A NORMALISING FLOW BASED BAYESIAN OSCILLATION ANALYSIS PIPELINE

JUNE 19TH 2026

ANDREW ATTA, ERIC THRANE, PHILIP URQUIJO

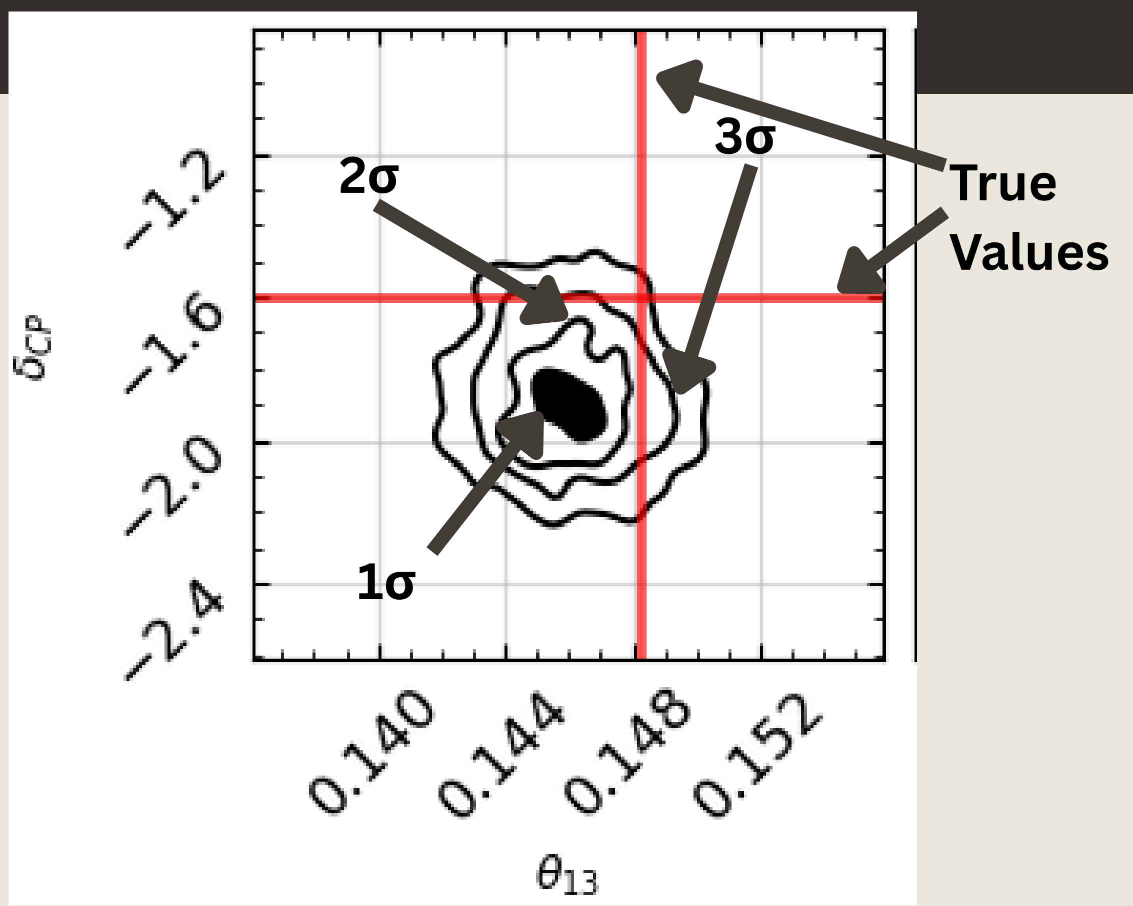
BAYESIAN ANALYSIS.

- This corner plot is just a visualisation of this infromation:

$$P(\theta_{13}, \delta_{CP} \mid data) \propto \mathcal{L}(data \mid \theta_{13}, \delta_{CP})\pi(\theta_{13}, \delta_{CP})$$

Posterior **Likelihood** **Prior**

Posterior



- θ_{13} and δ_{CP} determines the distribution of ***neutrino energies*** but our ***data*** is not the neutrino energies
- the data is *some measured quantity that relates to the neutrino energy*
- So our likelihood of seeing the data given some θ_{13} and δ_{CP} decomposes like so:

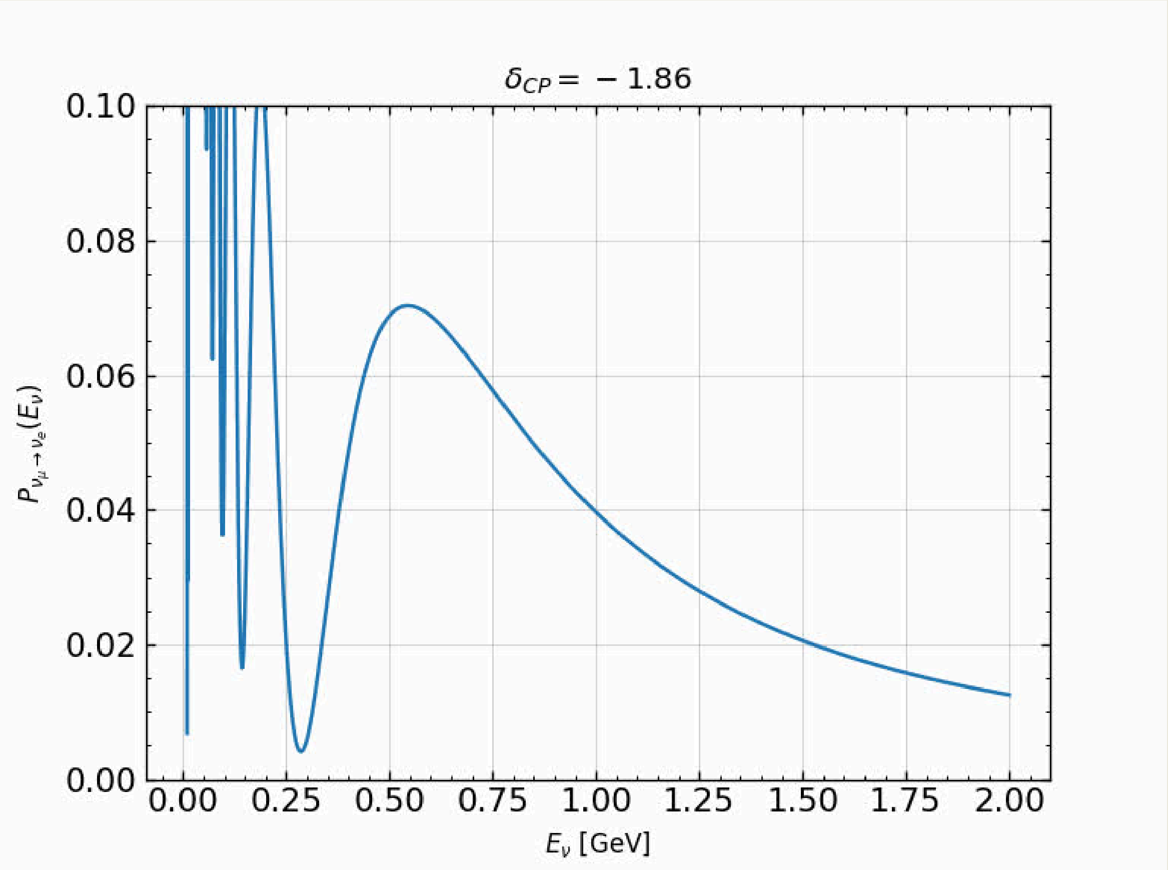
$$P(\theta_{13}, \delta_{CP} \mid data) \propto \int \mathcal{L}(data \mid E_\nu)\pi(E_\nu \mid \theta_{13}, \delta_{CP})dE_\nu \cancel{\pi(\theta_{13}, \delta_{CP})}$$

constant

Detector Model

Oscillation + Flux + cross section physics

Uniform Prior

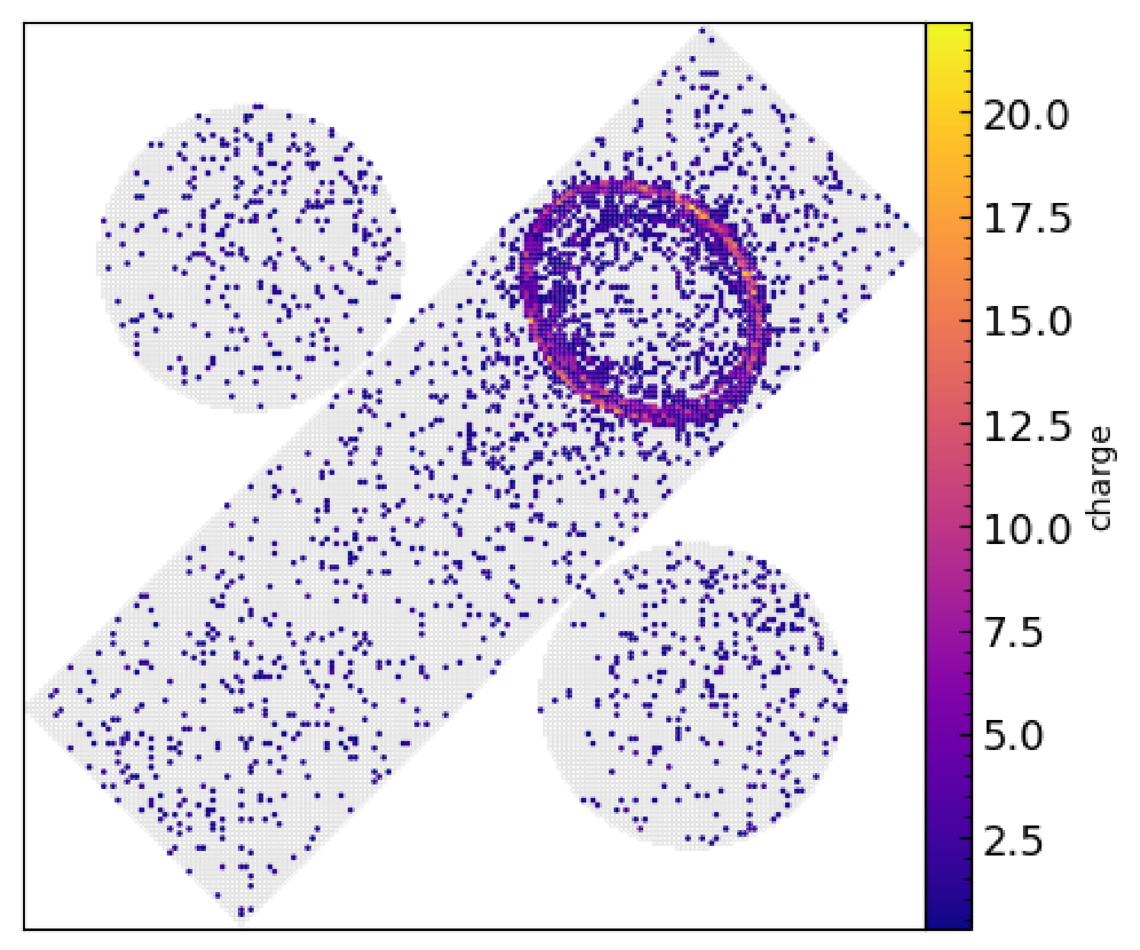


THE DATA

- The raw data in this case are the PMT hit charges and times – **up to** 40,000 dimensions at FD
- mathematically difficult to construct a probability density for all 40,000 data points given some true energy.

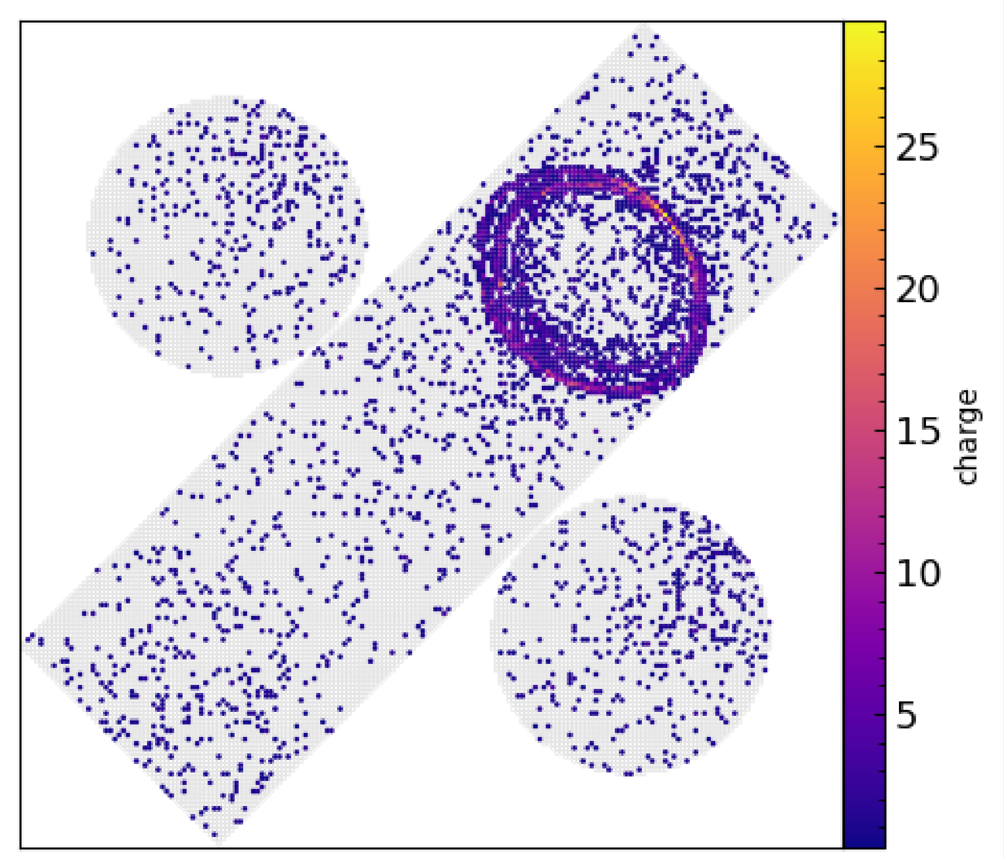
$$\mathcal{L}(data \mid E_\nu) \to \mathcal{L}(\mathbb{R}^{\leq 40{,}000} \mid E_\nu)$$

- What is the probability distribution of these two data realisations of the same true energy?
- How do we compare them? Which is more likely?
- How do we even begin to construct a likelihood to describe this?



$$\mathcal{L}(data_1 \mid E_\nu = 600\text{MeV})$$

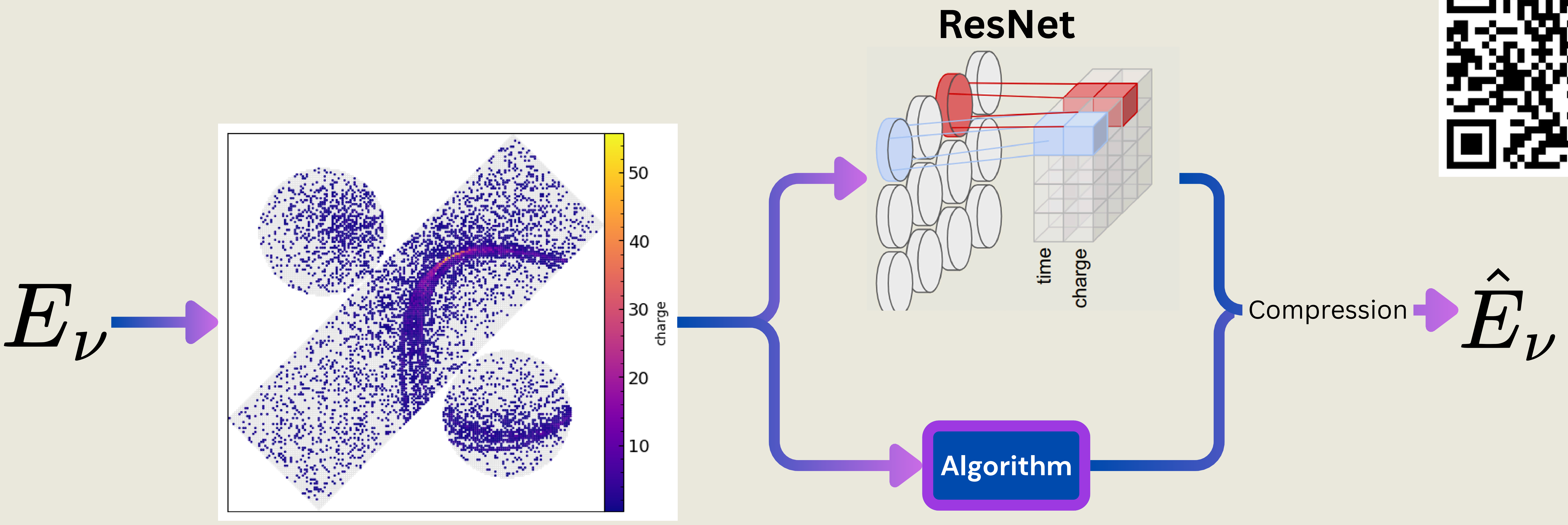
$$\mathcal{L}(data_2 \mid E_\nu = 600\text{MeV})$$



RECONSTRUCTION AS DIMENSIONAL REDUCTION

- ResNet (and fiTQun) act to dimensionally reduce this data from ≤40,000 down to just a few summary statistics (i.e. the direction, energy, position, pid).

