Current Statistical Practices at Long Baseline Accelerator Experiments

Asher Kaboth

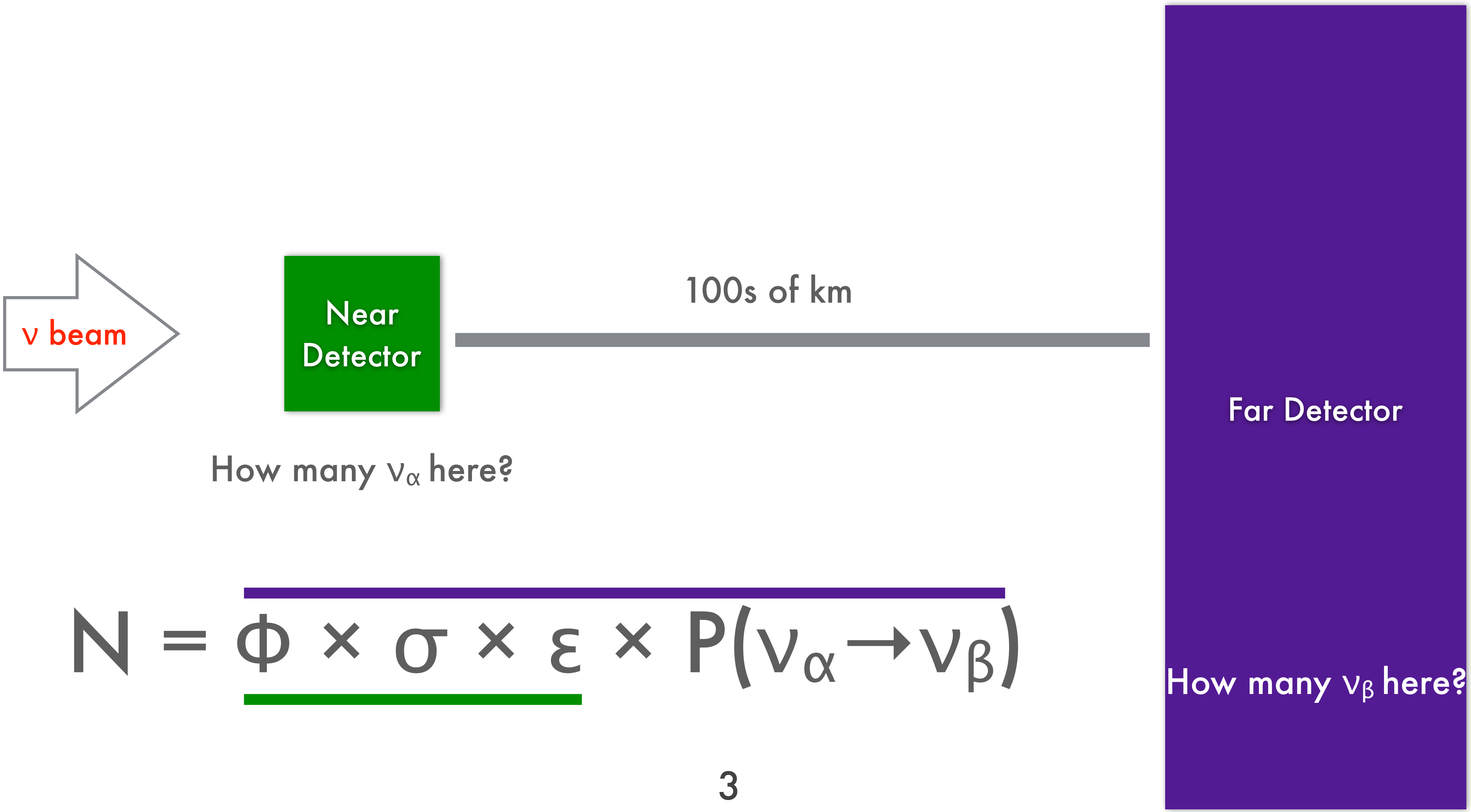
PhyStat-ν

30 May 2016

Outline

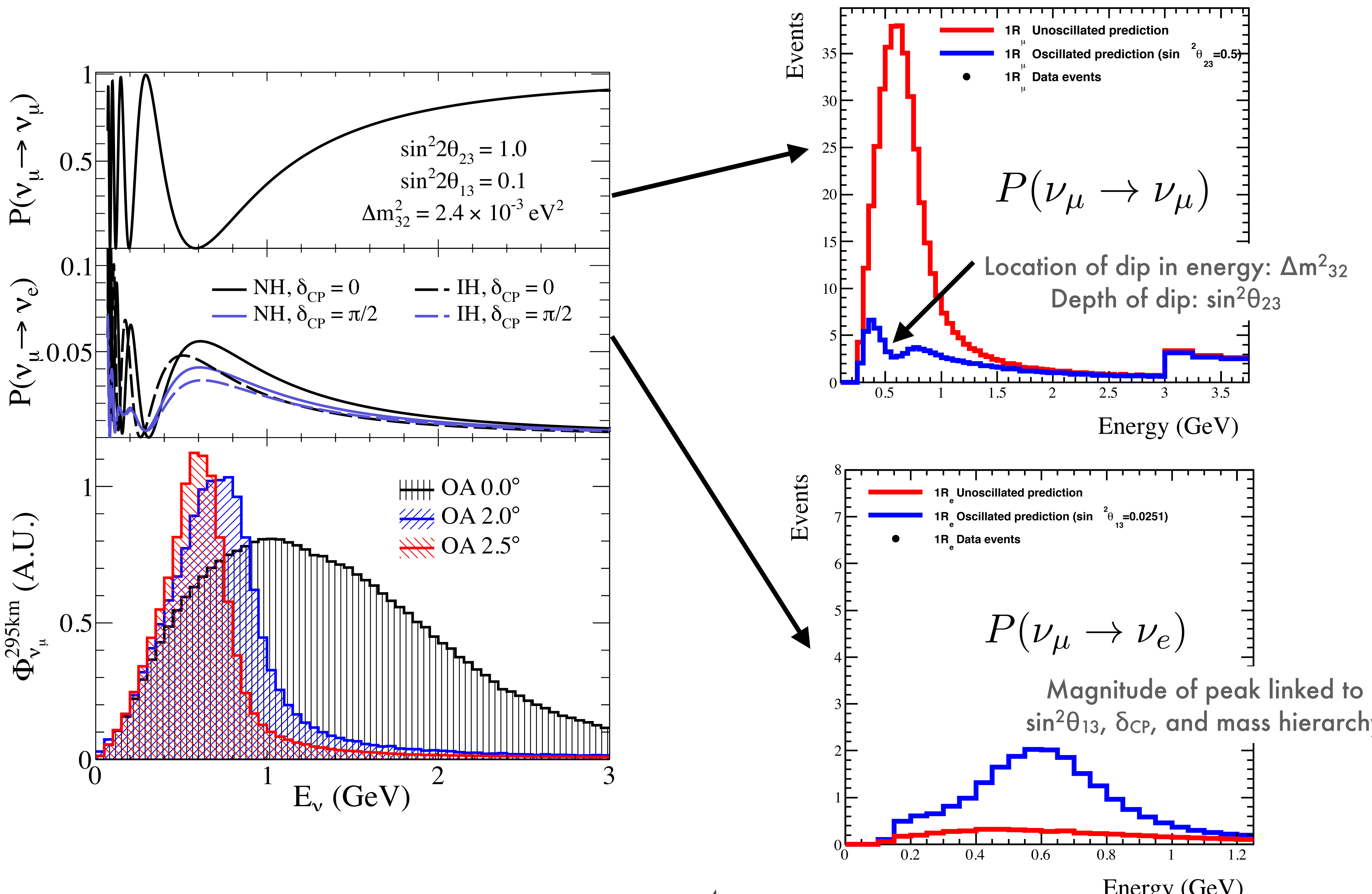
- Quick Review of LBL physics
- T2K
- NOvA
- Special Issues
- Will not talk about: MINOS, DUNE, HyperK

Long Baseline Experiments in a Nutshell



$$N = \overline{\underline{\Phi \times \sigma \times \varepsilon} \times P(\nu_\alpha \rightarrow \nu_\beta)}$$

