Neutrino Oscillation

Mar 10, 2023 NeutrinoNet, NTHU Physics AI

Yuan-Yen Peng, Siang-Yuan Lin
PHYS591000 Hands-on Artificial Intelligence for Physics
National Tsing Hua University (NTHU)
Hsinchu, Taiwan

Outline

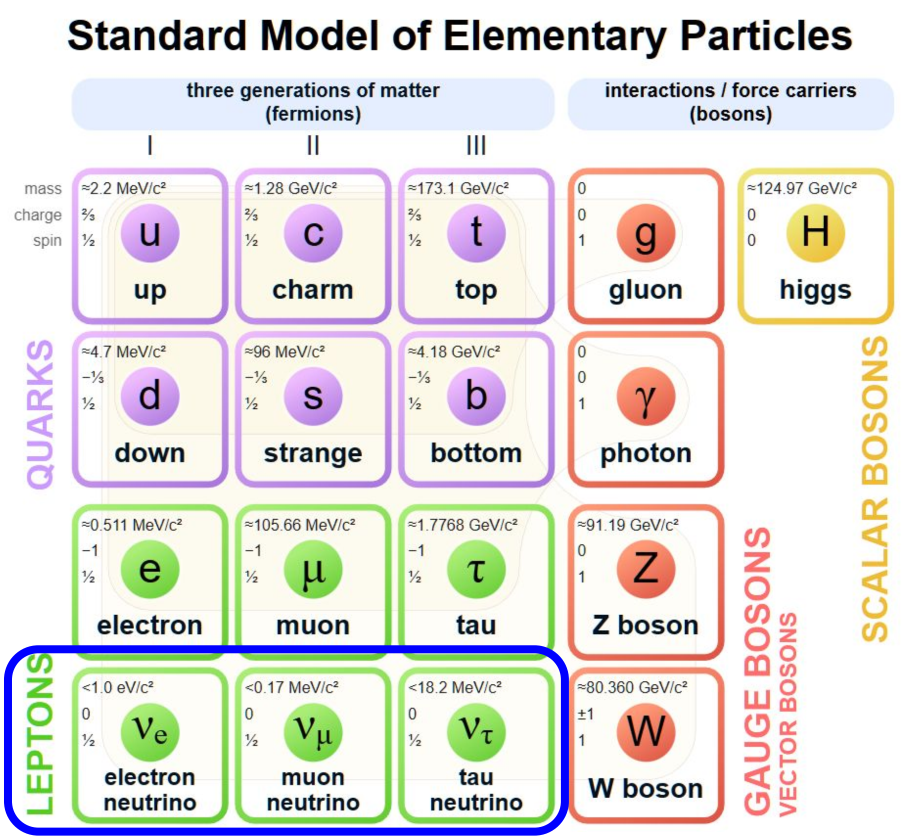
- Introduction
  - Theoretical Background
  - Experiments
- Methodology
  - Data exploration and inspired models
- Discussion
  - A preliminary evaluation on our models
- Summary
  - Outlook of our own model

What are neutrinos?

- Fermion: spin-½
- Lepton: no strong interaction
- Charge neutral

Where do they come from?

- Weak interactions
- Nuclear decays
- Nuclear reactions:
    - fusion, spallation, etc.



Source: Wikipedia

What is oscillating?

- Mass Eigenstates:
  - The eigenstates to the free Hamiltonian are mass eigenstates
- Flavor Eigenstates:
  - The eigenstates to the interaction Hamiltonian are flavor eigenstates
- Superposition:
  - The standard model neutrinos are theorized to be a superposition of mass eigenstates

$$|\nu_i\rangle = \sum_{\alpha} U^*_{\alpha i} \, |\nu_\alpha\rangle \, ,$$

$$|\nu_\alpha\rangle = \sum_{i} U_{\alpha i} \, |\nu_i\rangle \, ,$$

- Time evolution:

$$|\,\nu_j(t)\,\rangle = e^{-i\left(E_j t - \vec{p}_j \cdot \vec{x}\right)} \, |\,\nu_j(0)\,\rangle$$

NATIONAL TSING HUA UNIVERSITY 國立清華大學

A. Donini, M.B. Gavela, P. Hernández, S. Rigolin, Neutrino mixing and CP-violation, Nuclear Physics B. 574 (2000) 23–42. https://doi.org/10.1016/S0550-3213(00)00029-8.

How does it oscillate?

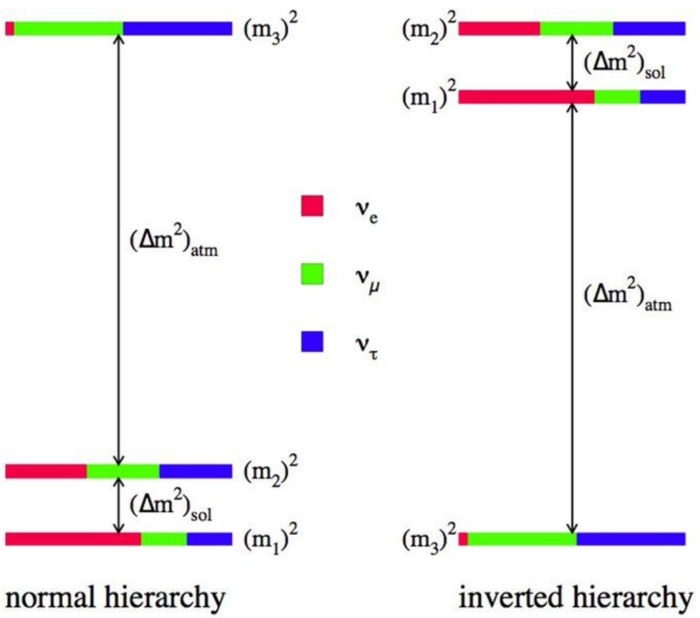
- The **P**ontecorvo–**M**aki–**N**akagawa–**S**akata (PMNS) Matrix

$$
\begin{aligned}
U &= \begin{bmatrix} U_{e1} & U_{e2} & U_{e3} \\ U_{\mu 1} & U_{\mu 2} & U_{\mu 3} \\ U_{\tau 1} & U_{\tau 2} & U_{\tau 3} \end{bmatrix} \\
&= \begin{bmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{bmatrix} \begin{bmatrix} c_{13} & 0 & s_{13}e^{-i\delta} \\ 0 & 1 & 0 \\ -s_{13}e^{i\delta} & 0 & c_{13} \end{bmatrix} \begin{bmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} e^{i\alpha_1/2} & 0 & 0 \\ 0 & e^{i\alpha_2/2} & 0 \\ 0 & 0 & 1 \end{bmatrix} \\
&= \begin{bmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{bmatrix} \begin{bmatrix} e^{i\alpha_1/2} & 0 & 0 \\ 0 & e^{i\alpha_2/2} & 0 \\ 0 & 0 & 1 \end{bmatrix},
\end{aligned}
$$

- The unitary transformation between the flavor and mass eigenstates can be determined with 4 parameters. θ_{12}, θ_{13}, θ_{23}, and δ_{CP}
- With this information, we can calculate the interference between *waves*. Given a sufficient amount of time, the phase difference becomes measurable. ***Long baseline experiments*** are designed to observe this phenomenon.

The mass hierarchy and measurements

As we know, the quantum states is not affected by some ***overall phase***, it is the ***relative phase*** that matters. Which is why we are only able to know the **difference** but not the **ordering.**



Source: Wikipedia

Source: DUNE FD TDR vol 1

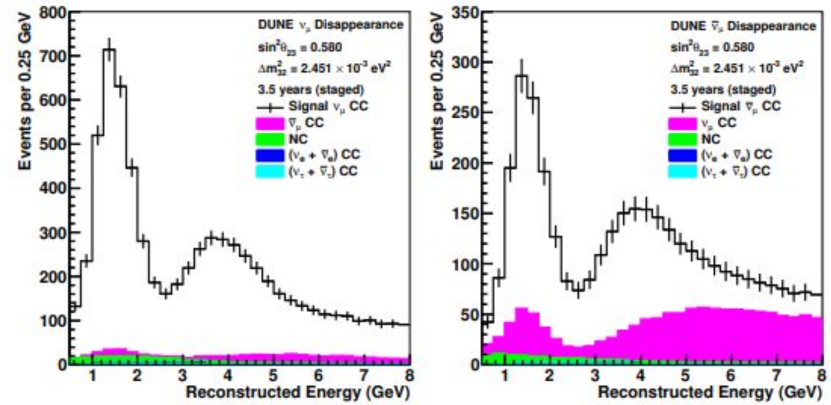


Figure 2.5. ν_μ and $\bar{\nu}_\mu$ disappearance spectra: reconstructed energy distribution of selected ν_μ CC-like events assuming 3.5 years (staged) running in the neutrino-beam mode (left) and antineutrino-beam mode (right), for a total of seven years (staged) exposure. The plots assume normal mass ordering.

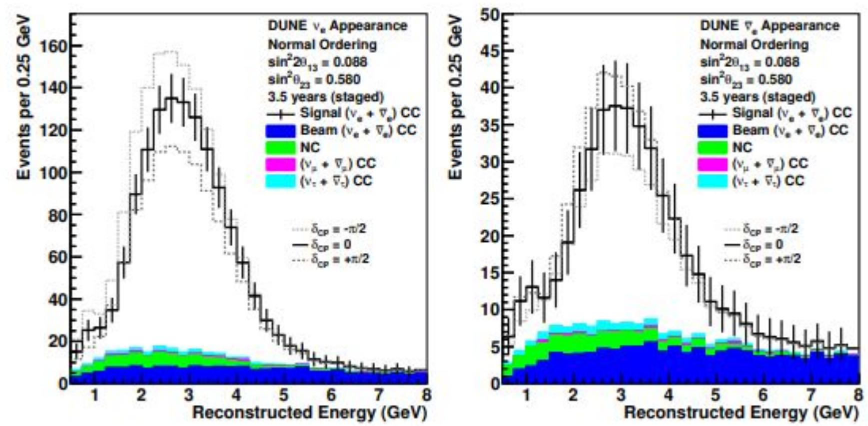


Figure 2.4. ν_e and $\bar{\nu}_e$ appearance spectra: reconstructed energy distribution of selected ν_e CC-like events assuming 3.5 years (staged) running in the neutrino-beam mode (left) and antineutrino-beam mode (right), for a total of seven years (staged) exposure. The plots assume normal mass ordering and include curves for $\delta_{CP} = -\pi/2, 0$, and $\pi/2$.

