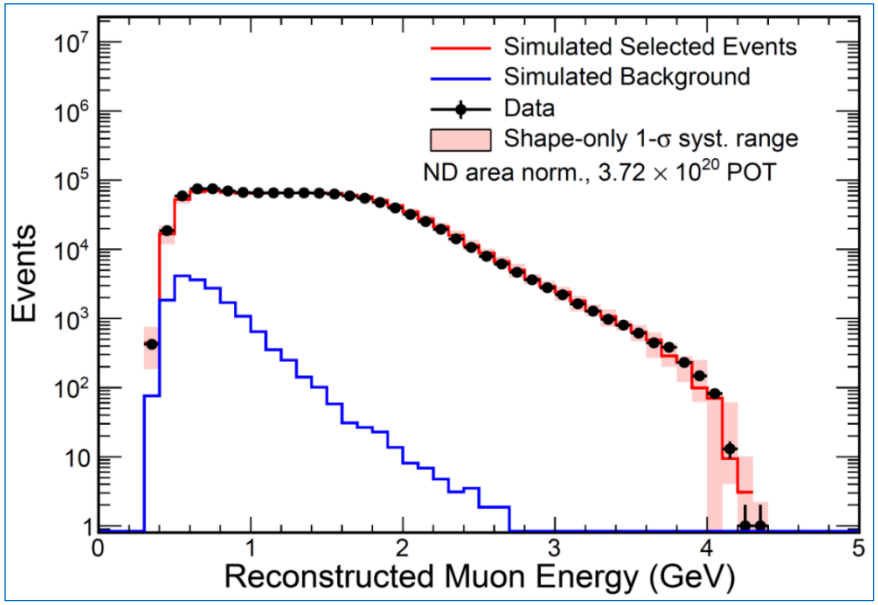


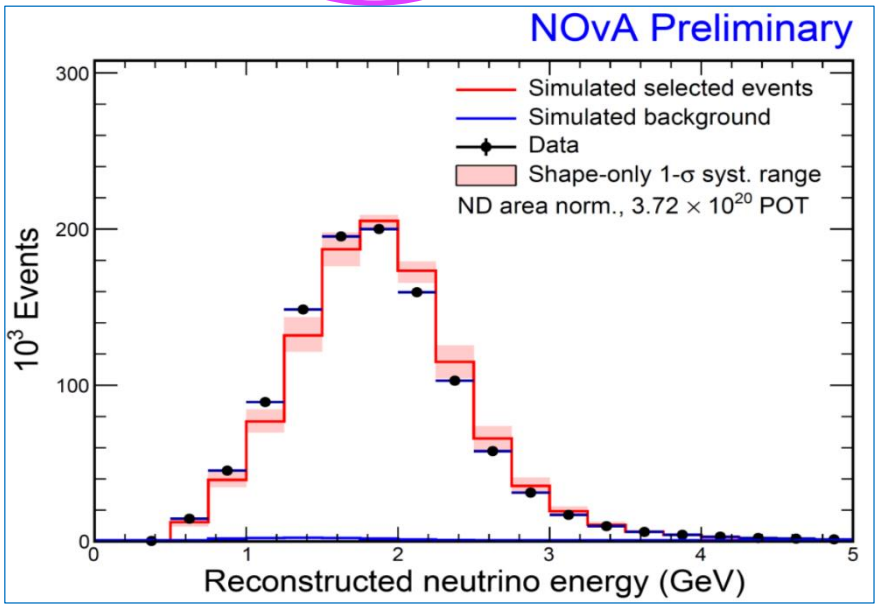
- Using a k-nearest neighbor (KNN) algorithm and a neural network, NOvA developed a muon identification (ID) algorithm:
  - Muon ID produces a score from 0 to 1
  - Higher scores mean event is very likely to have a muon
  - Low scores mean event is likely to not have a muon, i.e., event is background

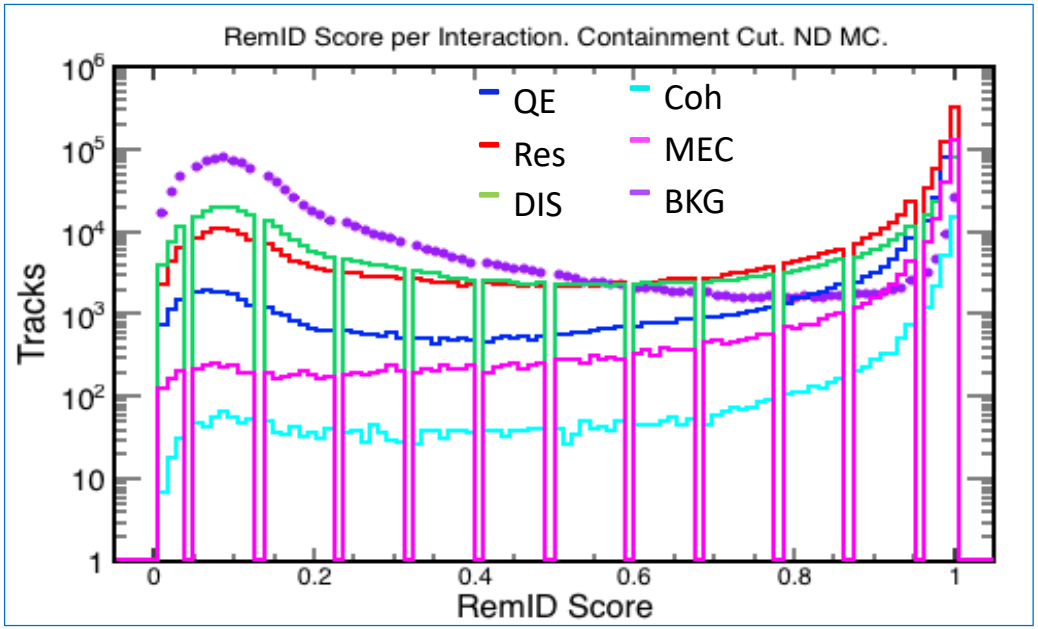


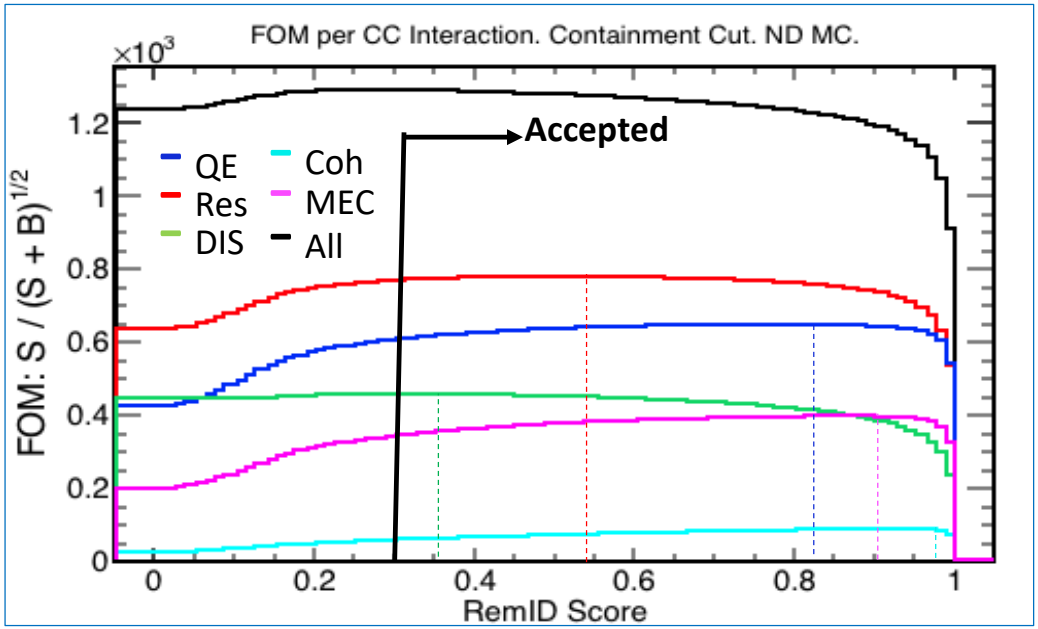


$E_n = E_{had} + E_{muon}$





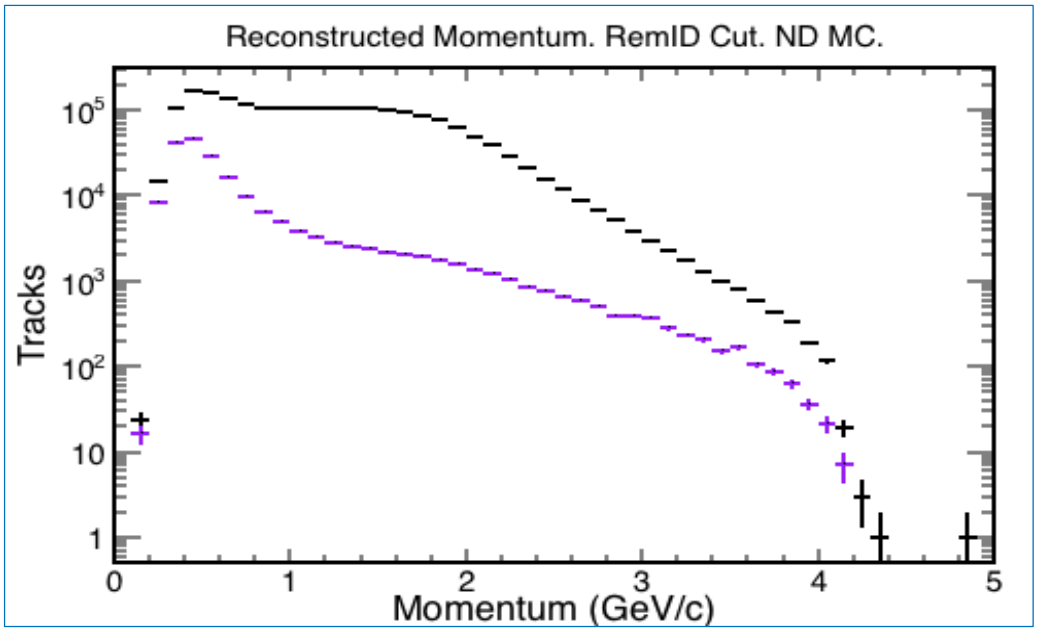


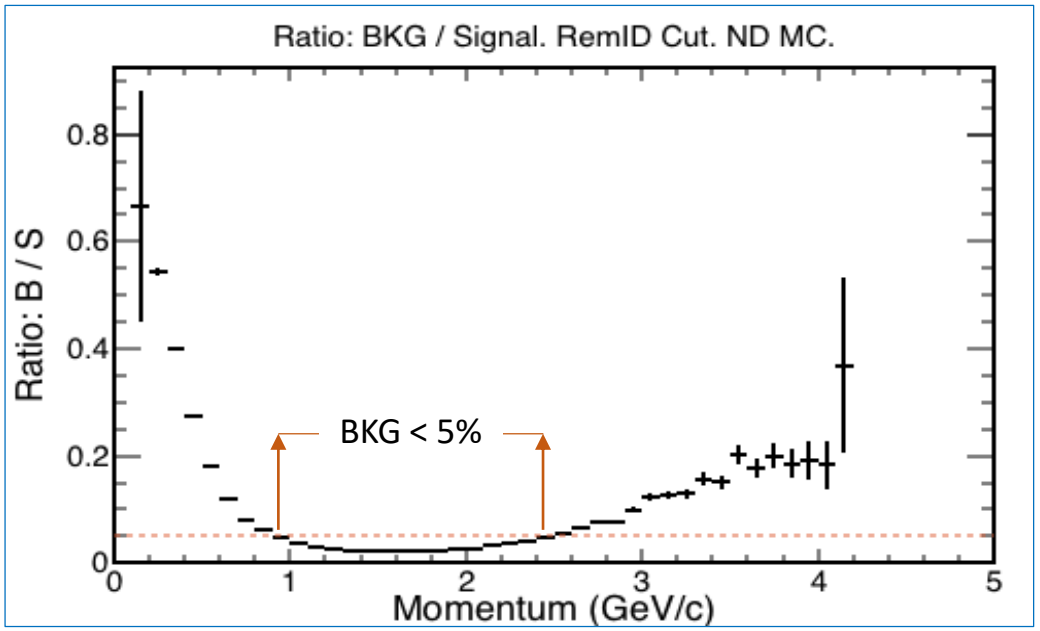


- $FOM = \frac{S}{\sqrt{S+B}}$, Figure Of Merit
- Each CC interaction is S
- B: NC + μ^+ + e, for each S
  - μ^+ events are not rejected by the RemID cut
    - In the to-do list

