Study of muon neutrino oscillations using MicroBooNE data

$$\begin{pmatrix} \nu_e \\ \nu_\mu \\ \nu_\tau \end{pmatrix} = U^{PMNS} \begin{pmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \end{pmatrix}$$

↳ Flavour eigenstates ↳ Mass eigenstates

$$\Rightarrow \begin{pmatrix} \nu_\alpha \\ \nu_\beta \end{pmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{pmatrix} \nu_1 \\ \nu_2 \end{pmatrix}$$

$$|\nu(t=0)\rangle = |\nu_\alpha\rangle = \cos\theta\,|\nu_1\rangle + \sin\theta\,|\nu_2\rangle$$

$$\Rightarrow |\nu(t)\rangle = e^{iq_1\cdot x}\cos\theta\,|\nu_1\rangle + e^{iq_2\cdot x}\sin\theta\,|\nu_2\rangle$$

↳ $q_{1,2}$ are time propagated 4-momenta of two states : $E_{1,2}t - P\cdot x$

$$P(\nu_\alpha \to \nu_\beta) = |\langle \nu_\beta | \nu(t)\rangle|^2$$

$$\Rightarrow \langle \nu_\beta | \nu(t)\rangle = (-\sin\theta\,\langle \nu_1| + \cos\theta\,\langle \nu_2|)(e^{iq_1\cdot x}\cos\theta\,|\nu_1\rangle + e^{iq_2\cdot x}\sin\theta\,|\nu_2\rangle) \longrightarrow \langle \nu_n | \nu_m\rangle = \delta_{nm}$$

$$= -e^{iq_1\cdot x}\sin\theta\cos\theta\,\langle \nu_1|\nu_1\rangle + e^{iq_2\cdot x}\sin\theta\cos\theta\,\langle \nu_2|\nu_2\rangle$$

$$= \sin\theta\cos\theta\,(-e^{iq_1\cdot x} + e^{iq_2\cdot x})$$

$$E^2 = p^2 + m^2$$

↳ Show that

$$q_i\cdot x = \frac{m_i^2 L}{2E}$$

m = Mass

L = length ??

↳ what length??

E = energy

$e^{ik\cdot x} \to$ Plane wave

↳ $p = \hbar k$

$$P\cdot x = p_x + p_y + p_z$$

$$\langle \nu_\beta | \nu(t)\rangle = \sin\theta\cos\theta\,(-e^{iq_1\cdot x} + e^{iq_2\cdot x})$$

$$= \sin\theta\cos\theta\left(-\exp\left[\frac{im_1^2 L}{2E}\right] + \exp\left[\frac{im_2^2 L}{2E}\right]\right)$$

$$= \sin\theta\cos\theta\left(-\left\{1 + \frac{im_1^2 L}{2E} + \left(\frac{im_1^2 L}{2E}\right)^2 + \left(\frac{im_1^2 L}{2E}\right)^3\right\} + \left\{1 + \frac{im_2^2 L}{2E} + \left(\frac{im_2^2 L}{2E}\right)^2 + \left(\frac{im_2^2 L}{2E}\right)^3\right\}\right)$$

↳ $\sin x \simeq x - \frac{x^3}{3!} \longrightarrow \sin\left(\frac{\Delta m_{21}^2 L}{4E}\right) \simeq \underbrace{\frac{L}{4E}(m_2^2 - m_1^2)} - \left(\frac{L}{4E}\right)^3 (m_2^2 - m_1^2)^3$

↳ $e^x \simeq 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!}$

Stop here ?? $\longrightarrow (m_2^2 - m_1^2)^3 \ll (m_2^2 - m_1^2)$

$$= \frac{1}{2}\sin 2\theta\left(i\frac{L}{2E}(m_2^2 - m_1^2) \underbrace{- \left(\frac{L}{2E}\right)^2 (m_2^4 - m_1^4) - i\left(\frac{L}{2E}\right)^3 (m_2^6 - m_1^6)}_{\text{Negligible ??}}\right)$$

$$= \sin 2\theta\left\{i\frac{L}{4E}\Delta m_{21}^2\right\}$$

$$\simeq i\sin 2\theta\,\sin\left(\frac{L}{4E}\Delta m_{21}^2\right) \Rightarrow P = |\langle \nu_\mu | \nu(t)\rangle|^2 = \left|i\sin 2\theta\,\sin\left(\frac{L}{4E}\Delta m_{21}^2\right)\right|^2 = \sin^2 2\theta\,\sin^2\left(\frac{\Delta m_{21}^2 L}{4E}\right)$$

$$P(\nu_\alpha \to \nu_\beta) = \sin^2(2\theta)\,\sin^2\left(\frac{\Delta m_{21}^2 L}{4E}\right)$$

↳ Natural units: $\hbar = c = 1$

↳ In natural units: $P(\nu_\alpha \to \nu_\beta) = \sin^2(2\theta)\,\sin^2\left(1.27\,\frac{\Delta m_{21}^2\,[eV]^2\,L\,[km]}{E\,[GeV]}\right)$

↳ θ: Mixing angle ⇒ Defines how different flavour states are from mass state

↳ $\theta = 0$: Identical; no oscillations

$\theta = \frac{\pi}{4}$: Maximal oscillations; ALL $\nu_\alpha \to \nu_\beta$

↳ Δm_{21}^2: Mass squared difference ⇒ For neutrino oscillation to occur, at least one mass state MUST be non-zero...

∴ Neutrinos **MUST have mass!!**

AND the masses of mass states are different!!

↳ Masses control relative phase of two mass wavefunctions

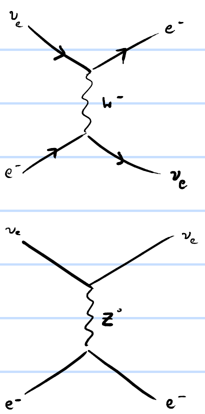
↳ L: Distance between source & detector

E: Energy of neutrino

Exercise 2: Feynman diagrams of other neutrino interactions...

Elastic scattering (ES)

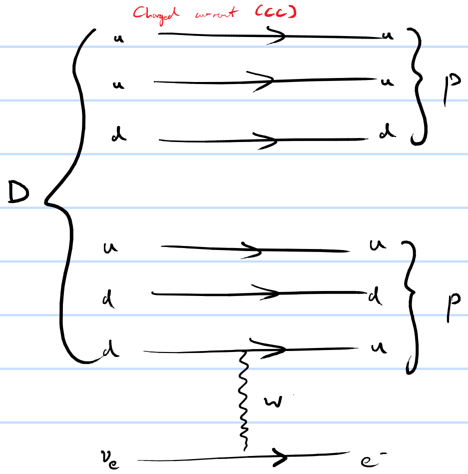
$e^- \nu_e \to e^- \nu_e$



Deuteron: pn

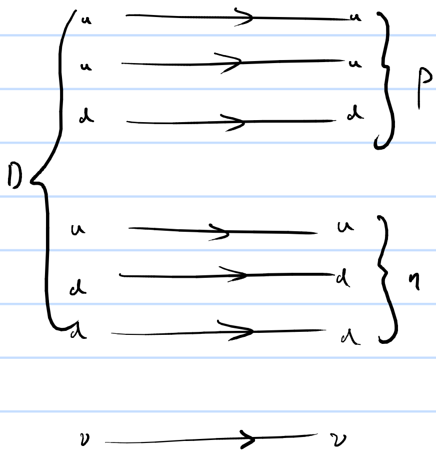
$\nu_e D \to p p e^-$

Charged current (CC)



Neutral current (NC)

$\nu D \to n p \nu$



⇒ Use all three as a cross-check.

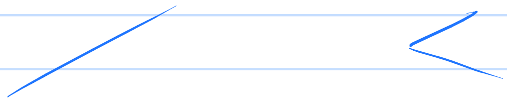
Day 1 : 07/03/2023

[11:21] Still installing all packages and freeing up storage in PC.

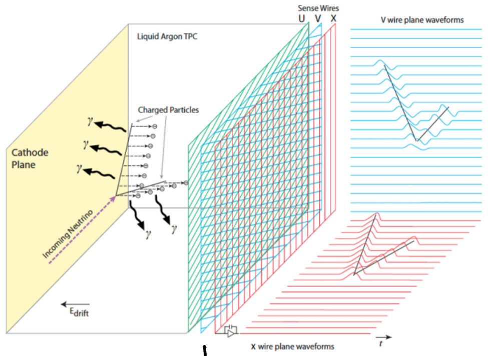
[11:55] Event displays

↳ Cosmic muons : Muons from the cosmos

→ Straight lines ??



Cosmic muons ?   Produced charged particles ??

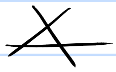


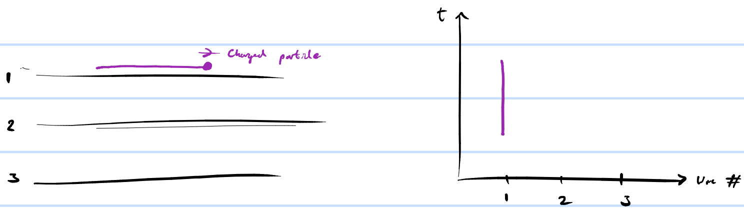
↳ 3 wires as a failsafe (?)

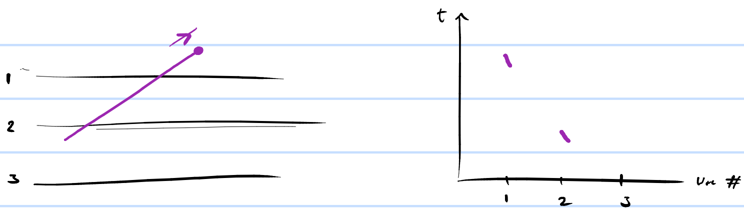
↳ Recreation of particle path.

↳ How do the wire planes detect ??

↳ Consider a wire plane ⇒ Wire planes 60° from each other





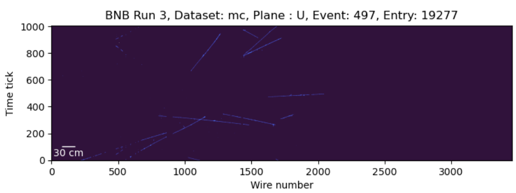


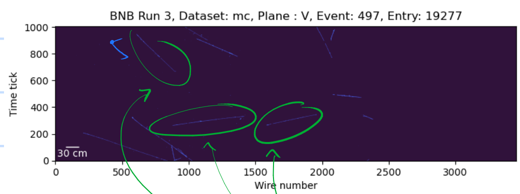
[13:47] ⇒ How can we tell it's a muon?

↳ Lifetime ?? → ~ 2 μs

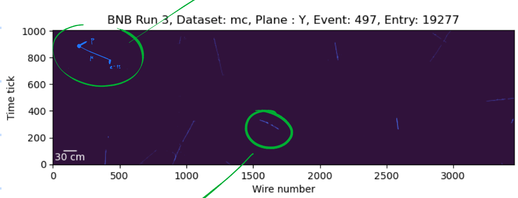
↳ Charge : − 1e

[13:54] : What characteristics are common ??





Proton → High energy (bright) ; short along line

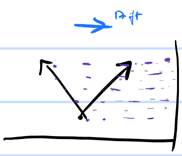


Long straight lines.

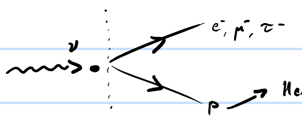
↳ By itself → Cosmic muons!

Cosmic muon ??