Probability and symmetry

- Equations for transition probabilities:

$$P_{\mathrm{CP\ even}}(\nu_\mu \to \nu_e) = \sin^2(\theta_{13}) \sin^2(2\theta_{13}) \sin^2\left(\frac{\Delta m_{23}{}^2 L}{4E}\right)$$

$$P_{\mathrm{CP\ even}}(\nu_e \to \nu_\tau) = \cos^2(\theta_{23}) \sin^2(2\theta_{13}) \sin^2\left(\frac{\Delta m_{23}{}^2 L}{4E}\right)$$

$$P_{\mathrm{CP\ even}}(\nu_\mu \to \nu_\tau) = \cos^4(\theta_{13}) \sin^2(2\theta_{23}) \sin^2\left(\frac{\Delta m_{23}{}^2 L}{4E}\right)$$

$$P_{\mathrm{CP\ odd}}(\nu_e \to \nu_\mu) = -P_{\mathrm{CP\ odd}}(\nu_\mu \to \nu_\tau) = -P_{\mathrm{CP\ odd}}(\nu_e \to \nu_\tau) =$$

$$-8\cos(\theta_{12})\cos^2(\theta_{13})\sin(\theta_{12})\sin(\theta_{13})\sin(\theta_{23})\sin\left(\delta\frac{\Delta m_{13}{}^2 L}{4E}\right)\sin^2\left(\frac{\Delta m_{23}{}^2 L}{4E}\right)$$

Table 4.5 Scalars and vectors under parity

Scalar	:	$P(s) = s$
Pseudoscalar	:	$P(p) = -p$
Vector (or polar vector)	:	$P(\mathbf{v}) = -\mathbf{v}$
Pseudovector (or axial vector)	:	$P(\mathbf{a}) = \mathbf{a}$

- Transition Amplitudes:
If a phase is picked up from the PMNS matrix, the interference terms will cause difference between the squared amplitude.

$$\mathcal{M} = |\mathcal{M}|e^{i\theta}e^{+i\delta}$$

$$\bar{\mathcal{M}} = |\bar{\mathcal{M}}|e^{i\theta}e^{-i\delta}$$

$$\mathcal{M} = |\mathcal{M}_1|e^{i\theta_1}e^{+i\delta_1} + |\mathcal{M}_2|e^{i\theta_2}e^{+i\delta_2}$$

$$\bar{\mathcal{M}} = |\bar{\mathcal{M}}_1|e^{i\theta_1}e^{-i\delta_1} + |\bar{\mathcal{M}}_2|e^{i\theta_2}e^{-i\delta_2}$$

What is being measured?

- The experiment aims to probe the effects of neutrino oscillations and CP violation measurements. Consider the CP asymmetry observable:

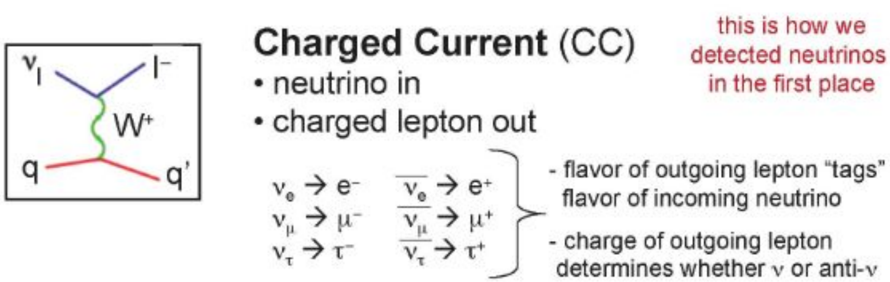
$$A_{ab}^{CP}(\delta) \equiv \frac{P(\nu_a \to \nu_b) - P(\bar{\nu}_a \to \bar{\nu}_b)}{P(\nu_a \to \nu_b) + P(\nu_a \to \nu_b)}$$

- In practice, this observable is constructed from the integrated spectrum of decaying leptons $N_{\ell\pm}$ and wrong-sign leptons $N_{\mathrm{wrong}\ell\pm}$

$$A_{e\mu}^{\mathrm{CP}}(\delta) = \frac{N_{\mu^-}/N_{e^-}^{\mathrm{wrong}} - N_{\mu^+}/N_{e^+}^{\mathrm{wrong}}}{N_{\mu^-}/N_{e^-}^{\mathrm{wrong}} + N_{\mu^+}/N_{e^+}^{\mathrm{wrong}}}$$

How is it measured?

- Appearance Experiments:
    - Measure transitions between different neutrino flavors. If the final flavor to be searched for in the detector is not present in the initial beam, it could be evidence of oscillation.
- Disappearance Experiments
    - Measure the survival probability of a neutrino by comparing the number of interactions in the detector with expected value.
- ***Wrong-sign leptons***: (an Example)
    - In a muon μ^+ decay process $\mu^+ \rightarrow e^+ \underline{\nu}_e \nu_\mu$, as the neutrino charged current interactions produce μ^-, termed as a wrong-sign muon.



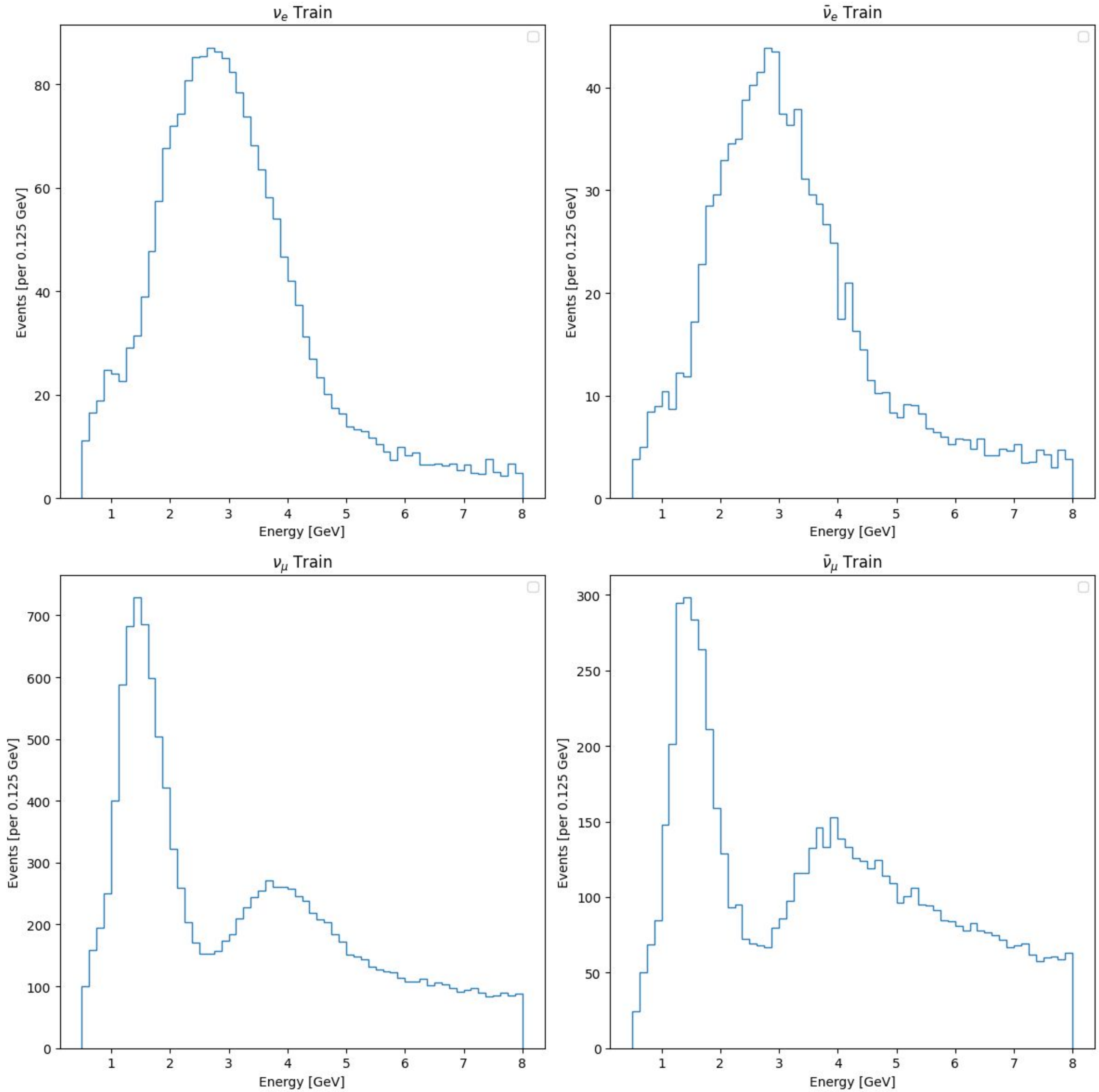
Given Data

- Training:

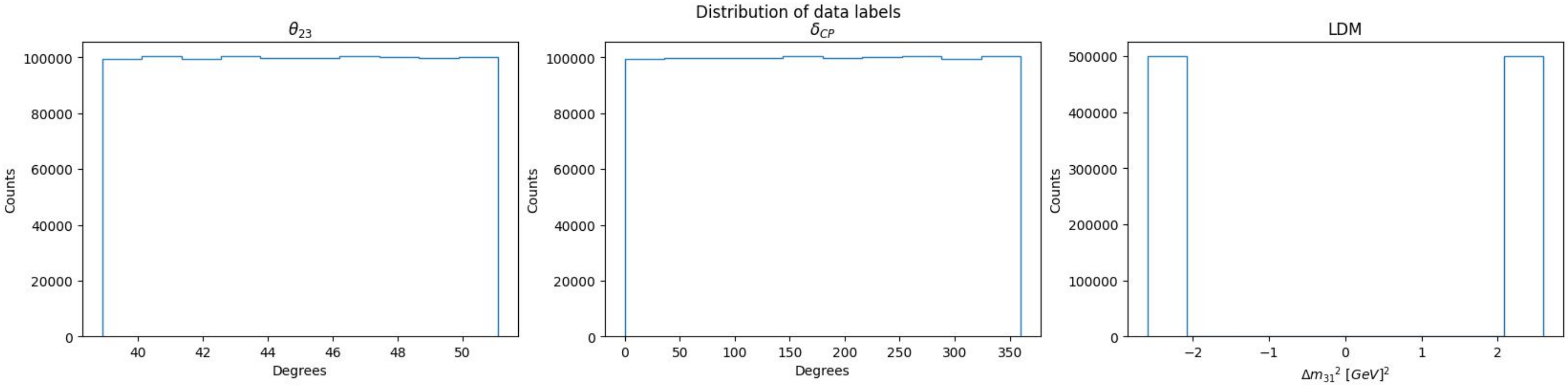
  ν_e energy spectrum
  $\underline{\nu}_e$ energy spectrum
  ν_μ energy spectrum
  $\underline{\nu}_\mu$ energy spectrum
  θ_{23}
  δ_{CP}

- Testing:

  ν_e energy spectrum
  $\underline{\nu}_e$ energy spectrum
  ν_μ energy spectrum
  $\underline{\nu}_\mu$ energy spectrum