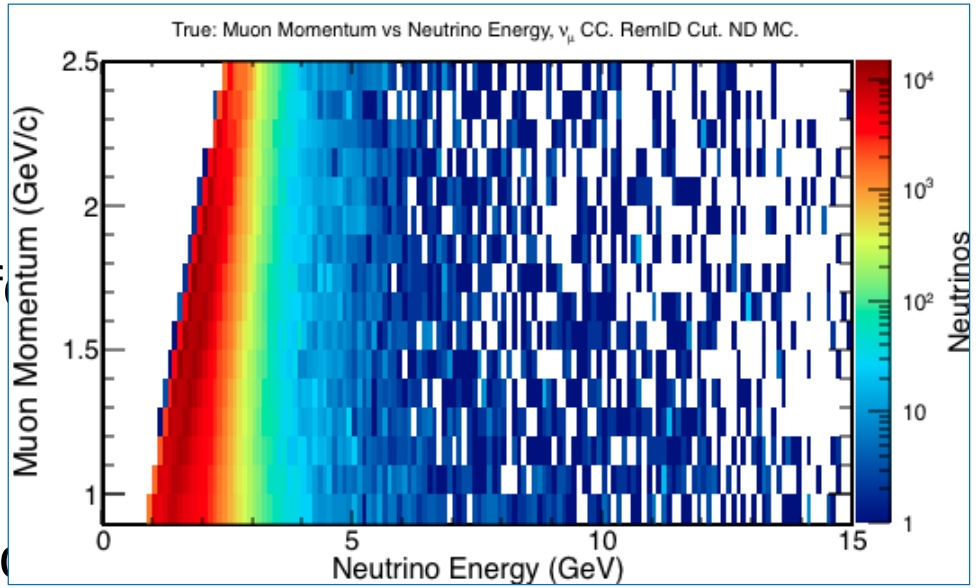
- FOM_{QE} = 651, at RemID Score = 0.8305
- FOM_{Res} = 781, at RemID Score = 0.5445
- FOM_{DIS} = 460, at RemID Score = 0.3575
- FOM_{Coh} = 92, at RemID Score = 0.9845
- FOM_{MEC} = 400, at RemID Score = 0.9075

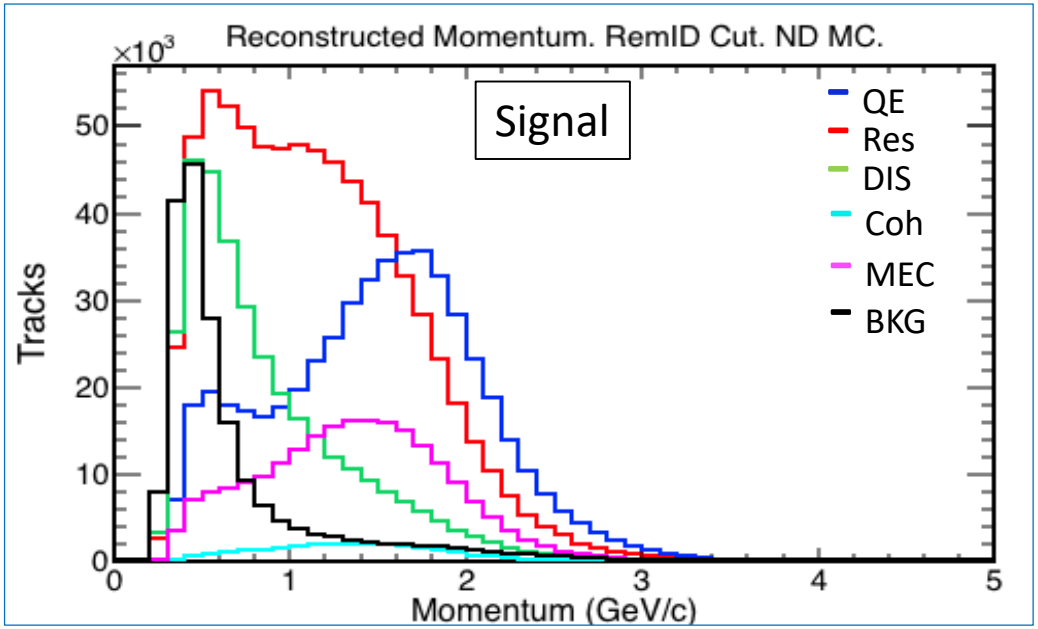
- FOM = 1292, at RemID Score = 0.3
  - **RemID Score = 0.3**
  - Everything with RemID Score > 0.3 is a μ^-

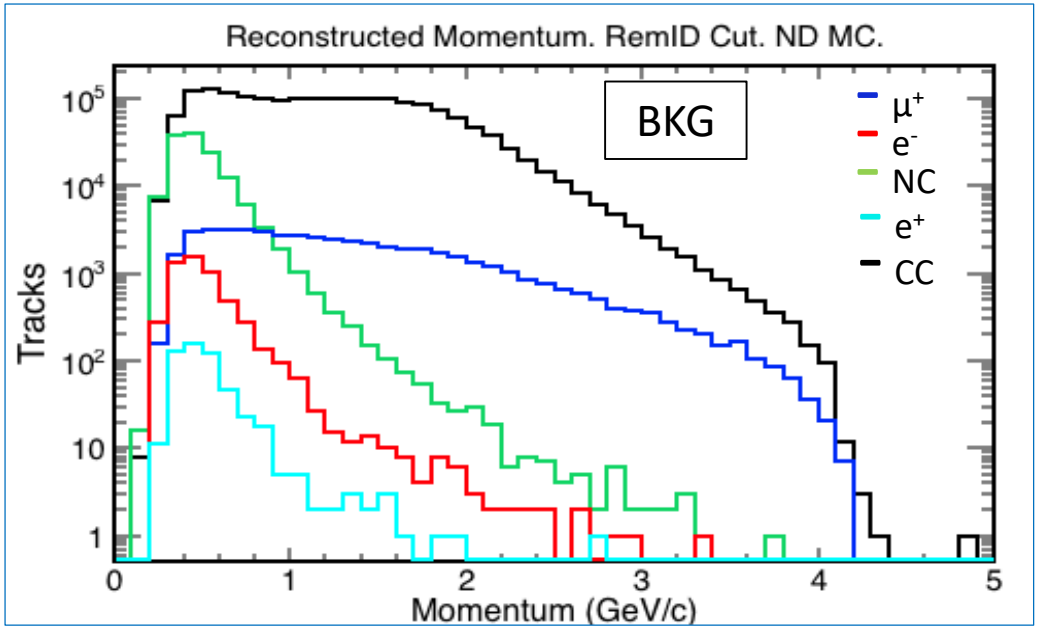




- After applying the RemID cut, the tracks passing the cut are presented:
  - Signal: CC interaction + BKG
  - BKG
- The BKG to Signal ratio is less than 5% f
  - 0.9 GeV/c < P < 2.5 GeV/c
  - For this region, only 0.25% of neutrin have E > 5 GeV







- Signal:
  - Selected events after reconstruction
  - Discriminated by CC type (Truth)

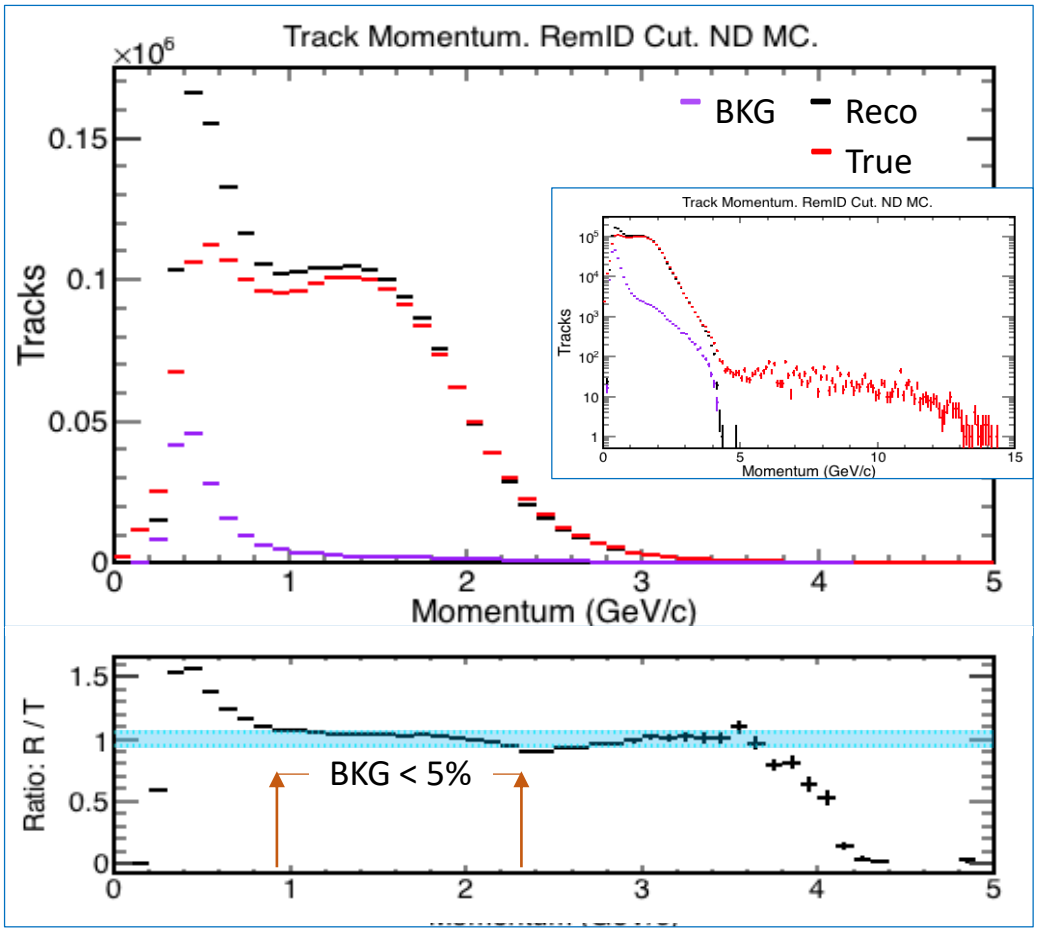
90.4% CC in Signal

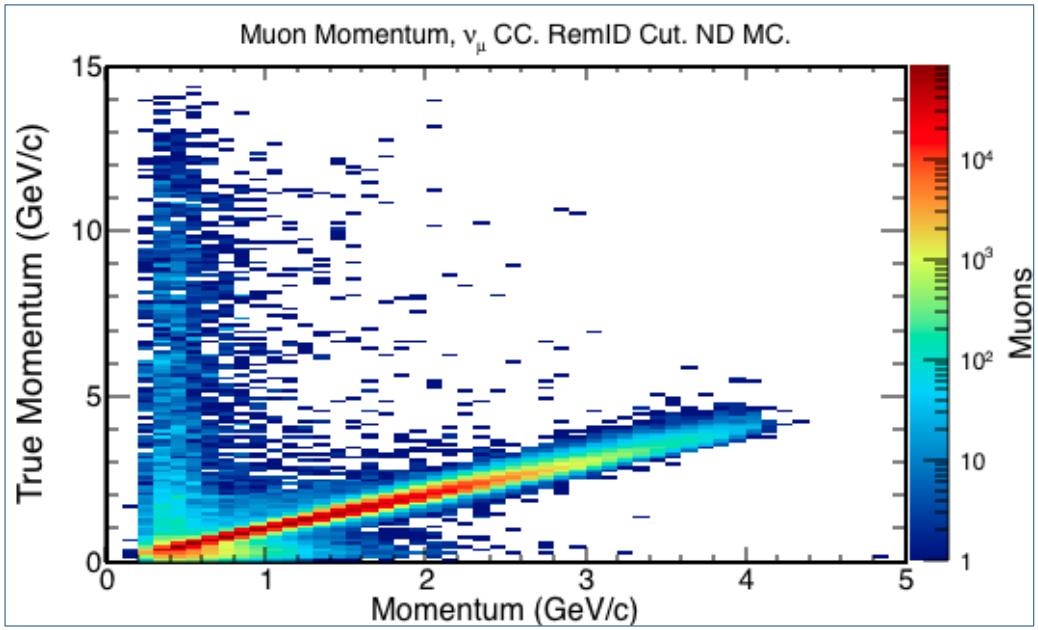
	QE	Res	DIS	Coh	MEC
CC %	27.7	40.6	18.1	1.6	12
Signal %	25	36.7	16.4	1.4	10.9

- BKG:
  - Selected events after reconstruction
  - Discriminated by NC, EM & μ^+ (Truth)