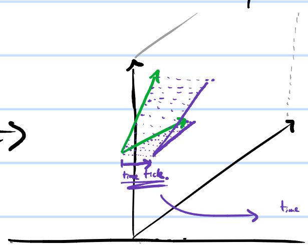
↳ Or is it a decay



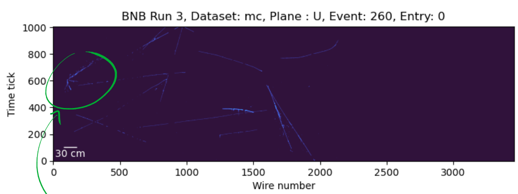
Normally :



↳ Cosmic muons ONLY pass through, charging particles.



time taken since interaction for charged particle to drift



What are these?? → Jets ?? → What kind of interaction.

↳ why fuzzy ??.

↳ $\nu_\mu D \to pp\mu$??.

[14:29] Variables we're working with: All in python program...

Variable Name	Description
$\mathrm{vertex}_{x,y,z}$ reco_nu_vtx_sce_x,y,z	Reconstructed position of the vertex of an interaction in the detector (cm).
$\mathrm{track}^{start}_{x,y,z}$ trk_sce_start_x,y,z_v	Reconstructed position of the start of an identified track in the detector. (cm)
$\mathrm{track}^{end}_{x,y,z}$ trk_sce_end_x,y,z_v	Reconstructed position of the end of an identified track in the detector (cm).
L_{track} trk_len_v	The distance between the longest track in an event and the reconstructed neutrino vertex (cm).
d_{track} trk_llr_pid_score_v	Log likelihood score used to find the ID of the particle. For example, if it is a muon or a proton.
$\mathrm{L}_{\nu,cosmic}$ _closestNuCosmicDist	Distance between our detected Neutrino track and closest cosmic muon track within the detector.
$\mathrm{S}_{topological}$ topological_score	A score which determines how much a signal in the detector looks like a track. This is also a variable that is determined by machine learning, where the machine has been trained with track-like events.
S_{track} trk_score_v	A score which determines how much a signal in the detector looks like a track. This is also a variable that is determined by machine learning, where the machine has been trained with track-like events.
E_{n} trk_energy_tot	Reconstructed energy of the neutrino in the interaction.

2 data files:

↳ "MC_EXT": Predicted distribution of neutrino events, developed using Monte Carlo techniques. AND real data. from MicroBooNE when muon neutrino beam is off

"data": Real data from MicroBooNE

[15:39] Background events

Cosmic → Cosmic muons

Out Fid. Vol. → Outside Fiducial Volume (Volume of absolute certainty)

EXT → External particles → why different from cosmic??

$\nu_e CC$ → Electron charged currents

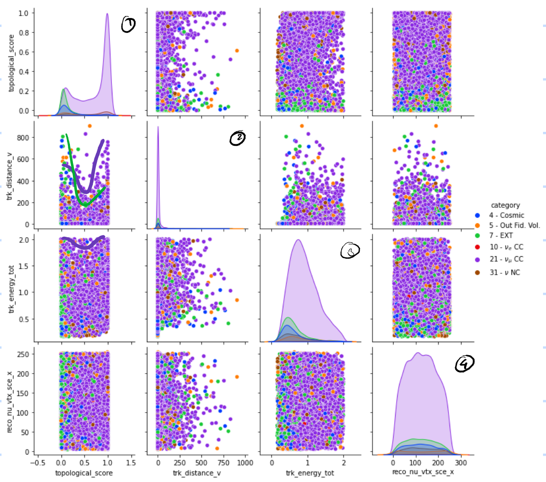
νNC → Neutrino neutral current.

[15:43] What does .pairplot do??

↳ Plots pairs of categories against each other

↳ IF a category is plotted against itself, it shows a univariate distribution plot for the marginal distribution of data. in each column

Proportional sizings of area ??



$\mathrm{S}_{topological}$ topological_score	A score which determines how much a signal in the detector looks like a track. This is also a variable that is determined by machine learning, where the machine has been trained with track-like events.
L_{track} trk_len_v	The distance between the longest track in an event and the reconstructed neutrino vertex (cm).
E_{n} trk_energy_tot	Reconstructed energy of the neutrino in the interaction.
$\mathrm{vertex}_{x,y,z}$ reco_nu_vtx_sce_x,y,z	Reconstructed position of the vertex of an interaction in the detector (cm).

1 High topological score for $CC\nu_\mu$ BUT low topological scores for background → Difficult to distinguish ??

↳ Too short of track length ??

↳ Fuzzy track length → Why is it fuzzy??

2 ALL particles have really short track length → Short lifetimes ?? → Nothing too stable...

3 What units is energy ?? [E] = ~~MeV~~ OR GeV ?

↳ [E] = GeV

↳ What significance is ~0.8 [E] ??

4 Uniform distribution ?? (Somewhat) → No special area in the tank to create interactions..

↳ Middle better distributed due to better sensors ??

[15:54] Deep copy of dataset WITHOUT signals

↳ How to deep copy

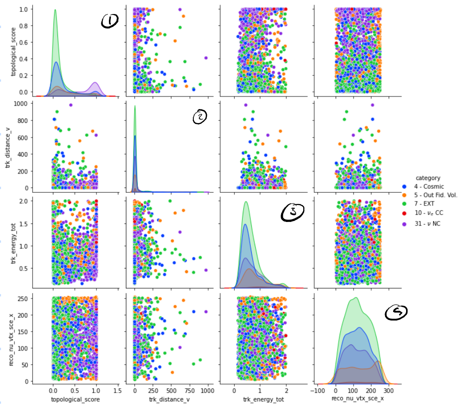
↳ What is a deep copy?? : Separate variable that doesn't get changed

↳ df.copy(deep = True)

↳ How to delete rows?

↳ Use df.drop(indices, inplace = True)

No signals. (ν_μ CC)



Day 2: 09/03

[10:40] Why remove signal CCν_μ? → Look at trends of background

↳ Cosmic AND EXT have similar distribution

↳ Well, cosmic muons ARE external particles

① All fairly low topological score EXCEPT for NC ν → Measures total flux: $\phi(\nu_e) + \phi(\nu_\mu) + \phi(\nu_\tau)$

low high ??

↳ Why??

↳ Why is CCν_e low, yet CC ν_μ high??

↳ Sun ONLY produces electron neutrinos ∴ Neutrinos MUST change flavour FROM ν_e

↳ Receive less ν_e than ν_μ; still hard to recognise ν_e tracks...

② ALL still have low track distance

③ ~ 0.5 [E], is overall less avg energy than ν_μ CC.

=> Distribution has a sharp decline for energies greater than 1 [E]

④ Again, similar to uniform distribution EXCEPT Out Fid. Vol.

↳ Fiducial volume: Volume of detector where a specified number (~ 90%) of events are accepted; smaller than total volume

=> Expect tighter distribution to a certain region for Out Fid. Vol.

[11:07] Bonus 1: Make trends clearer to visualise??

Bonus 2: Add new variables to the plot; look for new trends.

[11:26] Decision trees

Gini criterion : $G_{1,2} = 1 - \sum_{class} p^2_{class}$ where P_{class} is probability of finding one class of one side of a partition (sides 1 & 2)

↳ Gini minimises its own value for each partition

↳ 100% accuracy IF $G_{1,2} = 0$.

Problems : Overfitting → Partitions tailored specifically for ONLY one dataset

↳ Too many partitions allowed

∴ Less accurate for a different dataset...

Solution / fix : Bootstrap aggregation (or "bagging") called random decision trees

→ Produce multiple, independent decision trees

→ Different random choice for each tree across different features to ensure they're distinct

⇒ Aggregates results...

[11:35] Exercise 6:

Exclude CC events (ν_eCC & ν_μCC) from database

⇒ Form decision trees → Requires 'training data' 'truth data' } create training and testing data sets with sklearn.model_selection.train_test_split()

↳ Split data into 80% training, 20% testing

⇒ Create a random forest:

sklearn.ensemble

What if we change these??

↳ RandomForestClassifier() → n = 1000 trees, max_depth = 8

↳ rf.fit ↳ Gini criterion...

Testing accuracies:

↳ y_pred = rf.predict(X_train)

⇒ Predictions of classification based on training dataset.

y_pred = rf.predict(X_test)

⇒ Predictions of classification based on testing dataset.

See accuracy via sklearn.metrics.accuracy_score

↳ $\frac{\#\text{ correct classifications}}{\#\text{ total datapoints.}}$ → Normalized...

⇒ IF accuracy of training dataset AND testing dataset are similar, data has NOT been overfit

↳ Similar enough:

```
Accuracy on training dataset: 0.663612050656808
Accuracy on testing dataset: 0.6509359760893503
```

⇒ Successful decision tree ??

→ Accuracy NOT best metric...

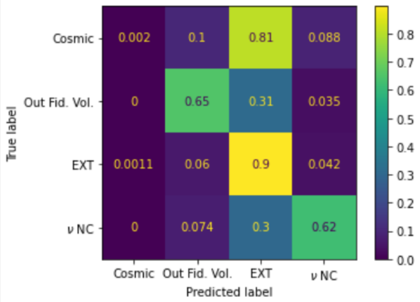
↳ Use a confusion matrix: Plots probability that the truth data (y) matches what the model predicts...

⇒ IF 100% accurate: Diagonal would be 1:

P_1	P_2	P_3
P_4	P_5	P_6
P_7	P_8	P_9

→ $P_1 = P_5 = P_9 = 1$

[14:16] Produced a confusion matrix:



=> Cosmic has VERY low probability... Why??.

↳ Is cosmic being misidentified?? → Should be most common...

=> Meanwhile EXT has high probability

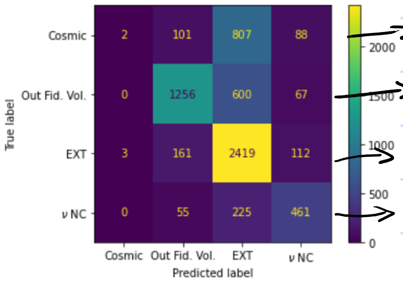
↳ Does cosmic fall under EXT??. → Technically true...

↳ How is it differentiated??.

=> Out Fid. Vol. and νNC have similar probabilities...

↳ Matches accuracy found earlier in our model!!

=> Look at unnormalised table



→ ≃ 1000 → ONLY 5 classified as cosmic, 2 of which were correct...

↳ Majority classified as EXT...

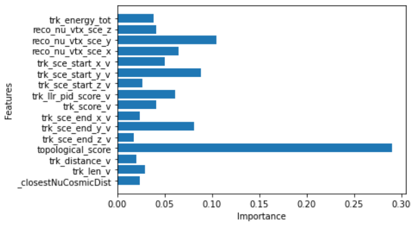
→ ≃ 1900

→ ≃ 2700

→ ≃ 700

Can also plot importance of each variable

↳ Which variable determined categories...



↳ Most important: Topological score → How likely a certain detection is a track.

[14:51] Selection

Want to reduce background events induced by background AND retain as many ν_μ CC signals.

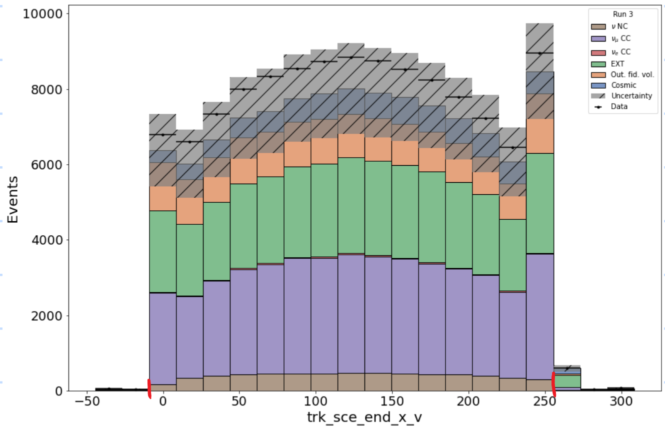
↳ Major contributor: Cosmic rays.