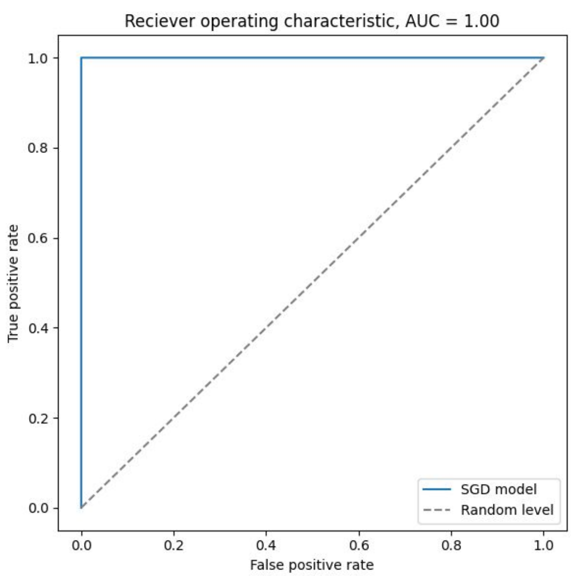
Given Data

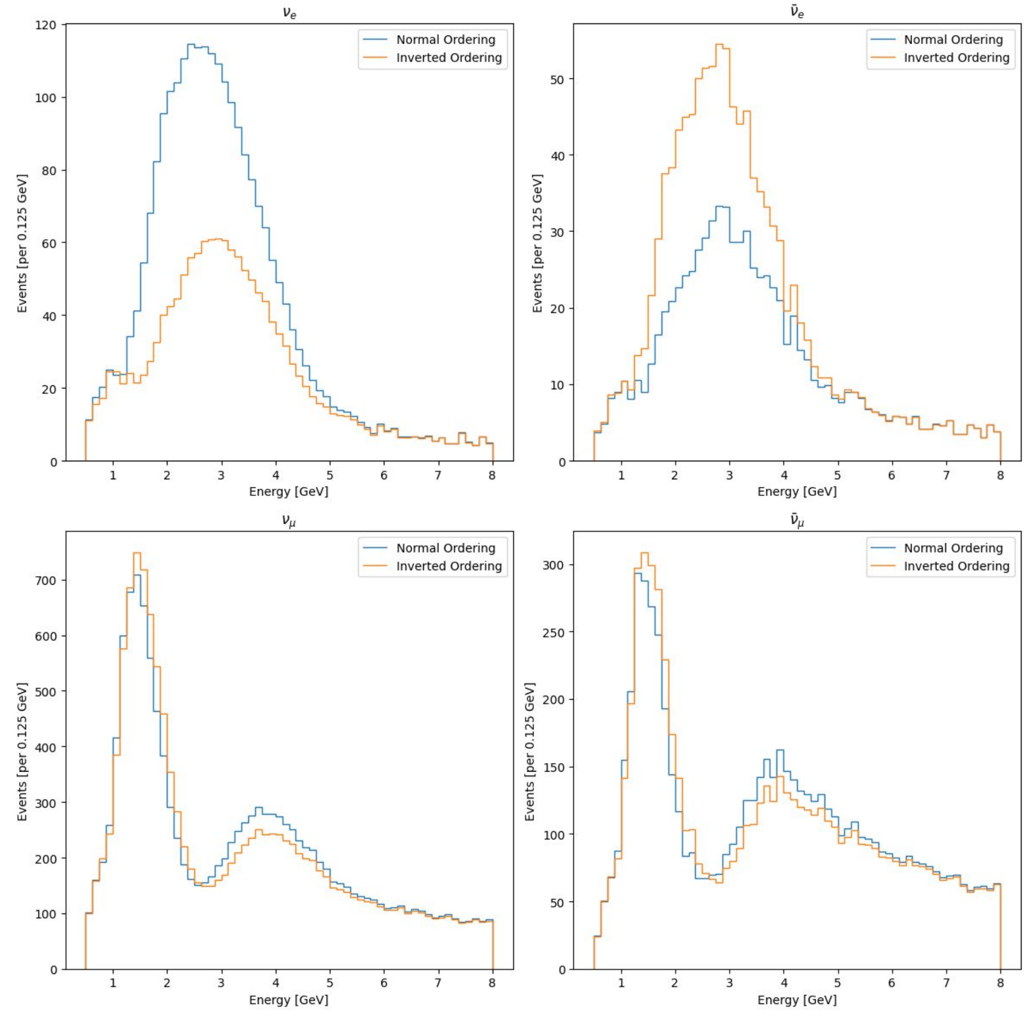


- Tasked to find the correlation between angles and spectrum plots.
- θ_{23} ranges from approximately 39 to 51 degrees,
  δ_{CP} ranges from 0 to 360
- Immediately noticeable, is the very distinct separation in ***LDM***. One on the positive side, and the other on the negative end. This corresponds to the mass hierarchy mentioned in the referenced literature.
- We will make this distinction for all following analysis.

Given Data

Here we plot the averaged spectrum of the normal and inverted ordering data. In the electron neutrino spectrum, we can see a distinct difference in the distribution.
The first model implemented is a support vector machine to make this binary classification.





The θ_{23} mixing angle

Theoretically, θ_{23} in (with non-trivial leading order):

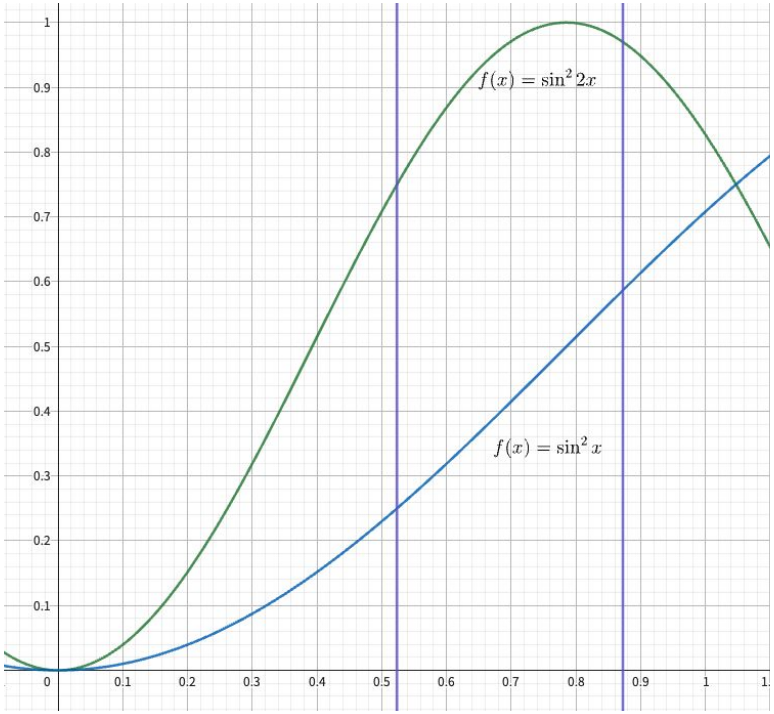
appearance channel: $sin^2\ \theta_{23}$

disappearance channel: $sin^2\ 2\theta_{23}$

$$P_{\mu e} = \underbrace{4 s_{13}^2 s_{23}^2 \frac{\sin^2 (A-1)\Delta}{(A-1)^2}}_{\mathcal{O}_o} + \underbrace{\alpha^2 \cos^2\theta_{23} \sin^2 2\theta_{12} \frac{\sin^2 A\Delta}{A^2}}_{\mathcal{O}_2}$$

$$+ \underbrace{\alpha s_{13} \sin 2\theta_{12} \sin 2\theta_{23} \cos(\Delta + \delta_{cp}) \frac{\sin(A-1)\Delta}{(A-1)} \frac{\sin A\Delta}{A}}_{\mathcal{O}_1}$$

$$P_{\mu\mu} = 1 - \sin^2 2\theta_{23} \sin^2\Delta + \mathcal{O}(\alpha, s_{13})$$



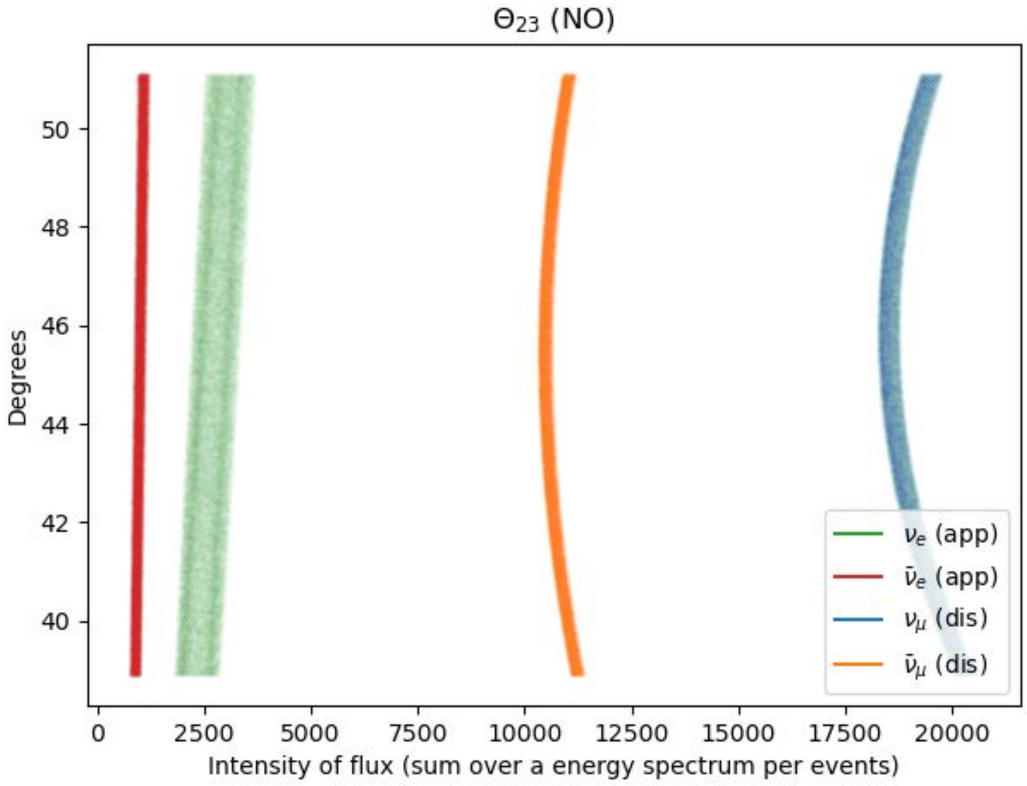
In this interval (30 ~ 50 degrees), the probability can be determined, due to the monotonicity of the function (for appearance channels).

Ref: N. Nath, M. Ghosh, and S. Goswami, *The Physics of Antineutrinos in DUNE and Determination of Octant and* δ_{CP}, Nuclear Physics B **913**, 381 (2016).