ResNet reconstruction pre-print

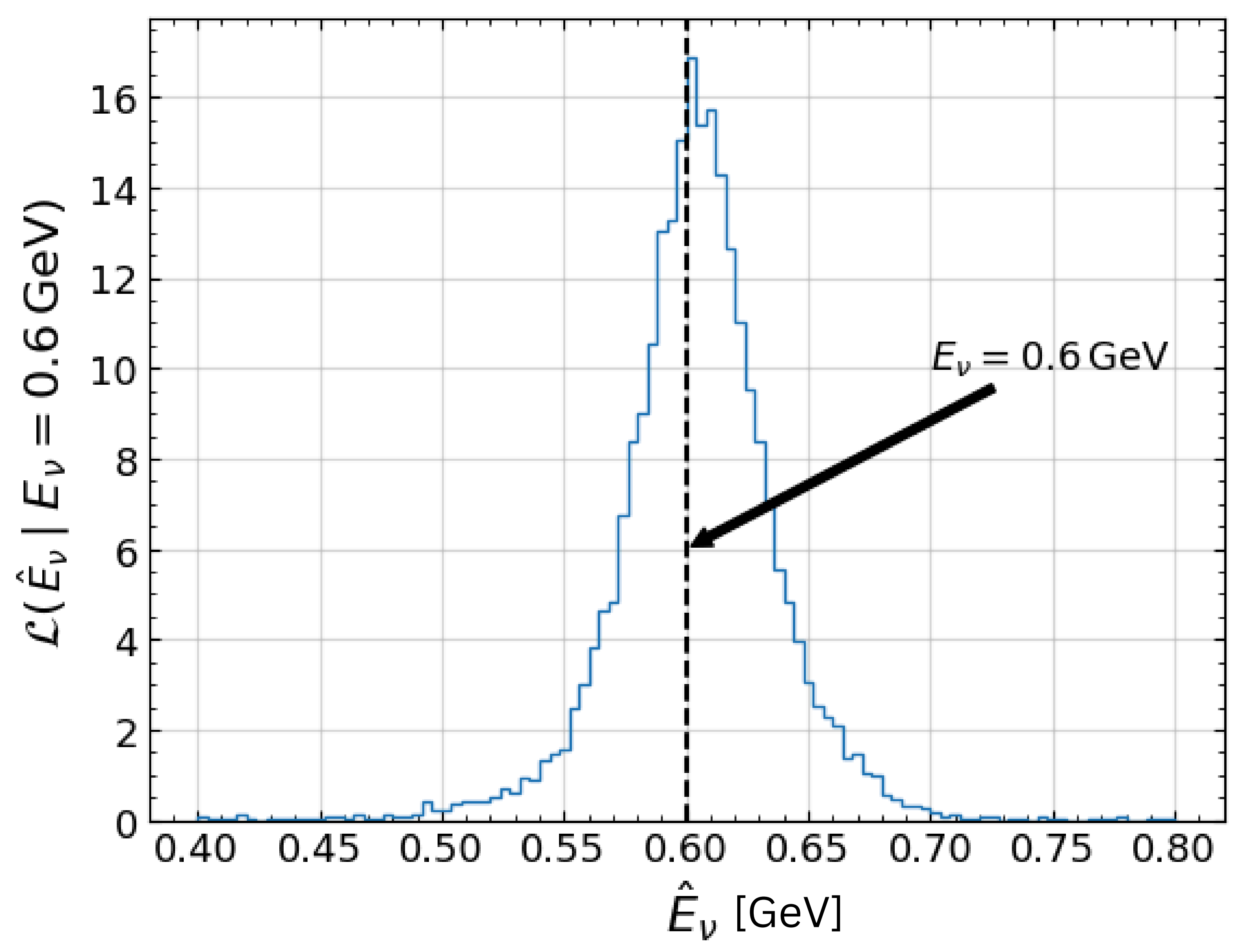


THE VARIANCE IN THE DATA.

$E_\nu = 300\ \mathrm{MeV}$

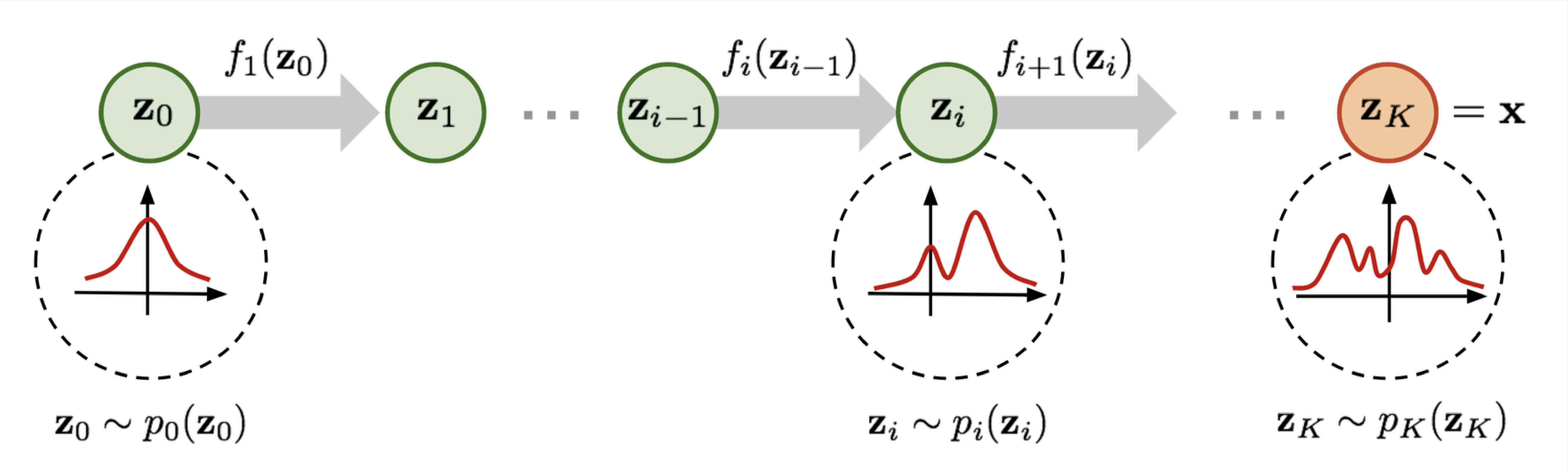
ESTIMATING $\mathcal{L}(\hat{E}_\nu \mid E_\nu)$

- By compressing the data down into a few summary statistics we make the likelihood much simpler while still capturing most of the information.
  $\mathcal{L}(\mathbb{R}^{\leq 40{,}000} \mid E_\nu) \rightarrow \mathcal{L}(\hat{E}_\nu \mid E_\nu)$
- Importantly, this makes building the likelihood empirically significantly more feasible!
- Well within the domain of what simulation based inference can do
- One approach we could take is using a KDE to estimate $\mathcal{L}(\hat{E}_\nu \mid E_\nu)$
- KDEs scale extremely poorly with dimensionality and data [$O(n^3)$] - curse of dimensionality :(
- What if we want a distribution over more dimensions? Eg $\mathcal{L}(E_\ell, \theta_\ell, \phi_\ell, x, y, z, PID \mid E_\nu)$



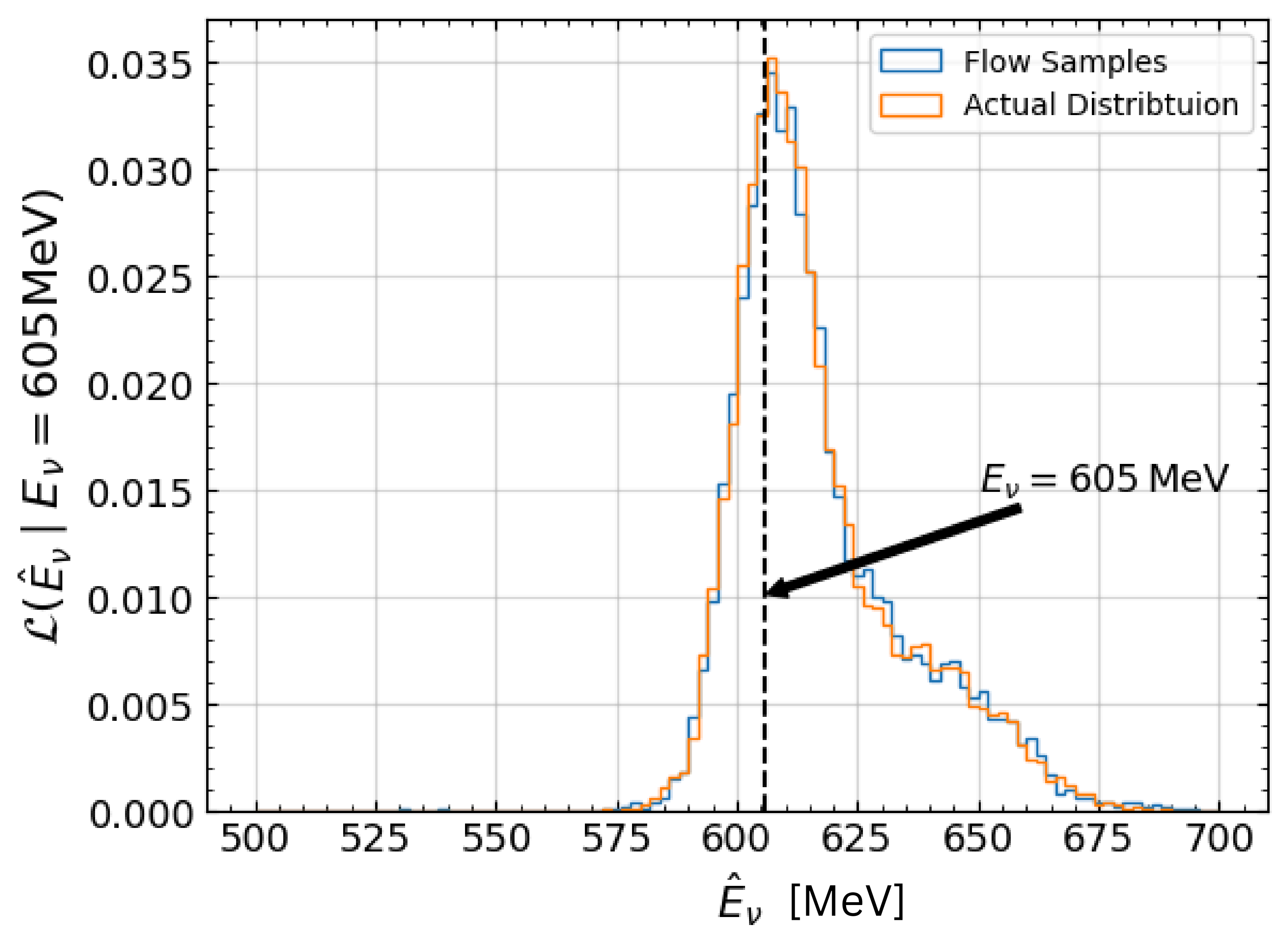
NORMALISING FLOWS.

- A ***normalising flow*** can easily handle these higher dimensions
- A normalising flow is a type of neural network that attempts to learn a series of bijective *transformations* which allows you to transform some base distribution (often a Gaussian) into a more complicated distribution (eg $\mathcal{L}(\hat{E}_\nu \mid E_\nu)$ or $\mathcal{L}(E_\ell, \theta_\ell, \phi_\ell, x, y, z \mid E_\nu)$).
- “Neural simulation based inference” - Used ubiquitously throughout data inference world



FLOW AS A DENSITY ESTIMATOR.

- The normalising flow allows us to estimate the probability density of $\mathcal{L}(\hat{E}_\nu \mid E_\nu)$
- Trained a flow on pairs of $(\hat{E}_\nu, E_\nu)$ such that the flow learned $\mathcal{L}(\hat{E}_\nu \mid E_\nu)$
- Importantly, this gives you an ***unbinned event-by-event*** likelihood we can evaluate to ultimately sample the posterior
- **All information conserving past the initial compression!**
- Doesn't even matter if ResNet or fiTQun is right or unbiased! Likelihood will just broaden or shift



Detector Model (flow)

Oscillation + Flux + cross section physics

$$P(\theta_{13}, \delta_{CP} \mid \hat{E}_\nu) \propto \int \mathcal{L}(\hat{E}_\nu \mid E_\nu)\pi(E_\nu \mid \theta_{13}, \delta_{CP})dE_\nu$$

- **Note: We actually use $\mathcal{L}(\hat{E}_\nu, \mathrm{towall} \mid \mathrm{E}_\nu)$ but for now we will keep it simple as 1D**

WHY DO THIS?

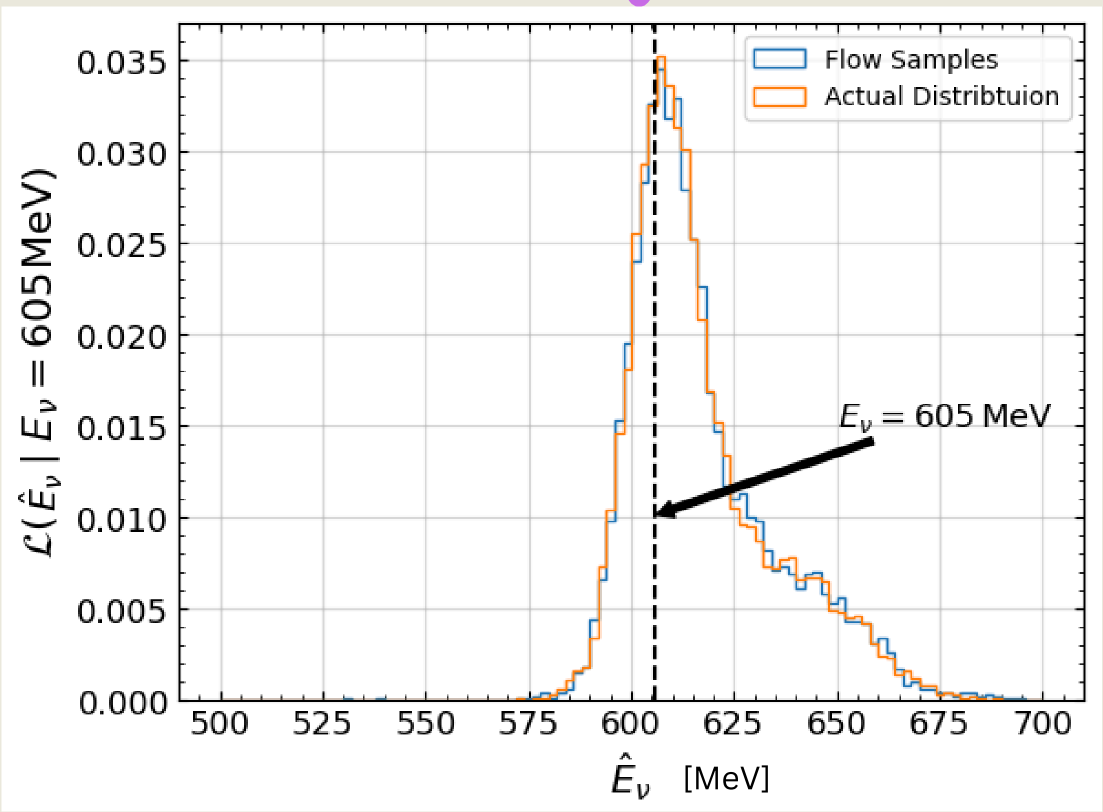
Each event has a measured energy and uncertainty

COMING BACK TO THE GOAL.