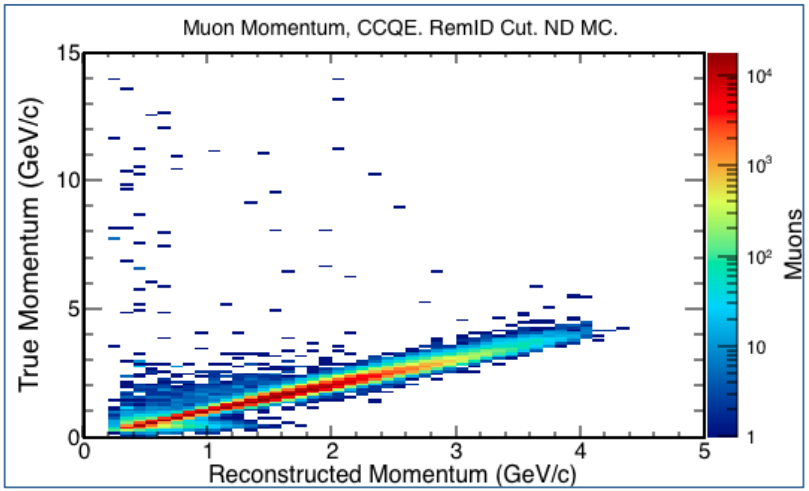
9.6% BKG in Signal

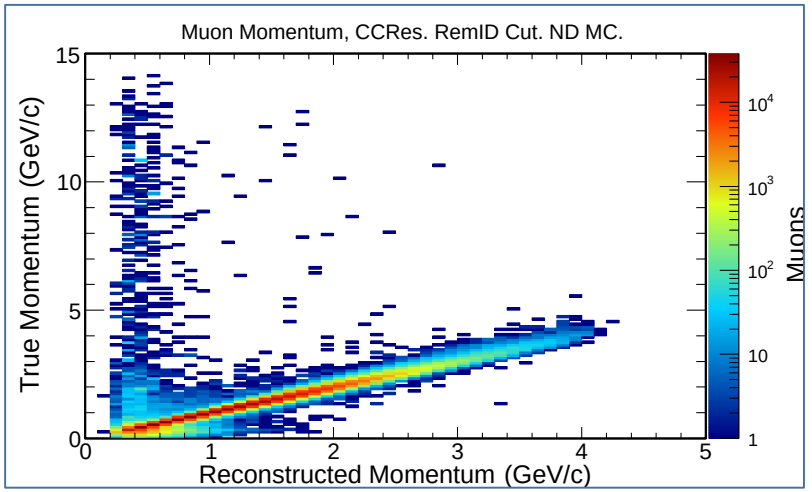
	NC	μ^+	e^-	e^+
BKG %	70.7	26.2	2.8	0.3
Signal %	6.7	2.5	0.3	0.1

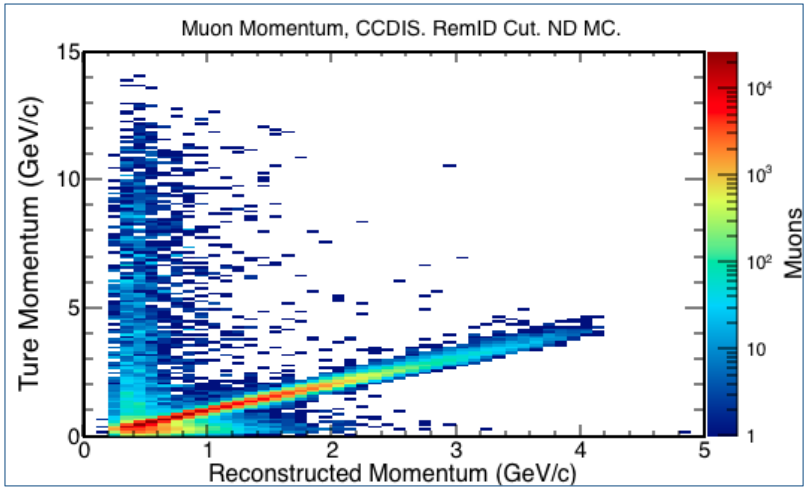


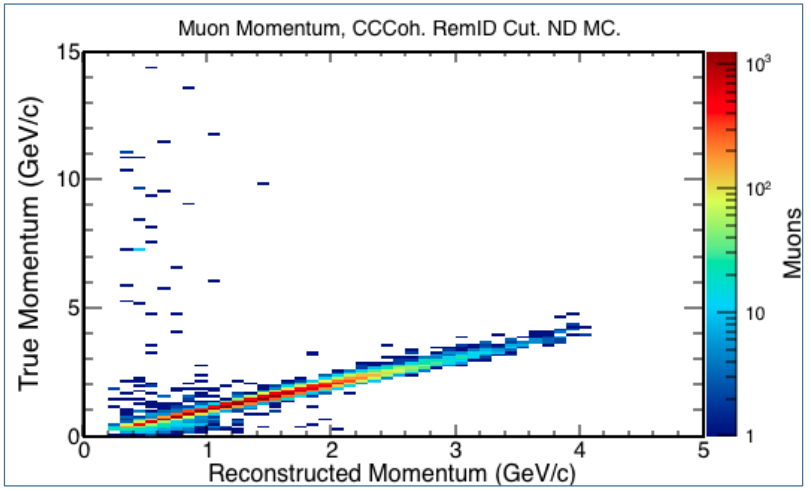


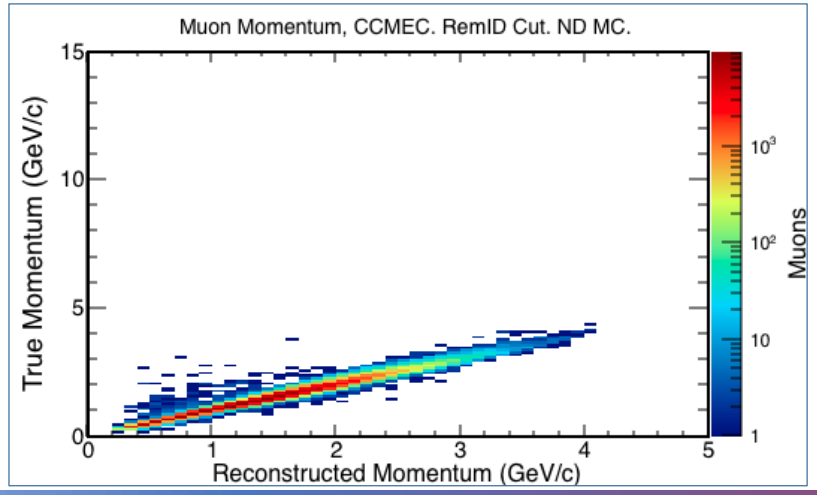
- Reconstructed μ momentum up to 5 GeV/c
- True μ momentum goes up to 15 GeV/c
  - This is mostly for low reco momentum
- 5% blue band in Reco / True ratio agrees with low BKG region
- Peak in Reco at 0.5 GeV/c mostly from BKG



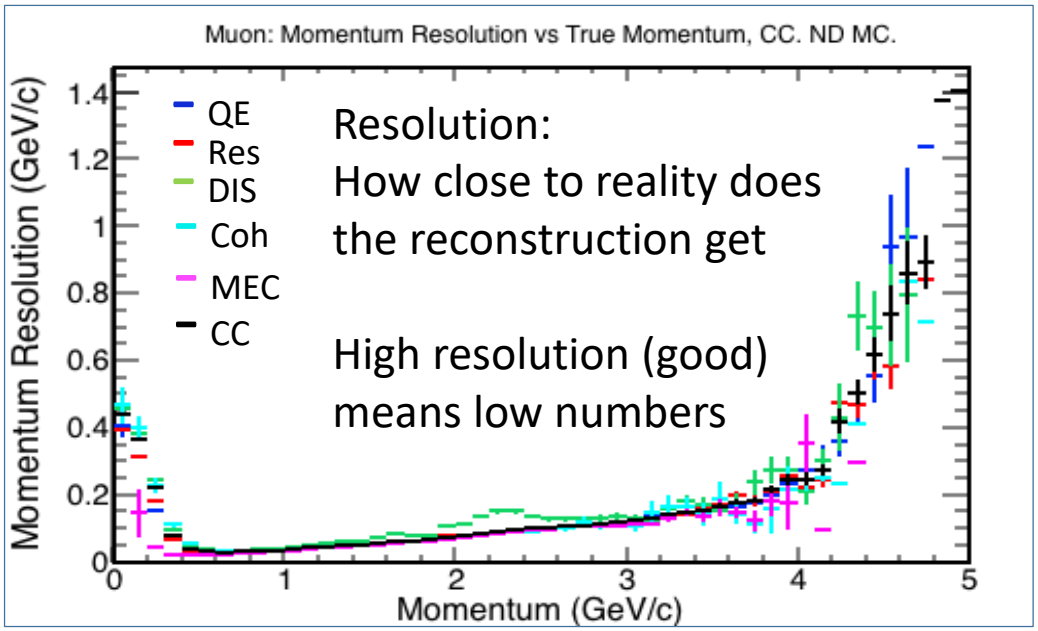


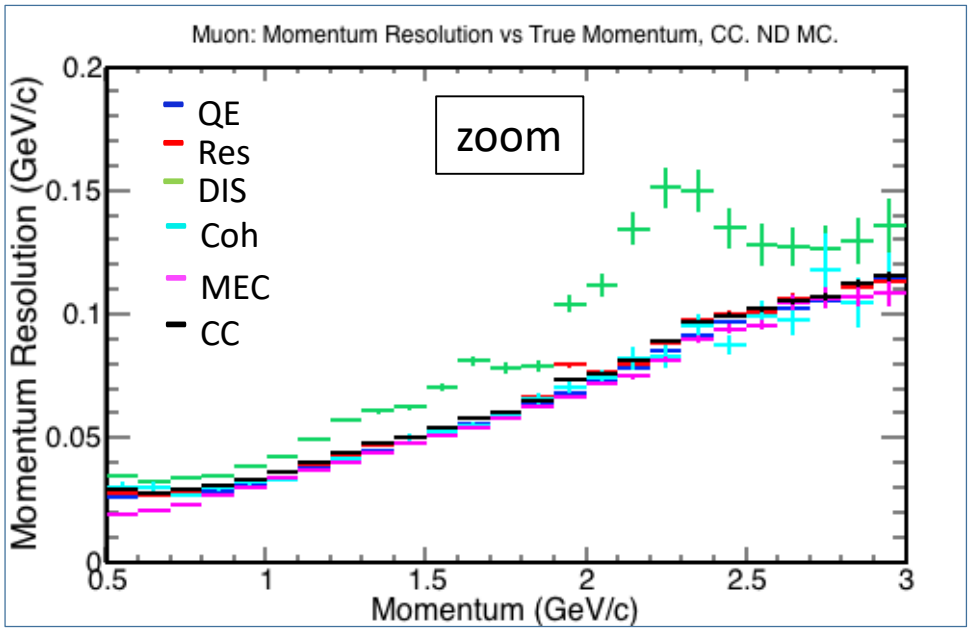






- The existence of a hadronic system in the interaction:
  Res & DIS, also contributes to the tale on True momentum for low Reco momentum
  - The best RemID Score track is not the muon,
    which is likely escaping the ND