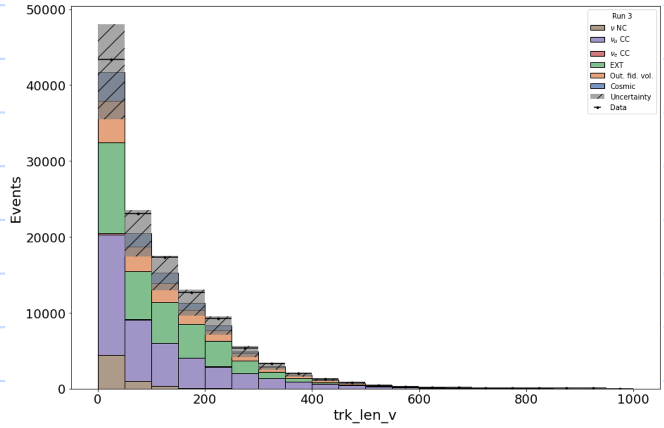
↳ High-energy cosmic radiation

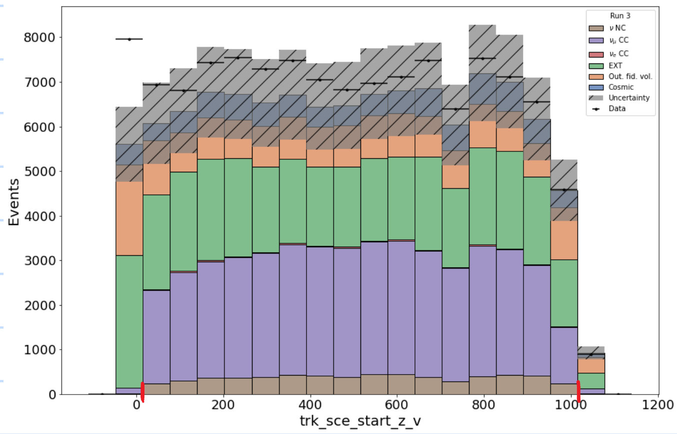
↳ Traverse most of detector; usually detected near edge of detector coordinates

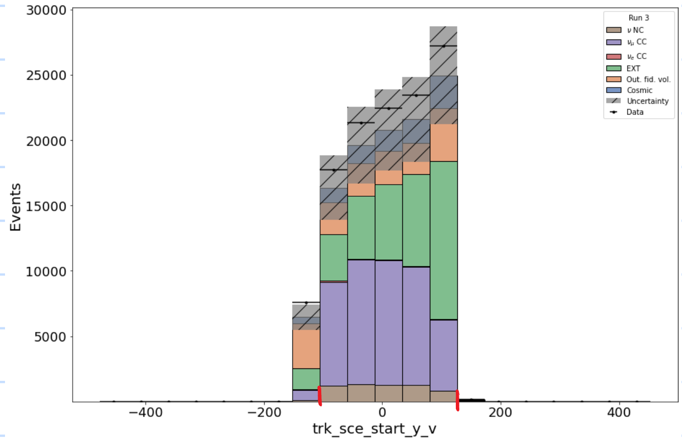
=> How can we tell?

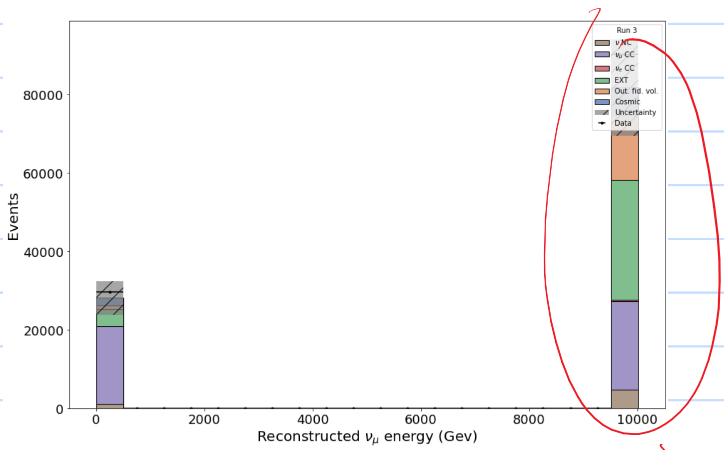
→ Look at ALL histograms & apply conditions

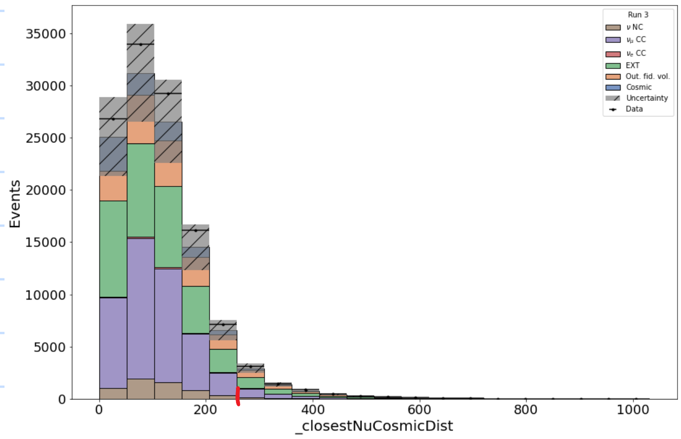




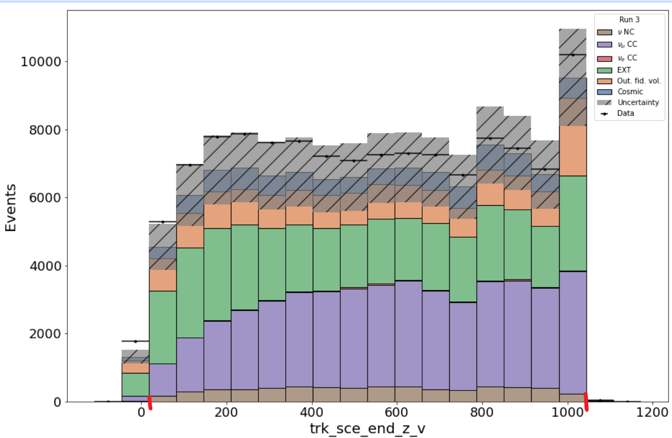


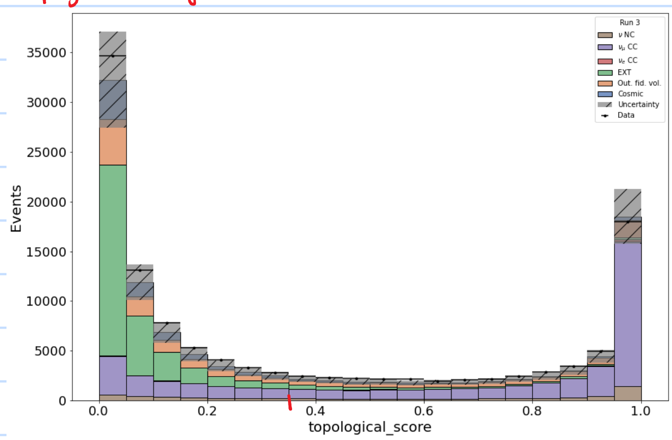


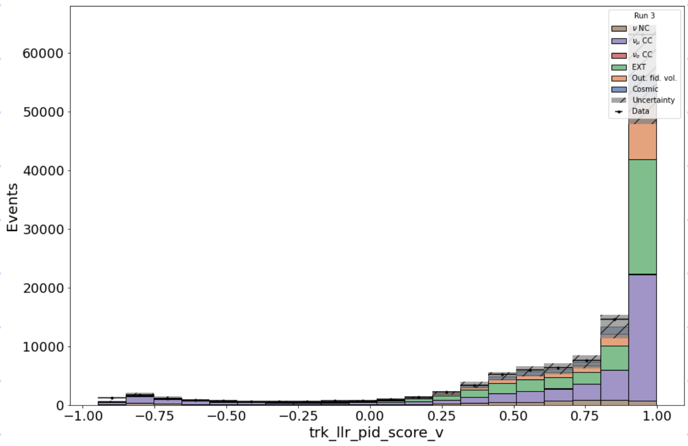


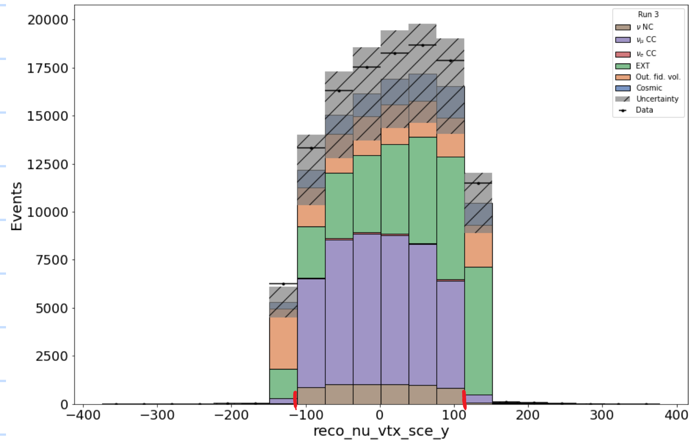


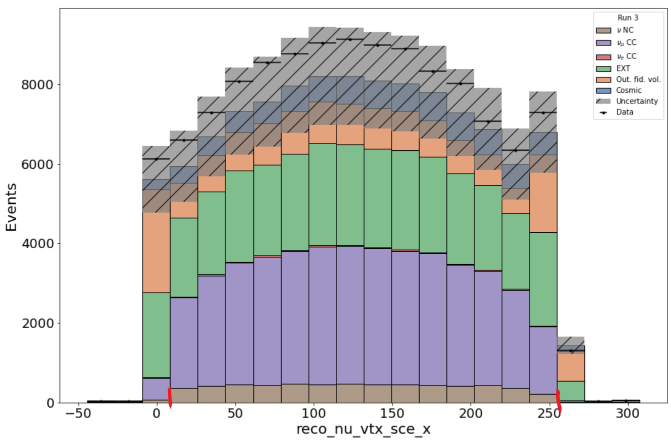
Unphysical energies!!