Detector Model (flow)

Oscillation + Flux + cross section physics

$$P(\theta_{13}, \delta_{CP} \mid \hat{E}_\nu) \propto \int \mathcal{L}(\hat{E}_\nu \mid E_\nu)\pi(E_\nu \mid \theta_{13}, \delta_{CP})dE_\nu$$



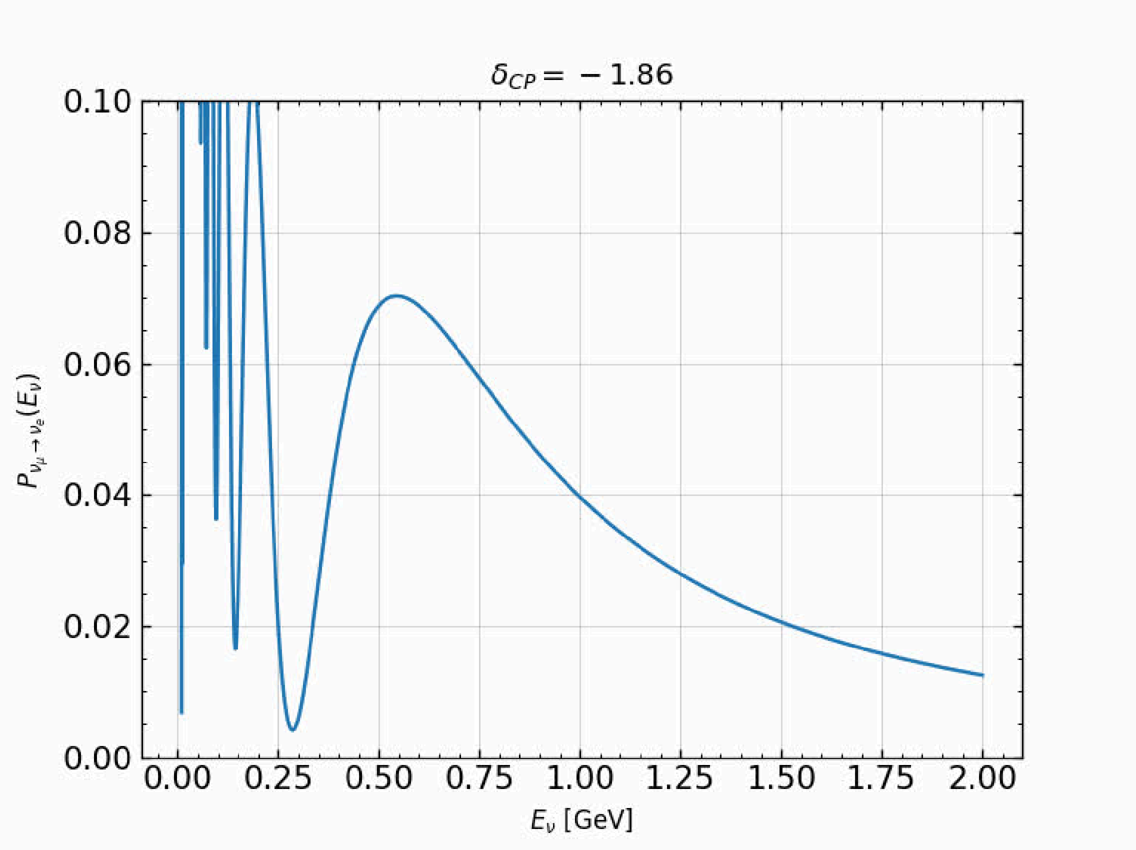
?

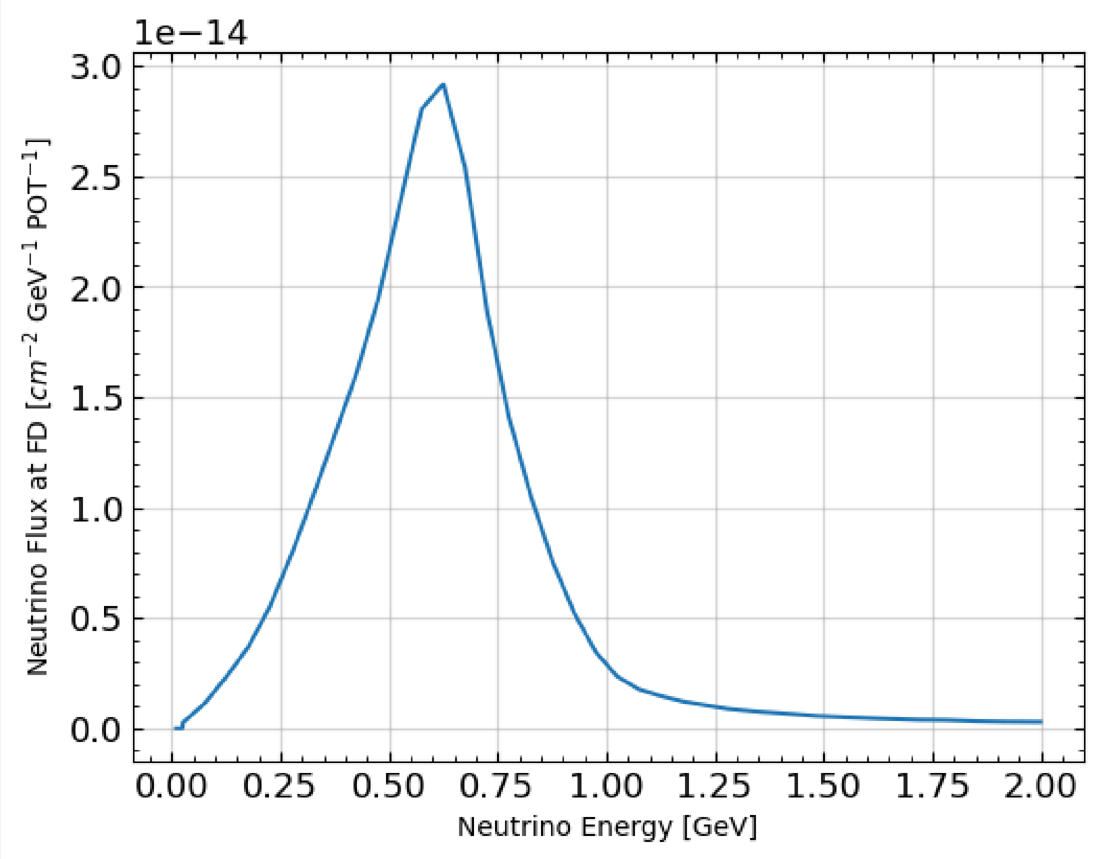
THE PHYSICS PRIOR.

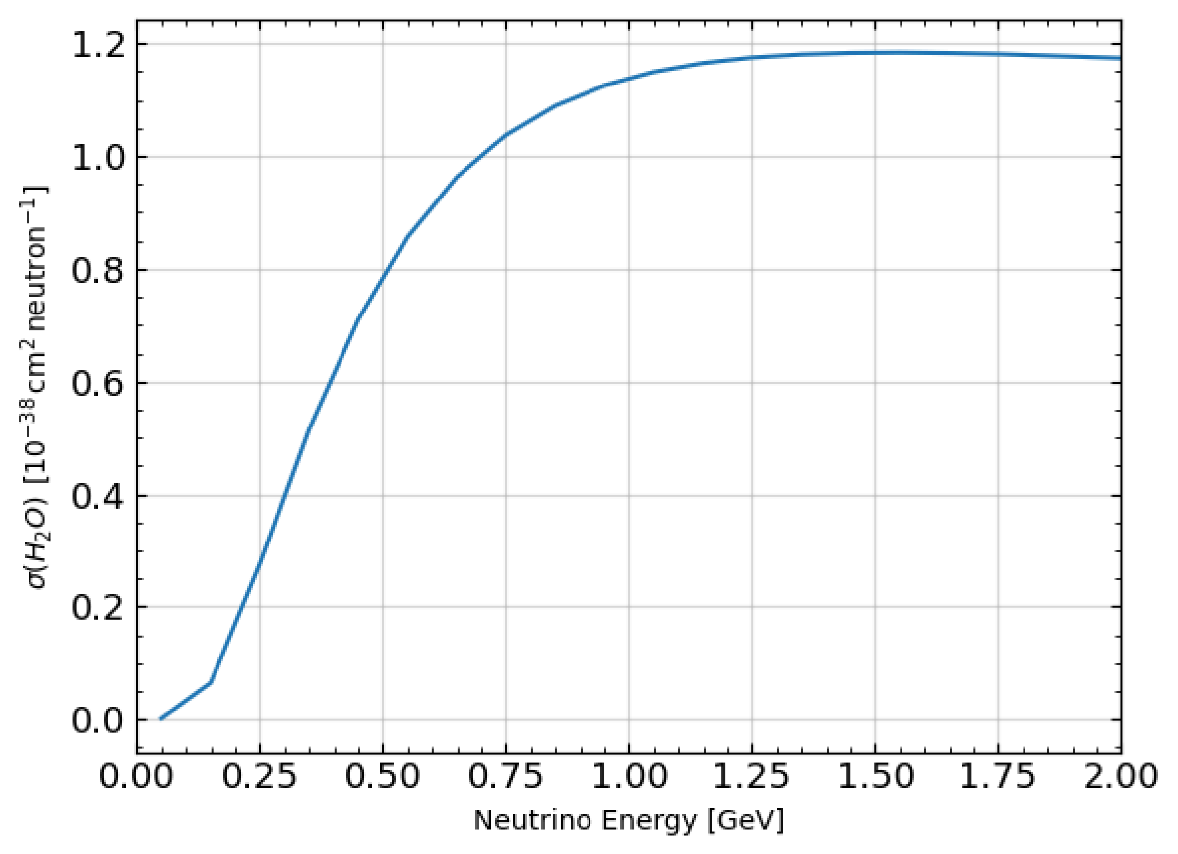
0.6??

$5\text{x}10^{34}$

$$\pi(E_\nu \mid \theta_{13}, \delta_{CP}) \propto P_{\nu_\mu \to \nu_e}(E_\nu \mid \theta_{13}, \delta_{CP}) \times \phi(E_\nu) \times \epsilon(E_\nu) \times \sigma(E_\nu) \times N_{\text{targets}}$$

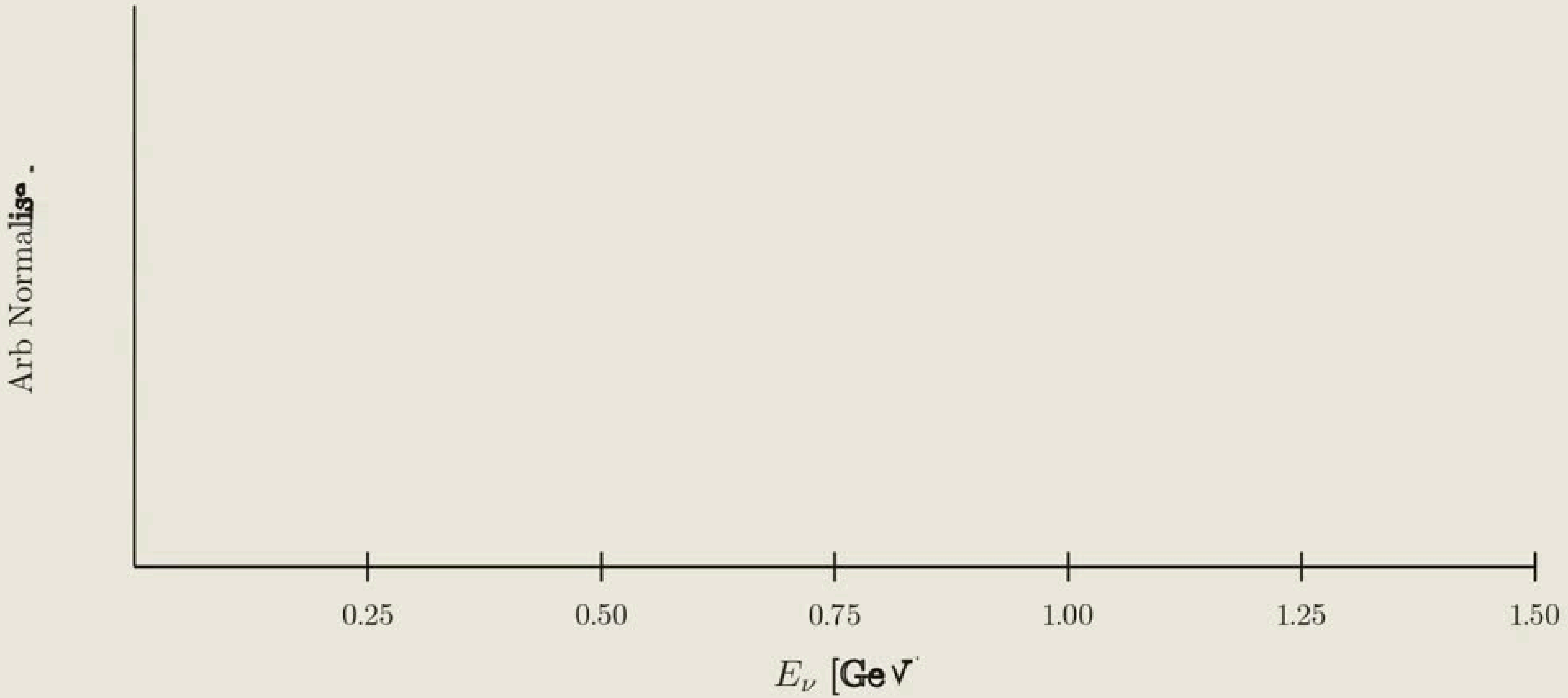






Toy Data - Not Official

THE PHYSICS PRIOR.