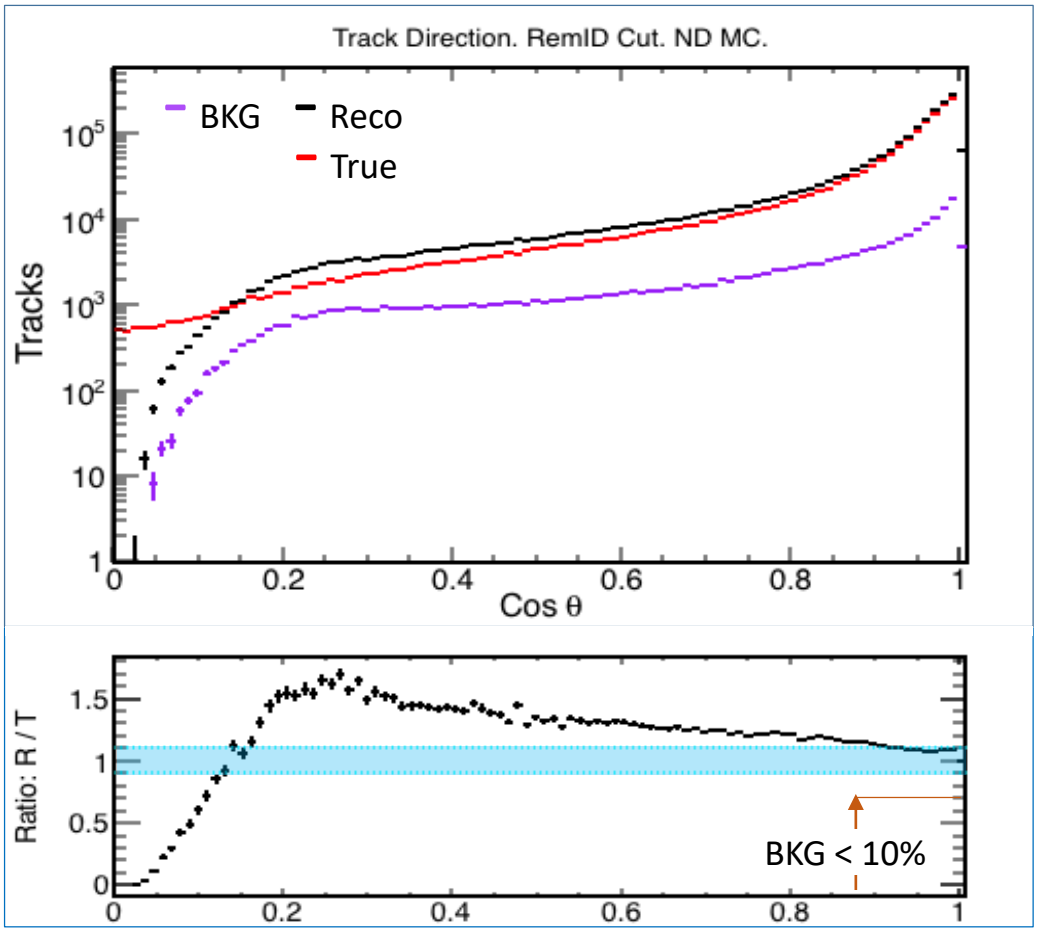


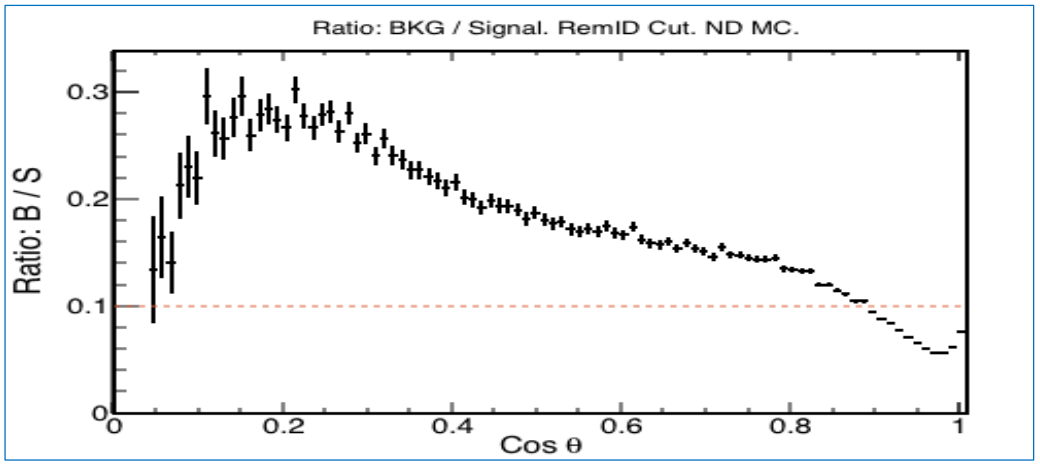


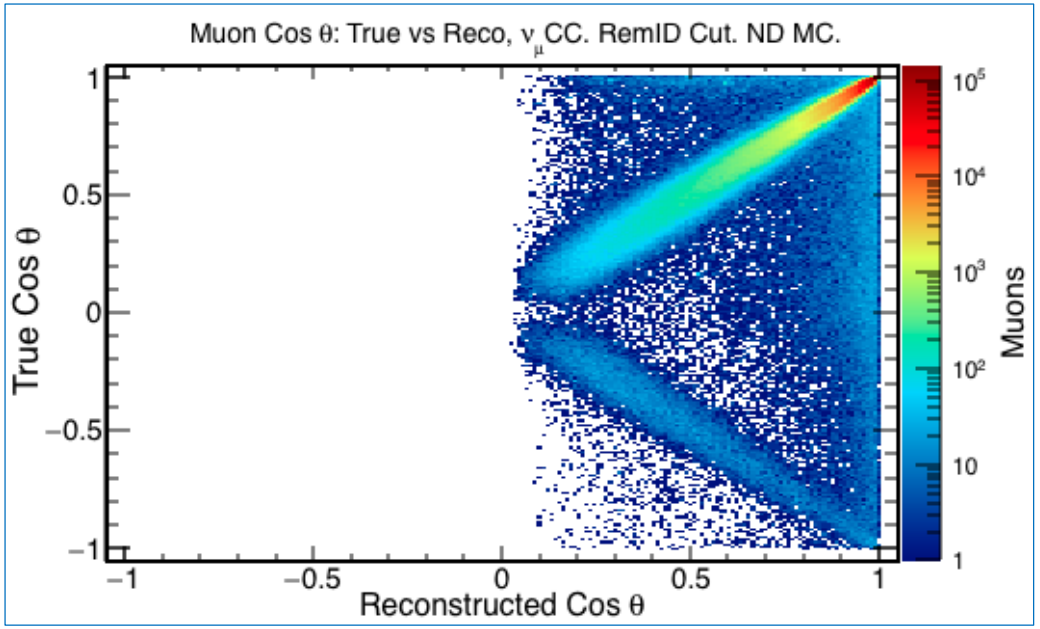
- Muon momentum resolution is very good for:
  - 0.7 GeV/c < P < 2.1 GeV/c
  - DIS is ~28% worse on average

P (GeV/c)	0.5	1.0	1.5	2.0	2.5
Resolution %	7.5	3.3	3.4	3.7	4.0

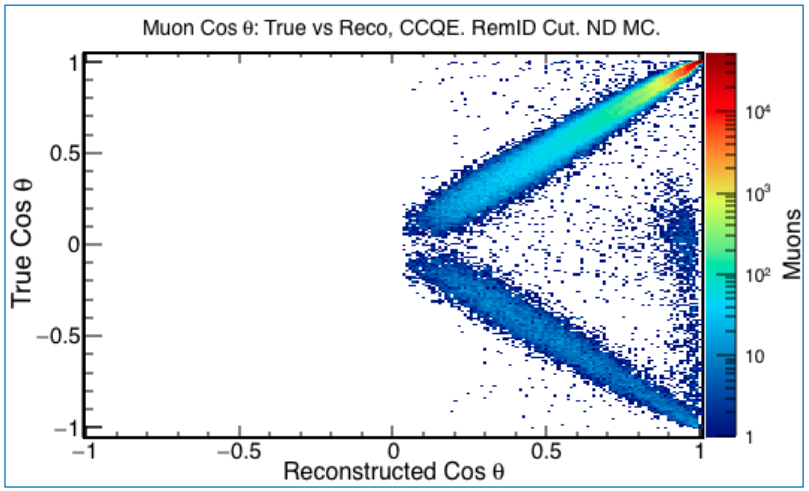
- With this momentum resolution, bins of **0.1 GeV/c** are set for use
- The BKG for P = 0.5 GeV/c is 18%, and goes down to 6% for P = 0.8 GeV/c
- **16 bins** from 0.9 GeV/c to 2.5 GeV/c are reasonable for a double differential cross section measurement with **5% BKG**

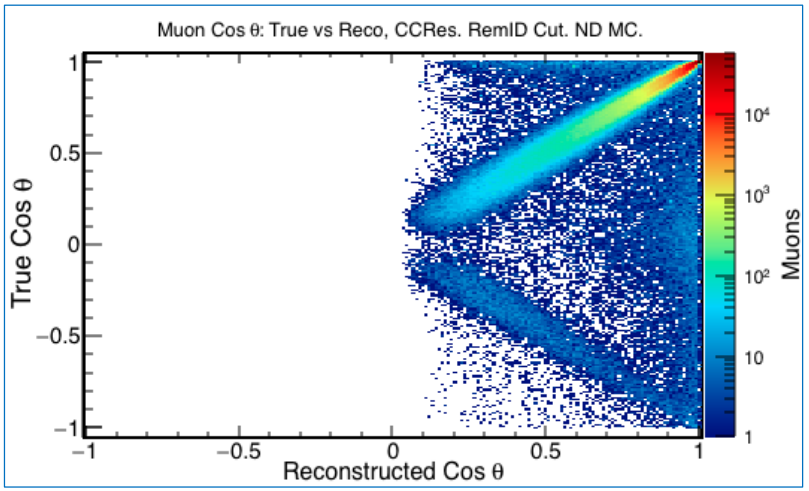


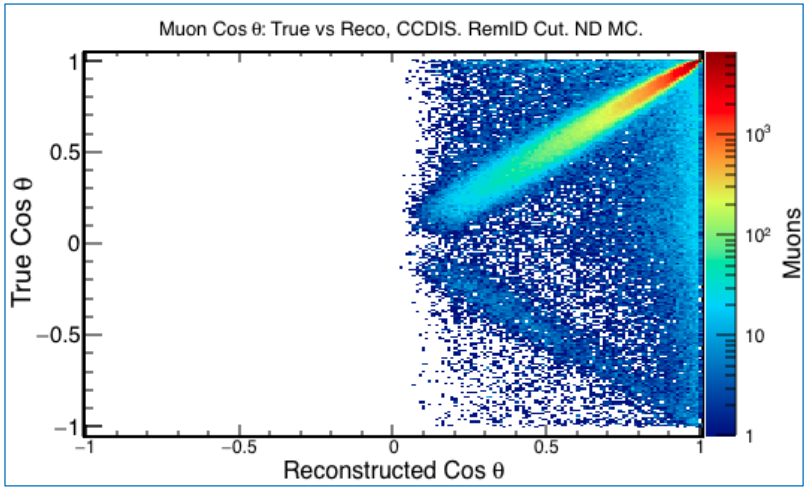


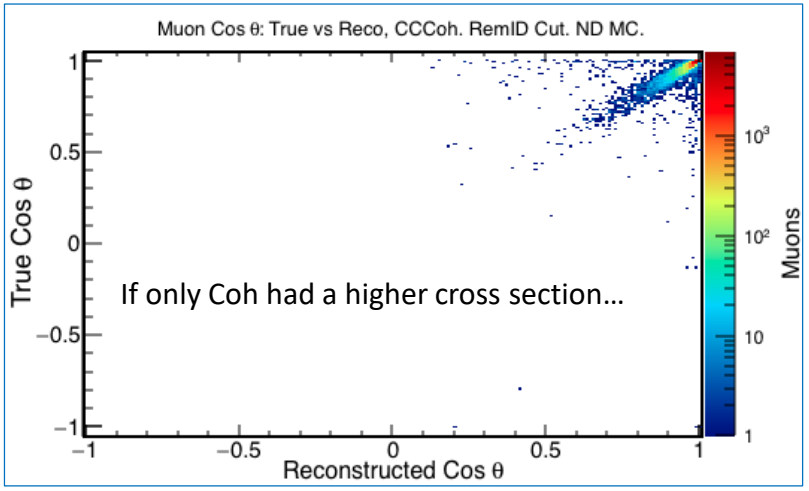


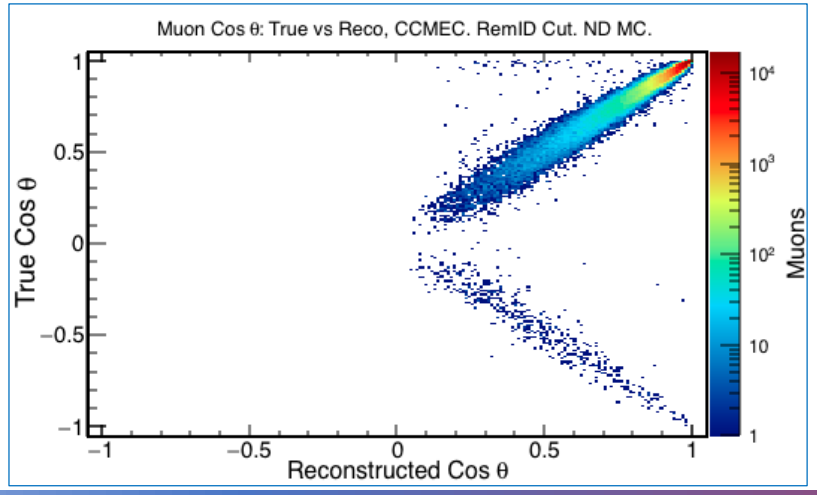
- Reconstructed Cos θ goes from 0 to 1
- True Cos θ goes from -1 to 1: 1.5% of tracks with True Cos θ < 0
- 10% blue band in Reco / True ratio agrees with low BKG region (67% of Reco here)
- No cuts in muon momentum yet