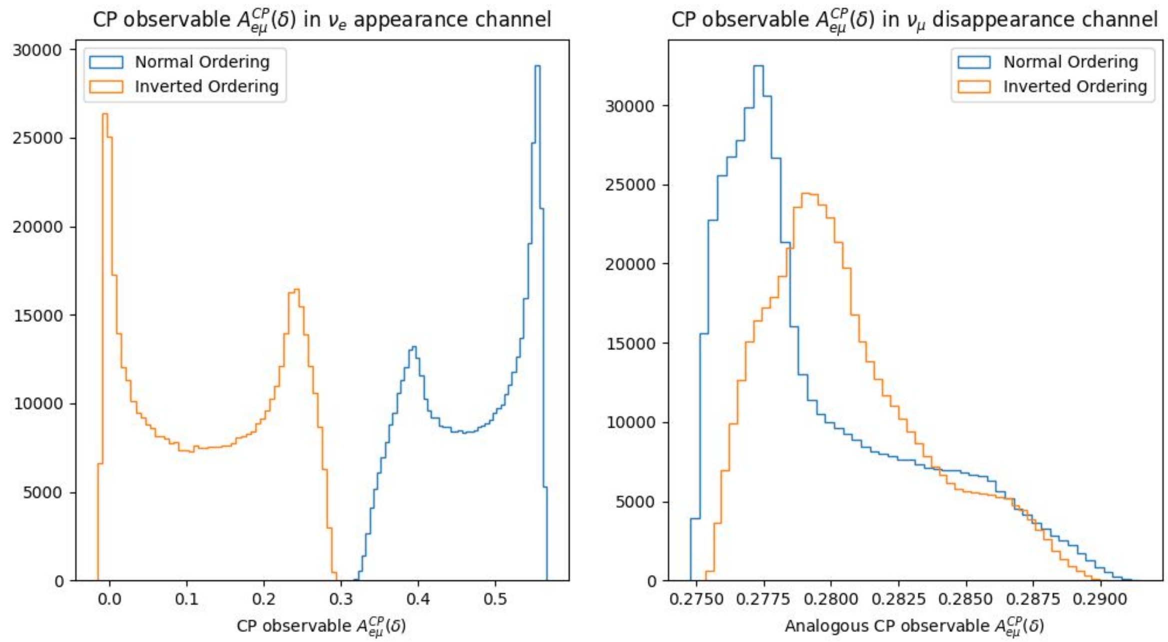
The θ_{23} Regression

As demonstrated in the formula and this plot, a one-to-many relationship between the total events and θ_{23} poses as a challenge in the regression task. But not in the appearance channel.

Combining these features, and with aid of machine learning methods, we can expect good predictability in θ_{23}

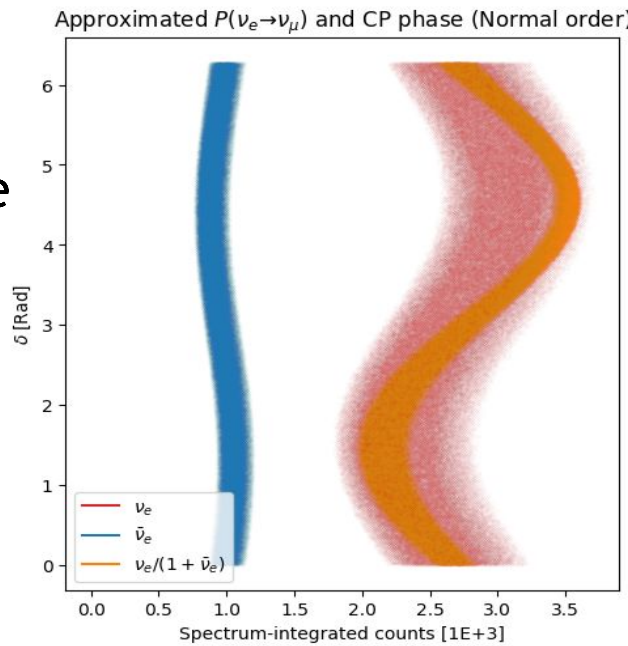


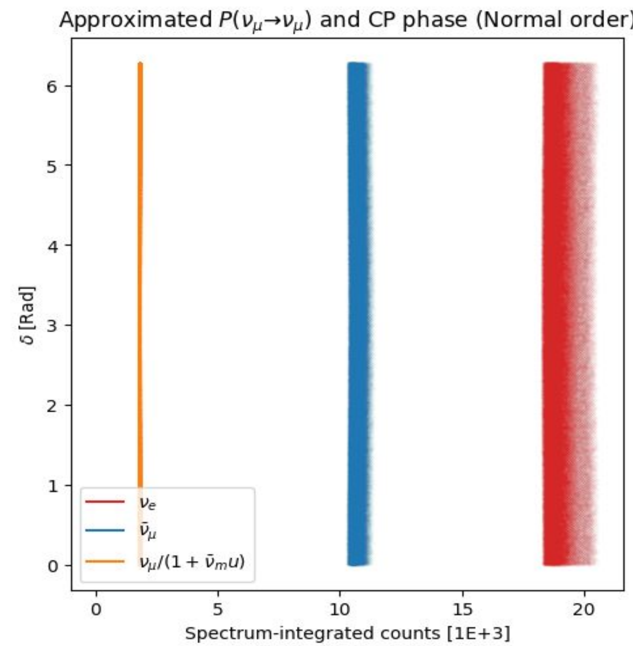
The CP phase and asymmetry observable



- As mentioned earlier, the CP phase can be quantified by the level of asymmetry.
- Here we present the distribution of CP phase and this observable in both the appearance channel and disappearance channel.

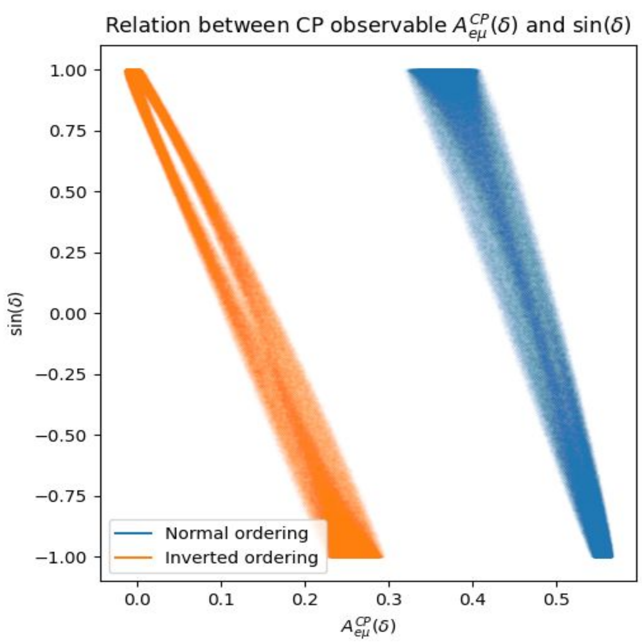
The upper and lower plots present the distribution of CP phase and its relationship with total event counts and the CP asymmetry observable respectively.

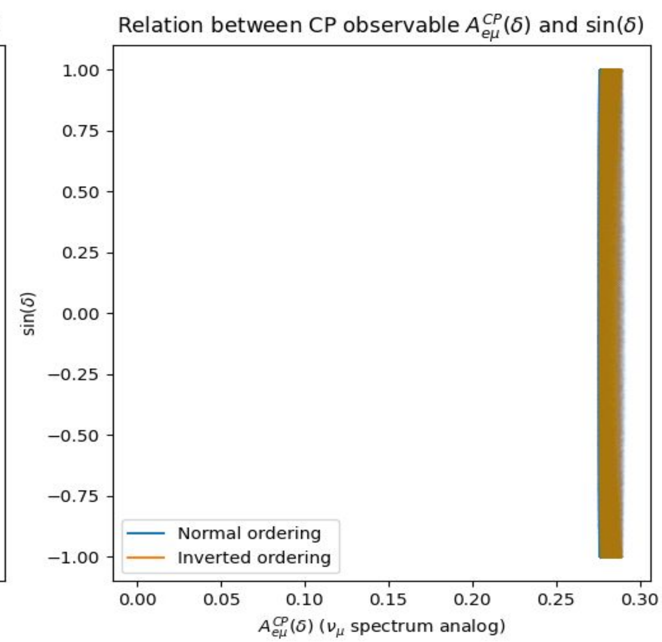


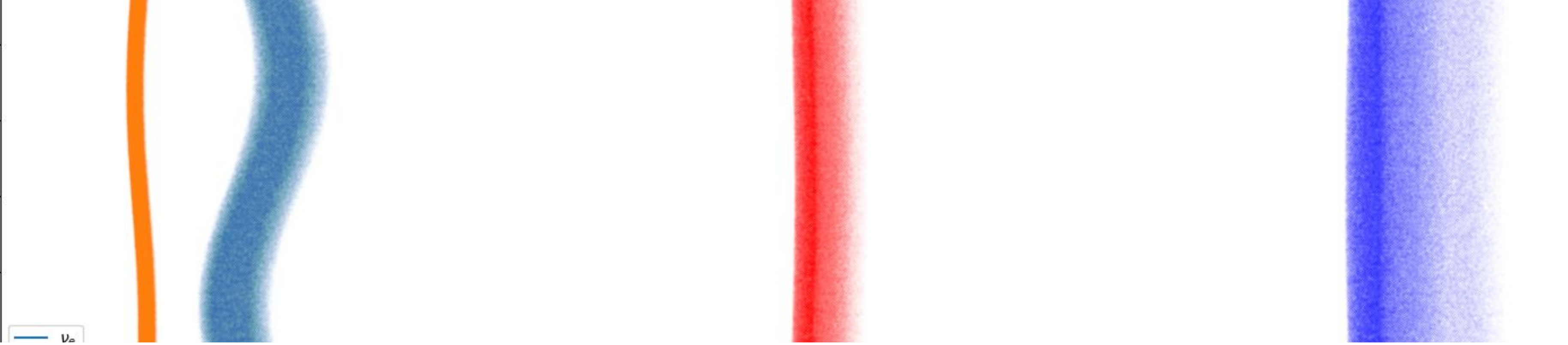


Plots of the ν_e and ν_μ spectrums are on the left and right side respectively.