USING ALL THE DATA

$$\underbrace{P(\theta_{13}, \delta_{CP} \mid \hat{E}_\nu)}_{\text{Posterior}} \propto \int \underbrace{\mathcal{L}(\hat{E}_\nu \mid E_\nu)}_{\text{Detector Model (flow)}} \underbrace{\pi(E_\nu \mid \theta_{13}, \delta_{CP})}_{\text{Oscillation + Flux + cross section physics}} \, dE_\nu$$

- Above expression is for a ***single*** accelerator neutrino event
- Including all available accelerator neutrino data would result in a produt over the marginalised likelihoods

$$P(\theta_{13}, \delta_{CP} \mid \{\hat{E}_\nu^i\}) \propto \prod_{i=1}^{N_{\text{obs}}} \int \mathcal{L}(\hat{E}_\nu^i \mid E_\nu) \, \pi(E_\nu \mid \theta_{13}, \delta_{CP}) \, dE_\nu$$

INCLUDING RATE INFO

Posterior — Detector Model (flow) — Oscillation + Flux + cross section physics

$$P(\theta_{13}, \delta_{CP} \mid \{\hat{E}_\nu^i\}) \propto \prod_{i=1}^{N_{\text{obs}}} \int \mathcal{L}(\hat{E}_\nu^i \mid E_\nu)\, \pi(E_\nu \mid \theta_{13},\, \delta_{CP})\, dE_\nu$$

- This expression will tell you how the likely the *shape* of the observed neutrino energy distribution is for given oscillation parameters.
- We also want to include information about how the *expected number* of neutrinos differs from the *observed number.*

$$P(\theta_{13}, \delta_{CP} \mid \{\hat{E}_\nu^i\}, N_{\text{obs}}) \propto \frac{N_{\text{exp}}^{N_{\text{obs}}} e^{N_{\text{exp}}}}{N_{\text{obs}}!} \prod_{i=1}^{N_{\text{obs}}} \int \mathcal{L}(\hat{E}_\nu^i \mid E_\nu)\, \pi(E_\nu \mid \theta_{13},\, \delta_{CP})\, dE_\nu$$

ADDING ANTINEUTRINO INFO

$$\underbrace{P(\theta_{13}, \delta_{CP} \mid \{\hat{E}_\nu^i\}, N_{\text{obs}})}_{\text{Posterior}} \propto \underbrace{\frac{N_{\text{exp}}^{N_{\text{obs}}} e^{N_{\text{exp}}}}{N_{\text{obs}}!}}_{\text{Count Information}} \prod_{i=1}^{N_{\text{obs}}} \int \underbrace{\mathcal{L}(\hat{E}_\nu^i \mid E_\nu)}_{\text{Detector Model (flow)}} \underbrace{\pi(E_\nu \mid \theta_{13}, \delta_{CP})}_{\text{Oscillation + Flux + cross section physics}} \, dE_\nu$$

- Information about the total number of neutrinos by itself is not informative about oscillation parameters
- The inclusion of Anti-nuetrino information is where information lies

$$P(\theta_{13}, \delta_{CP} \mid \{\hat{E}_\nu^i\}, \{\hat{E}_{\bar{\nu}}^i\}, N_{\text{obs}}, \bar{N}_{\text{obs}}) \propto \underbrace{\frac{N_{\text{exp}}^{N_{\text{obs}}} e^{N_{\text{exp}}}}{N_{\text{obs}}!} \prod_{i=1}^{N_{\text{obs}}} \int \mathcal{L}(\hat{E}_\nu^i \mid E_\nu) \, \pi(E_\nu \mid \theta_{13}, \delta_{CP}) \, dE_\nu}_{\text{Neutrino Info}}$$

$$\times \underbrace{\frac{\bar{N}_{\text{exp}}^{\bar{N}_{\text{obs}}} e^{\bar{N}_{\text{exp}}}}{\bar{N}_{\text{obs}}!} \prod_{i=1}^{\bar{N}_{\text{obs}}} \int \mathcal{L}(\hat{E}_{\bar{\nu}}^i \mid E_{\bar{\nu}}) \, \pi(E_{\bar{\nu}} \mid \theta_{13}, \delta_{CP}) \, dE_{\bar{\nu}}}_{\text{Anti-Neutrino Info}}$$

THE FINAL POSTERIOR

Posterior | Count Information | Detector Model (flow) | Oscillation + Flux + cross section physics

$$P(\theta_{13}, \delta_{CP} \mid \{\hat{E}^i_\nu\}, \{\hat{E}^i_{\bar{\nu}}\}, N_{\text{obs}}, \bar{N}_{\text{obs}}) \propto \frac{N_{\text{exp}}^{N_{\text{obs}}} e^{N_{\text{exp}}}}{N_{\text{obs}}!} \prod_{i=1}^{N_{\text{obs}}} \int \mathcal{L}(\hat{E}^i_\nu \mid E_\nu)\, \pi(E_\nu \mid \theta_{13}, \delta_{CP})\, dE_\nu$$

Neutrino Info

Anti-Neutrino Info

$$\times \frac{\bar{N}_{\text{exp}}^{\bar{N}_{\text{obs}}} e^{\bar{N}_{\text{exp}}}}{\bar{N}_{\text{obs}}!} \prod_{i=1}^{\bar{N}_{\text{obs}}} \int \mathcal{L}(\hat{E}^i_{\bar{\nu}} \mid E_{\bar{\nu}})\, \pi(E_{\bar{\nu}} \mid \theta_{13}, \delta_{CP})\, dE_{\bar{\nu}}\, \pi(\theta_{13}, \delta_{CP})$$

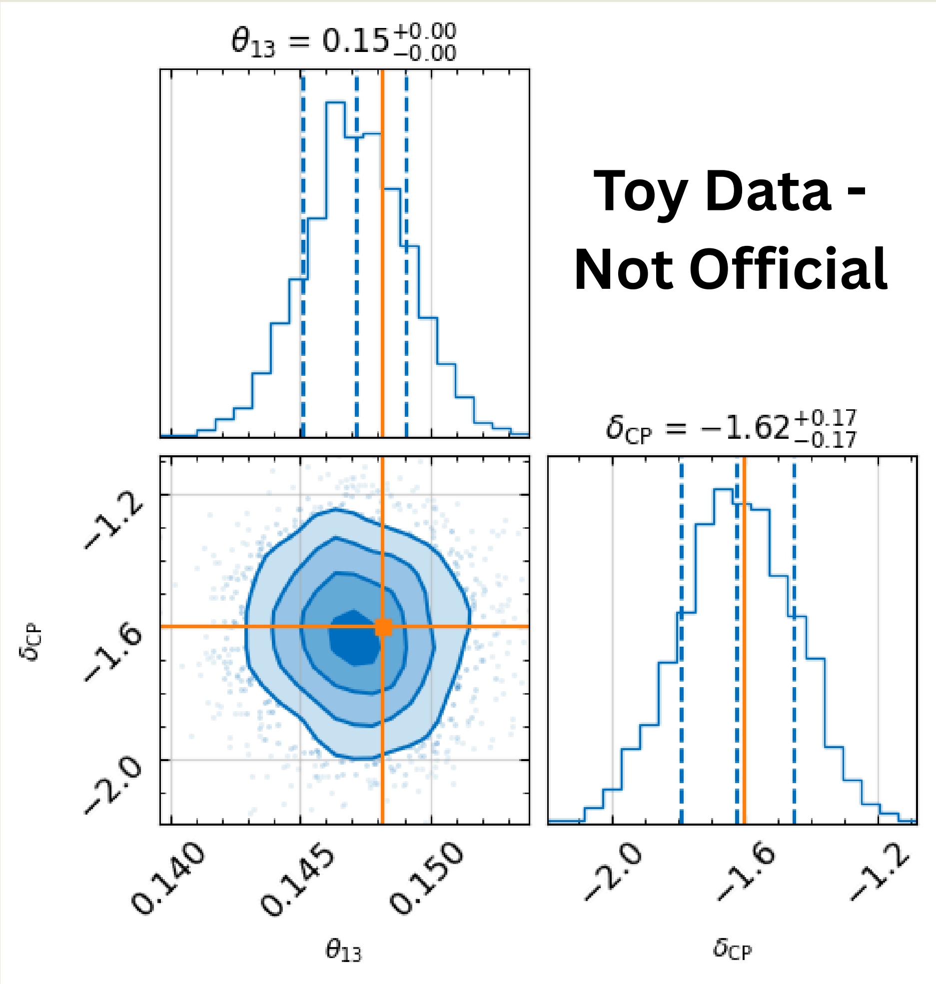
Marginalised Likelihood | Prior

- Can continue to play this game by adding more terms for more physics eg ν_μ survival, interaction modes, beam contaminants, **more oscillation parameters**, etc but we stopped here for this analysis.
- From this construction we can sample the posterior by evaluating the marginal likelihood at different values of θ_{13} and δ_{CP} under some prior belief about them (uniform in both)
- We use nested sampling to do so, but any sampler is fine

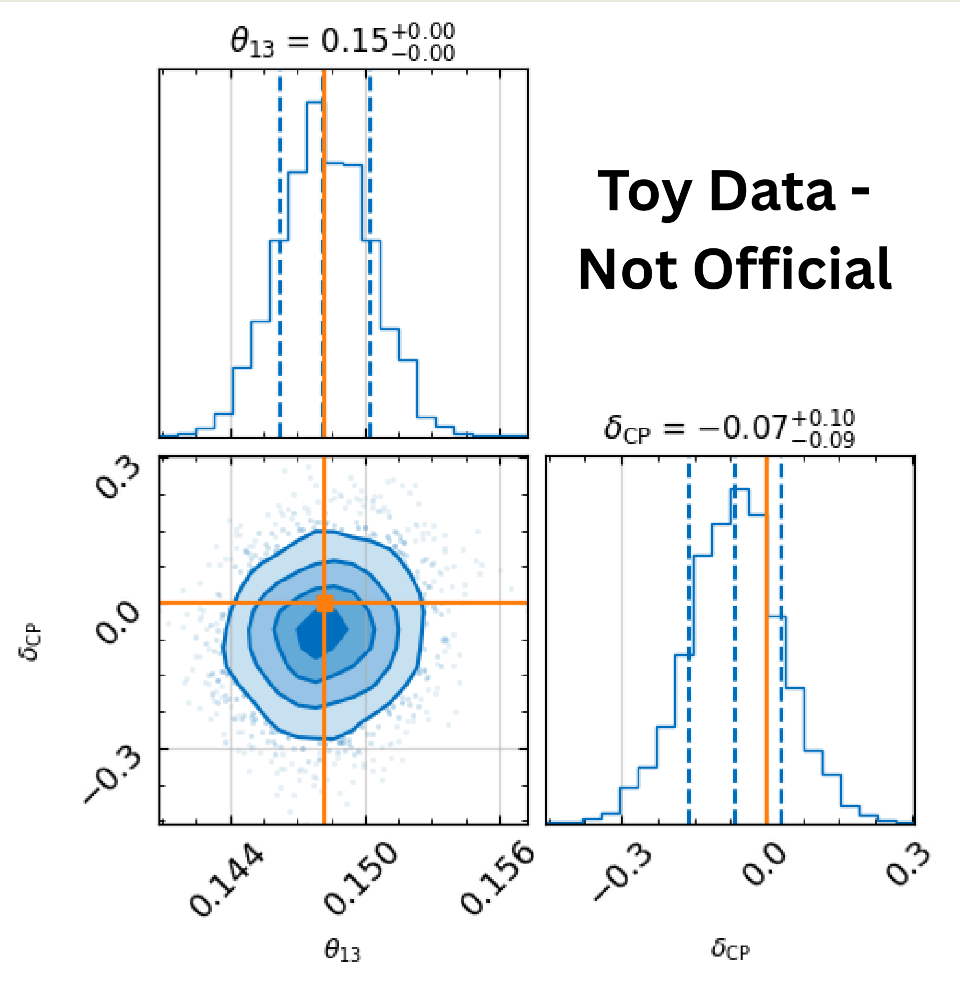
PERFORMANCE

$\mathcal{L}(\hat{E}_\nu, \text{towall} \mid \text{E}_\nu)$

$\delta_{\text{CP}} = -1.601$



$\delta_{\text{CP}} = 0$



Neutrino Events: 1143
Anti Neutrino Events: 440

Neutrino Events: 902
Anti Neutrino Events: 609

$\sin^2\theta_{12}$	$\sin^2\theta_{13}$	$\sin^2\theta_{23}$	Δm^2_{32}	Δm^2_{21}	**MO**	ρ_e
0.307	0.0218	0.528	0.002509	0.0000753	normal	2.6 g/cm^3

- $2.7\text{x}10^{22}$ POT with 1:3 FHC to RHC ratio
- In this toy model can exclude CP conservation to 5σ when only statistics are used
- Posteriors will undoubtedly broaden when we include systematics, generators, etc but for now it looks good

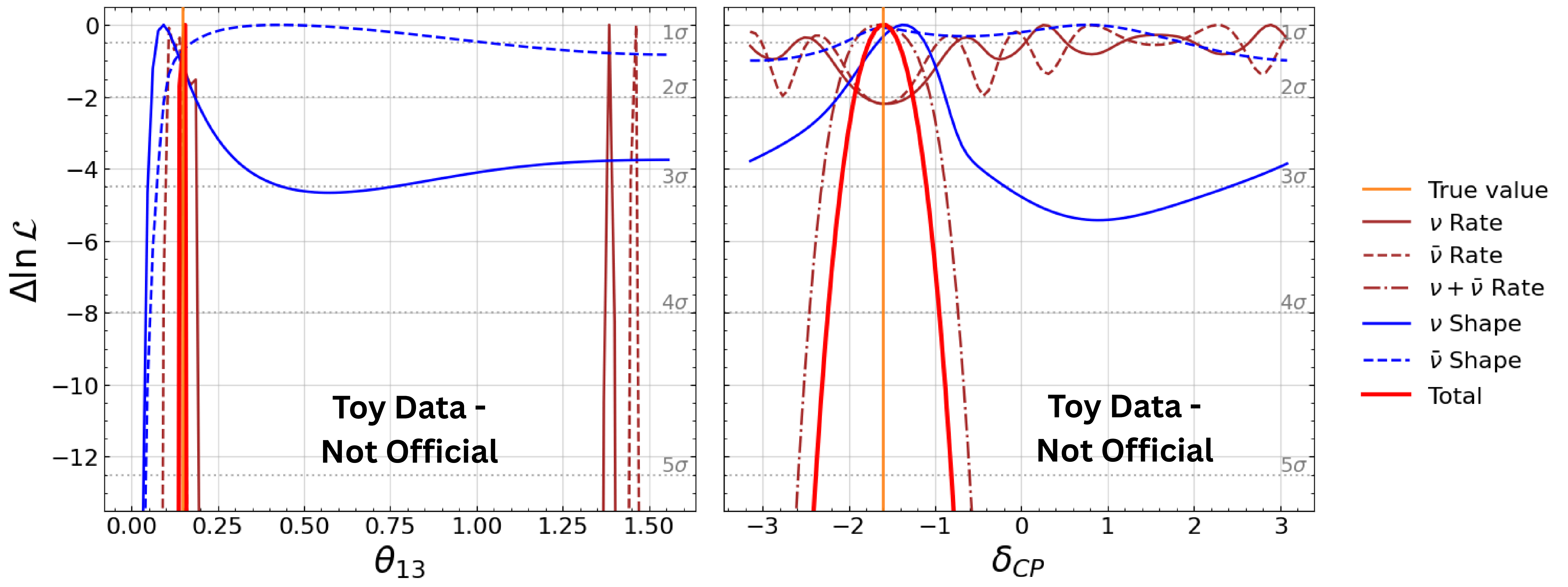
This is toy data - doesn't represent final results

INDIVIDUAL TERMS

"Total" "Rate" "Shape"

$$P(\theta_{13}, \delta_{CP} \mid \{\hat{E}_\nu^i\}, \{\hat{E}_{\bar{\nu}}^i\}, N_{\rm obs}, \bar{N}_{\rm obs}) \propto \frac{N_{\rm exp}^{N_{\rm obs}} e^{N_{\rm exp}}}{N_{\rm obs}!} \prod_{i=1}^{N_{\rm obs}} \int \mathcal{L}(\hat{E}_\nu^i \mid E_\nu)\, \pi(E_\nu \mid \theta_{13}, \delta_{CP})\, dE_\nu \quad \} \nu$$

$$\times \frac{\bar{N}_{\rm exp}^{\bar{N}_{\rm obs}} e^{\bar{N}_{\rm exp}}}{\bar{N}_{\rm obs}!} \prod_{i=1}^{\bar{N}_{\rm obs}} \int \mathcal{L}(\hat{E}_{\bar{\nu}}^i \mid E_{\bar{\nu}})\, \pi(E_{\bar{\nu}} \mid \theta_{13}, \delta_{CP})\, dE_{\bar{\nu}} \quad \} \bar{\nu}$$