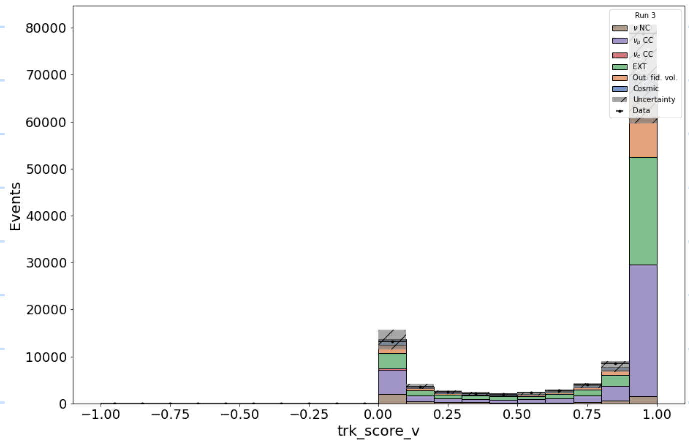


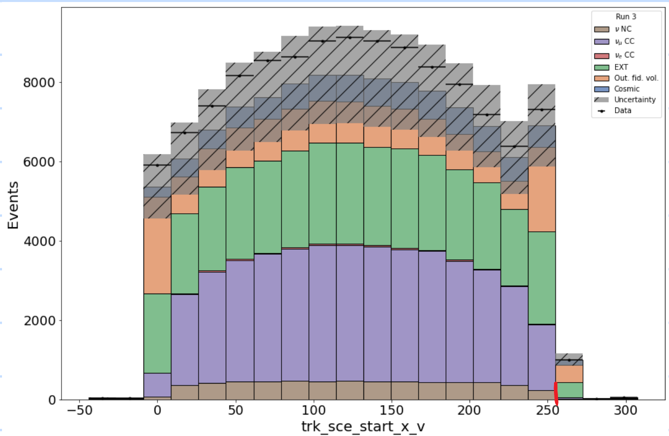


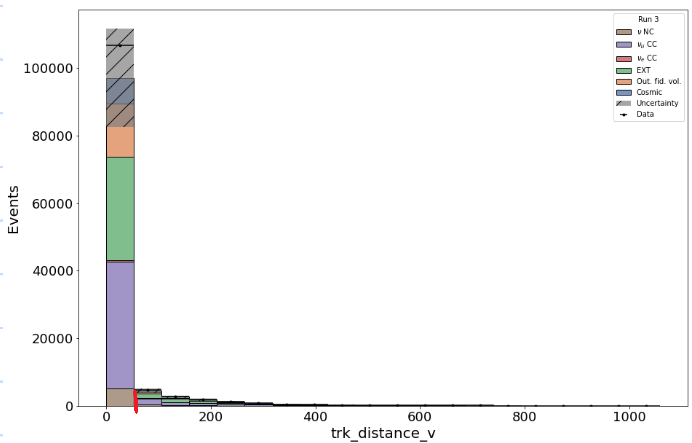

Run 3
ν NC
ν_μ CC
ν_e CC
EXT
Out. fid. vol.
Cosmic
Uncertainty
Data
Events
trk_llr_pid_score_v

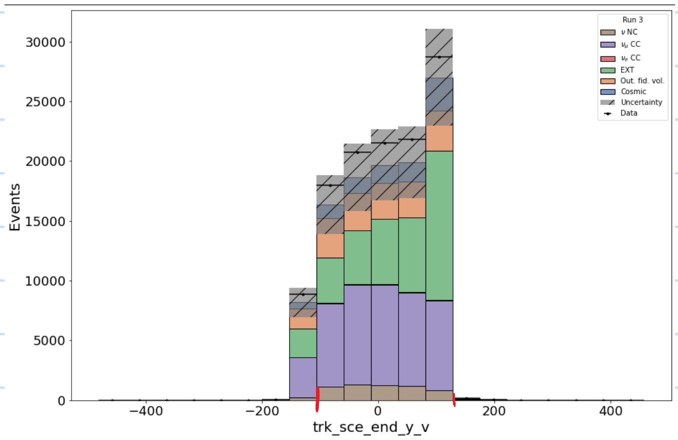

Run 3
ν NC
ν_μ CC
ν_e CC
EXT
Out. fid. vol.
Cosmic
Uncertainty
Data
Events
reco_nu_vtx_sce_y


Run 3
ν NC
ν_μ CC
ν_e CC
EXT
Out. fid. vol.
Cosmic
Uncertainty
Data
Events
reco_nu_vtx_sce_x


Run 3
ν NC
ν_μ CC
ν_e CC
EXT
Out. fid. vol.
Cosmic
Uncertainty
Data
Events
trk_score_v


Run 3
ν NC
ν_μ CC
ν_e CC
EXT
Out. fid. vol.
Cosmic
Uncertainty
Data
Events
trk_sce_start_x_v


Run 3
ν NC
ν_μ CC
ν_e CC
EXT
Out. fid. vol.
Cosmic
Uncertainty
Data
Events
trk_distance_v


Run 3
ν NC
ν_μ CC
ν_e CC
EXT
Out. fid. vol.
Cosmic
Uncertainty
Data
Events
trk_sce_end_y_v

Day 3 - 14/03

11.1 Efficiency

[09:00] After each selection:

$$\text{Efficiency} = \frac{\text{Events surviving the selections}}{\text{Total number of events.}}$$

=> Retain high efficiency

11.2 Purity

Fraction of events as the signal we're looking for

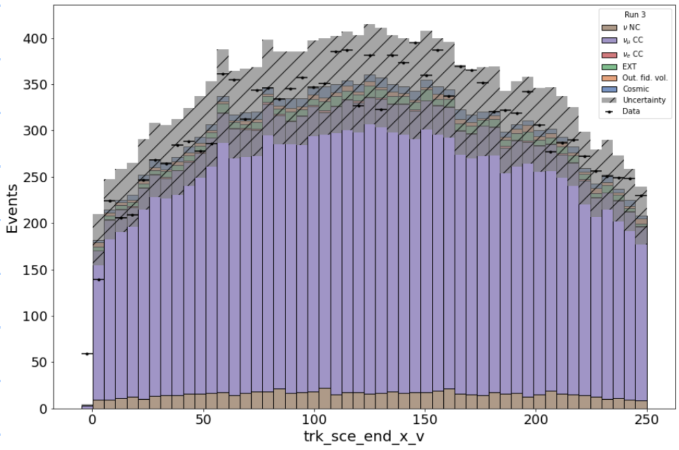
$$\text{Purity} = \frac{\text{Total number of signal events passing selection cuts}}{\text{Total events in selected sample.}}$$

=> Plot purity & efficiency onto a graph... Use smaller bins to make cuts more precise...

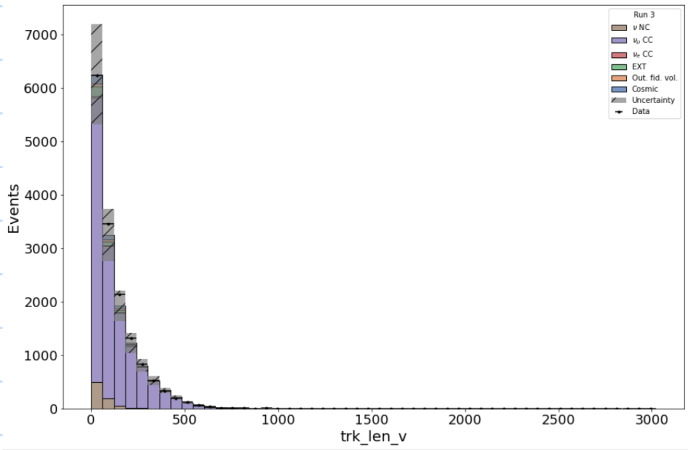
[10:40]

Selections	Cuts
trk_len_v	> -1000, < 1000
trk_end_x_v	> ~~-10~~, < 250, > -5
trk_start_z_v	> ~~0~~, < 1000, > 20
trk_start_y_v	> -100, < ~~150~~, < ~~120~~, < 110
topological	> 0.4
trk_end_z_v	> 50.
reco_y	> -110, < ~~120~~, < 110
reco_x	> 10, < 250
trk_start_x_v	> 5, < 250
reco_z	> 20, < 980
trk_energy_tot	< ~~10~~. < 4

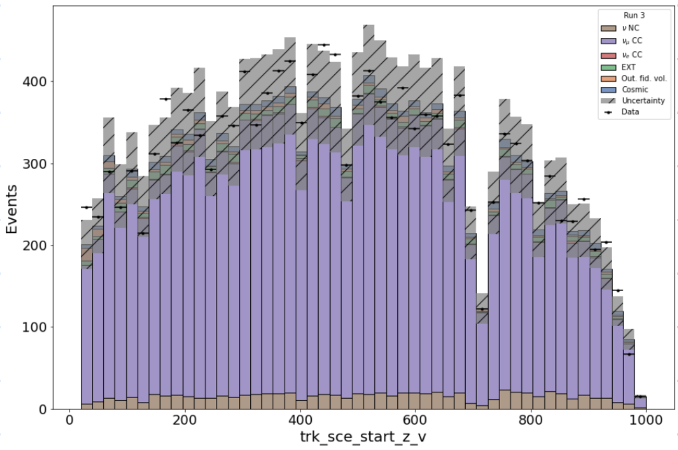
[11:20]



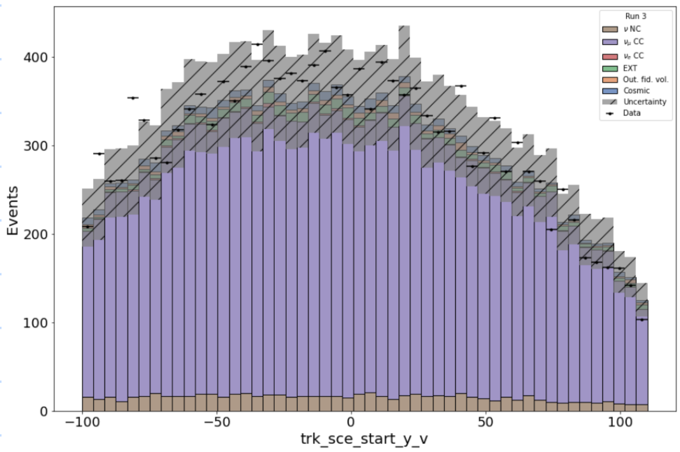
Run 3
ν NC
ν_μ CC
ν_e CC
EXT
Out. fid. vol.
Cosmic
Uncertainty
Data
Events
trk_sce_end_x_v



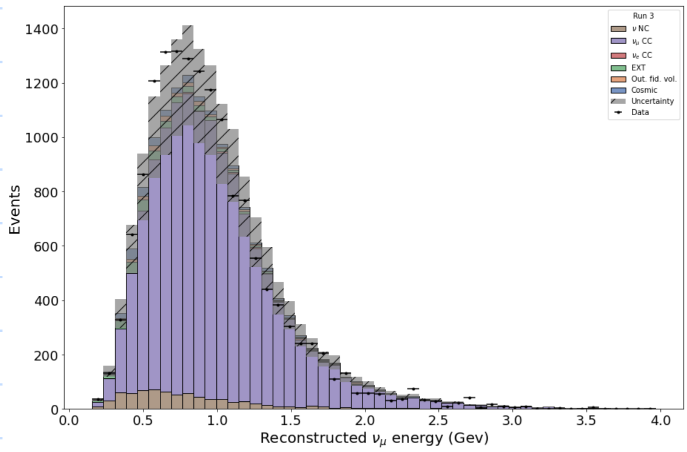
Run 3
ν NC
ν_μ CC
ν_e CC
EXT
Out. fid. vol.
Cosmic
Uncertainty
Data
Events
trk_len_v



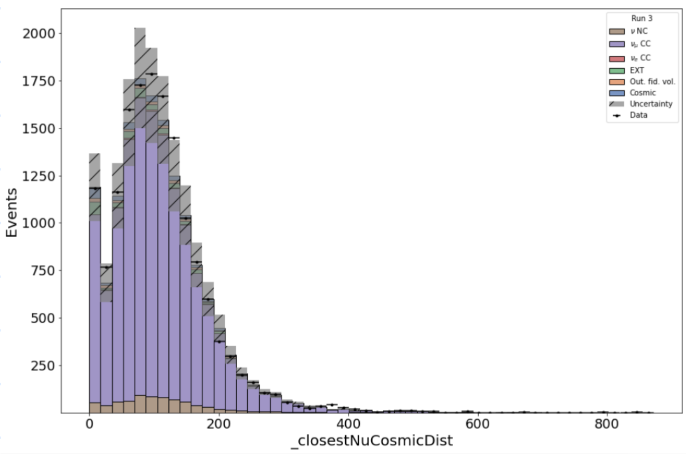
Run 3
ν NC
ν_μ CC
ν_e CC
EXT
Out. fid. vol.
Cosmic
Uncertainty
Data
Events
trk_sce_start_z_v