Long Baseline Neutrino Oscillation



Long Baseline Neutrino Oscillation

$$P(\nu_\mu \to \nu_\mu) \approx 1 - 4\boxed{\cos^2(\theta_{13})}\boxed{\sin^2(\theta_{23})}[1 - \cos^2(\theta_{13})\sin^2(\theta_{23})]\sin^2\left(\frac{\boxed{\Delta m^2_{31}}L}{E_\nu}\right)$$

$$P(\nu_\mu \to \nu_e) \approx \boxed{\sin^2\theta_{23}}\boxed{\sin^2 2\theta_{13}}\frac{\sin^2(\boxed{\Delta}(1-A))}{(1-A)^2} + \alpha^2\cos^2\theta_{23}\sin^2 2\theta_{12}\frac{\sin^2 A\Delta}{A^2}$$

Expanded under small θ_{13},α

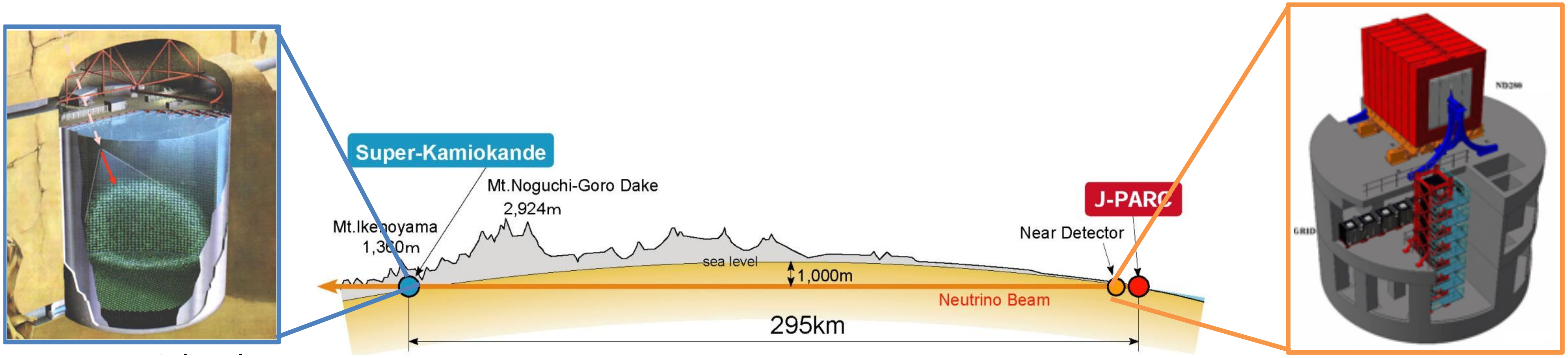
$$+\alpha\cos\theta_{13}\sin 2\theta_{12}\sin 2\theta_{23}\sin 2\theta_{13}\frac{\sin((1-A)\Delta)}{1-A}\frac{\sin A\Delta}{A}\cos(\boxed{\delta}+\Delta)$$

$$A = \pm 2\sqrt{2}G_F n_e E \Delta m^2_{31} \qquad \Delta = \frac{\boxed{\Delta m^2_{31}}L}{E} \qquad \alpha = \frac{\Delta m^2_{21}}{\Delta m^2_{31}} \qquad \Delta m^2_{ij} = m^2_i - m^2_j$$