SUMMARY AND FUTURE WORK

- We developed a Bayesian pipeline to do oscillation analysis
- We use reconstruction as a data compression tool to map $\mathbf{R}^{40,000} \rightarrow \mathbf{R}^{<10}$
- Trained a normalising flow and constructed a simple conditional physics prior that allows us to build a marginal likelihood and connect the flow to the parameters we care about (θ_{13} and δ_{CP})
- Much work still needs to be done before any serious analysis can be attempted but the scaffolding is in place and results are already very promising
  - Primary want to incorporate as much physics as possible (ν_{μ} survival, interaction modes, beam contaminants, etc)
  - Including systematics will be critical and discussion is ongoing about the best way to do this (merging with existing tools, creating something bespoke

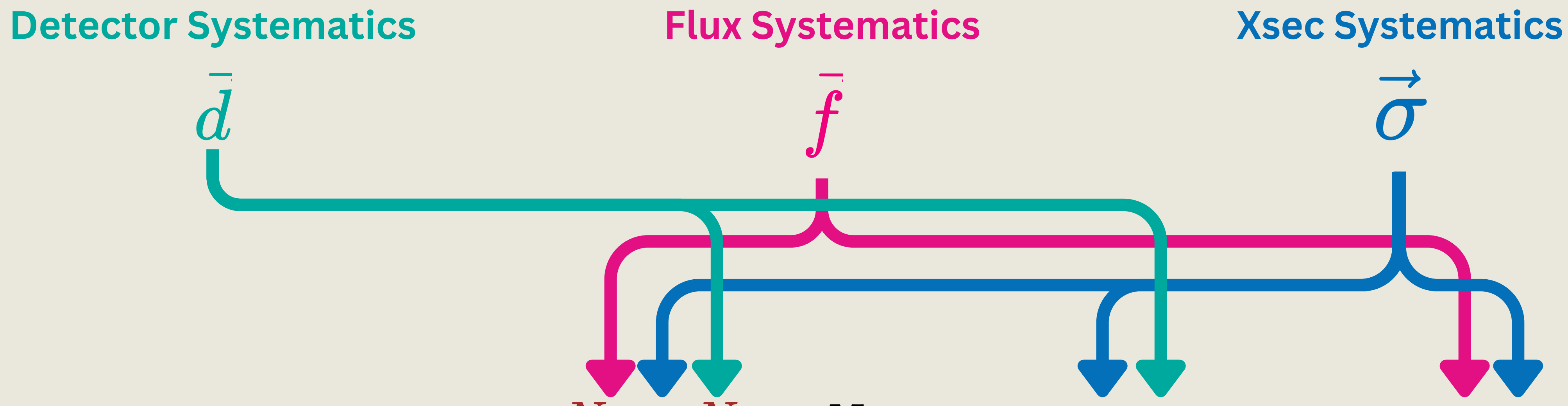
Would love to hear your thoughts or suggestions for improvements!

Email: andrew.atta1@monash.edu

APPENDIX

INLCUDING SYSTEMATICS

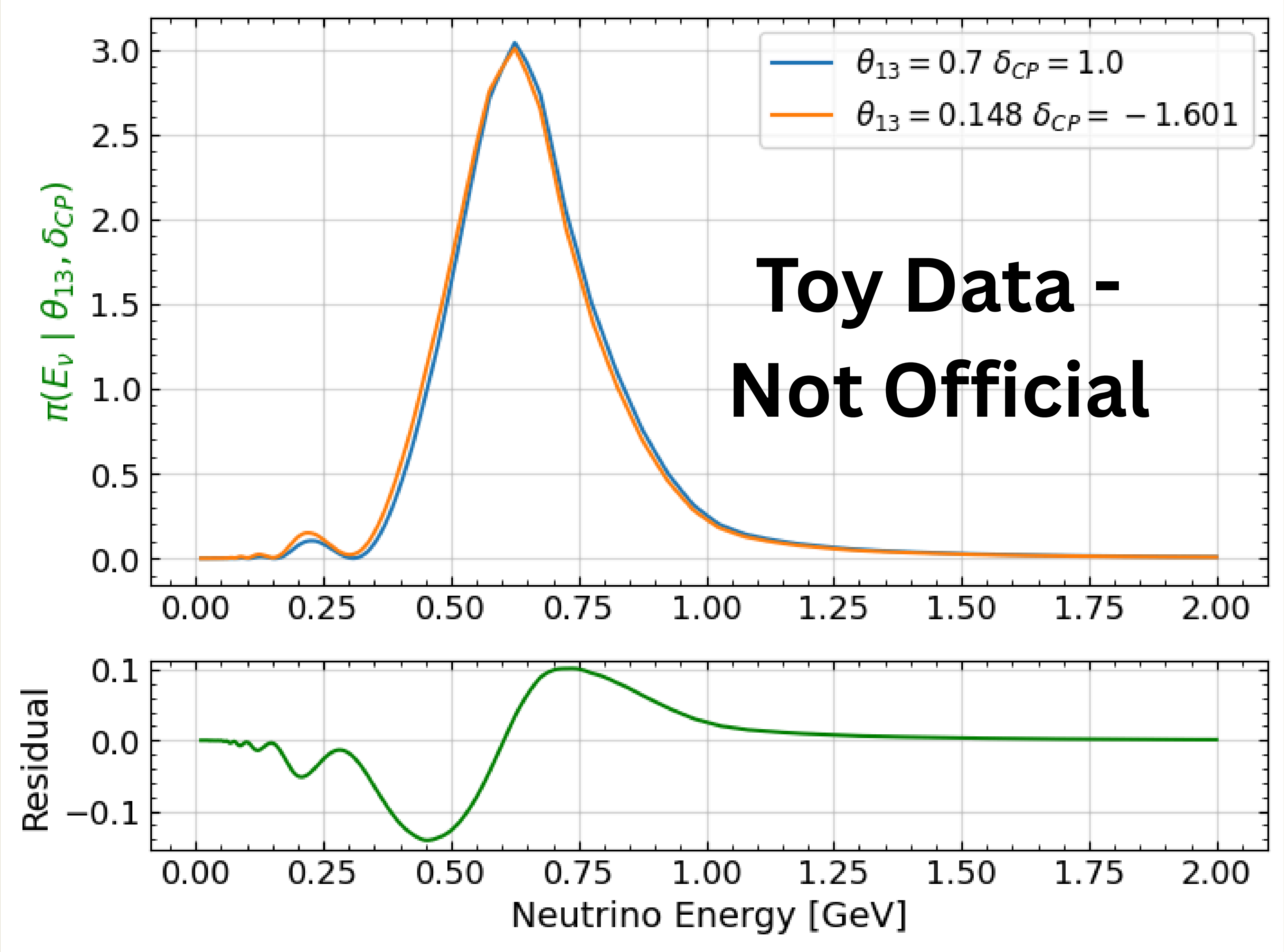
- Have not fully implemented yet, but look to do so in the near future
- Some systematics can be marginalised over
- Depending on the systematic, this can either be:
  - Forward simulated - the nuisance parameter is baked simulation and marginalisation happens by pooling the simulated data together before training flow
  - Those that aren't easily forward simulated can still be marginalised over normally
- **Should not be handwaved away. This will be mission critical** but does not effect the efficacy of the method itself
- Systematics can broadly divided into three categories:



$$P(\theta_{13}, \delta_{CP} \mid \{\hat{E}_\nu^i\}, N_{\text{obs}}) \propto \frac{N_{\text{exp}}^{N_{\text{obs}}} e^{N_{\text{exp}}}}{N_{\text{obs}}!} \prod_{i=1}^{N_{\text{obs}}} \int \mathcal{L}(\hat{E}_\nu^i \mid E_\nu)\, \pi(E_\nu \mid \theta_{13},\, \delta_{CP})\, dE_\nu$$

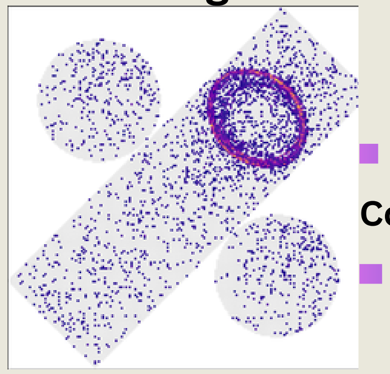
SMALL DIFFERENCES

$$\pi(E_\nu \mid \theta_{13}, \delta_{CP}) \propto P_{\nu_\mu \to \nu_e}(E_\nu \mid \theta_{13}, \delta_{CP}) \times \phi(E_\nu) \times \epsilon(E_\nu) \times \sigma(E_\nu) \times N_{\text{targets}}$$



COMPARED TO BINNING

charge



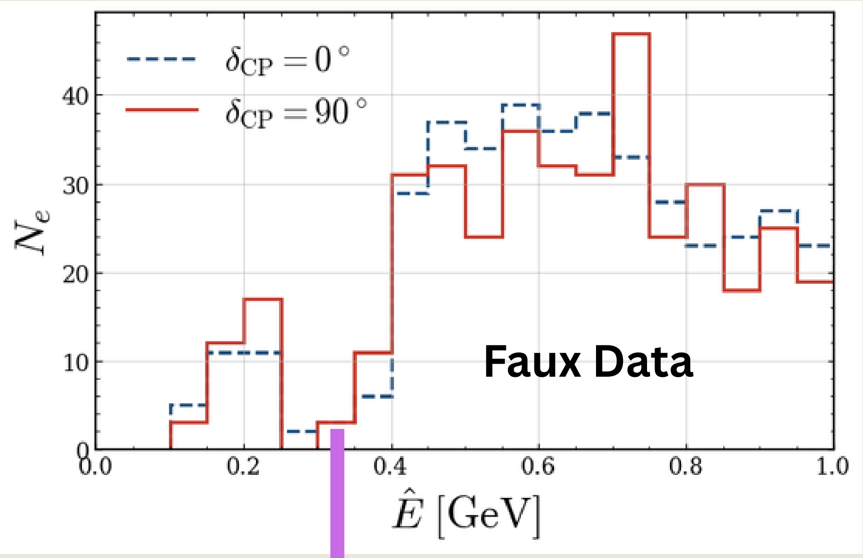
time

ResNet

Compression

fiTQun

$\hat{E}_\nu$ — traditional →



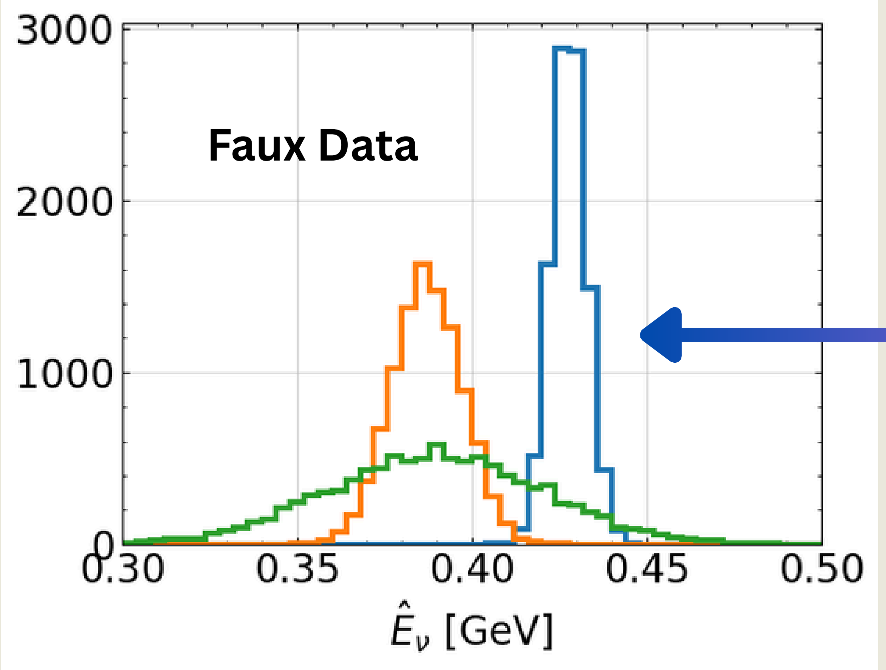
$$\sum_{i}^{N_{\text{bins}}} \mathcal{L}(N^i_{\text{obs}} \mid N^i_{\text{exp}})\pi(N^i_{\text{exp}} \mid \theta_{13}, \delta_{CP}) \quad ?$$

Flow

$$P(\theta_{13}, \delta_{CP} \mid \hat{E}_\nu) \propto \int \mathcal{L}(\hat{E}_\nu \mid E_\nu)\, \pi(E_\nu \mid \theta_{13},\, \delta_{CP})\, dE_\nu$$

the data

pure physics

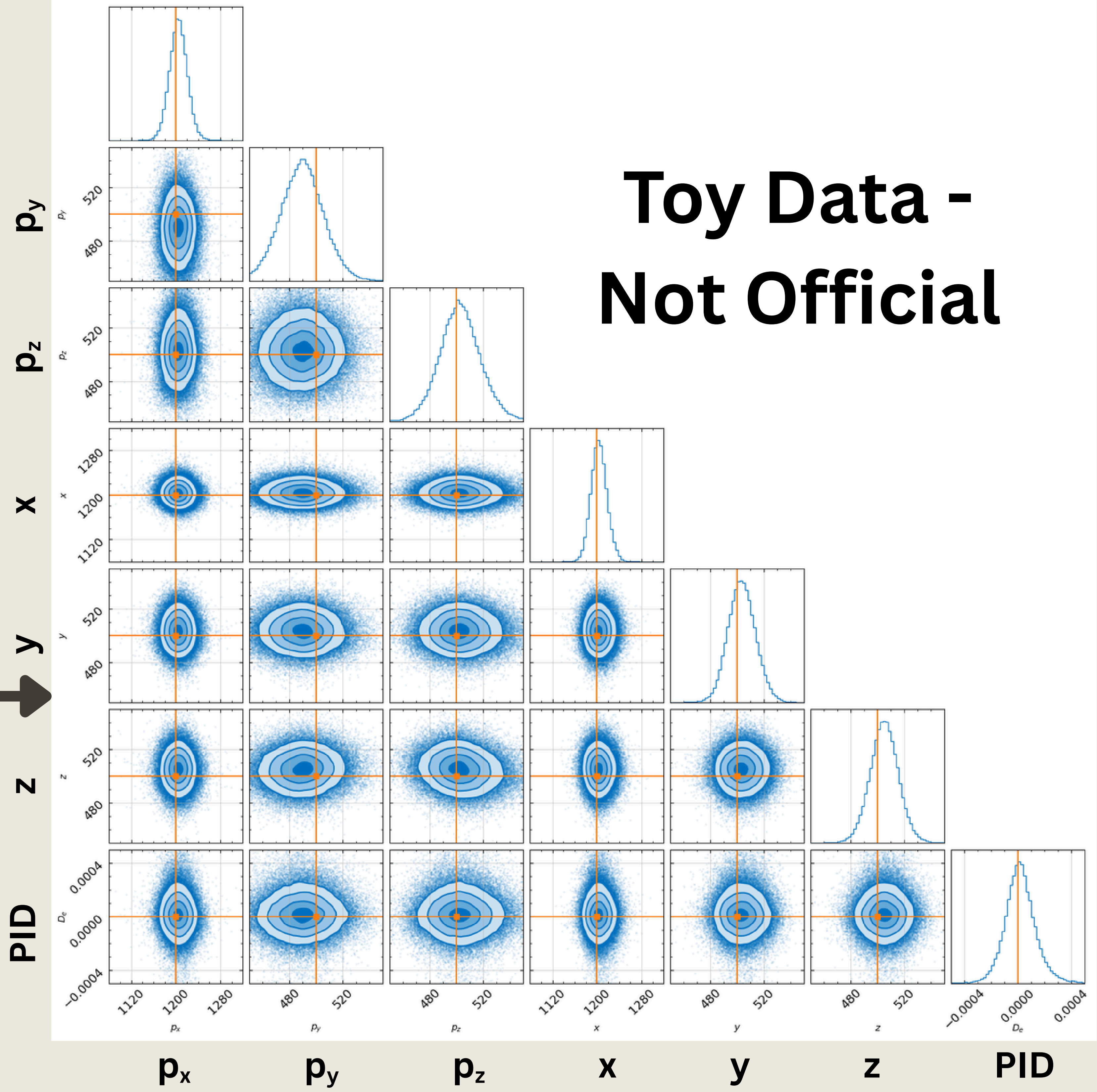


EXPANDING THE LIKELIHOOD.