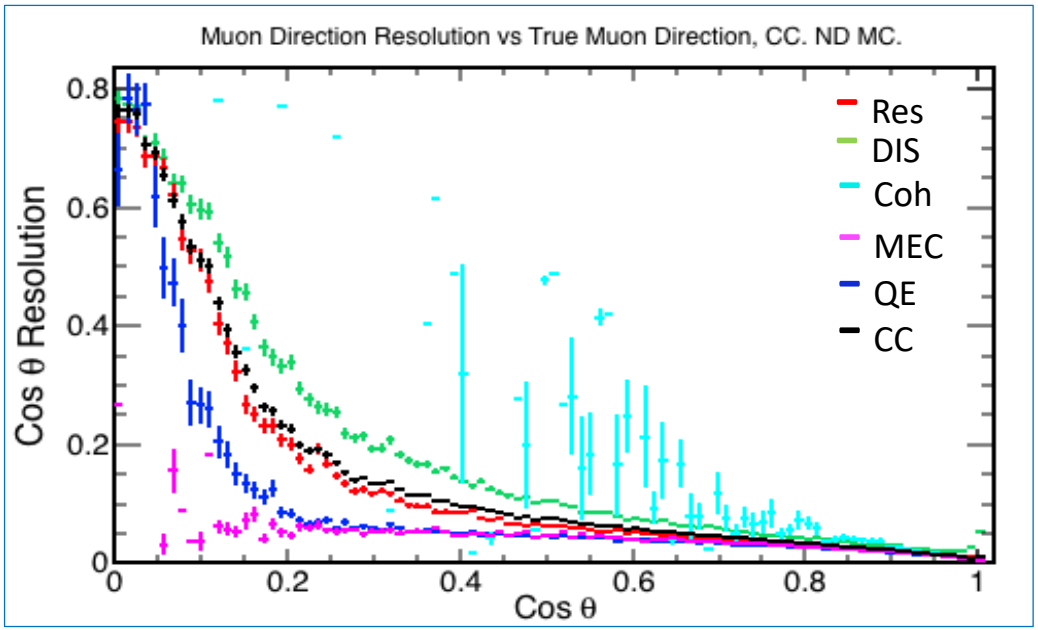


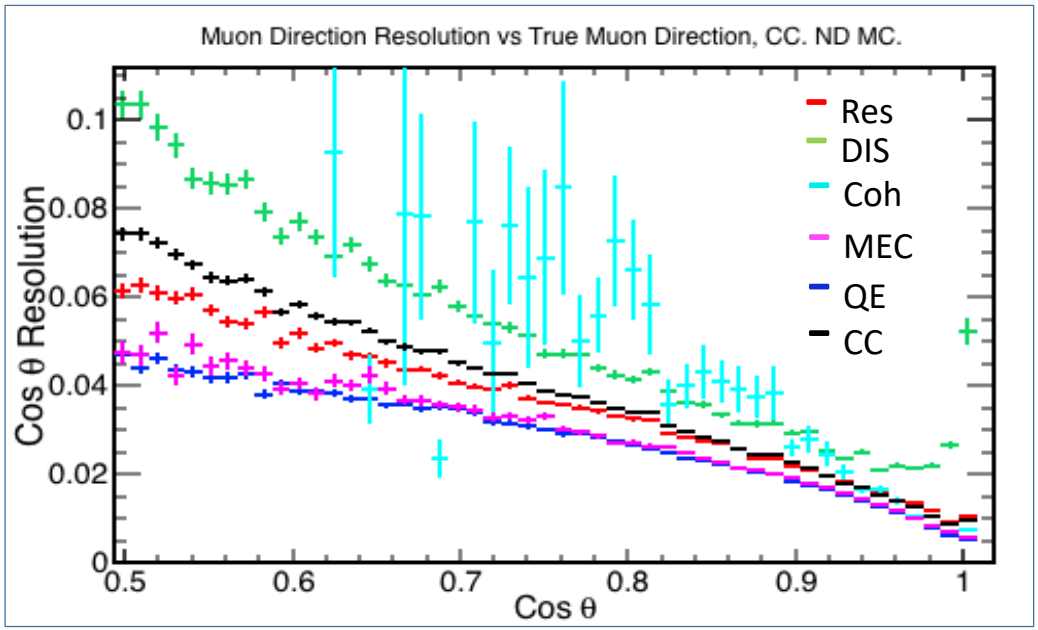




- Muon misidentification in: Res & DIS due to a strong
  hadronic system manifests in two places:
  - Vertical band at Reco Cos θ ≃ 1
  - Horizontal band at True Cos θ ≃ 1 (less freq.)

Properties of forward going muons in MEC



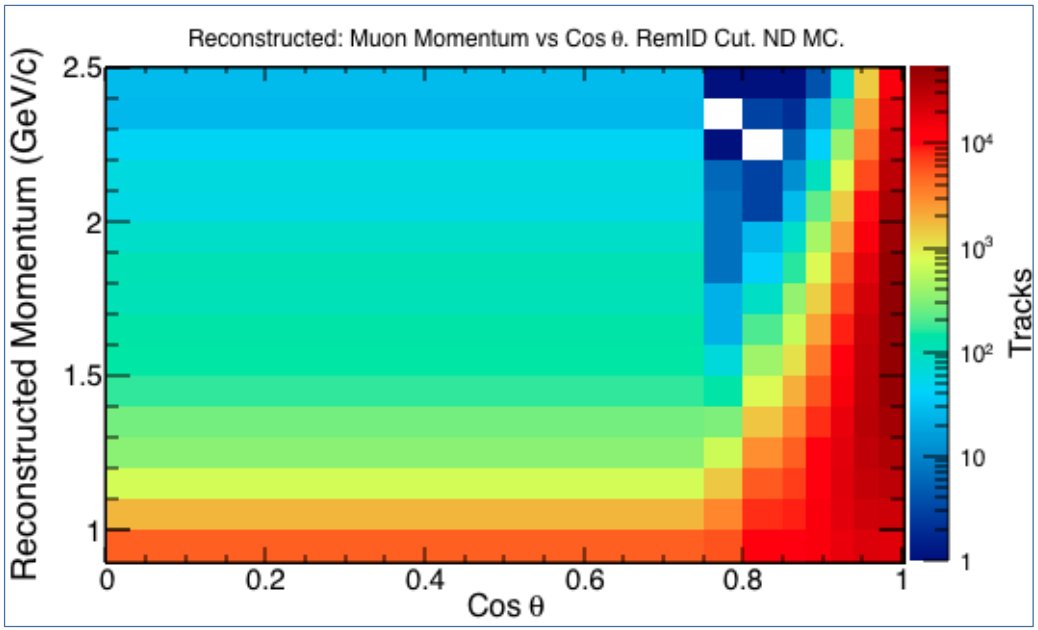


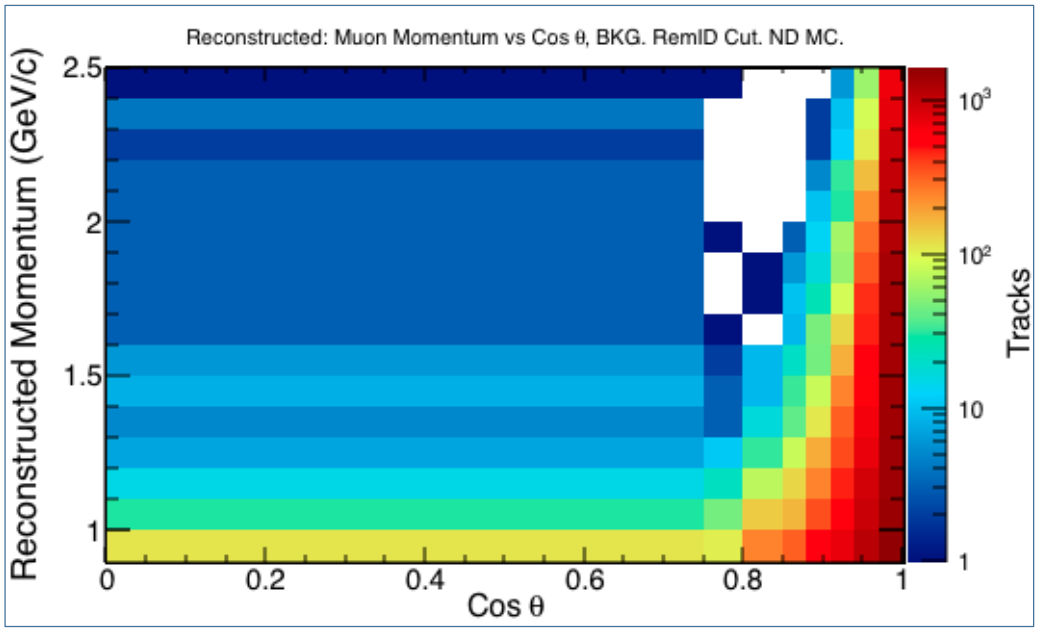
- Muon momentum resolution is very good for:
  - 0.75 < Cos θ < 1
  - DIS is ~125% worse on average

Cos θ	0.95	0.90	0.85	0.75	0.50
Resolution %	1.5	2.5	3.4	5.2	15

- With this direction resolution the following binning is set for use:
  - 5 bins of 0.03: 0.85 < Cos θ < 1
  - 2 bins of 0.03: 0.75 < Cos θ < 0.85
  - 1 bin: 0 < Cos θ < 0.75

- **8 bins** in Cos θ for a double differential cross section measurement

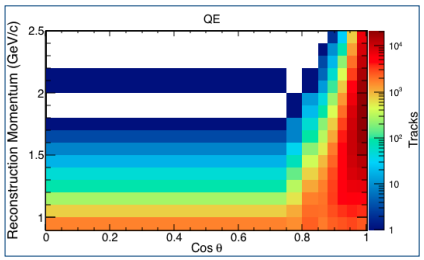


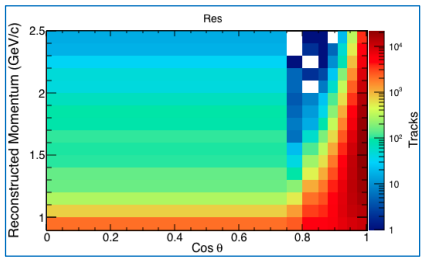


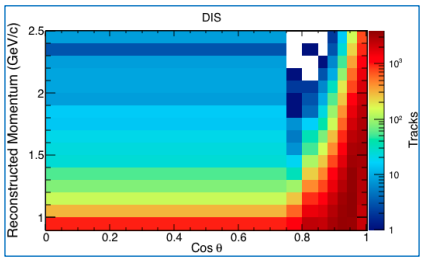
- On the left, the reconstructed variables: Momentum vs Cos θ, signal
- On the right, the reconstructed variables of the BKG interactions

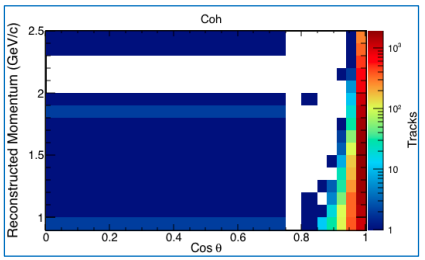
4×10^{20} POT

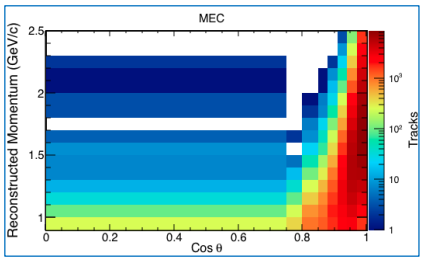
- The CC portion of the signal is presented below
  - Note that the colored scales are different