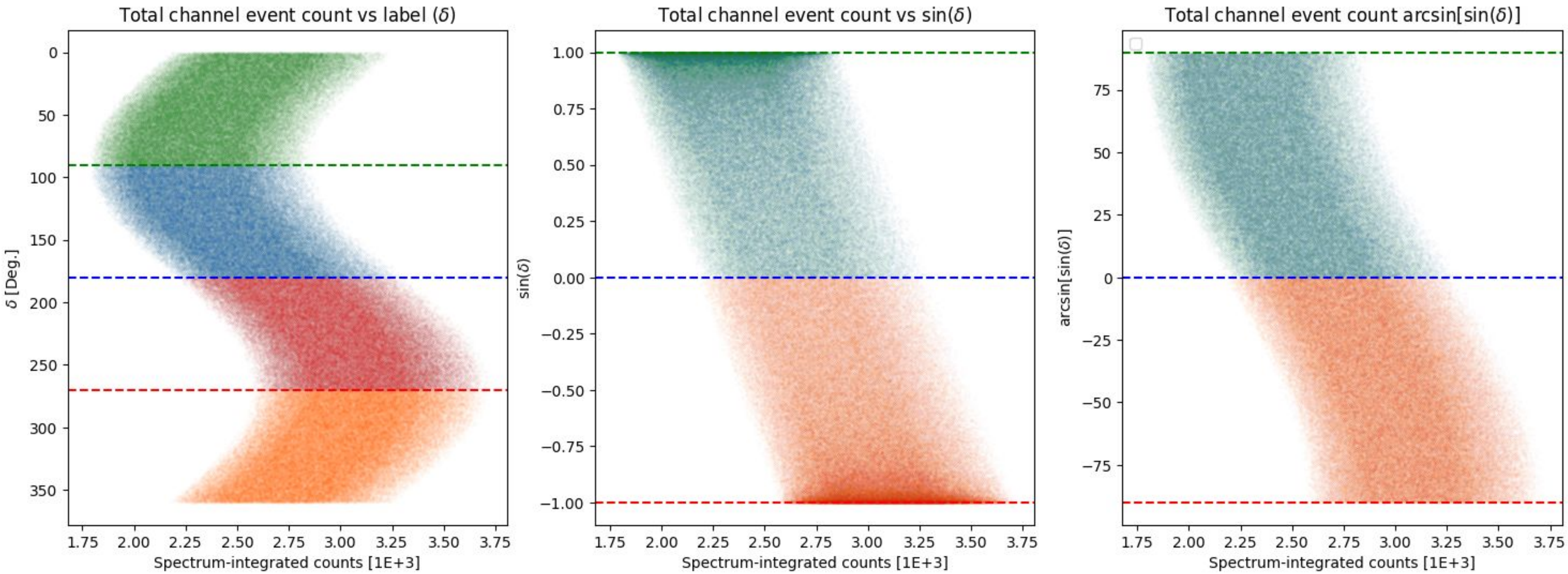
Our observations show that the event counts in ν_μ spectrums are not correlated with the CP phase.



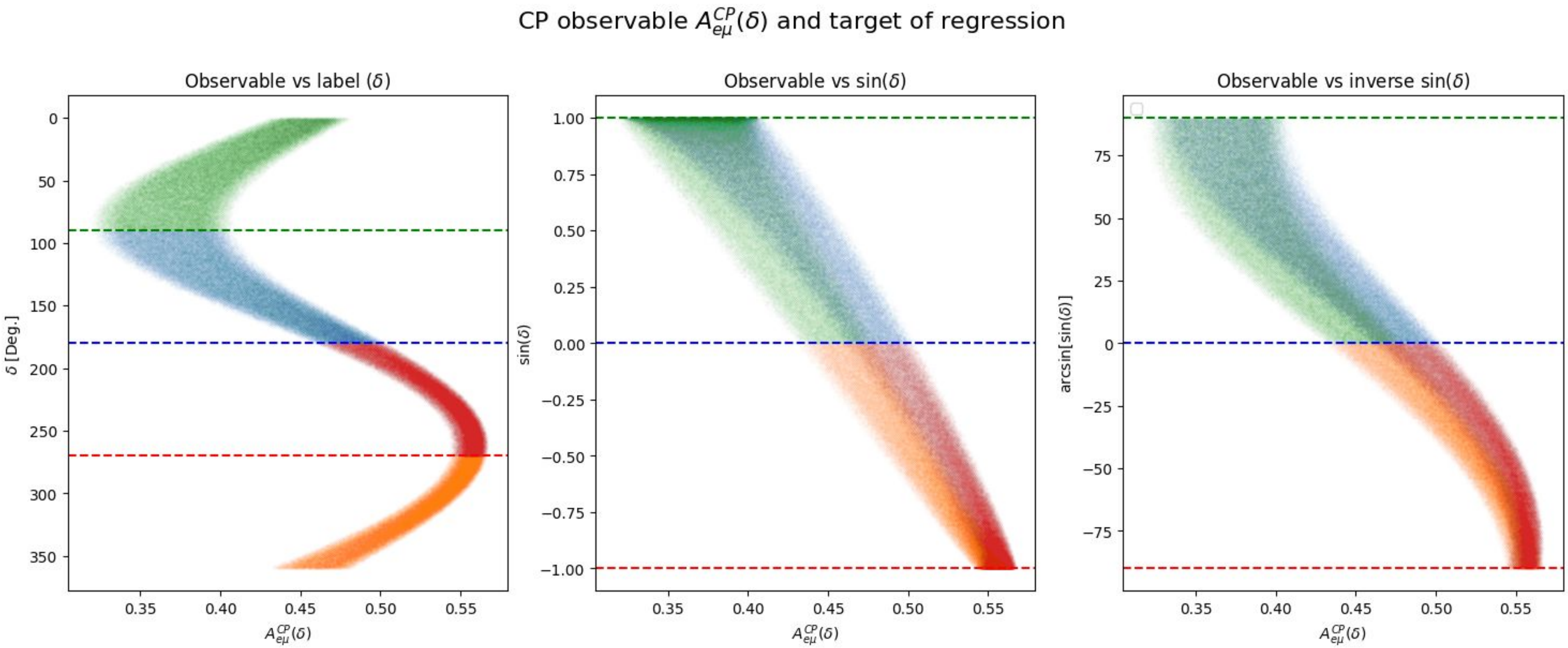




ν_e channel event count and target of regression

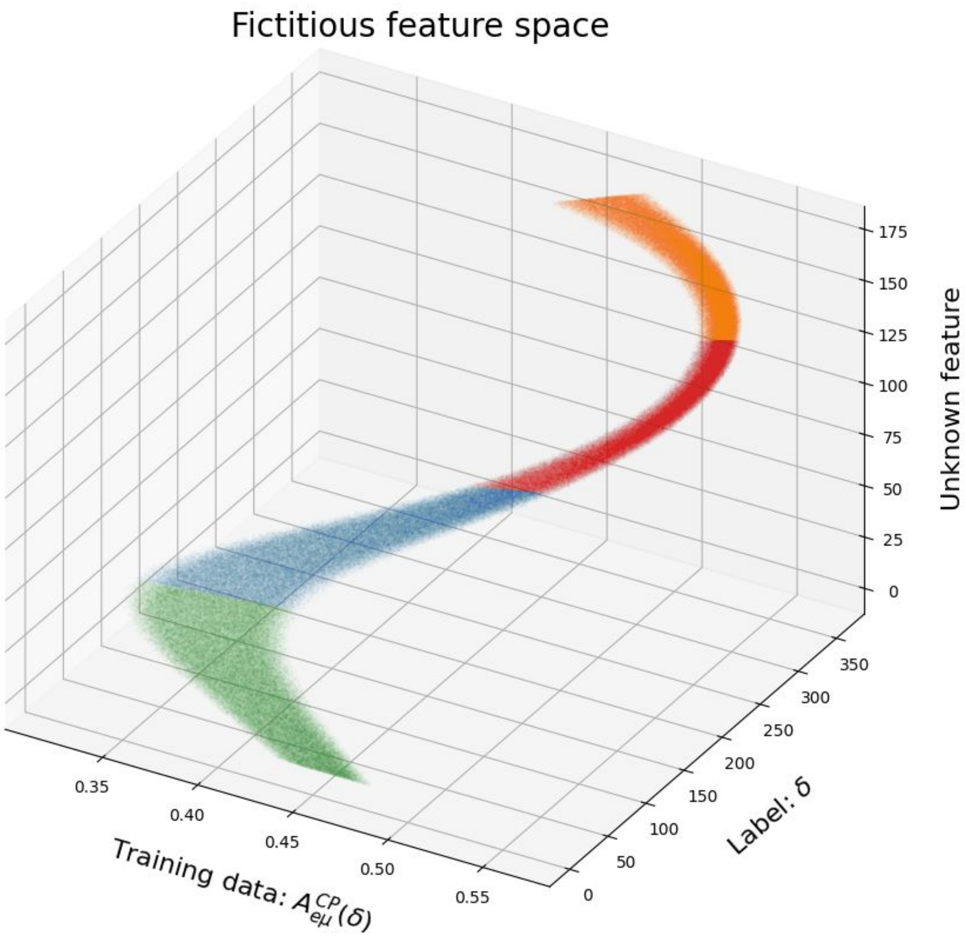


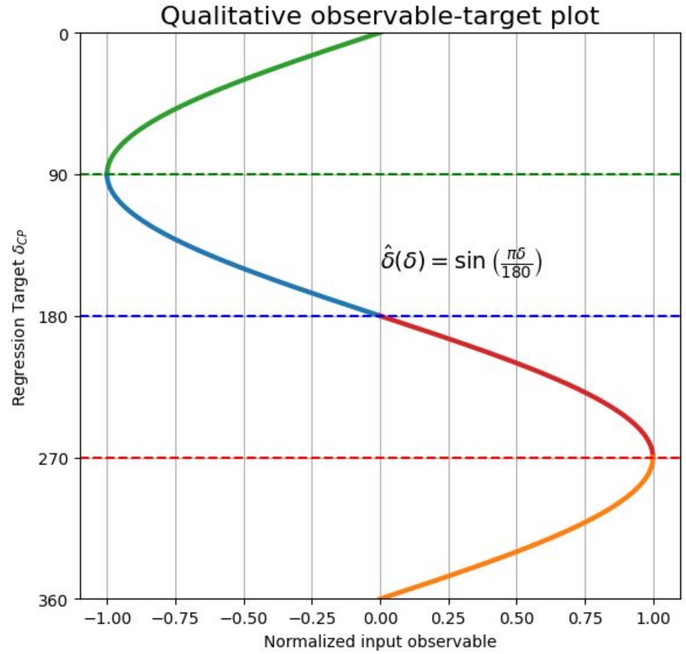
The CP phase Regression

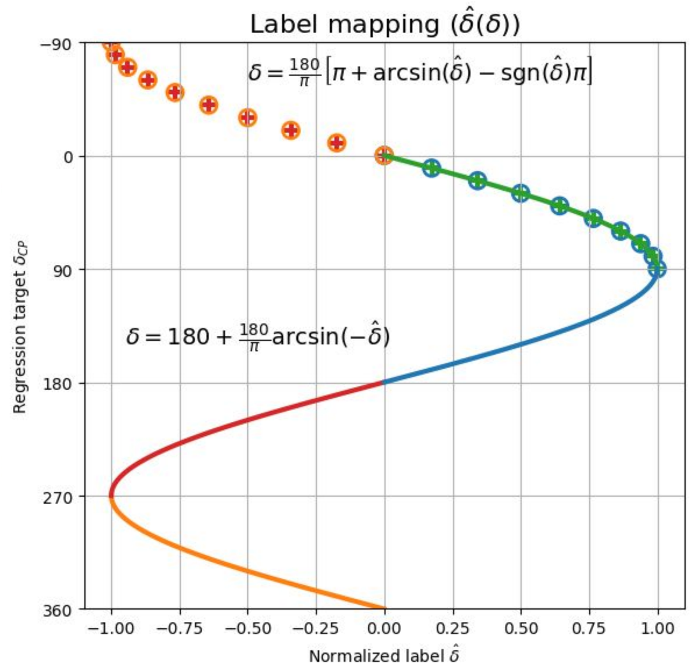


- By taking the sine value of the CP phase, we see a clear one-to-one relationship between the asymmetry observable and the sine value.
- However, transforming back using the arcsine function will only return values in the interval [-90,90].