- With a flow we can include all the information of event without facing dimensionality issues

$$\mathcal{L}(\hat{E}_\nu \mid E_\nu) \to \mathcal{L}(E_\ell, \theta_\ell, \phi_\ell, x, y, z, PID \mid E_\nu)$$

- All we have to do is include the information when training the normalising flow. Allows us to simultaneously fit additional data that might hold important information
- For example information about the distance the particle is the to the wall (towall)
- Testing different datasets that have the best results
- Others we've tried are:
  - $\mathcal{L}(\hat{E}_\nu, \mathrm{towall} \mid \mathrm{E}_\nu)$
  - $\mathcal{L}(E_\ell, \theta_\ell, \phi_\ell, x, y, z \mid E_\nu)$
  - $\mathcal{L}(E_\ell, \theta_\ell, \phi_\ell, x, y, z, PID \mid E_\nu)$
  - $\mathcal{L}(p_{\ell\,x}, p_{\ell\,y}, p_{\ell\,z}, x, y, z, PID \mid E_\nu)$
- Found that $\mathcal{L}(\hat{E}_\nu, \mathrm{towall} \mid \mathrm{E}_\nu)$ worked well for particles close to the detector wall and was nice and interpretable



GETTING N_{OBS} AND N_{EXP}

Posterior — Detector Model (flow) — Oscillation + Flux + cross section physics

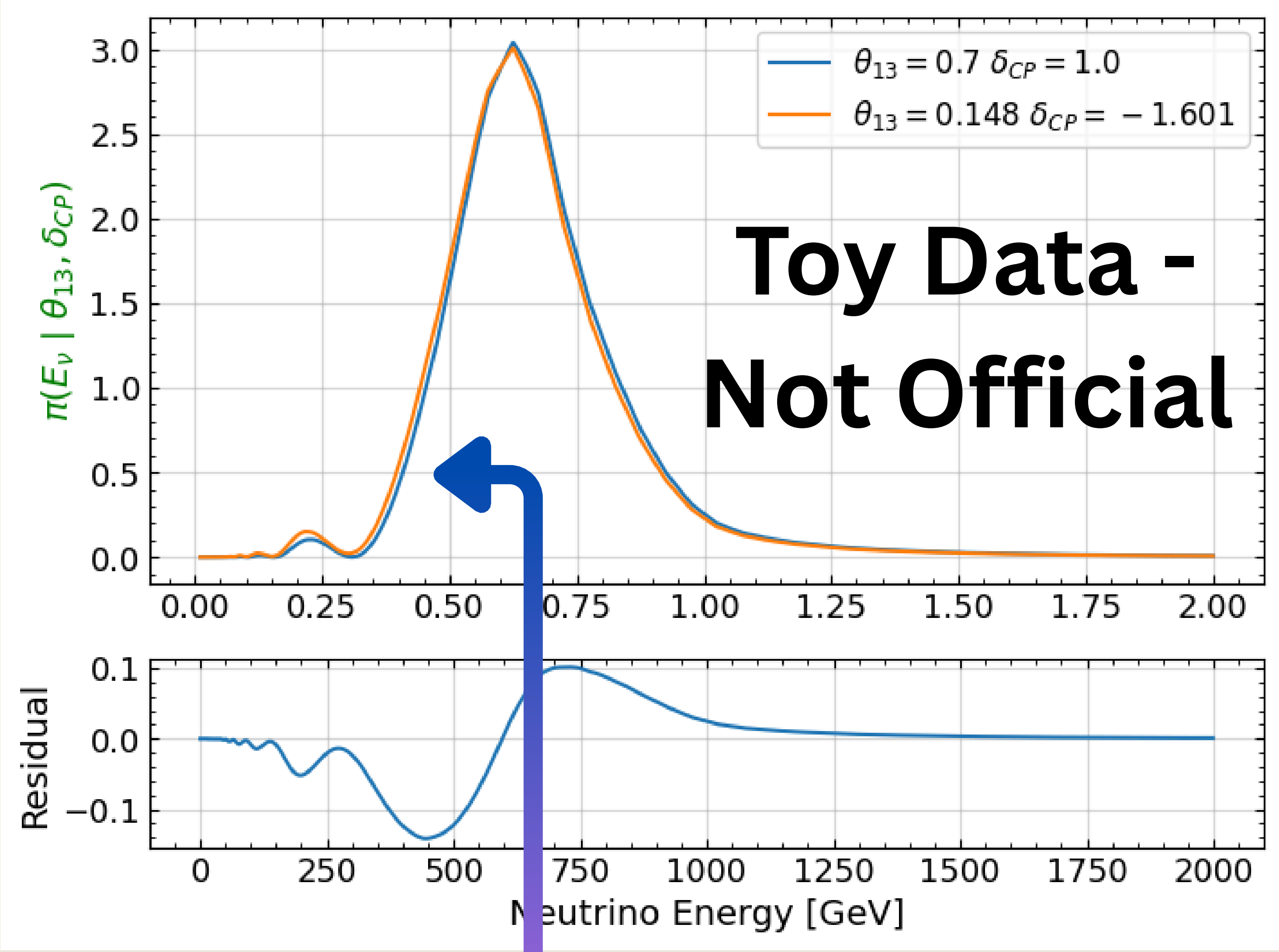
$$P(\theta_{13}, \delta_{CP} \mid \{E_\ell^i\}) \propto \prod_{i=1}^{N_{obs}} \int \mathcal{L}(E_\ell^i \mid E_\nu)\, \pi(E_\nu \mid \theta_{13},\, \delta_{CP})\, dE_\nu$$

- This expression will tell you how the likely the *shape* of the observed neutrino energy distribution is for given oscillation parameters.
- We also want to include information about how the *expected number* of neutrinos differs from the *observed number.*

$$P(\theta_{13}, \delta_{CP} \mid \{E_\ell^i\}, N_{obs}) \propto \frac{N_{exp}^{N_{obs}} e^{N_{exp}}}{N_{obs}!} \prod_{i=1}^{N_{obs}} \int \mathcal{L}(E_\ell^i \mid E_\nu)\, \pi(E_\nu \mid \theta_{13},\, \delta_{CP})\, dE_\nu$$

- Where:

$$N_{exp}(\theta_{13}, \delta_{CP}) = \mathrm{POT} \int \mathrm{P}_{\nu_\mu \to \nu_e}(\mathrm{E}_\nu \mid \theta_{13}\,, \delta_{CP}) \times \phi(\mathrm{E}_\nu) \times \epsilon(\mathrm{E}_\nu) \times \sigma(\mathrm{E}_\nu) \times \mathrm{N}_{targets}\, \mathrm{dE}_\nu$$



- **Also added some Poisson noise to N_{exp} and N_{obs}**

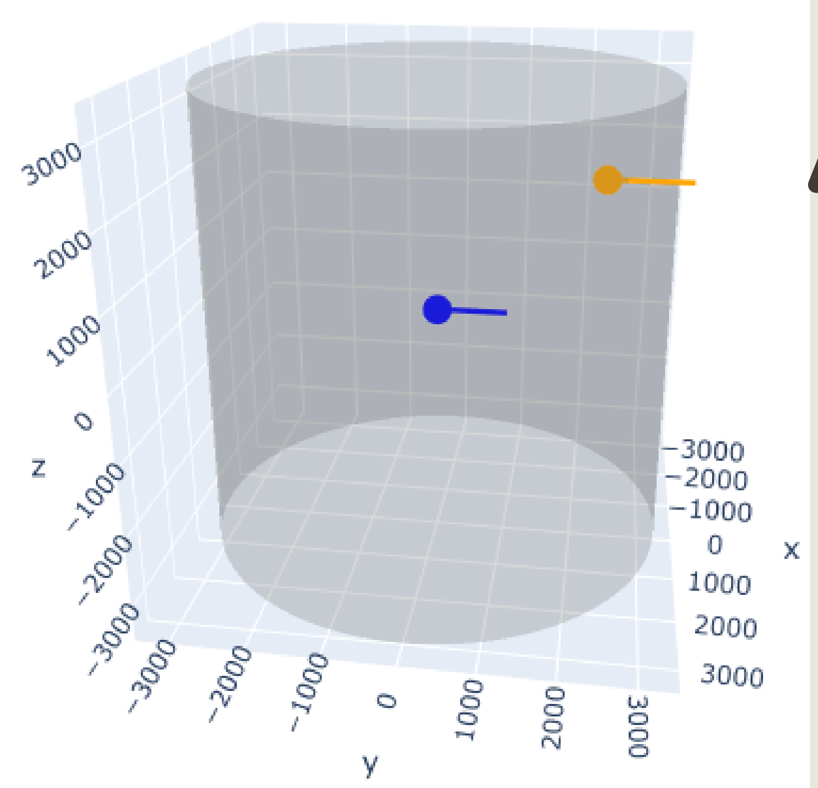
DEFINING E_ν

- ResNet reconstructs $\hat{E}_\ell$ and $\hat{\theta}_\ell$ that tries to predict E_ℓ and θ_ℓ
- We use CCQE approximation for E_ν and $\hat{E}_\nu$

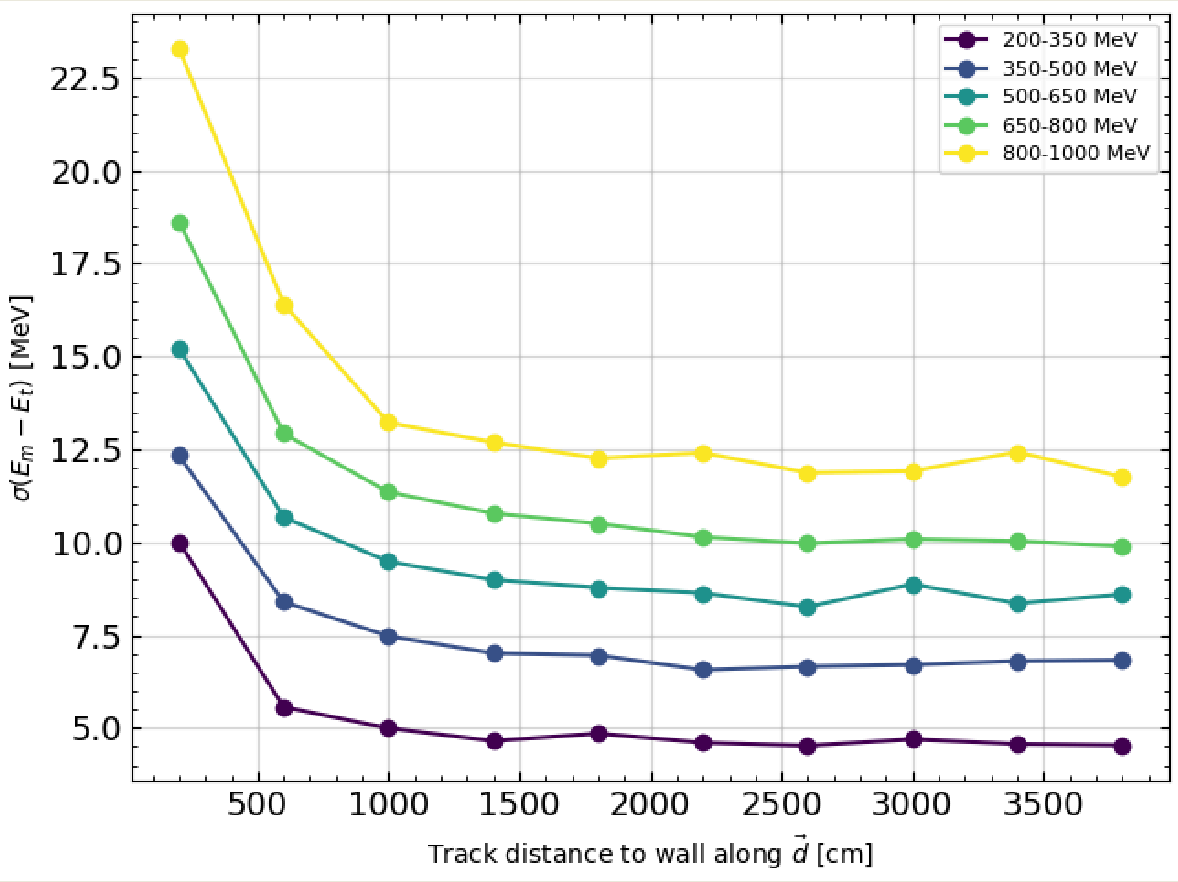
$$E_\nu = \frac{2(m_n - V)E_\ell + m_p^2 - (m_n - V)^2 - m_\ell^2}{2(m_n - V - E_\ell + p_\ell \cos\theta_\ell)} \qquad \hat{E}_\nu^{rec} = \frac{2(m_n - V)\hat{E}_\ell + m_p^2 - (m_n - V)^2 - m_\ell^2}{2(m_n - V - \hat{E}_\ell + \hat{p}_\ell \cos\hat{\theta}_\ell)}$$

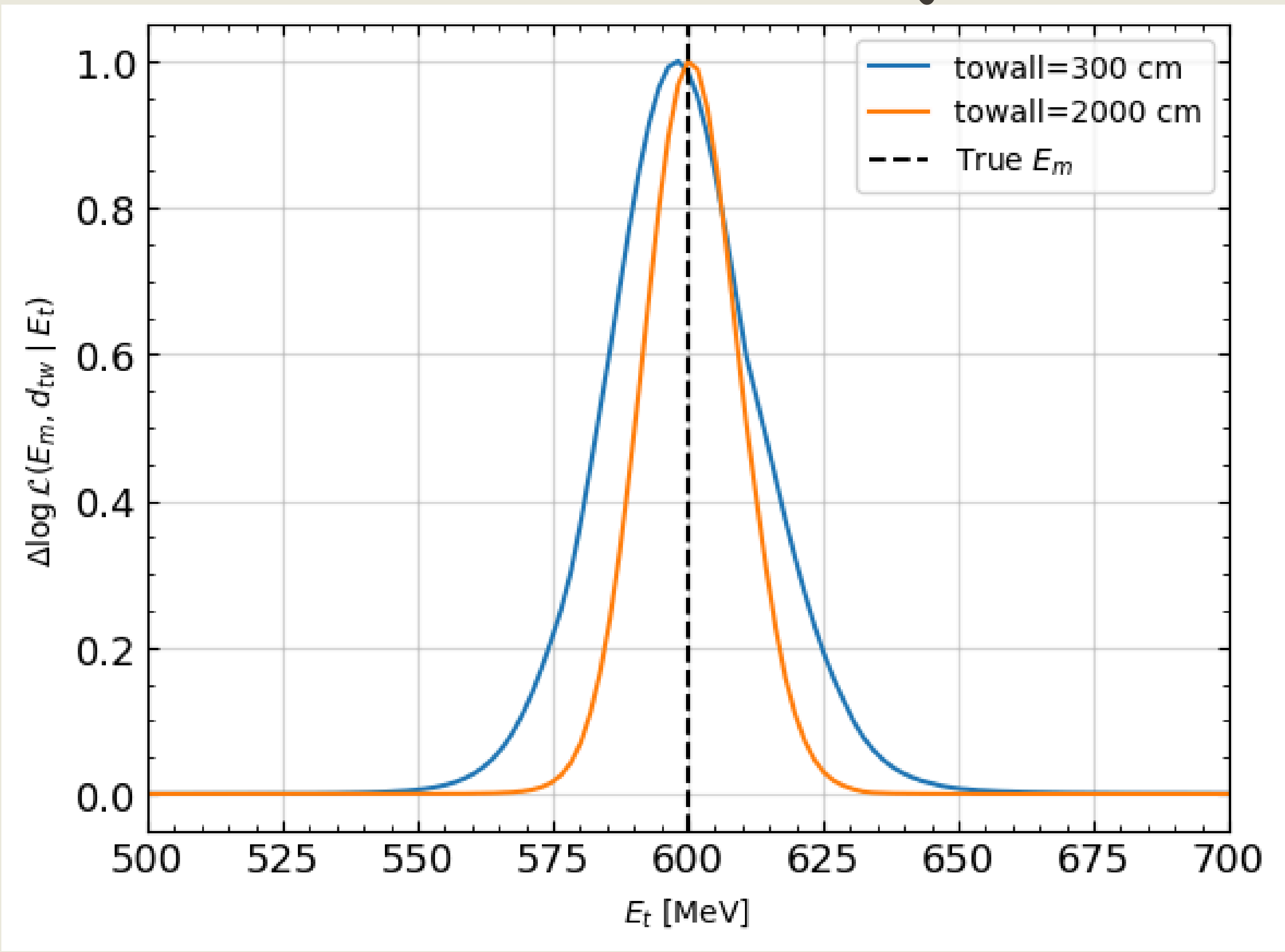
- We can and want to use NEUT to generate CCQE only events
- Could get E_ν form NEUT and calculate $\hat{E}_\nu$ from $\hat{E}_\ell$ and $\hat{\theta}_\ell$ using above approximation
  - The flow would have to be updated to include more systematics around neutrino interactions
  - Will be done with further development