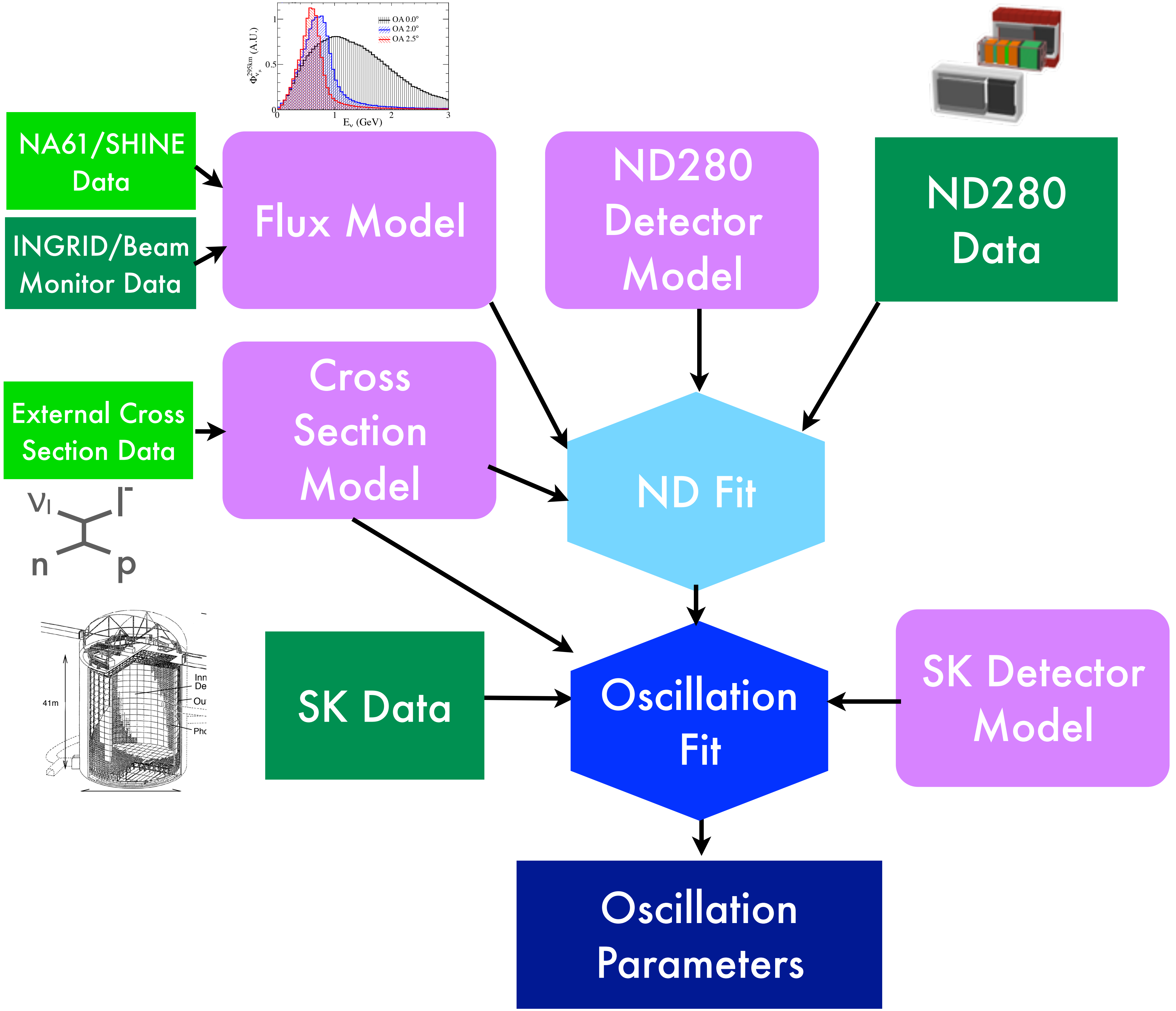
Open Questions

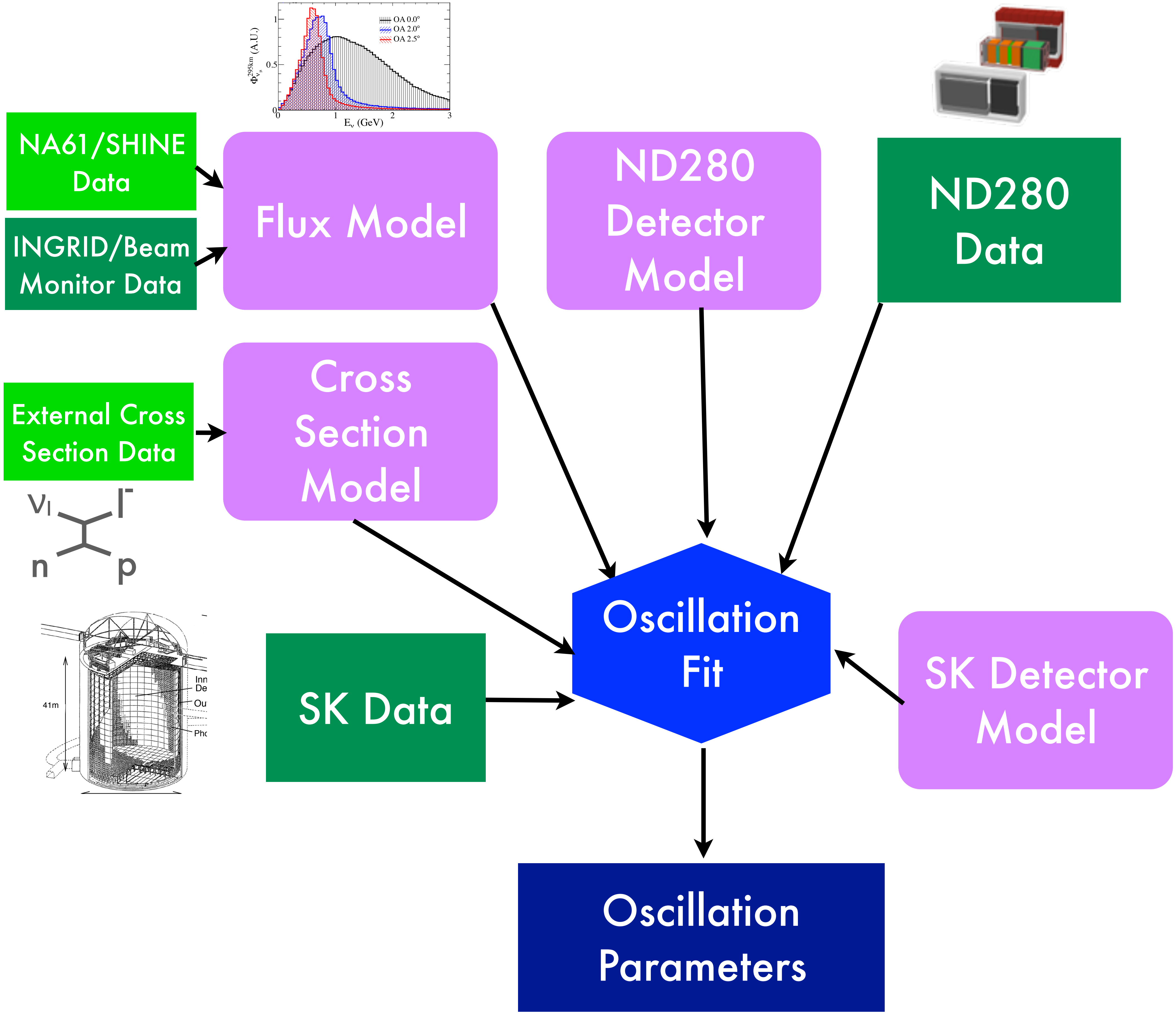
- How well can we estimate the values of θ_{13}, θ_{23}, and Δm^2_{32}?
- Is δ_{CP} different from 0 and π?
- Do neutrinos and antineutrinos behave differently?
- What is the mass hierarchy?