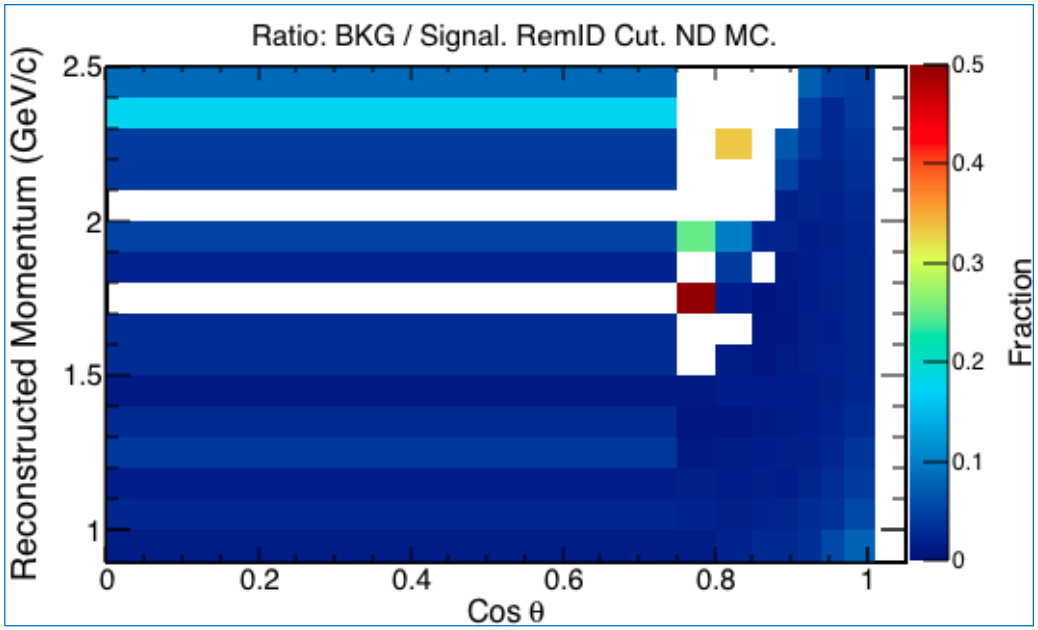


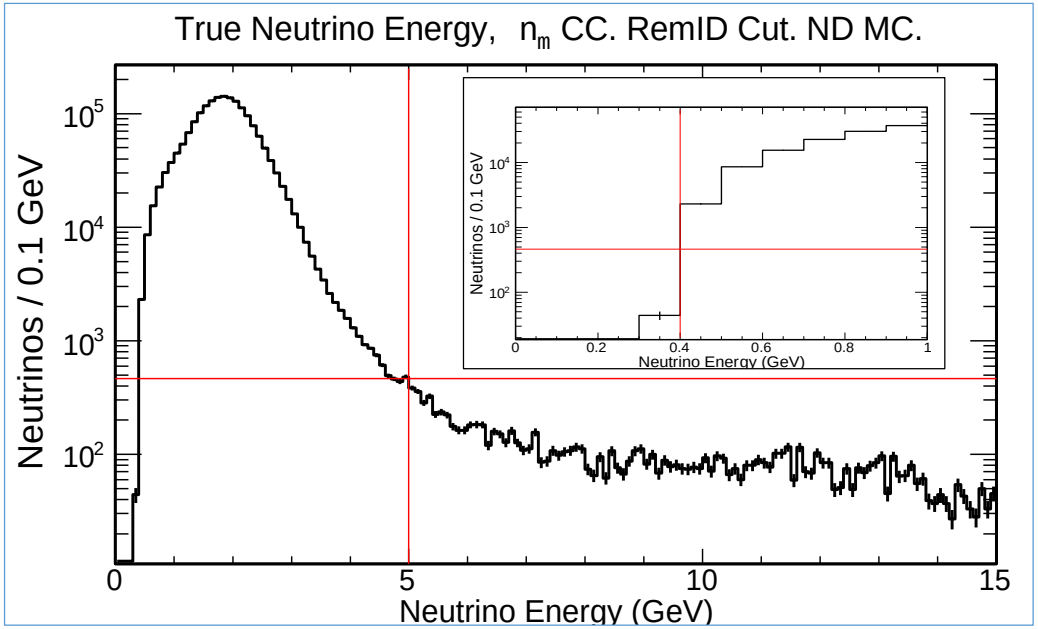


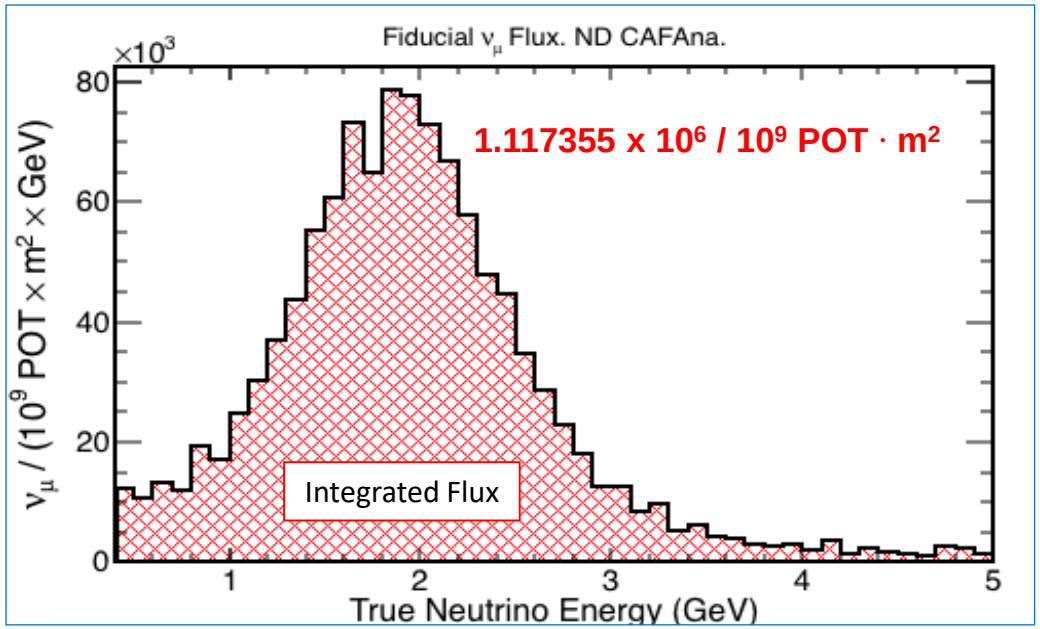


Percentage (%) of BKG Per Bin

GeV/c \ Cos θ	0 - 0.75	0.75 - 0.80	0.80 - 0.85	0.85 - 0.88	0.88 - 0.91	0.91 - 0.94	0.94 - 0.97	0.97 - 1.0
2.5 – 2.4	8.3	0.0	0.0	0.0	0.0	7.5	5.0	4.6
2.4 – 2.3	17.6	0.0	0.0	0.0	0.0	4.9	2.8	4.4
2.3 – 2.2	4.3	0.0	33.3	0.0	6.7	4.1	2.8	3.6
2.2 – 2.1	4.2	0.0	0.0	0.0	5.3	2.6	2.5	3.3
2.1 – 2.0	0.0	0.0	0.0	0.0	1.8	2.4	2.1	2.8
2.0 – 1.9	5.3	25.0	10.0	2.3	2.5	1.6	1.9	2.6
1.9 – 1.8	2.2	0.0	4.3	0.0	1.2	1.7	2.1	2.5
1.8 – 1.7	0.0	50.0	1.7	0.6	1.1	1.7	1.8	2.4
1.7 – 1.6	3.1	0.0	0.0	0.6	0.6	1.8	1.7	2.5
1.6 – 1.5	3.1	0.0	1.4	0.7	1.5	1.8	2.0	2.5
1.5 – 1.4	1.1	1.1	1.6	1.7	1.6	1.7	1.9	2.7
1.4 – 1.3	2.9	0.7	0.6	1.2	1.5	1.7	2.0	3.1
1.3 – 1.2	4.1	1.1	1.5	1.6	1.8	1.9	2.6	3.8
1.2 – 1.1	1.7	1.9	1.7	1.8	1.7	2.5	3.2	4.5
1.1 – 1.0	2.5	2.2	1.9	2.3	2.5	3.0	3.7	5.5
1.0 – 0.9	1.6	1.5	2.3	2.9	2.9	3.5	5.6	7.8

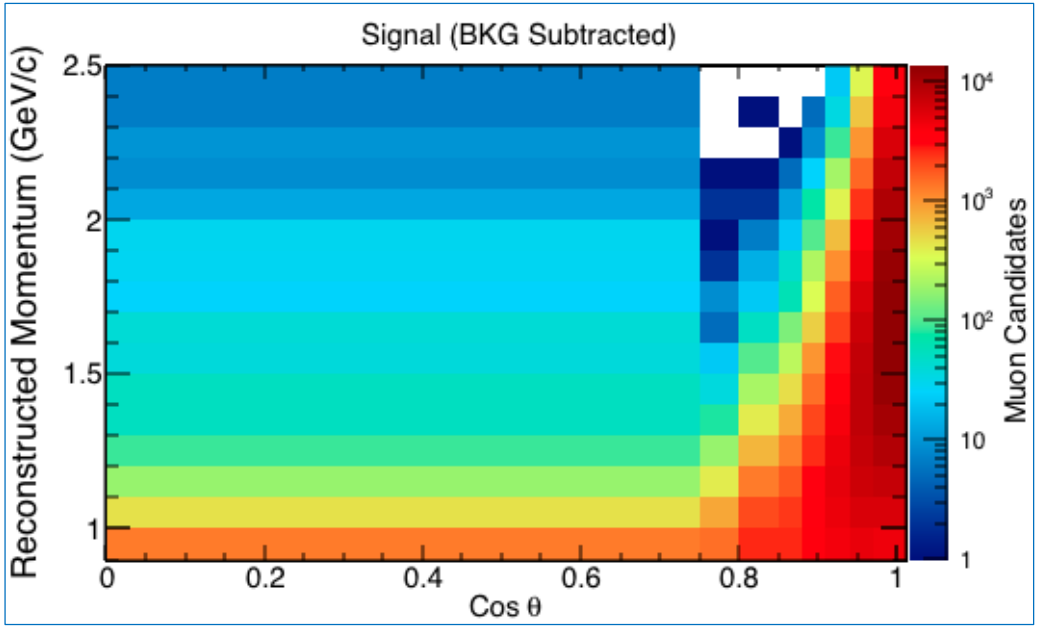






- The true neutrino energy spectrum after selection procedure
  - Cut off at 0.3 GeV
  - If E > 5 GeV is not taken into account for low statistics, then the low E boundary is at 0.4 GeV
  - The integrated region is: (0.4, 5) GeV
- The integrated ν_μ flux is then used to calculate the double differential cross section

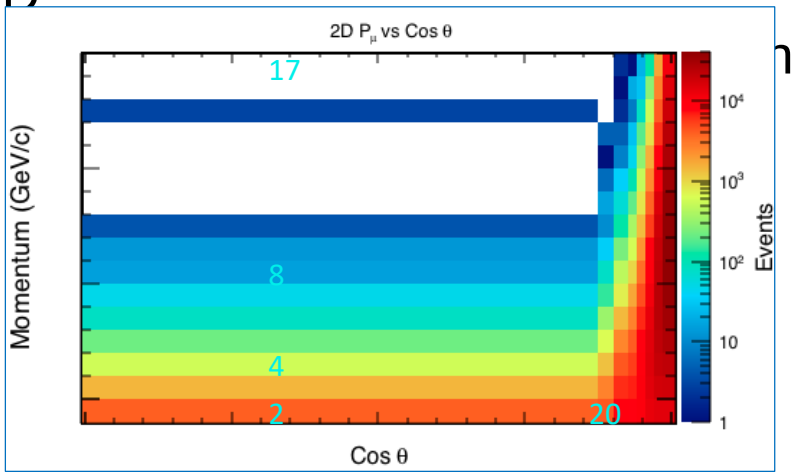
- Fake data sample with
  - 1.9863×10^{20} POT
- Table presents candidate events minus BKG
  - Number of muon candidates: 302151.9
- No rock muons BKG considered



GeV/c \ Cos θ	0 - 0.75	0.75 - 0.80	0.80 - 0.85	0.85 - 0.88	0.88 - 0.91	0.91 - 0.94	0.94 - 0.97	0.97 – 1.0
2.5 – 2.4	6.5	0.0	0.0	0.0	0.0	21.5	358.2	3484.9
2.4 – 2.3	6.5	0.0	1.0	0.0	5.0	36.0	604.1	4463.2
2.3 – 2.2	9.5	0.0	0.5	1.0	8.5	91.5	979.7	6022.9
2.2 – 2.1	8.5	1.0	1.0	5.0	25.5	196.5	1560.3	7707.2
2.1 – 2.0	13.0	2.0	2.0	12.0	76.0	377.1	2437.4	9669.3
2.0 - 1.9	29.0	-0.5	6.5	21.5	105.5	656.6	3446.1	11243.0
1.9 – 1.8	28.5	2.0	14.5	47.0	220.5	1059.2	4670.8	12846.3
1.8 – 1.7	25.0	8.5	21.5	63.5	342.0	1674.7	6036.9	13905.3
1.7 – 1.6	42.0	5.0	51.0	143.0	548.5	2205.3	7072.5	13900.9
1.6 – 1.5	38.5	22.0	106.5	248.0	987.1	2957.5	7880.8	13653.3
1.5 – 1.4	54.5	37.5	205.5	468.6	1436.2	3869.5	8213.3	12526.8
1.4 – 1.3	54.0	82.5	402.5	800.6	2071.8	4346.1	7985.9	10821.3
1.3 – 1.2	90.0	179.0	715.1	1254.7	2509.4	4729.3	7622.7	9245.5
1.2 – 1.1	179.0	403.6	1274.7	1801.3	3185.5	5062.6	7026.9	7540.4
1.1 – 1.0	452.6	816.2	2001.8	2323.0	3413.2	4777.7	5938.4	6033.9
1.0 – 0.9	1314.7	1442.2	2613.5	2648.6	3561.9	4424.3	5012.4	4650.7

- **TSVDUnfolding** requires the following input:
  - **xini**: true distribution
  - **bini**: reconstructed distribution
  - **Adet**: true vs reconstructed
  - **datatrue**: true distribution (fake data prediction)
  - **data**: distribution to be unfolded (from fake data)