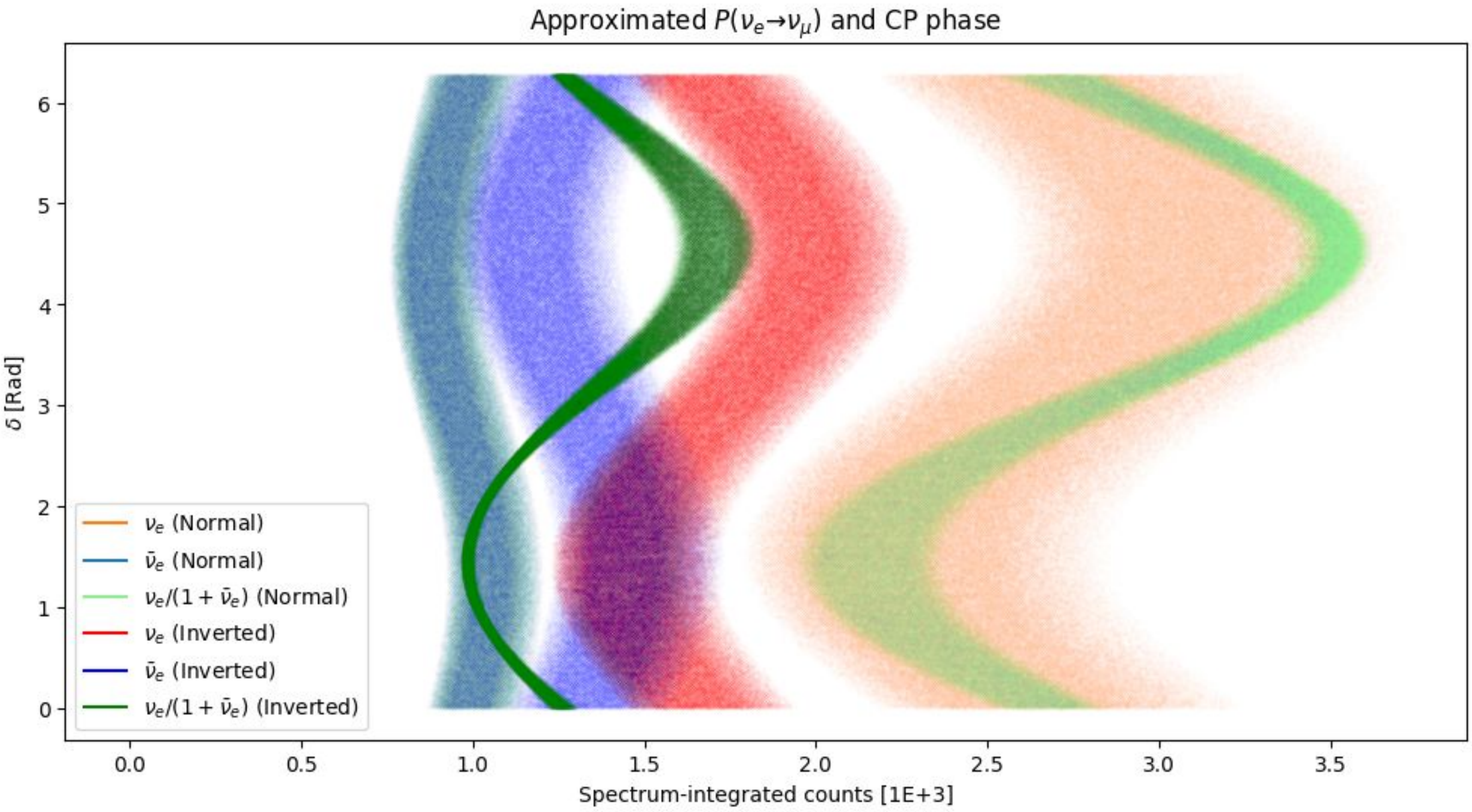
Idealized feature space







The CP phase Regression



Strategy in CP phase regression

There is more in these spectrum than meets the eye

Our proposed method is to separate this task into three parts:

- **Part 1: Regression**
  - The aim is to utilize the one-to-one relationship. From the plots, we are confident that the regression to the CP phase sine value should be quite accurate.
- **Part 2: Classification**
  - The domain of the sine value is between 1 and -1. Which separates the high and low end of the level of asymmetry.
  - What remains is to perform classification based on the hidden features that allows to separate [0,90] from [90,180], and [180,270] from [270,360]. The classification will help with restoring the actual degree from the arcsine function.

Model architecture

- **Mass hierarchy identification:** (takes one dimensional data)
  - SVM Binary classifier (using stochastic gradient descent)
- **θ_{23} Regression:** (takes multi-dimensional data)
  - A deep convolutional network followed by a multi-layer perceptron.
- **δ_{CP} Regressor → Classifier**
  - Regression: (Regression takes multi-dimensional data)
    - Deep convolutional network + multi-layer perceptron
  - Classification: (takes multi-dimensional data)
    - Deep convolutional network + multi-layer perceptron
- **δ_{CP} Classifier → Regressor (in development)**
  - Classification: (takes one dimensional data)
    - Boosted decision tree (using XGBoost)
  - Regression: (takes one dimensional data)
    - Single-layer perceptron

Model architecture

The current model performs regression before classification. The alternate method reverses the two stages.

- The Regressor → Classifier model
  - Relies on the arcsin function, which requires the regression estimate to be very precise and within the [-1,1] interval.
  - Only two binary classifier models required
- The Classifier → Regressor model (in development)
  - Performs a 4 quadrant classification, which is not easy to achieve
  - For each class (90 degree quadrants), the corresponding domain no longer overlaps, regression can be directly done on the CP phase in its original degree value (with suitable normalization)
  - Requires one regressor for each quadrant, though lightweight in comparison, requires individual tuning, which risks overfitting.

BACK UP

Code

General backup repo in GitHub:
git@github.com:gary20000915/Neutrino-Oscillation-DUNE.git

Submission backup in Kaggle:
https://www.kaggle.com/competitions/phys591000-2023-final-project-i

The physics of antineutrinos in DUNE and determination of octant and δ_{CP}

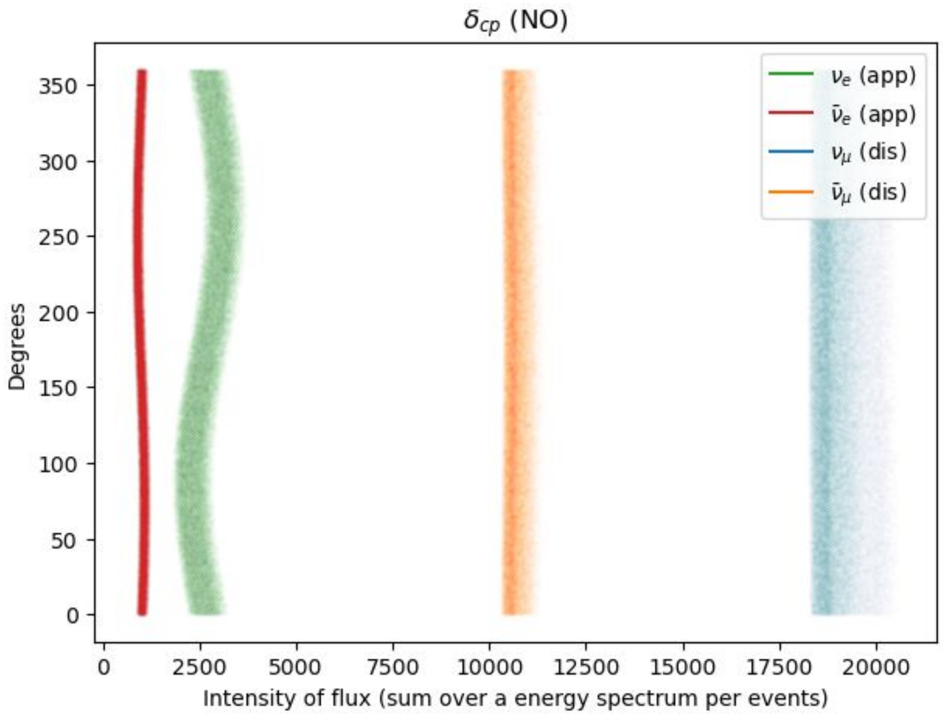
cos(π/2 + δ) = -sin(δ), so this formula is also proportional to sin(δ) if in the vacuum ideal situations. (ref page 386, 387)

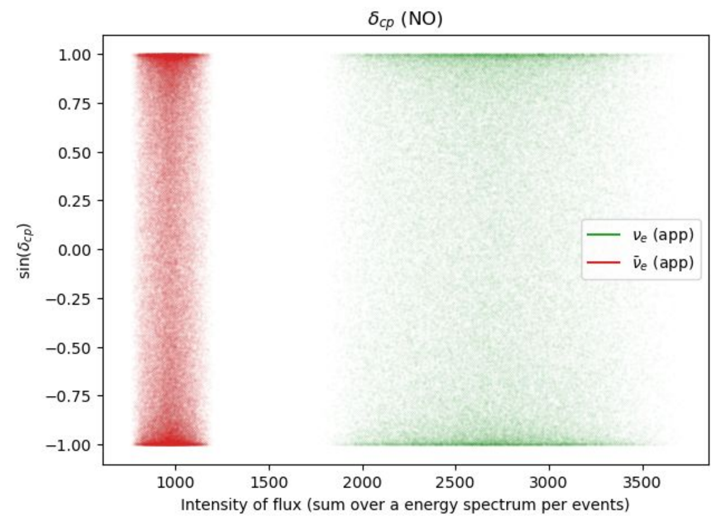
$$P_{\mu e} = \underbrace{4 s_{13}^2 s_{23}^2 \frac{\sin^2(A-1)\Delta}{(A-1)^2}}_{\mathcal{O}_o} + \underbrace{\alpha^2 \cos^2\theta_{23} \sin^2 2\theta_{12} \frac{\sin^2 A\Delta}{A^2}}_{\mathcal{O}_2} + \underbrace{\alpha s_{13} \sin 2\theta_{12} \sin 2\theta_{23} \cos(\Delta + \delta_{cp}) \frac{\sin(A-1)\Delta}{(A-1)} \frac{\sin A\Delta}{A}}_{\mathcal{O}_1}$$

$$P_{\mu\mu} = 1 - \sin^2 2\theta_{23} \sin^2 \Delta + \mathcal{O}(\alpha, s_{13})$$

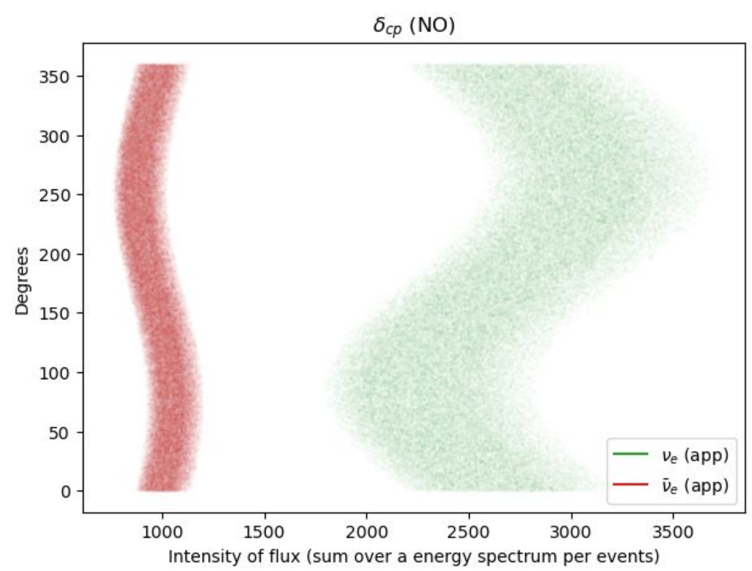
terms are α and α^2 suppressed respectively.