T2K

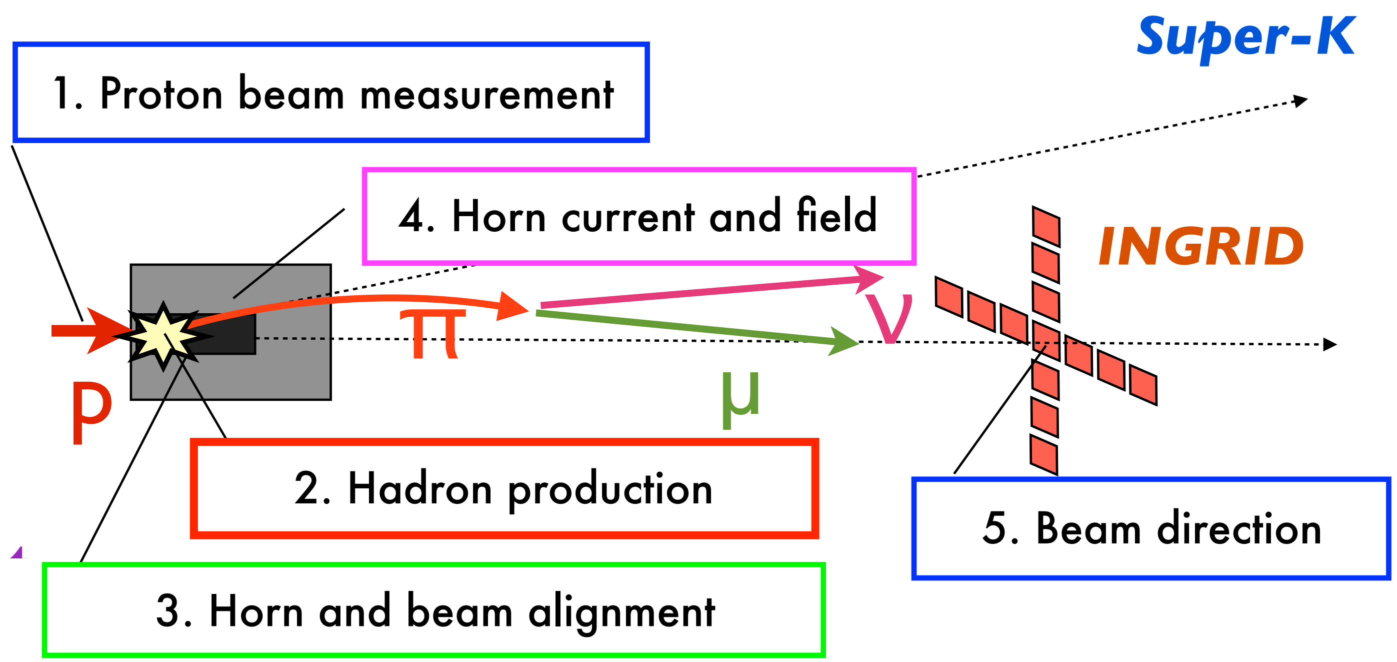


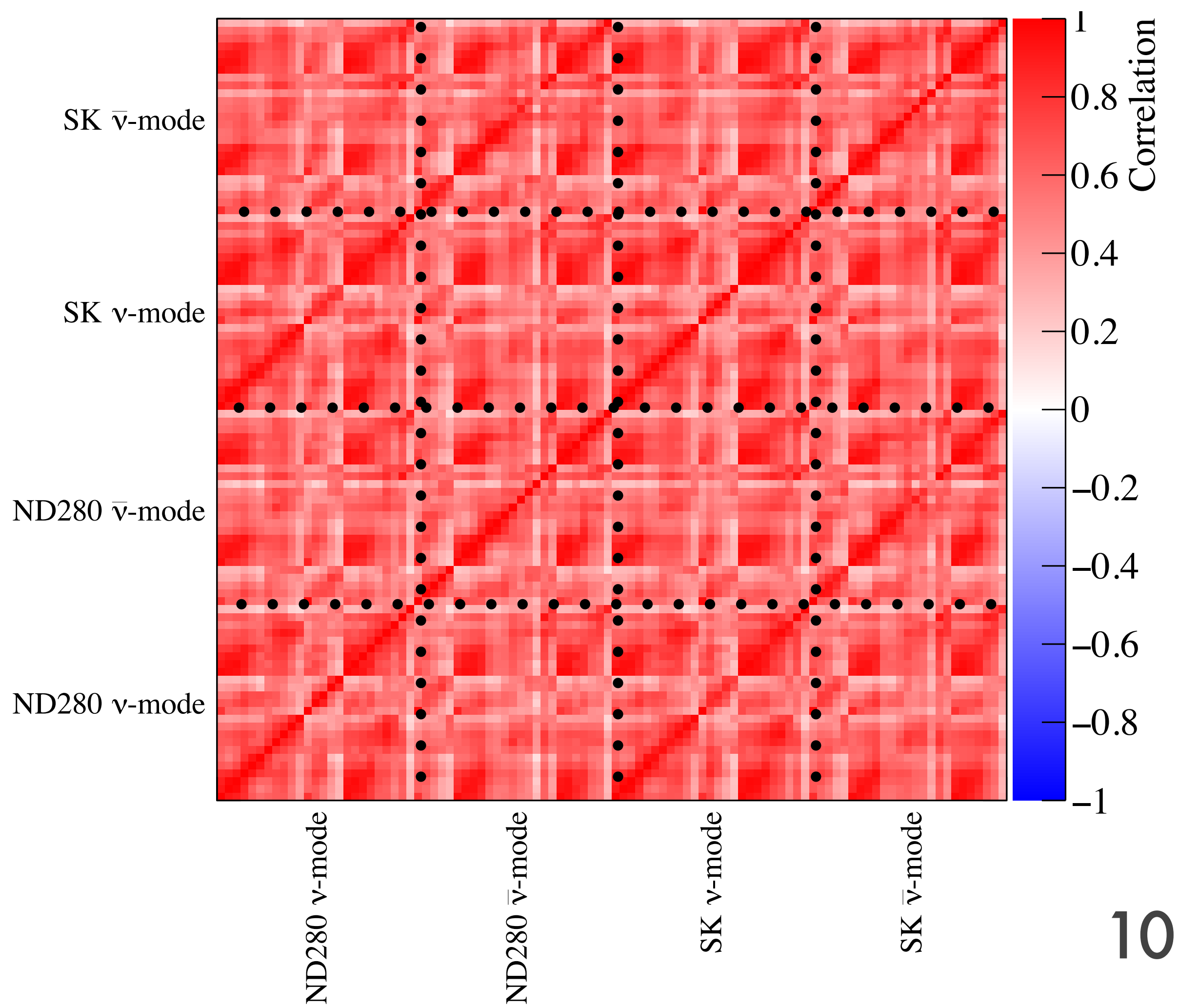
Super-Kamiokande
Mt.Noguchi-Goro Dake
2,924m
Mt.Ikenoyama
1,360m
sea level
1,000m
J-PARC
Near Detector
Neutrino Beam
295km
ND280
GRID