- The 2D P_μ vs Cos θ histogram is mapped into a 1D



nt variable is the

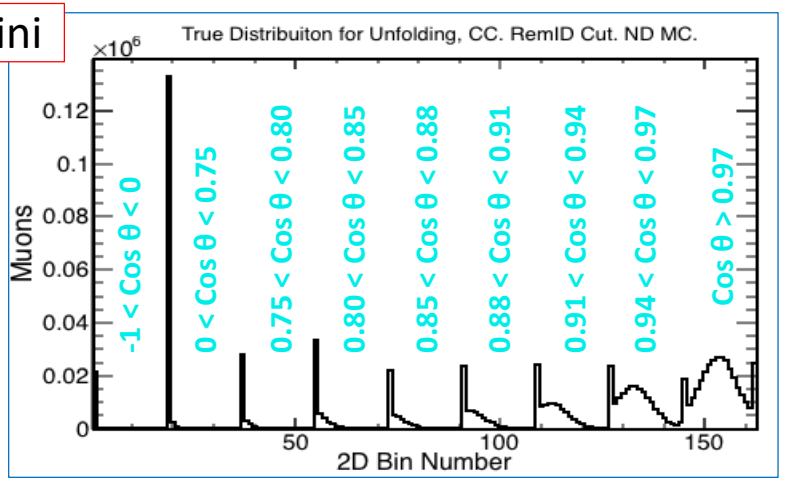
2

Dump Bins:
1, 18, 19, 36, …, 144, 162

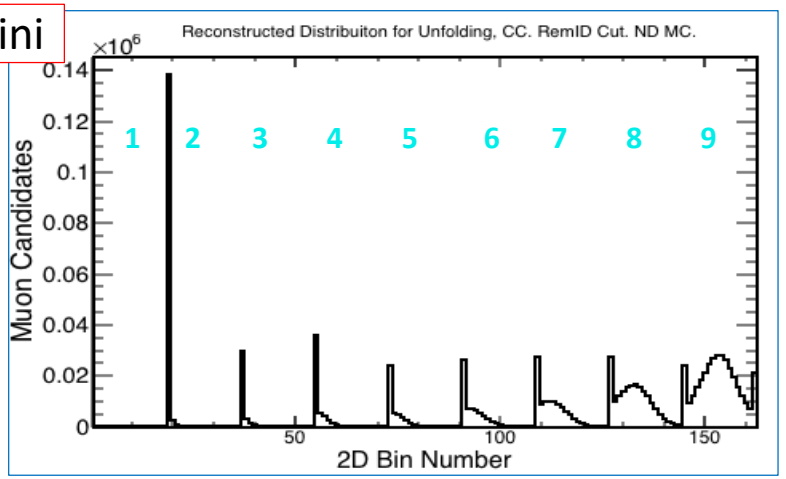
Signal Bins:
All the others

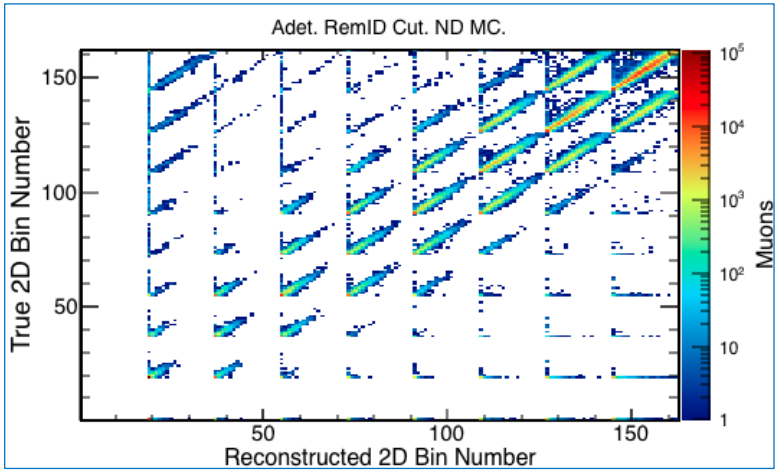
- Ni

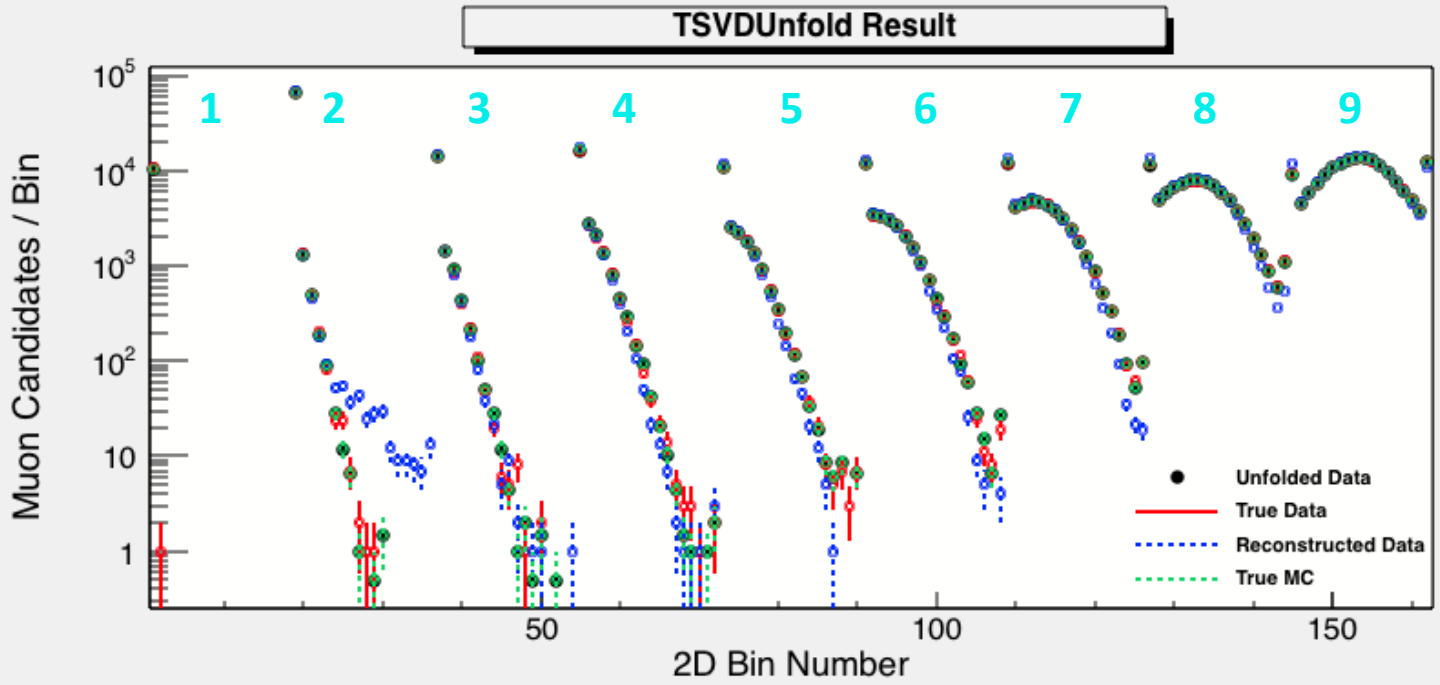
xini



bini

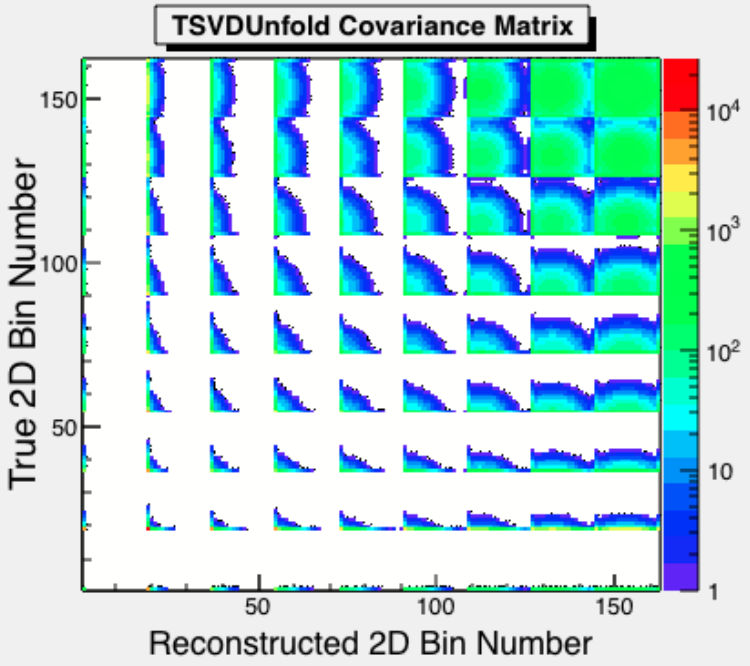


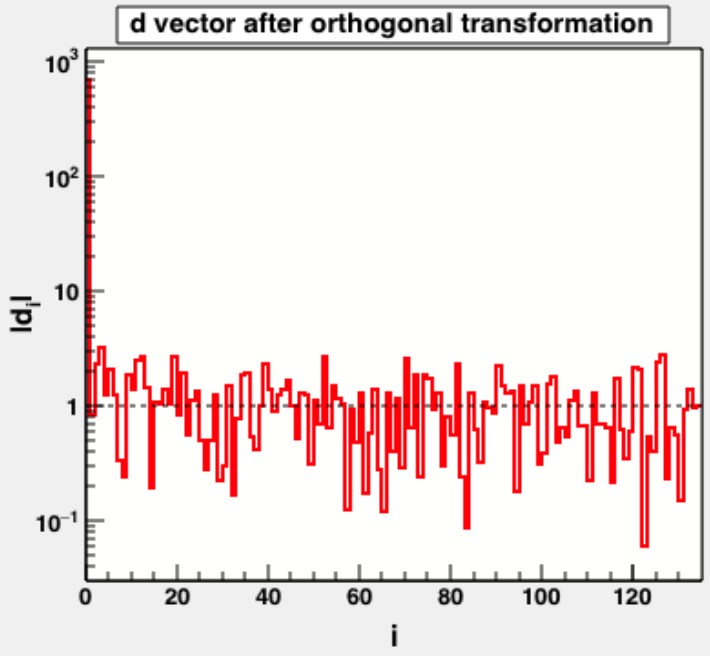




1: 2.2%
2: 14.2%
3: 3.6%
4: 5.1%
5: 4.4%
6: 6.4%
7: 10.3%
8: 18.7%
9: 35.0%

Percentage of events

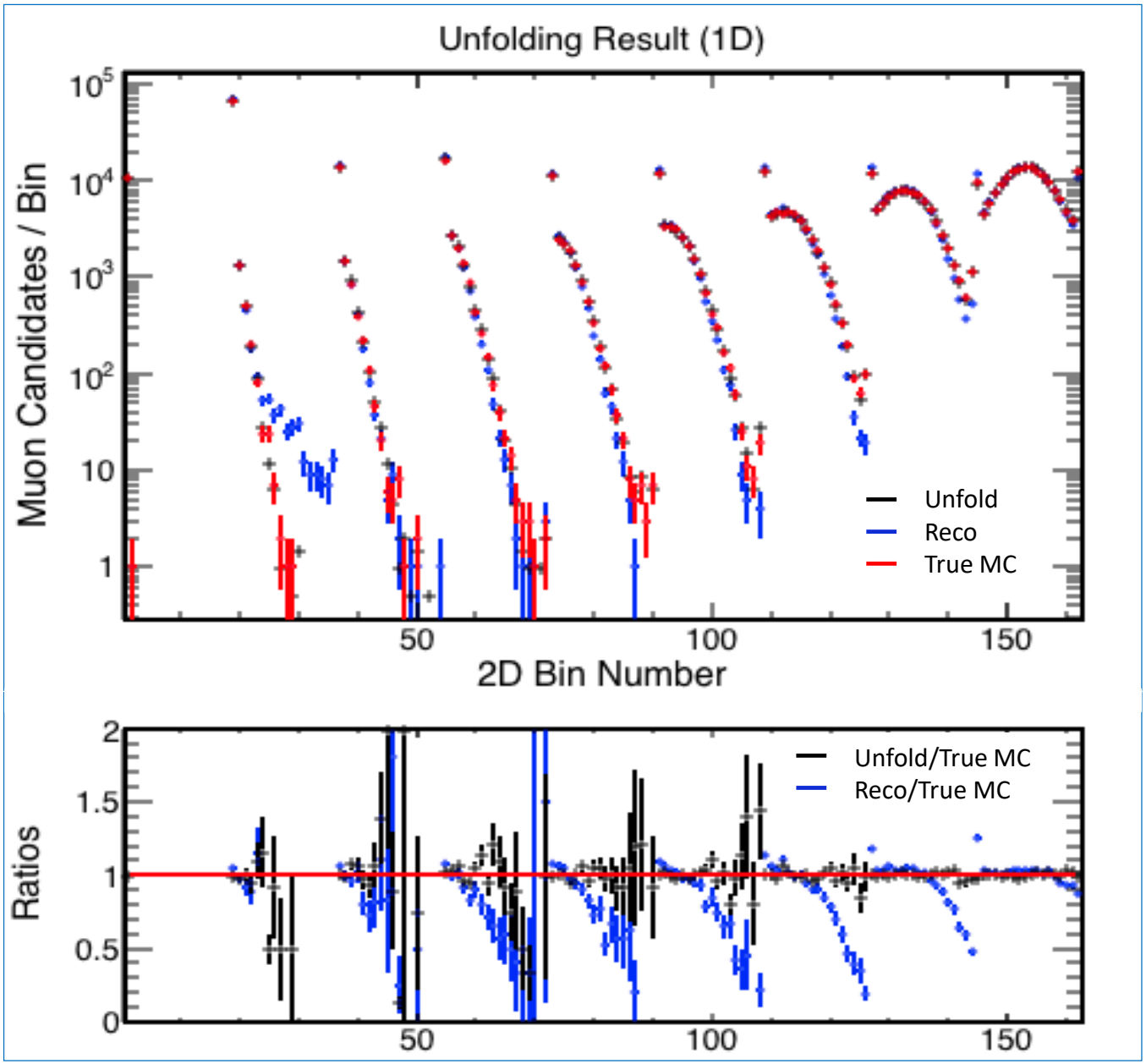




Recommended **i**:

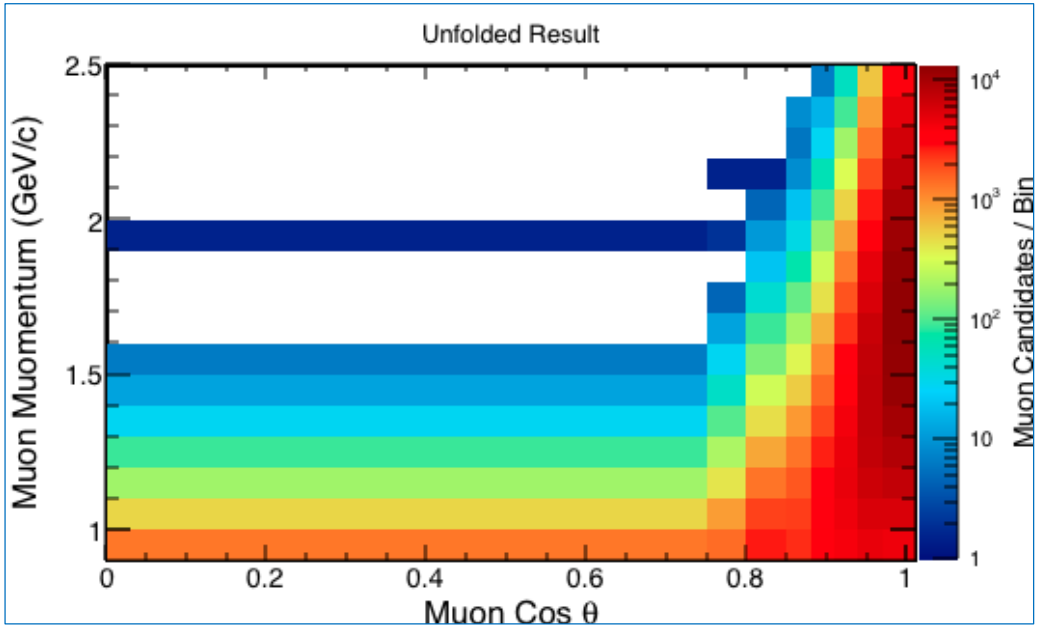
when $|d_i| < 1$
for the first time

i = 2, in this case



- Unfold / True MC:
  0.997 ± 0.002, χ^2/NDF = 599.4/128
- Reco / True MC:
  1.006 ± 0.002, χ^2/NDF = 2830.6/125