EFFECT OF TOWALL

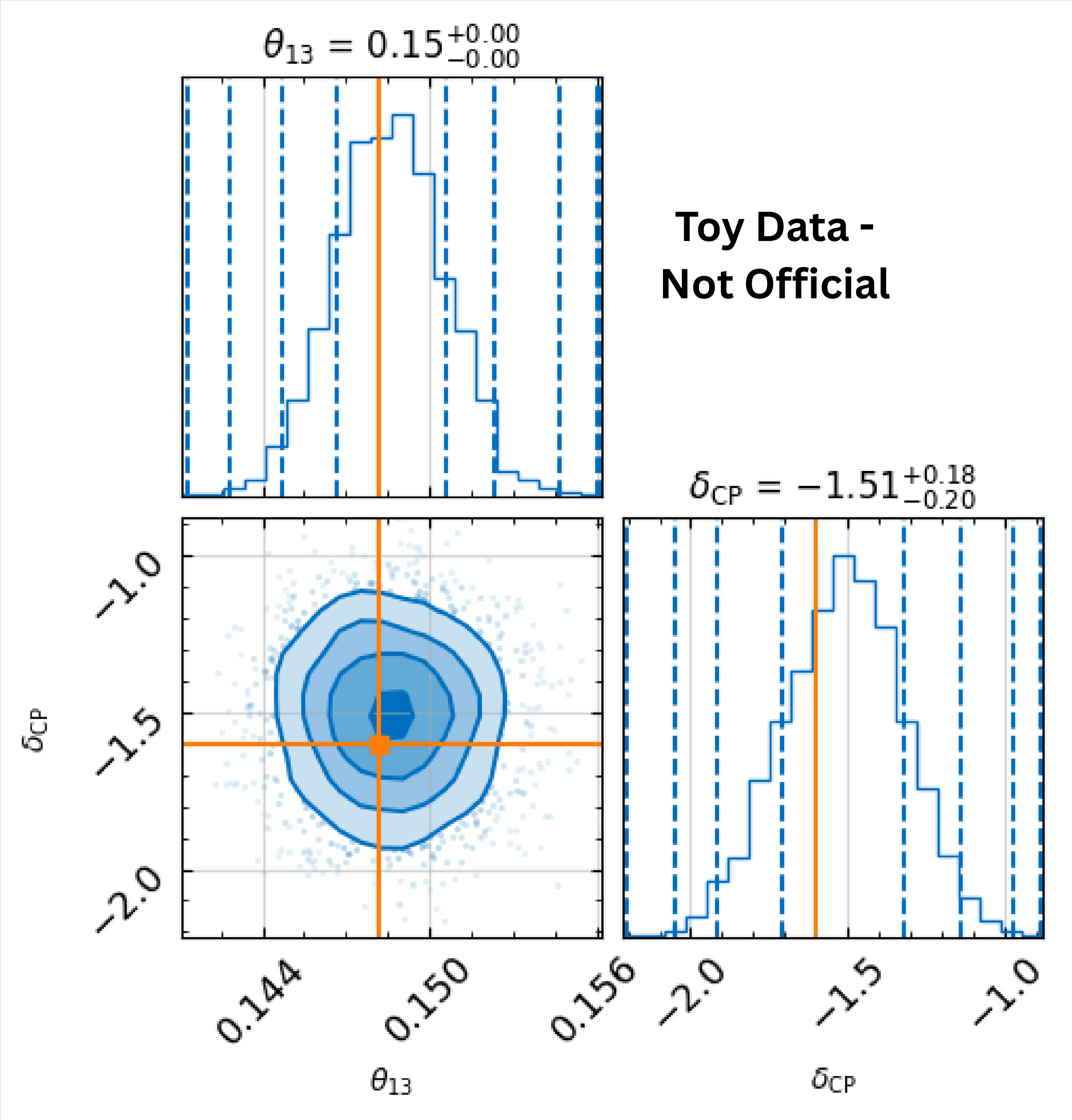


Colours flipped :(

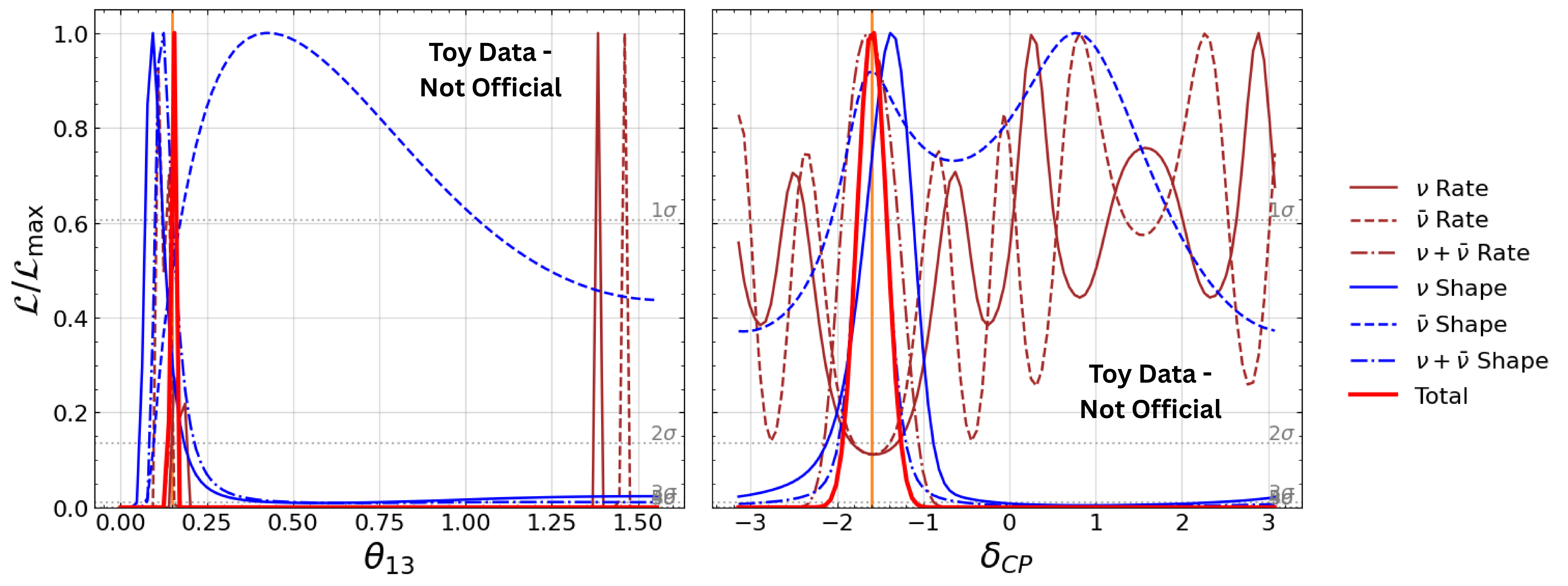




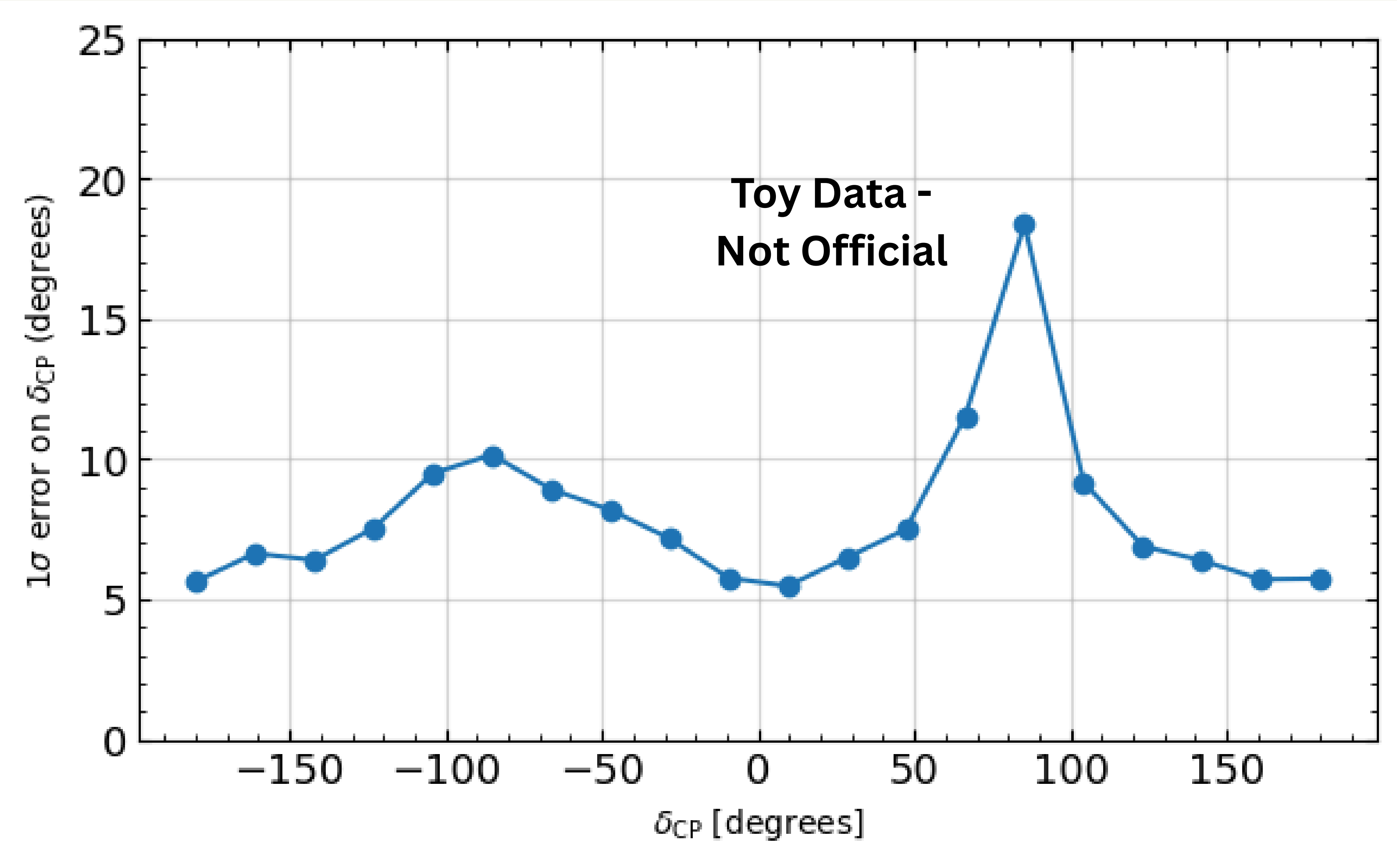
5-SIGMA CLAIM



COMPONENTS LINEAR VIEW

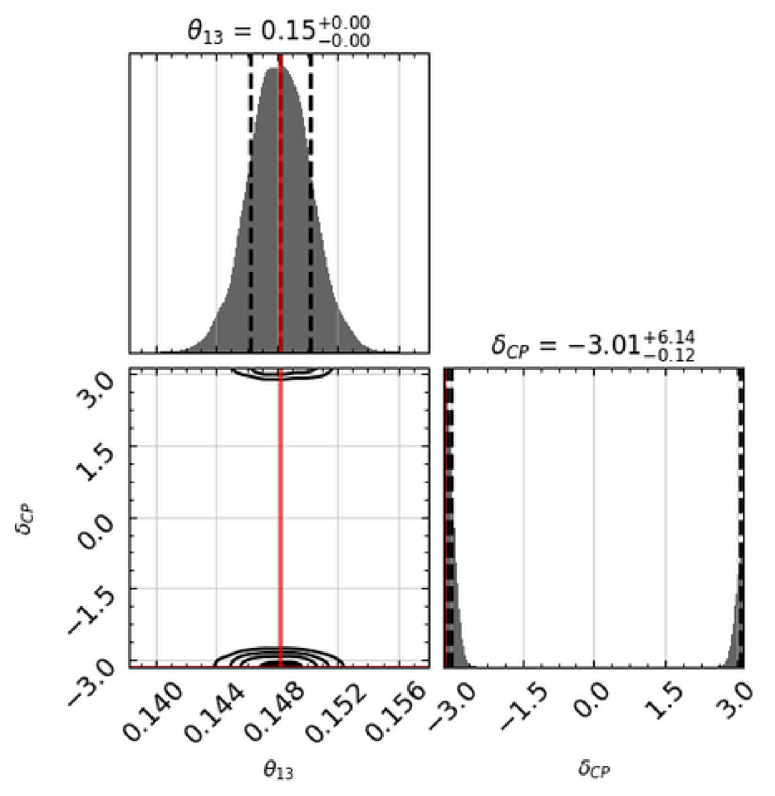


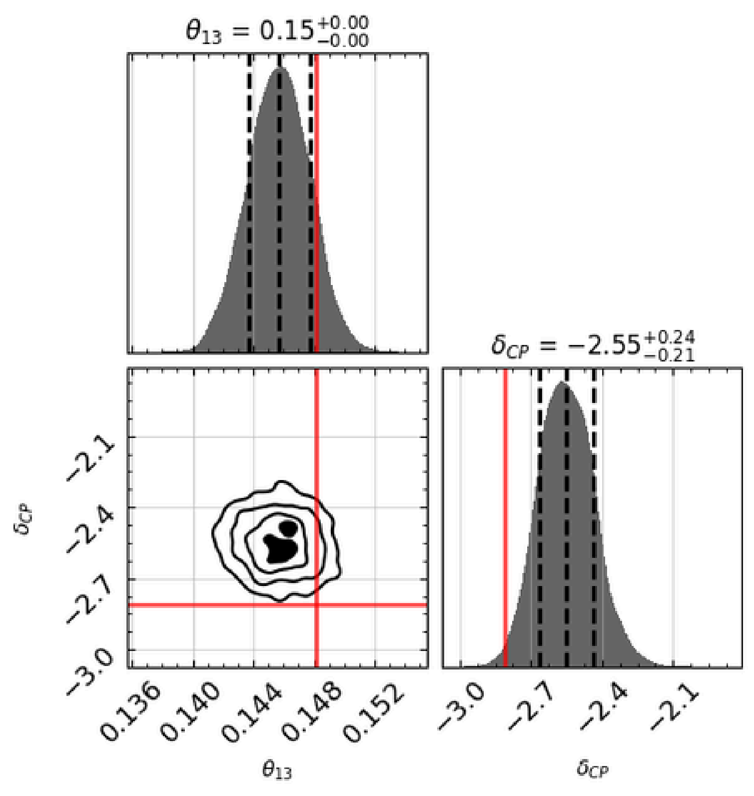
“SENSITIVITY”