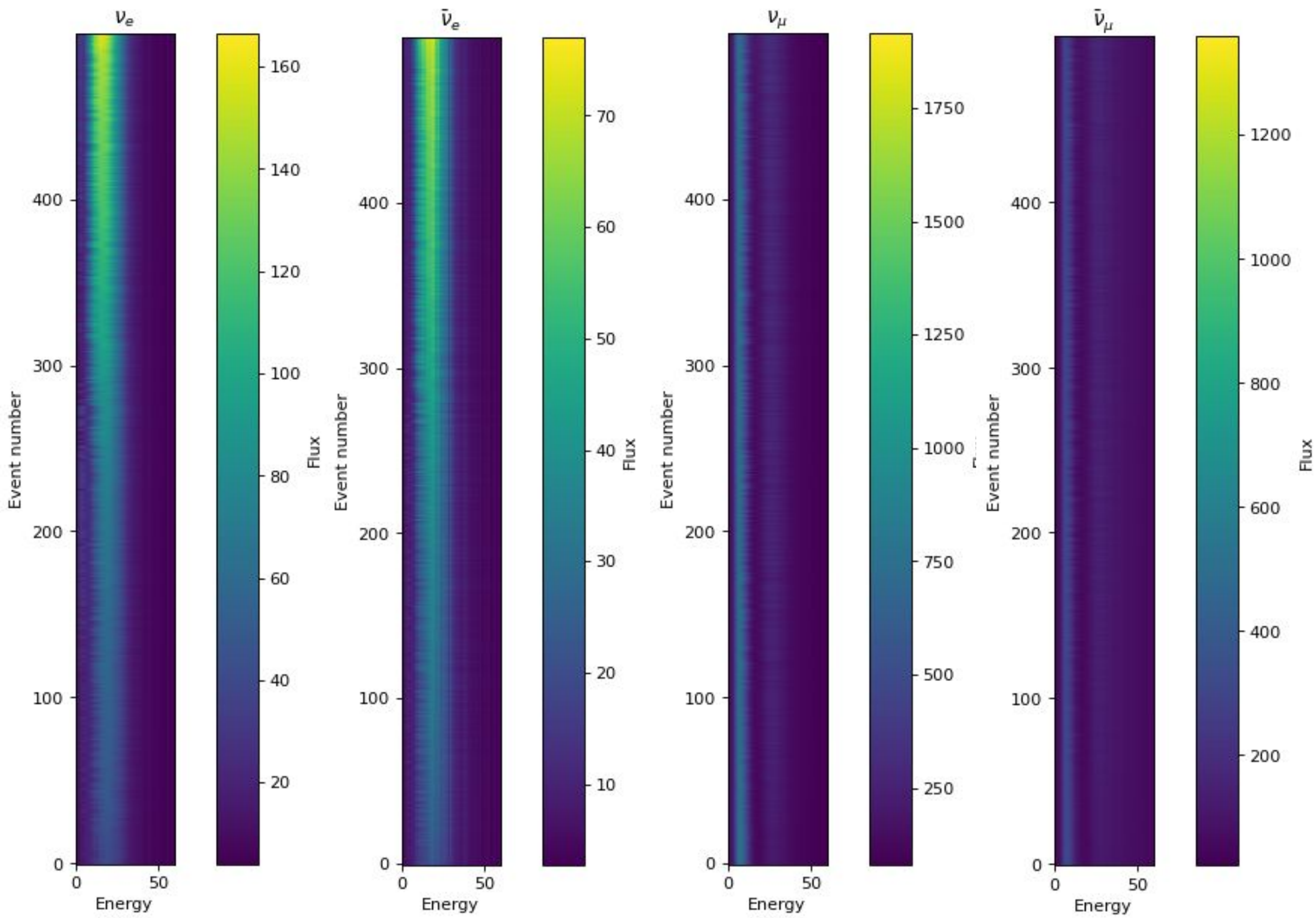
Note that for vacuum oscillation maxima, Δ corresponds to 90°. Thus in the appearance channel probability (cf. Eq. (2)), $\delta_{CP} = -90°(+90°)$ correspond to maximum (minimum) point in the probability for neutrinos. For antineutrinos it is the opposite. Thus, for these values of δ_{CP}, octant sensitivity is expected to be maximum if there is no degeneracy. Note that with the inclusion of matter effect, the appearance channel probability maxima does not coincide with the vacuum maxima and in that case the maximum and minimum points in the probability do not come exactly at ±90° but gets slightly shifted. This can be seen from Fig. 1. However for illustration, we will take $\delta_{CP} = \pm 90°$ as the reference points to describe the physics of octant in DUNE.





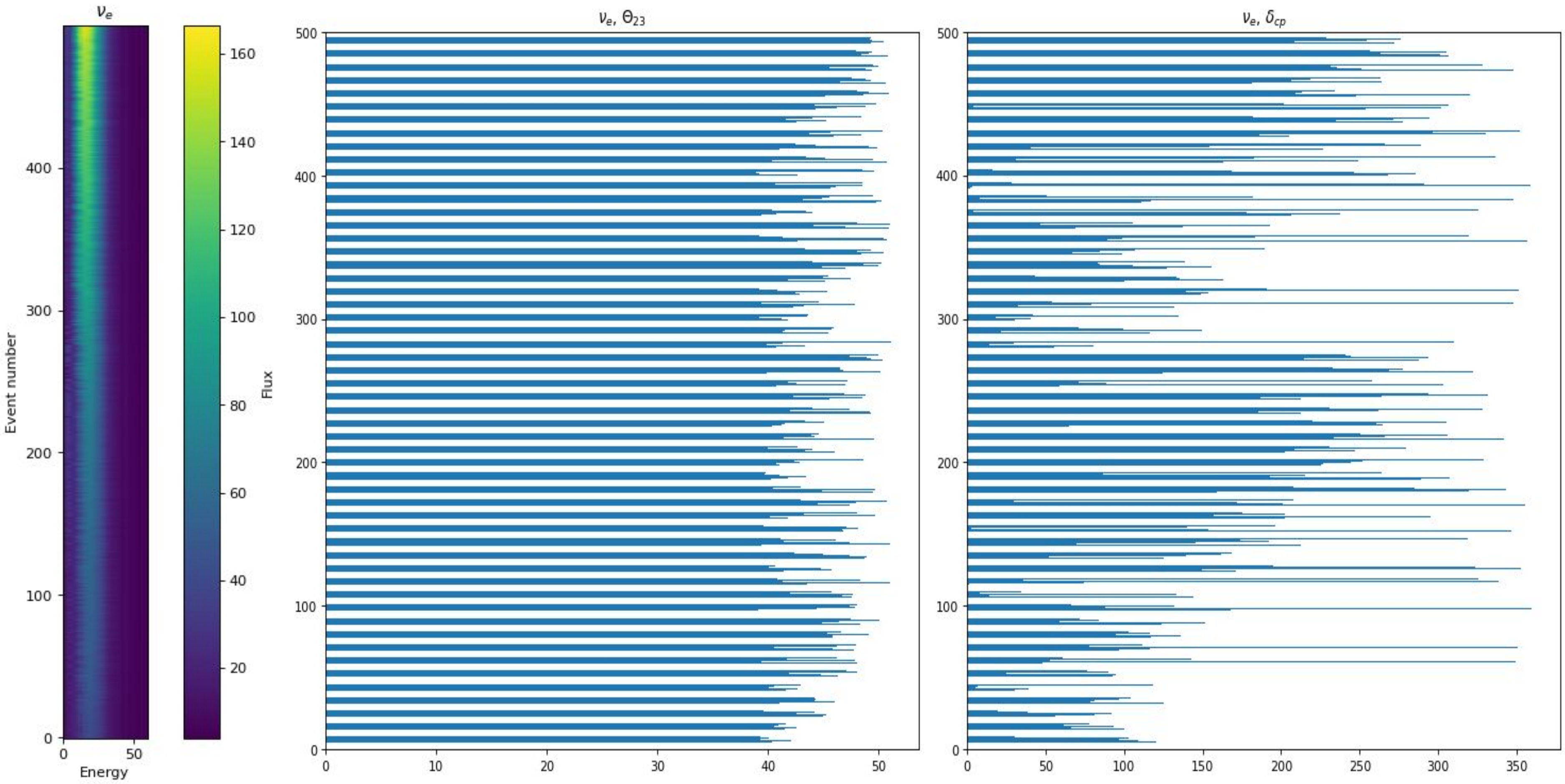
The monotonic relationship is hard to observe with only the channel event count and the sine of CP phase.