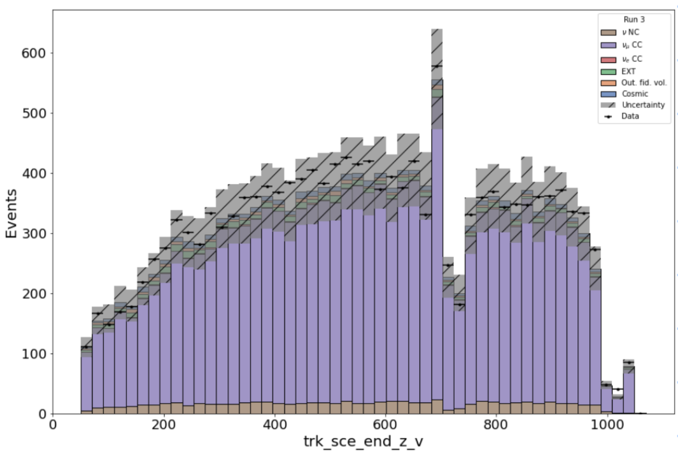
Run 3
ν NC
ν_μ CC
ν_e CC
EXT
Out. fid. vol.
Cosmic
Uncertainty
Data
Events
trk_sce_start_y_v



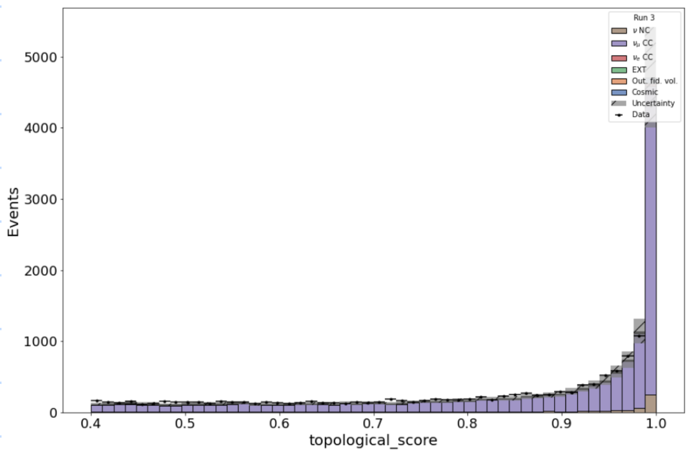
Run 3
ν NC
ν_μ CC
ν_e CC
EXT
Out. fid. vol.
Cosmic
Uncertainty
Data
Events
Reconstructed ν_μ energy (Gev)



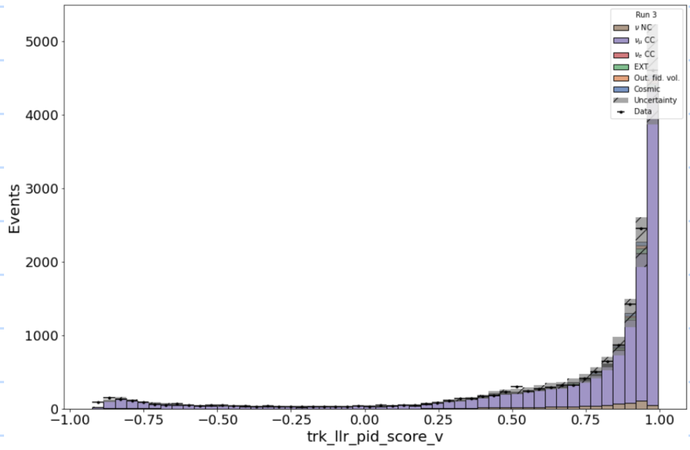
Run 3
ν NC
ν_μ CC
ν_e CC
EXT
Out. fid. vol.
Cosmic
Uncertainty
Data
Events
_closestNuCosmicDist

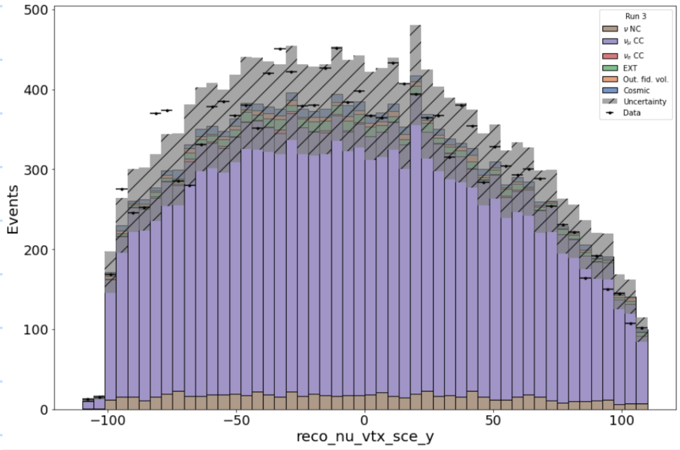


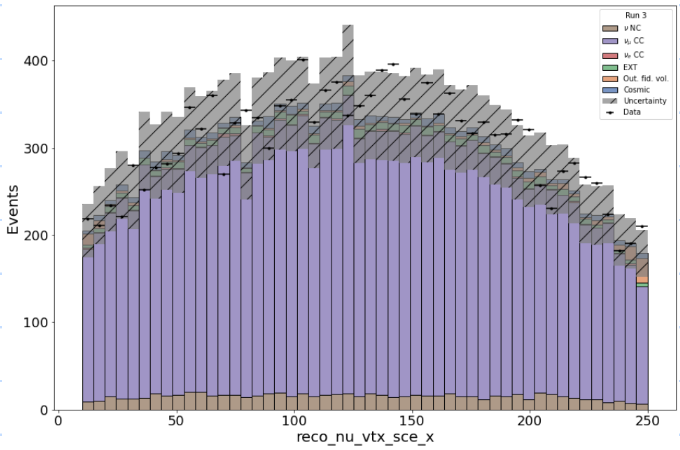
Run 3
ν NC
ν_μ CC
ν_e CC
EXT
Out. fid. vol.
Cosmic
Uncertainty
Data
Events
trk_sce_end_z_v

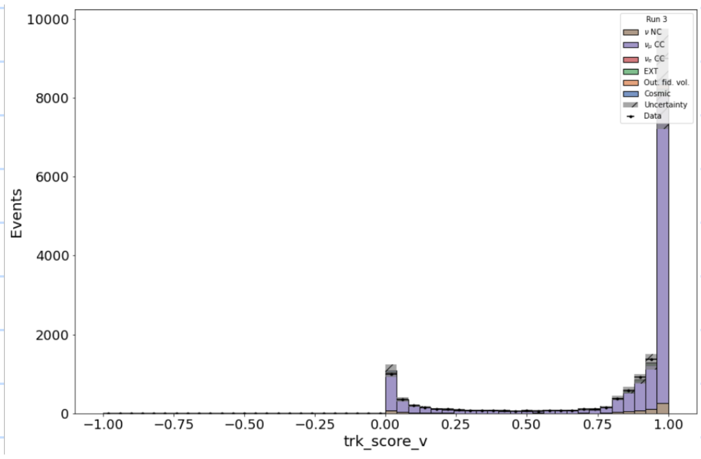


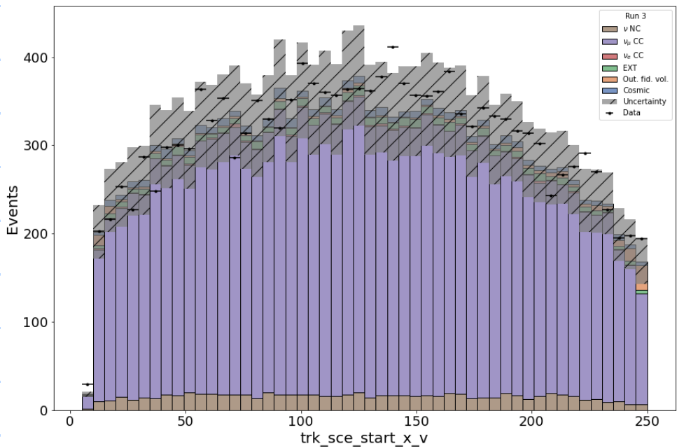
Run 3
ν NC
ν_μ CC
ν_e CC
EXT
Out. fid. vol.
Cosmic
Uncertainty
Data
Events
topological_score

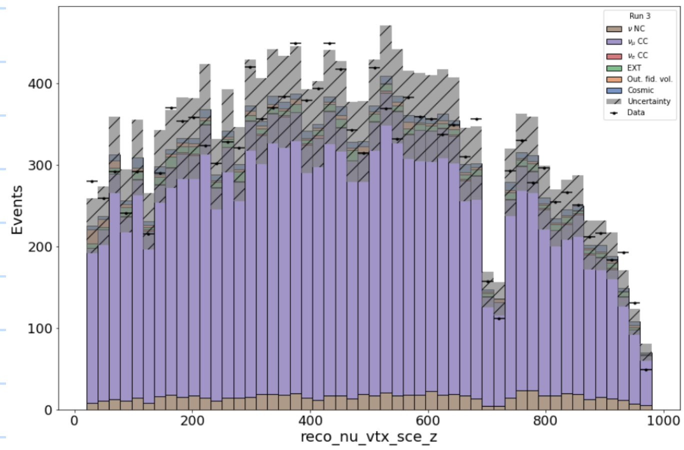

Run 3
ν NC
ν_μ CC
ν_e CC
EXT
Out. fid. vol.
Cosmic
Uncertainty
Data
Events
trk_llr_pid_score_v

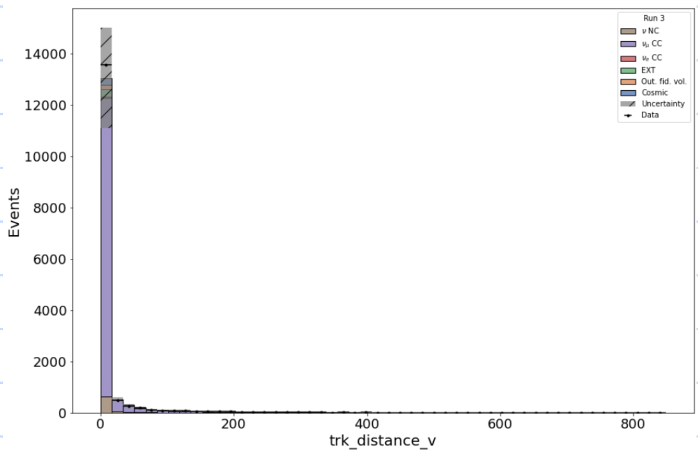

Run 3
ν NC
ν_μ CC
ν_e CC
EXT
Out. fid. vol.
Cosmic
Uncertainty
Data
Events
reco_nu_vtx_sce_y

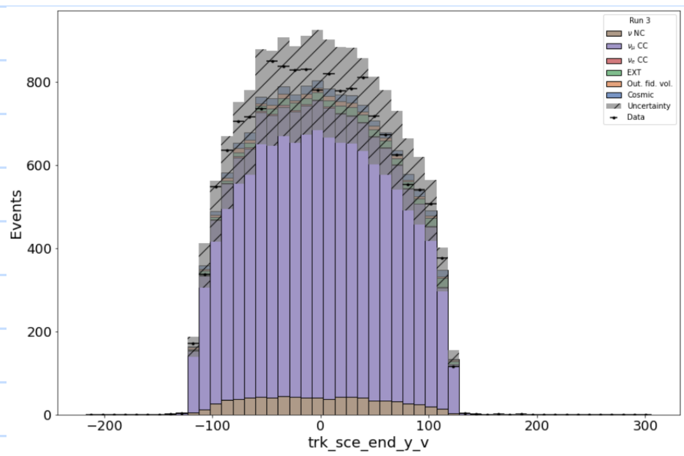

Run 3
ν NC
ν_μ CC
ν_e CC
EXT
Out. fid. vol.
Cosmic
Uncertainty
Data
Events
reco_nu_vtx_sce_x