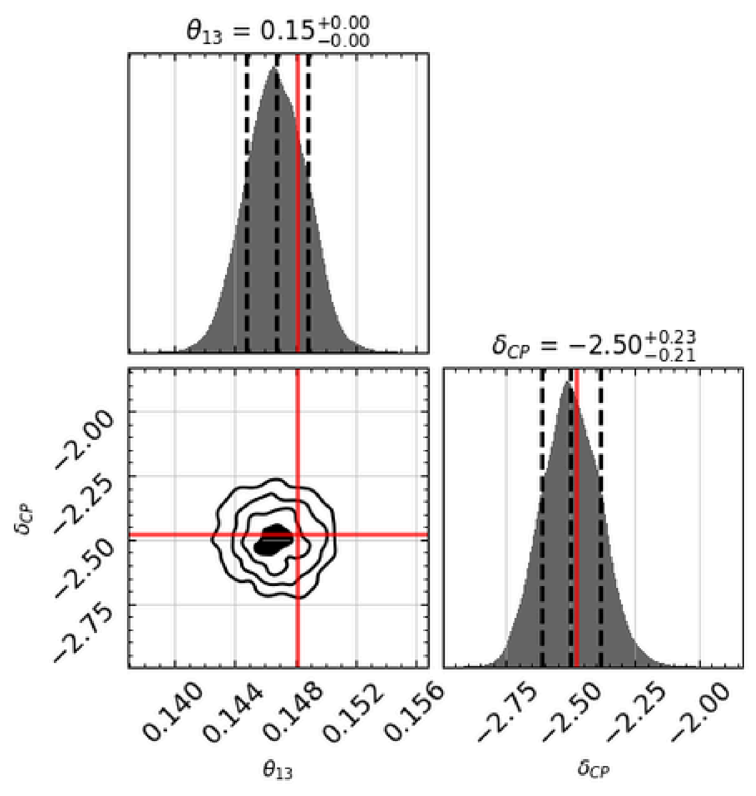


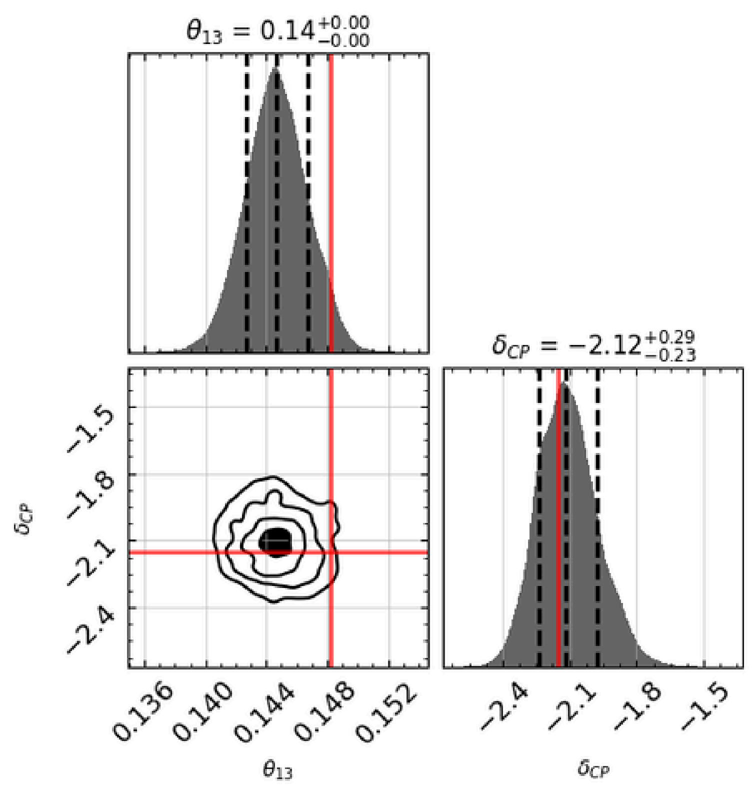
- **Not an actual sensitivity analysis. Take this with a heavy grain of salt**
- Only one posterior per point so expect some large fluctuations (eg ~60 and 80 degrees) where sampler struggled to converge for example
- Regardless, the shape we expect to see is there so that’s encouraging.

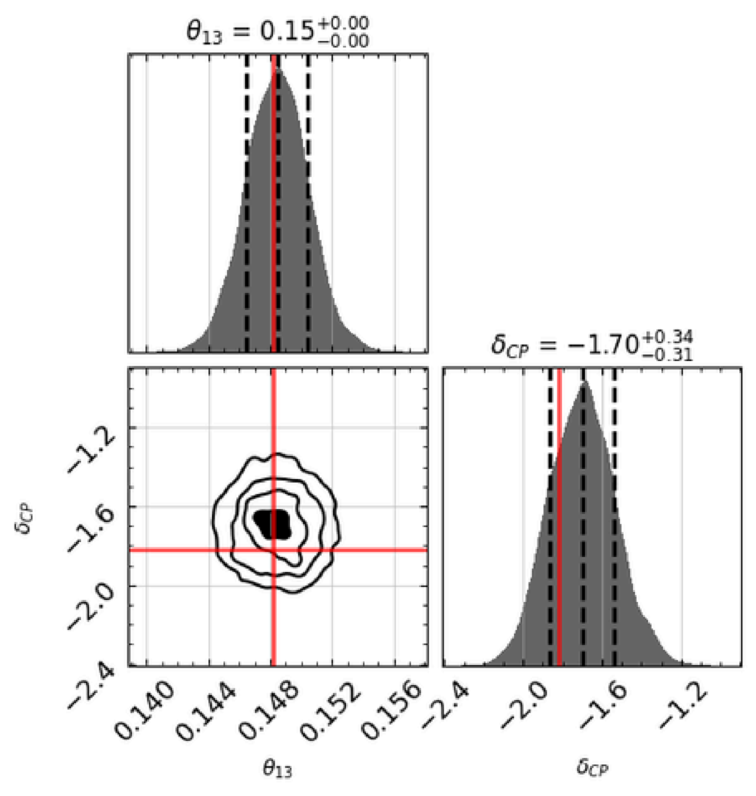
CORNER PLOTS

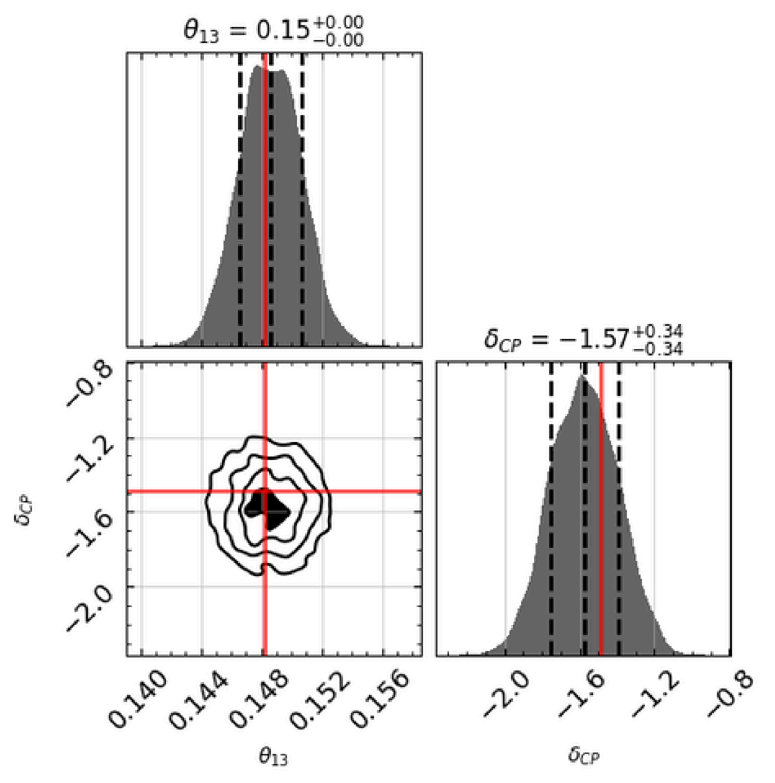


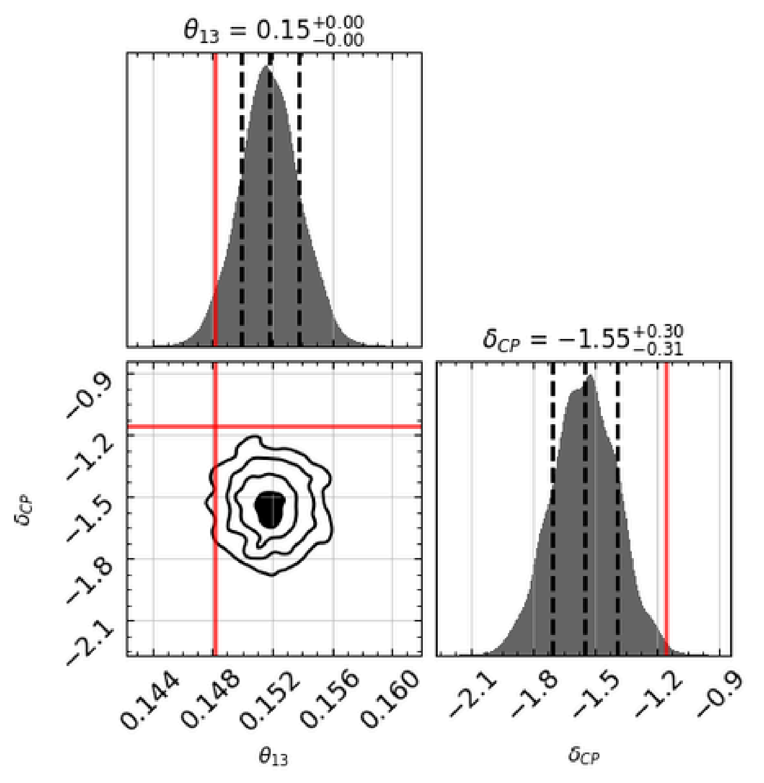


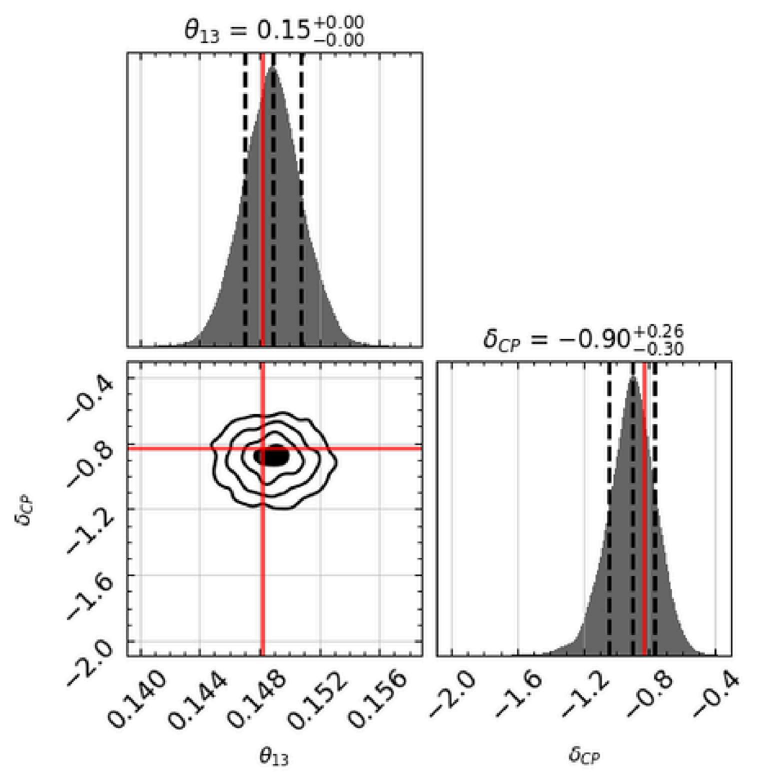


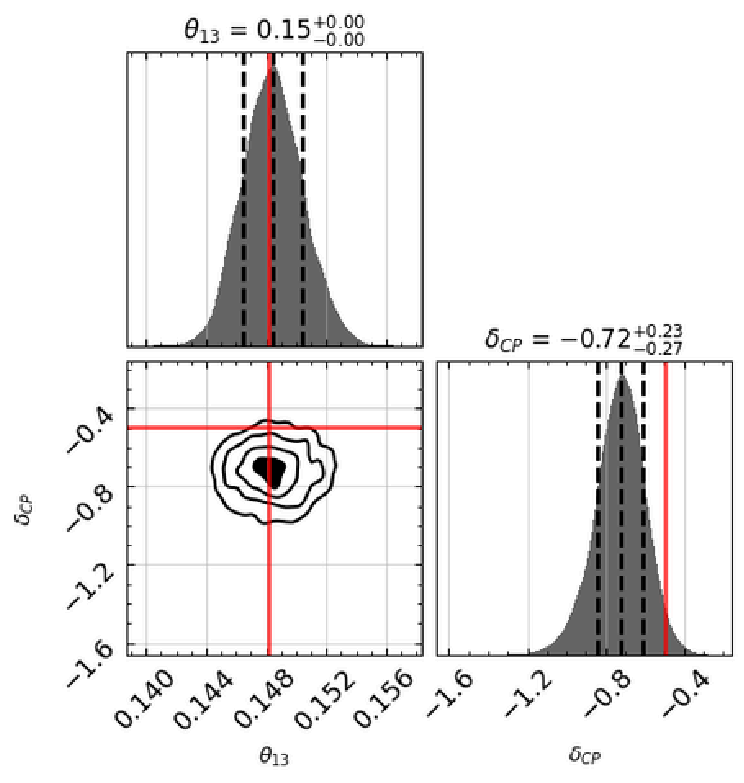


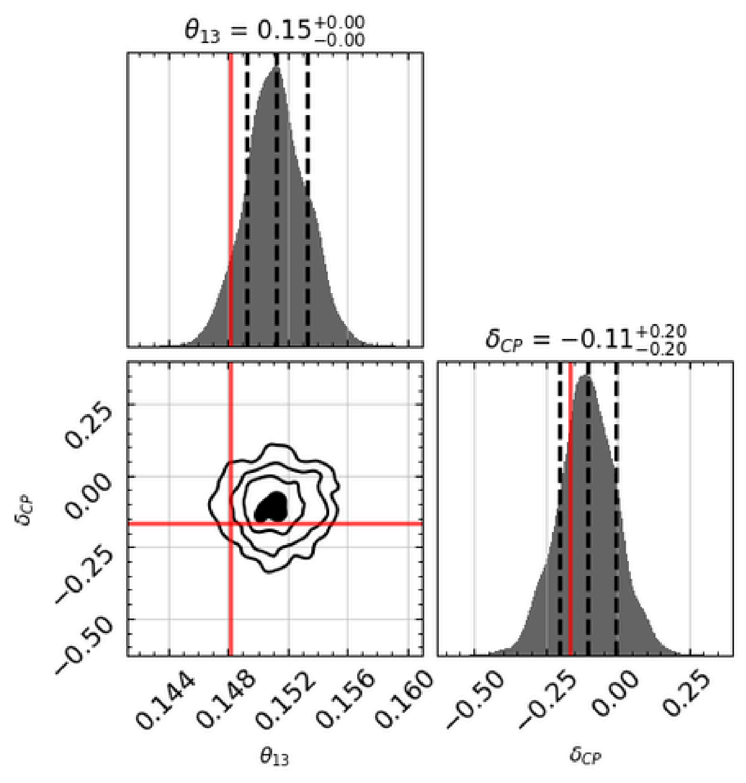












CORNER PLOTS