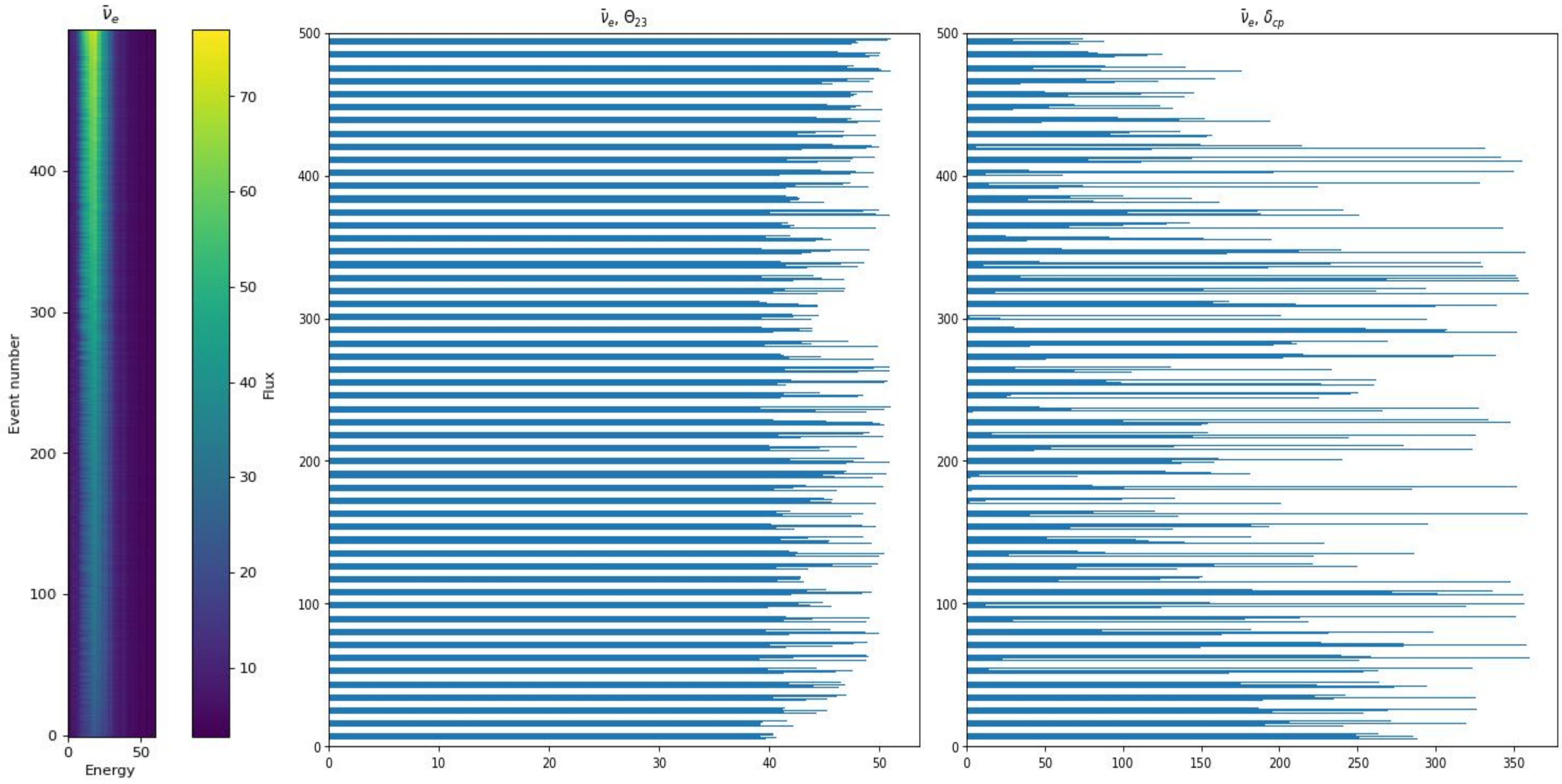




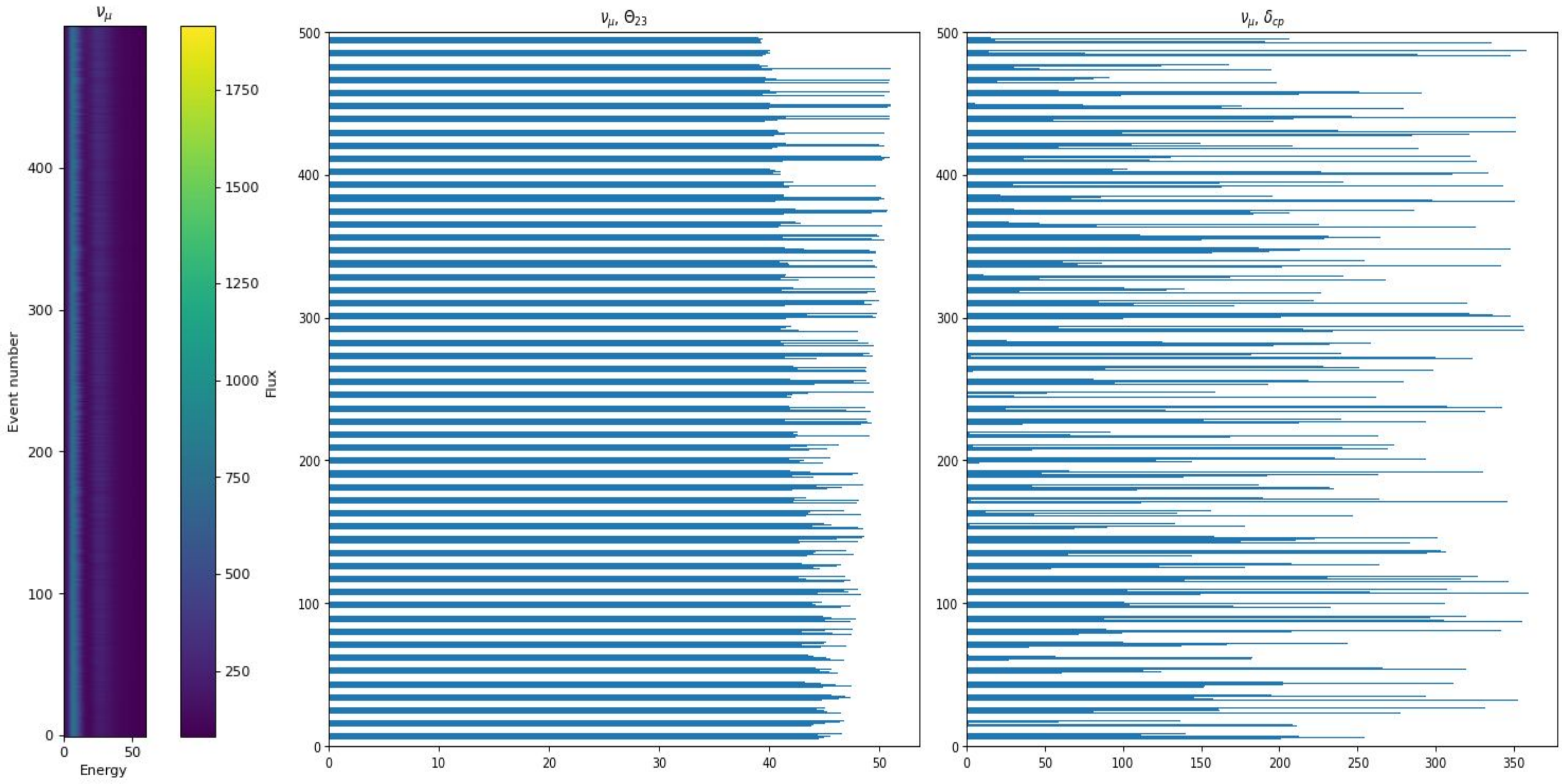
ν_e
$\bar{\nu}_e$
ν_μ
$\bar{\nu}_\mu$
Event number
Energy
Flux

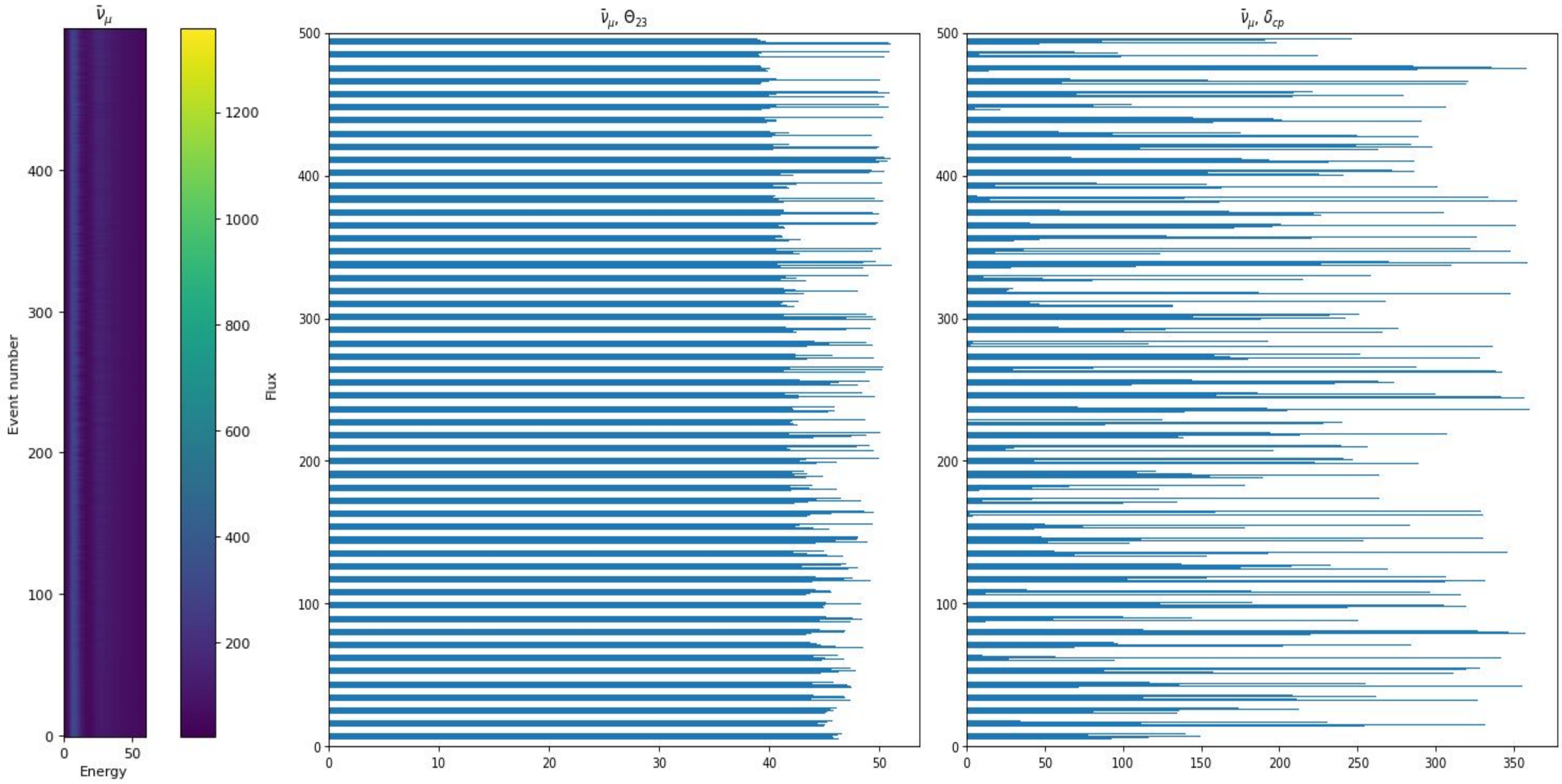
The results are sort by the sum of the specific channel.
(In this page, the y-axis is sorted by the sum of “event number ” by ve channel)





$\bar{\nu}_e$
Event number
Energy
Flux
$\bar{\nu}_e$, Θ_{23}
$\bar{\nu}_e$, δ_{cp}


ν_μ
Event number
Energy
Flux
ν_μ, Θ_{23}
ν_μ, δ_{cp}


$\bar{\nu}_\mu$
Event number
Energy
Flux
$\bar{\nu}_\mu$, Θ_{23}
$\bar{\nu}_\mu$, δ_{cp}