Run 3
ν NC
ν_μ CC
ν_e CC
EXT
Out. fid. vol.
Cosmic
Uncertainty
Data
Events
trk_score_v

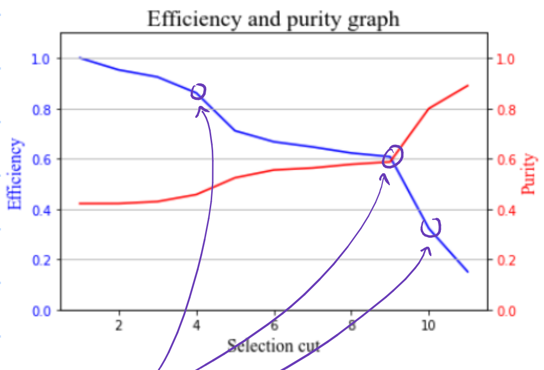

Run 3
ν NC
ν_μ CC
ν_e CC
EXT
Out. fid. vol.
Cosmic
Uncertainty
Data
Events
trk_sce_start_x_v


Run 3
ν NC
ν_μ CC
ν_e CC
EXT
Out. fid. vol.
Cosmic
Uncertainty
Data
Events
reco_nu_vtx_sce_z


Run 3
ν NC
ν_μ CC
ν_e CC
EXT
Out. fid. vol.
Cosmic
Uncertainty
Data
Events
trk_distance_v


Run 3
ν NC
ν_μ CC
ν_e CC
EXT
Out. fid. vol.
Cosmic
Uncertainty
Data
Events
trk_sce_end_y_v

[11:41] Efficiency & purity graph:



Final efficiency : 15.06%

Final purity : 88.99%

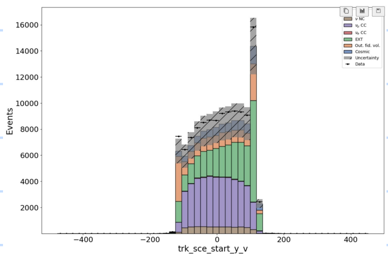
Order of cuts :

1. trk_end_x_v
2. trk_end_y_v
3. trk_start_x_v
4. trk_start_y_v
5. trk_start_z_v
6. reco_x
7. reco_y
8. reco_z
9. topological
10. trk_energy_tot.

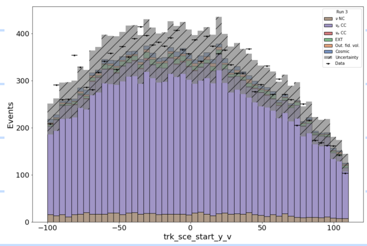
Biggest changes : Cuts 4, 9, 10

=> Why ??

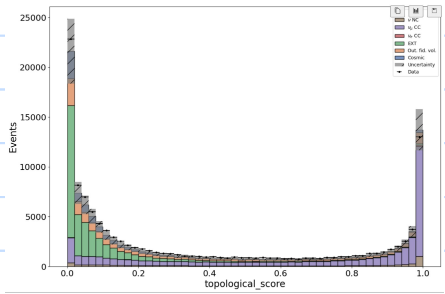
↳ 5 trk_start_y_v : Filtered LOTS of background (actual) events ∴ Efficiency drops...



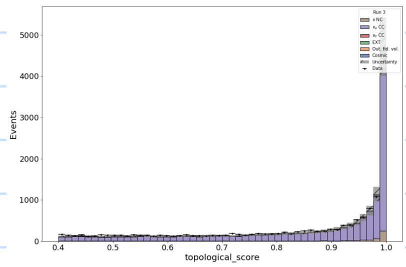
=>



↳ 9 topological : # of background events increase at low topological scores → Can filter easily - BUT lose efficiency

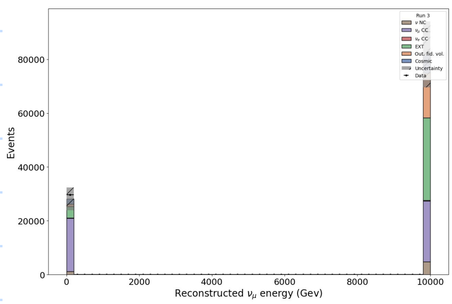


=>

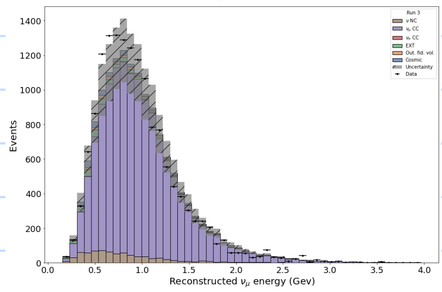


↳ 10 trk_energy_tot : Unphysical energies ~ 10 000 GeV → How did it get there ??

↳ Lots of data lost...



=>



[13:51] From exercise 1:

$$P(\nu_\alpha \rightarrow \nu_\beta) = \sin^2(2\theta)\,\sin^2\left(\frac{\Delta m^2_{21}\, L}{4E}\right)$$

$$= \sin^2(2\theta)\,\sin^2\left(1.27\,\frac{\Delta m^2_{21}\,[eV]^2\, L\,[km]}{E\,[GeV]}\right) \quad \text{in natural units } \hbar = c = 1$$

=> Consider the occurrence of muon neutrino disappearance: muon neutrino (ν_μ) oscillates into another neutrino flavour that MicroBooNE couldn't detect: $P(\nu_\mu \rightarrow \nu_\beta)$

=> Apply probability to data as a scaling factor

=> Apply 'goodness' of fit via minimisation of χ^2 between real & simulated data

parameter	best fit $\pm$ 1σ	2σ range	3σ range
Δm^2_{21} $[10^{-5}\text{eV}^2]$	$7.55^{+0.20}_{-0.16}$	7.20–7.94	7.05–8.14
$\lvert\Delta m^2_{31}\rvert$ $[10^{-3}\text{eV}^2]$ (NO)	2.50±0.03	2.44–2.57	2.41–2.60
$\lvert\Delta m^2_{31}\rvert$ $[10^{-3}\text{eV}^2]$ (IO)	$2.42^{+0.03}_{-0.04}$	2.34–2.47	2.31-2.51
$\sin^2\theta_{12}/10^{-1}$	$3.20^{+0.20}_{-0.16}$	2.89–3.59	2.73–3.79
$\theta_{12}/^\circ$	$34.5^{+1.2}_{-1.0}$	32.5–36.8	31.5–38.0
$\sin^2\theta_{23}/10^{-1}$ (NO)	$5.47^{+0.20}_{-0.30}$	4.67–5.83	4.45–5.99
$\theta_{23}/^\circ$	$47.7^{+1.2}_{-1.7}$	43.1–49.8	41.8–50.7
$\sin^2\theta_{23}/10^{-1}$ (IO)	$5.51^{+0.18}_{-0.30}$	4.91–5.84	4.53–5.98
$\theta_{23}/^\circ$	$47.9^{+1.0}_{-1.7}$	44.5–48.9	42.3–50.7
$\sin^2\theta_{13}/10^{-2}$ (NO)	$2.160^{+0.083}_{-0.069}$	2.03–2.34	1.96–2.41
$\theta_{13}/^\circ$	$8.45^{+0.16}_{-0.14}$	8.2–8.8	8.0–8.9
$\sin^2\theta_{13}/10^{-2}$ (IO)	$2.220^{+0.074}_{-0.076}$	2.07–2.36	1.99–2.44
$\theta_{13}/^\circ$	$8.53^{+0.14}_{-0.15}$	8.3–8.8	8.1–9.0
δ/π (NO)	$1.21^{+0.21}_{-0.15}$	1.01–1.75	0.87–1.94
$\delta/^\circ$	218^{+38}_{-27}	182–315	157–349
δ/π (IO)	$1.56^{+0.13}_{-0.15}$	1.27–1.82	1.12–1.94
$\delta/^\circ$	281^{+23}_{-27}	229–328	202–349

BUT we want to see the probability of "surviving"

$$\therefore\ P(\nu_\mu \rightarrow \nu_\mu) = 1 - \sin^2(2\theta)\,\sin^2\left(1.27\,\frac{\Delta m^2_{21}\,[eV]^2\, L\,[km]}{E\,[GeV]}\right)$$

↳ What is θ? What is Δm_{21}? (Value - one?)

↳ Consider θ_{21} & Δm_{21} from table:

$\therefore\ \theta_{12} \simeq 34.5^\circ$

$\Delta m^2_{21} \simeq 7.55 \times 10^{-5}\ [eV]^2$

=> $\sin^2 2\theta \simeq 0.9$

$$\sin^2\left(1.27\,\frac{\Delta m^2_{21}\, L}{E}\right) \longrightarrow L \simeq 10^{-1},\quad \Delta m^2_{21} \simeq 10^{-4}$$

$$E \simeq 1$$

$\simeq 10^{-14}$ VERY small!!

↳ $P(\nu_\mu \rightarrow \nu_\mu) \simeq 1$

Day 4 - 16/03

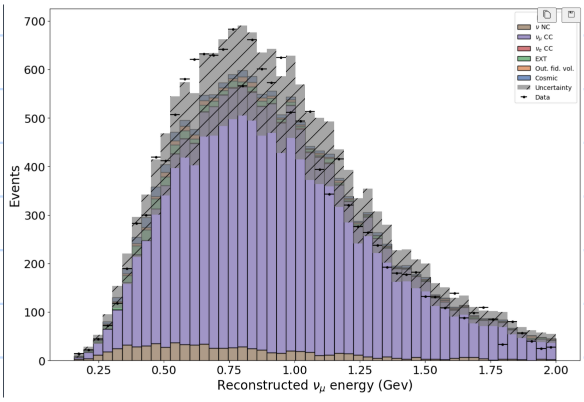
[10:04] Define probability of $\nu_\mu \to \nu_\mu$ ("surviving")

$$\Rightarrow P = 1 - \sin^2(2\theta)\,\sin^2\left(1.27\,\frac{\Delta m^2_{31} L}{E}\right)$$

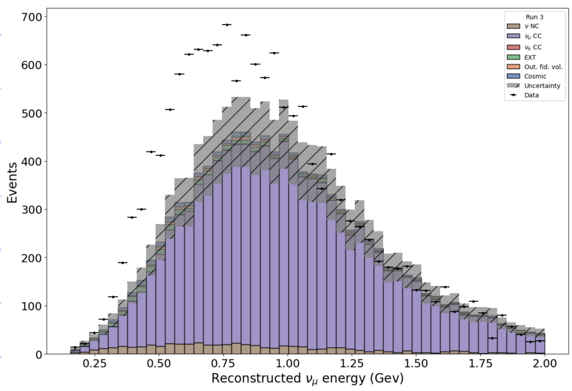
↳ Vary $\sin^2(2\theta)$ as a parameter; NOT θ

=> Applied corrections to histograms

↳ Treat P as a scaling factor → Apply to weight of dataset...



→ Apply changes to weighting.



=> Maximum correction of ~ 0.500.

[10:25] $$\chi^2 = \sum \left(\frac{\text{data} - \text{model}}{\sigma}\right)^2. \qquad \sigma = \text{Error on data}$$

=> $\chi^2 = 53.63$ (Before apply oscillation)

χ^2 of guess correction : 415.26 (After apply oscillation)

↳ Clear to be large : Oscillations result in reduced amount of surviving ν_μ

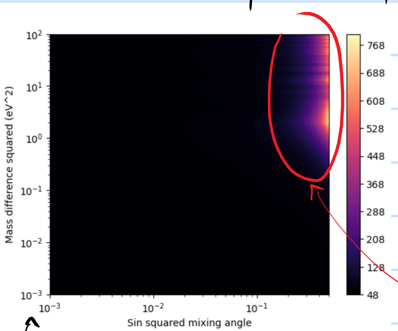
[12:30] Calculating χ^2 for each bin of varying energies takes too long...