- Number of muon candidates: **301897.9**
- Compare to: 302151.9 from reconstruction
- A difference of: 1%



GeV/c \ Cos θ	0 - 0.75	0.75 - 0.80	0.80 - 0.85	0.85 - 0.88	0.88 - 0.91	0.91 - 0.94	0.94 - 0.97	0.97 - 1.0
2.5 – 2.4	0.0	0.0	1.0	0.0	6.4	53.1	595.1	3799.3
2.4 – 2.3	0.0	0.5	0.0	8.4	15.4	91.8	884.0	4947.5
2.3 – 2.2	0.0	0.0	1.0	5.9	28.3	189.0	1326.7	6257.1
2.2 – 2.1	0.0	1.5	1.5	8.4	60.5	330.4	1958.6	7790.6
2.1 – 2.0	0.0	0.5	4.5	19.3	91.2	508.0	2732.8	9603.1
2.0 – 1.9	1.5	2.0	10.4	33.7	169.1	875.6	3763.1	11172.2
1.9 – 1.8	0.5	1.0	20.3	67.9	287.1	1273.9	4907.5	12697.7
1.8 – 1.7	0.0	4.5	42.6	117.4	448.7	1801.7	5879.2	13416.5
1.7 – 1.6	1.0	11.9	90.6	194.2	714.9	2381.5	6858.7	13419.3
1.6 – 1.5	6.4	27.7	141.2	348.8	1079.3	3166.7	7549.4	12903.5
1.5 – 1.4	11.9	50.0	291.2	546.5	1532.8	3771.7	7828.1	12036.6
1.4 – 1.3	27.7	103.0	459.6	934.4	2062.7	4346.5	7892.4	10690.7
1.3 - .12	88.6	213.9	811.7	1347.1	2609.9	4717.4	7417.8	9067.8
1.2 – 1.1	189.6	424.8	1345.1	1772.1	3050.0	4810.5	6612.8	7456.9
1.1 – 1.0	492.6	909.1	2109.7	2237.3	3287.8	4518.7	5785.4	5878.8
1.0 – 0.9	1303.2	1434.9	2742.6	2500.8	3398.2	4169.9	4848.9	4547.4

- In this case, Contained means within Fiducial Volume

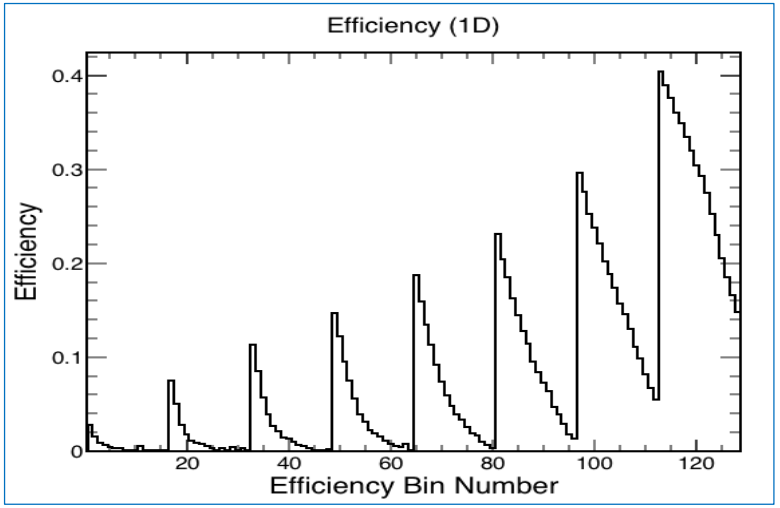
- The efficiency is: the fraction of selected events (after unfolding), *A*, over the events
  that have the truth vertex contained within the Fiducial Volume, *A* + *C*

$$\epsilon = \frac{A}{A+C}$$: efficiency ➞

-

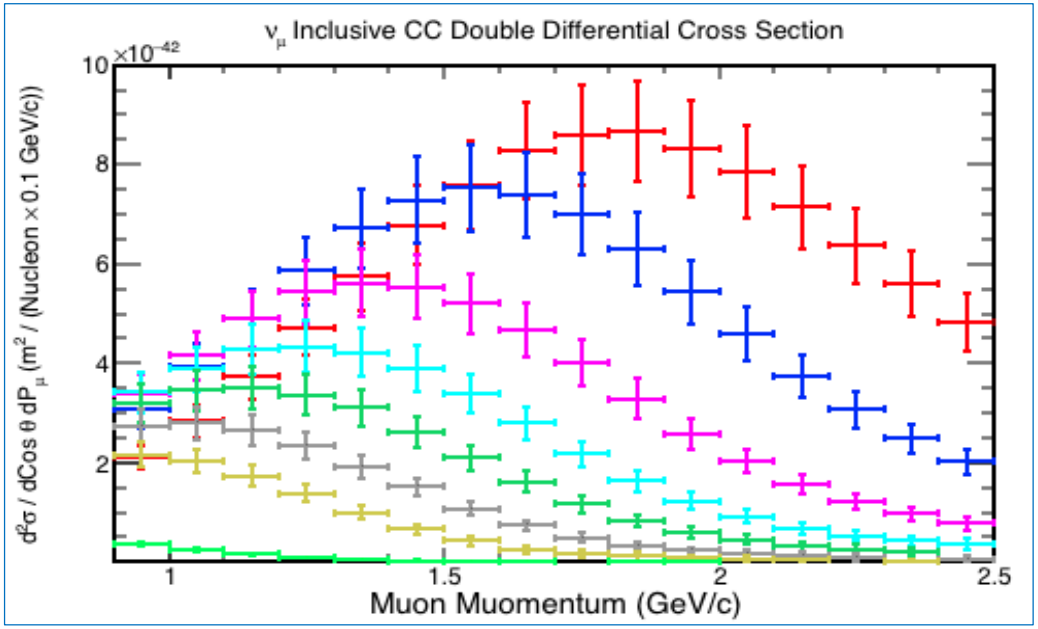
Simulation

Reconstructed	Contained	Uncontained
Contained	A	B
Uncontained	C	D

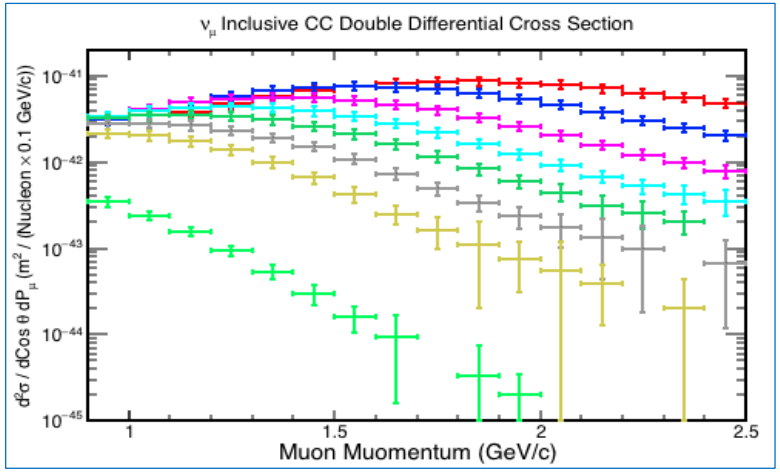


GeV/c \ Cos θ	0 - 0.75	0.75 - 0.80	0.80 - 0.85	0.85 - 0.88	0.88 - 0.91	0.91 - 0.94	0.94 - 0.97	0.97 - 1.0
2.5 – 2.4	0.0010	0.0010	0.0017	0.0010	0.0034	0.0129	0.0554	0.1476
2.4 – 2.3	0.0010	0.0027	0.0010	0.0077	0.0067	0.0177	0.0670	0.1661
2.3 – 2.2	0.0010	0.0010	0.0011	0.0043	0.0100	0.0293	0.0815	0.1848
2.2 – 2.1	0.0010	0.0043	0.0012	0.0050	0.0168	0.0395	0.0983	0.2048
2.1 – 2.0	0.0010	0.0010	0.0029	0.0081	0.0188	0.0469	0.1113	0.2296
2.0 – 1.9	0.0055	0.0030	0.0050	0.0107	0.0258	0.0638	0.1302	0.2520
1.9 – 1.8	0.0011	0.0010	0.0068	0.0154	0.0330	0.0726	0.1463	0.2754
1.8 – 1.7	0.0010	0.0031	0.0097	0.0189	0.0388	0.0844	0.1576	0.2924
1.7 – 1.6	0.0008	0.0053	0.0138	0.0226	0.0482	0.0957	0.1743	0.3045
1.6 – 1.5	0.0030	0.0074	0.0147	0.0314	0.0598	0.1144	0.1883	0.3199
1.5 – 1.4	0.0030	0.0086	0.0218	0.0391	0.0738	0.1278	0.2016	0.3338
1.4 – 1.3	0.0039	0.0117	0.0269	0.0566	0.0918	0.1451	0.2205	0.3491
1.3 – 1.2	0.0071	0.0175	0.0390	0.0750	0.1129	0.1632	0.2375	0.3606
1.2 – 1.1	0.0091	0.0276	0.0573	0.0950	0.1341	0.1848	0.2529	0.3760
1.1 – 1.0	0.0153	0.0504	0.0853	0.1217	0.1590	0.2039	0.2757	0.3899
1.0 – 0.9	0.0283	0.0748	0.1127	0.1473	0.1871	0.2315	0.2966	0.4039

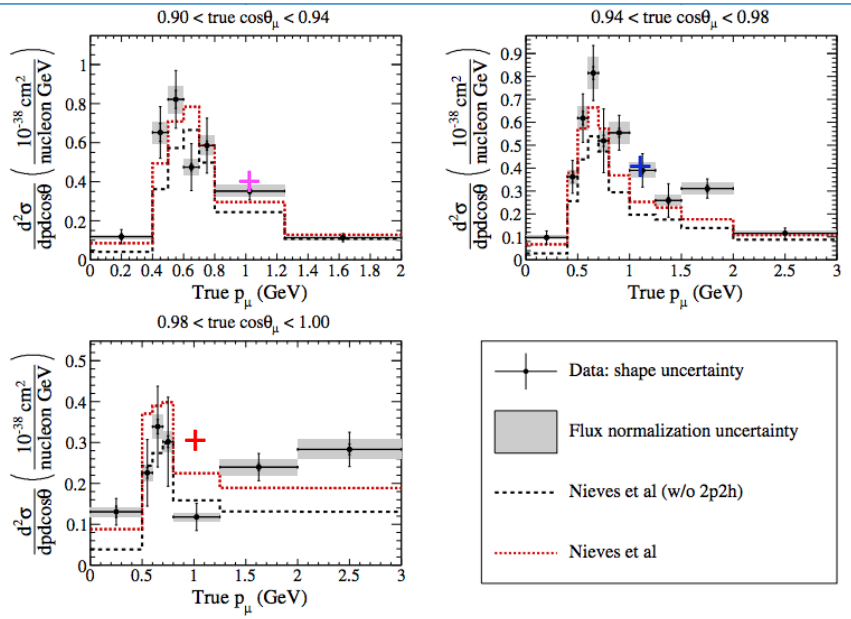
If $\epsilon = 0$ ➞ $\epsilon = 0.001$ & $\sigma = 0$, *i.e.* to get σ without *undetermined* output



0.97 < Cos θ < 1
0.94 < Cos θ < 0.97
0.91 < Cos θ < 0.94
0.88 < Cos θ < 0.91
0.85 < Cos θ < 0.88
0.80 < Cos θ < 0.85
0.75 < Cos θ < 0.80
0 < Cos θ < 0.75



- $\frac{d^2\sigma}{dCos\theta_i dP_j} = \frac{(N_{ij} - B_{ij})}{\epsilon_{ij}\Phi_j \Delta Cos\theta_i \Delta P_j T}$
- *T* for the Fiducial Volume is:
  - 7.99857×10^{31} nucleons
    - From CAFAna, Xsec
- Units: m^2/nucleon



PhysRevD.93.112012

FIG. 9. Measured cross section with shape uncertainties (error bars: internal systematics, external statistical) and fully correlated normalization uncertainty (gray band). The results from the fit to the data are compared to predictions from Nieves *et al.* without 2p2h (black dashed line), and with 2p2h (red dotted line).