→ Apply correction factor, P to EVERY bin height...

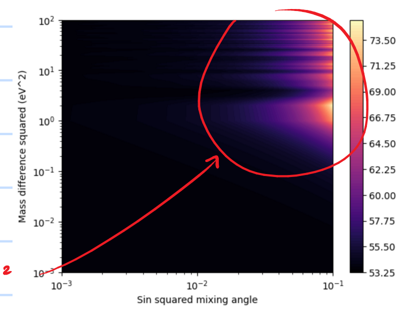
→ NOW: χ^2 values vary with changing parameters...

```
[[ 53.63254686  53.63254813  53.6325494  ...  53.63266995  53.63267122
   53.63267248]
 [ 53.73310407  54.25711461  54.81018672 ... 420.83323896 431.963133
  443.41524925]
 [ 53.72882254  54.24492485  54.81335134 ... 747.04548317 770.35082179
  794.408272  ]
 ...
 [ 53.62685992  53.61445043  53.62959395 ... 365.13805276 375.76644797
  386.72479819]
 [ 53.66353648  53.84102223  54.05465386 ... 533.8581161  550.72604091
  568.15865343]
 [ 53.68095199  53.94486802  54.24260824 ... 494.54360701 509.5420788
  525.02888613]]
```

[14:00] Can use plt.contourf & plt.contour to visualise varying χ^2 values



→ Smaller range of $\sin^2(2\theta)$



High χ^2 values...

Log scale...

Can plot interval levels...

↳ Regions where we can exclude data with a % of confidence.

→ Need associated χ^2 value for our data's d.o.f.

↳ D.o.F: 50 bins & 2 varying parameters

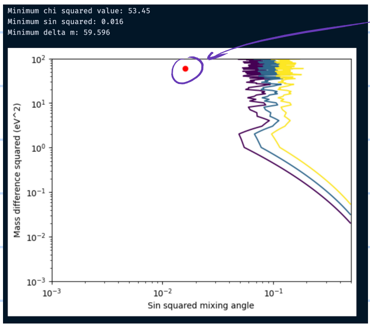
∴ D.o.F. = 50 − 2 = 48

=> Using https://www.itl.nist.gov/div898/handbook/eda/section3/eda3674.htm we find:

Degrees of freedom	Probability less than critical value				
	0.90	0.95	0.975	0.99	0.999
48	60.907	65.171	69.023	73.683	84.037

=> Can tabulate data into a dictionary...

∴ Can make a contour plot; drawing contour lines for the appropriate χ^2 values

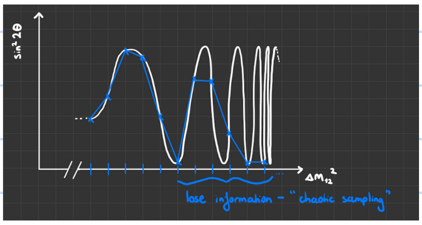


Minimum χ^2 ∴ Optimised param

↳ BOT unclear... => Sensitive to cuts in the dataset...

→ Needs more sensitivity...

=> Can see how samples are taken:



=> If taking samples of sinusoidal function whose frequency is increasing, we lose information about the shapes of the function at high frequencies...

→ Plot appears chaotic!

=> For small Δm^2: Sinusoidal

large Δm^2: Chaotic...

[15:34] Can minimise χ^2 for a certain $\sin^2(2\theta)$ and Δm^2 using curve-fit...

→ Results in $[\sin^2(2\theta)_{min}, \Delta m^2_{12\,min}]$ = [1e-3, 56.27]

↳ Smallest mixing angle as possible (boundary)

↳ Data points are ABOVE MC data

of MicroBooNE

=> Correction being in further from true data...

∴ Need small correction

=> $\sin^2(2\theta) \ll 1$.

Day 5 - 21/03.

[09:24] Do derivation of 3+1 neutrino model $\longrightarrow$ existence of a fourth sterile neutrino

↳ Larger mass than other neutrinos...

$$\begin{bmatrix} \nu_e \\ \nu_\mu \\ \nu_\tau \\ \nu_s \end{bmatrix} = \begin{bmatrix} U_{e1} & U_{e2} & U_{e3} & U_{e4} \\ U_{\mu 1} & U_{\mu 2} & U_{\mu 3} & U_{\mu 4} \\ U_{\tau 1} & U_{\tau 2} & U_{\tau 3} & U_{\tau 4} \\ U_{s1} & U_{s2} & U_{s3} & U_{s4} \end{bmatrix} \begin{bmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \\ \nu_4 \end{bmatrix}$$

↳ $|U_{e4}|^2 = \sin^2\theta_{14}$

$|U_{\mu 4}|^2 = \sin^2\theta_{24} \cos^2\theta_{14}$

$$\Rightarrow P(\alpha \to \beta) = \delta_{\alpha\beta} - 4(\delta_{\alpha\beta} - U_{\alpha 4}^* U_{\alpha 4}) U_{\beta 4} U_{\beta 4}^* \sin^2\left(1.27 \frac{\Delta m_{14}^2 L}{E} \frac{[eV^2][km]}{[GeV]}\right)$$

$$= \delta_{\alpha\beta} - \sin^2(2\theta_{\alpha\beta}) \sin^2\left(1.27 \frac{\Delta m_{14}^2 L}{E}\right)$$

$\rightarrow$ Value: $\sin^2 2\theta_{14} = \sin^2 2\theta_{ee} = 0.24$

↳ From before: Replace $\sin^2\theta$ with $\sin^2 2\theta_{\mu e}$

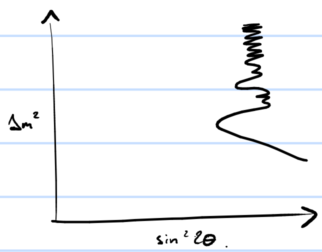
↳ Derive $\sin^2 2\theta_{\mu e} = 4 |U_{\mu 4}|^2 |U_{e4}|^2$ w.r.t. parameters of $\sin^2 2\theta_{\mu\mu}$ (i.e. original oscillation param range)

$\sin^2 2\theta_{ee} = 0.24$

$\rightarrow$ For now:

$$\sin^2 2\theta_{\mu e} = 4|U_{\mu 4}|^2 |U_{e4}|^2 = 4\cos^2\theta_{14} \sin^2\theta_{24} \sin^2\theta_{14}$$
$$= (2\cos\theta_{14}\sin\theta_{14})^2 \sin^2\theta_{24}$$
$$= \sin^2 2\theta_{14} \sin^2\theta_{24}$$

Current:



$\longrightarrow \sin^2 2\theta_{14} \sin^2\theta_{24} = 0.24 \sin^2\theta_{24}$. ??

↳ in term of $\sin^2 2\theta_{\mu\mu}$.

↳ From Ex 13:

$$\sin^2(2\theta_{\mu\mu}) = 4\cos^2\theta_{14}\sin^2\theta_{24} - 4\cos^4\theta_{14}\sin^4\theta_{24}$$

$$\rightarrow 4\cos^4\theta_{14}\sin^4\theta_{24} - 4\cos^2\theta_{14}\sin^2\theta_{24} + \sin^2(2\theta_{\mu\mu}) = 0$$

$$\sin^4\theta_{24} - \frac{\sin^2\theta_{24}}{\cos^2\theta_{14}} + \frac{\sin^2(2\theta_{\mu\mu})}{4\cos^4\theta_{14}} = 0$$

$\Rightarrow$ Let $x = \sin^2\theta_{24}$

$$\therefore \quad x^2 - \frac{1}{\cos^2\theta_{14}} x + \frac{\sin^2(2\theta_{\mu\mu})}{4\cos^4\theta_{14}} = 0$$

$$x = \frac{1}{2}\left\{ \frac{1}{\cos^2\theta_{14}} \pm \sqrt{\frac{1}{\cos^4\theta_{14}} - \frac{\sin^2(2\theta_{\mu\mu})}{\cos^4\theta_{14}}} \right\} = 0$$

$$= \frac{1}{2\cos^2\theta_{14}} \pm \frac{1}{2\cos^2\theta_{14}} \sqrt{1 - \sin^2(2\theta_{\mu\mu})}$$

$$\therefore \quad \sin^2\theta_{24} = \frac{1}{2\cos^2\theta_{14}}\left[1 \pm \sqrt{1 - \sin^2(2\theta_{\mu\mu})}\right]$$

$\sin^2 2\theta$

$$\therefore \quad \sin^2(2\theta_{\mu e}) = \frac{\sin^2(2\theta_{14})}{2\cos^2\theta_{14}}\left[1 \pm \sqrt{1 - \sin^2(2\theta_{\mu\mu})}\right]$$

$\theta \longrightarrow \sin^2 2\theta$

$$= \underbrace{\frac{\sin^2(2\theta_{14})}{1 + \sqrt{1 - \sin^2(2\theta_{14})}}}\left[1 \pm \sqrt{1 - \sin^2(2\theta_{\mu\mu})}\right]$$

$\hookrightarrow$ + or − $\Longrightarrow$ $|\sin^2(2\theta_{\mu e})| \leq 1$

$\hookrightarrow$ Do analysis!! $|\sin^2(2\theta_{\mu\mu})| \leq 1$

$$= \frac{0.24}{1 + \sqrt{1 - 0.24}} = 0.128.$$

$\hookrightarrow$ Consider $\sin^2(2\theta_{\mu\mu}) = -1$

$\therefore \quad \sin^2(2\theta_{\mu e}) \propto 1 \pm \sqrt{2}$

[10:59] Ex 13:

$$\begin{bmatrix} \nu_e \\ \nu_\mu \\ \nu_\tau \\ \nu_s \end{bmatrix} = \begin{bmatrix} U_{e1} & U_{e2} & U_{e3} & U_{e4} \\ U_{\mu 1} & U_{\mu 2} & U_{\mu 3} & U_{\mu 4} \\ U_{\tau 1} & U_{\tau 2} & U_{\tau 3} & U_{\tau 4} \\ U_{s1} & U_{s2} & U_{s3} & U_{s4} \end{bmatrix} \begin{bmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \\ \nu_4 \end{bmatrix}$$

$\hookrightarrow \quad |U_{e4}|^2 = \sin^2\theta_{14}$

$\quad\quad |U_{\mu 4}|^2 = \sin^2\theta_{24} \cos^2\theta_{14}$

$$\Rightarrow P(\alpha \to \beta) = \delta_{\alpha\beta} - 4(\delta_{\alpha\beta} - U_{\alpha 4}^* U_{\alpha 4}) U_{\beta 4} U_{\beta 4}^* \sin^2\left(1.27 \frac{\Delta m_{14}^2 L}{E} \frac{[eV^2][km]}{[GeV]}\right)$$

$$= \delta_{\alpha\beta} - \sin^2(2\theta_{\alpha\beta}) \underbrace{\sin^2\left(1.27 \frac{\Delta m_{14}^2 L}{E}\right)}_{\gamma}$$

$$\Rightarrow P(\mu \to \mu) = 1 - 4(1 - |U_{\mu 4}|^2)|U_{\mu 4}|^2 \gamma = 1 - \sin^2(2\theta_{\mu\mu})\gamma$$

$$\Rightarrow 4(1 - |U_{\mu 4}|^2)|U_{\mu 4}|^2 = \sin^2(2\theta_{\mu\mu})$$

$$\Rightarrow \sin^2(2\theta_{\mu\mu}) = 4|U_{\mu 4}|^2 - 4|U_{\mu 4}|^4$$

$$= 4\sin^2\theta_{24}\cos^2\theta_{14} - 4\sin^4\theta_{24}\cos^4\theta_{14}$$

$$\Rightarrow P(e \to e) = 1 - 4(1 - |U_{e4}|^2)|U_{e4}|^2 \gamma = 1 - \sin^2(2\theta_{ee})\gamma$$

$$\Rightarrow 4(1 - |U_{ee}|^2)|U_{ee}|^2 = \sin^2(2\theta_{ee})$$

$$\Rightarrow \sin^2(2\theta_{ee}) = 4|U_{e4}|^2 - 4|U_{e4}|^4$$

$$= 4\sin^2\theta_{14} - 4\sin^4\theta_{14}$$

$$\Rightarrow P(\mu \to e) = -4|U_{\mu 4}|^2|U_{e4}|^2\gamma = -\sin^2(2\theta_{\mu e})\gamma$$

$$\Rightarrow \sin^2(2\theta_{\mu e}) = 4(\sin^2\theta_{24}\cos^2\theta_{14})(\sin^2\theta_{14})$$

$$= \sin^2\theta_{24}(2\cos\theta_{14}\sin\theta_{14})^2$$

$$= \sin^2\theta_{24}\sin^2(2\theta_{14})$$

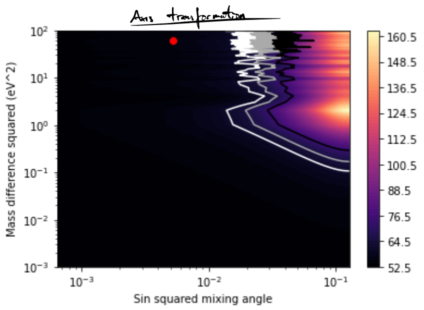
[11:38] Ex 14.

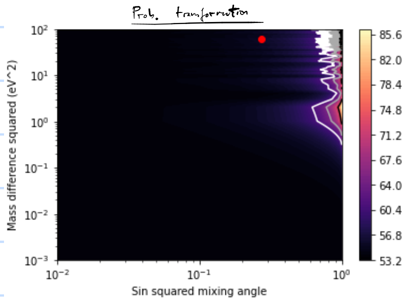
Apply transformation $\sin^2(2\theta_{\mu e}) = 0.128\left[1 - \sqrt{1 - \sin^2(2\theta_{\mu\mu})}\right]$

onto asrs OR probability ??

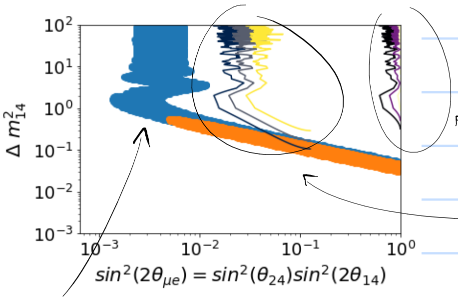
→ Return original χ^2 values IF applied to asrs.

→ Entire data transformed IF applied to probability





⇒ Which is better ?? → Compare to actual dataset...



Probability

Asrs

Possible values for sterile neutrinos to exist.

⇒ Where do we want the lines ??

→ Hypothesis testing

→ Is it possible for the data to exist??

→ Right of line : Doesn't exist with % confidence.

→ Not left enough :

- Not enough data (selection needs fixing)
- etc not sensitive enough

↳ Find balance between efficiency & purity.

⇒ Expected to exclude (majority)

$\sin^2(2\theta_{\mu e}) = \sin^2(\theta_{24})\ \sin^2(\theta_{14})$

$\nu_2 \to \nu_4$

⇒ $\mu \to s$

equivalent to

$\mu \to$ nothing

$= 1 - (\mu \to \mu)$

↳ Same as 2-flavour oscillation (μ disappearance)

Altered by a factor since its 4-flavour now

$\nu_1 \to \nu_4$

⇒ $e \to s \iff s \to e$

↳ Assumed to be 0.24

⇒ Scaling previous probability

→ Already calculated it all ; just needs scaling...

[13:47] Using pre-oscillated data.

-> Redo previous exercises -> Find minimum Δm^2 & $\sin^2 2\theta$.

TO DO:

Figure out why min $\sin^2 = 1$ in above exercise

- Work on selection cuts -> How to modify??
  - ↳ ML?
  - ↳ Theory?

Day 6 - 23/03

[9:30] Can we remodify contour calculation??

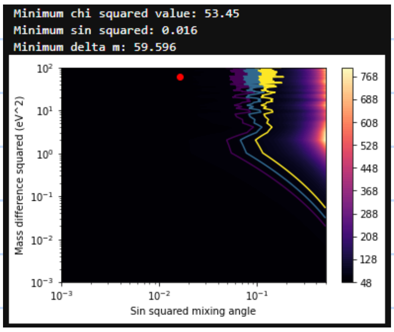
→ Use meshgrids ??

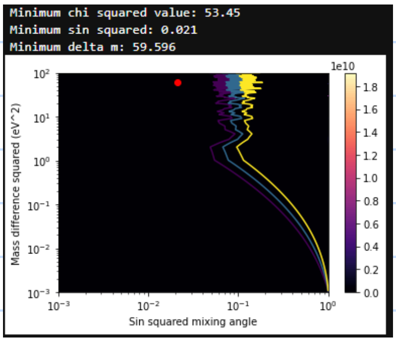
[10:07] Something went wrong...

→ Previous exercises have errors??

→ Doesn't like updated range of $\sin^2(2\theta)$ range:

(1e-3, 0.5, 100) ⟶ (1e-3, 1, 100)





↳ HUGE χ^2 value!!

Are we altering stuff in the original function??

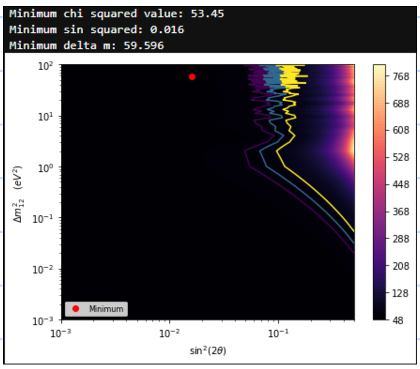
→ Change function / parameter names to be clearer / unique.

[14:20] All fixed!! → Graphs more consistent now

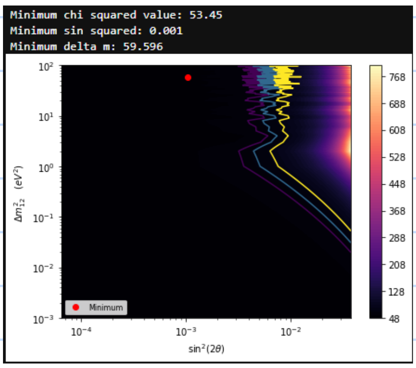
& made program more functional

New graphs:

2-neutrino models



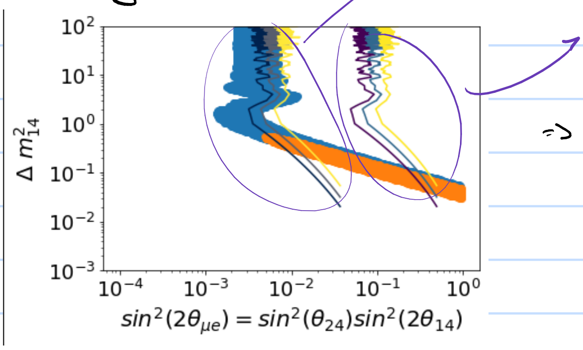
4-neutrino model.



Comparing both models

4-neutrino

2-neutrino



⇒ See that we start excluding some of the MicroBooNE data with 4-neutrino model

⇒ Still need to exclude more...

Closure test

=> Given an oscillated data set.

=> know that the true values of dataset are:

$\sin^2 2\theta = 0.556$

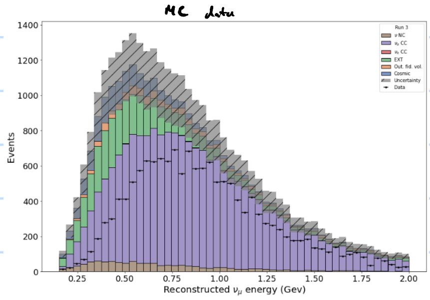
$\Delta m^2 = 11.12$.

=> Using defined functions from before → Find oscillation bin heights

↳ Treat equivalent to MicroBooNE dataset given

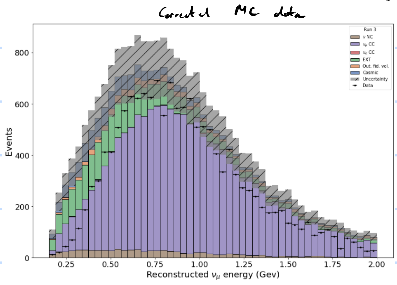
→ Apply oscillation corrections to MC dataset (NOTE: Ignore black points...)

MC data



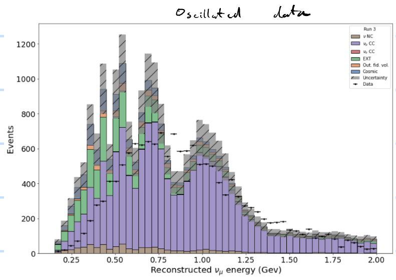
→ Oscillation correction

Corrected MC data

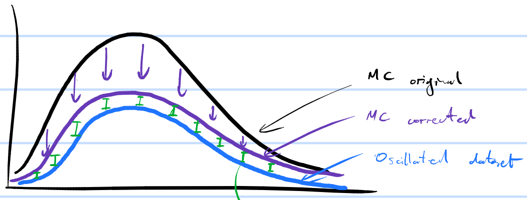


NOTE: ONLY selected $E < 2\,GeV$ => Not fully implemented for oscillated data set.

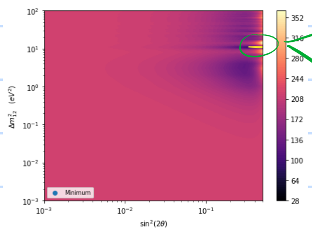
Oscillated data

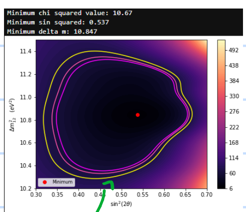


=> Should expect MC data to be larger THEN shrink towards oscillated data set:



=> Can find multivariate χ^2:





Confidence levels: 90%, 95%, 99%

=> Find that $\sin^2(2\theta)_{min} = 0.537$

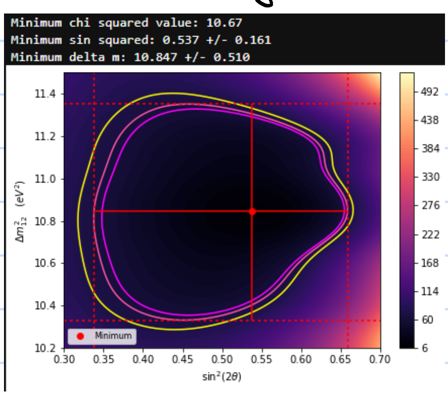
$\Delta m^2_{min} = 10.847$

=> BUT need to find uncertainties

=> Using 95% confidence level → Find max height $\sigma_{\Delta m^2}$

max width $\sigma_{\sin^2(2\theta)}$

=> Access contours directly...



=> Analyze 95% confidence level

→ Obtained values:

$\sin^2(2\theta) = 0.537 \pm 0.161$ (± 30.0%)

$\Delta m^2_{21} = 10.847 \pm 0.510$ (± 4.7%)

Why is $\sin^2 2\theta$ uncertainty so large??

-> Is it less significant than mass component??