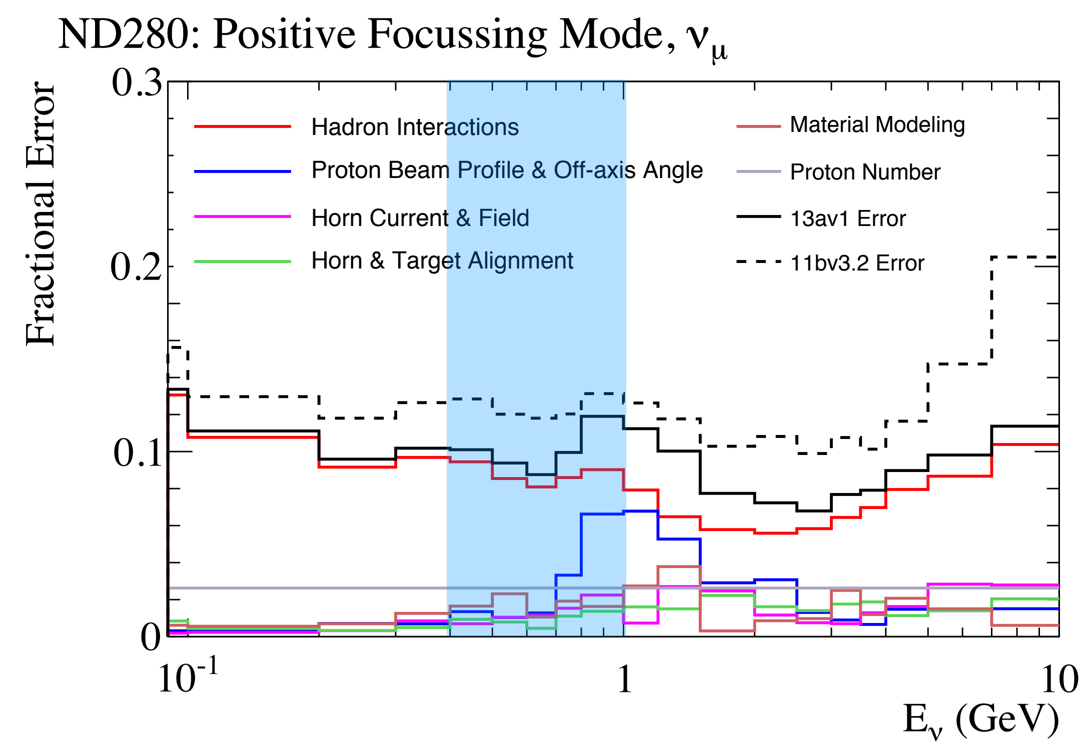

OA 0.0°
OA 2.0°
OA 2.5°
E_ν (GeV)
NA61/SHINE Data
INGRID/Beam Monitor Data
Flux Model
ND280 Detector Model
ND280 Data
External Cross Section Data
Cross Section Model
ν_l
l⁻
n
p
ND Fit
41m
SK Data
Oscillation Fit
SK Detector Model
Oscillation Parameters

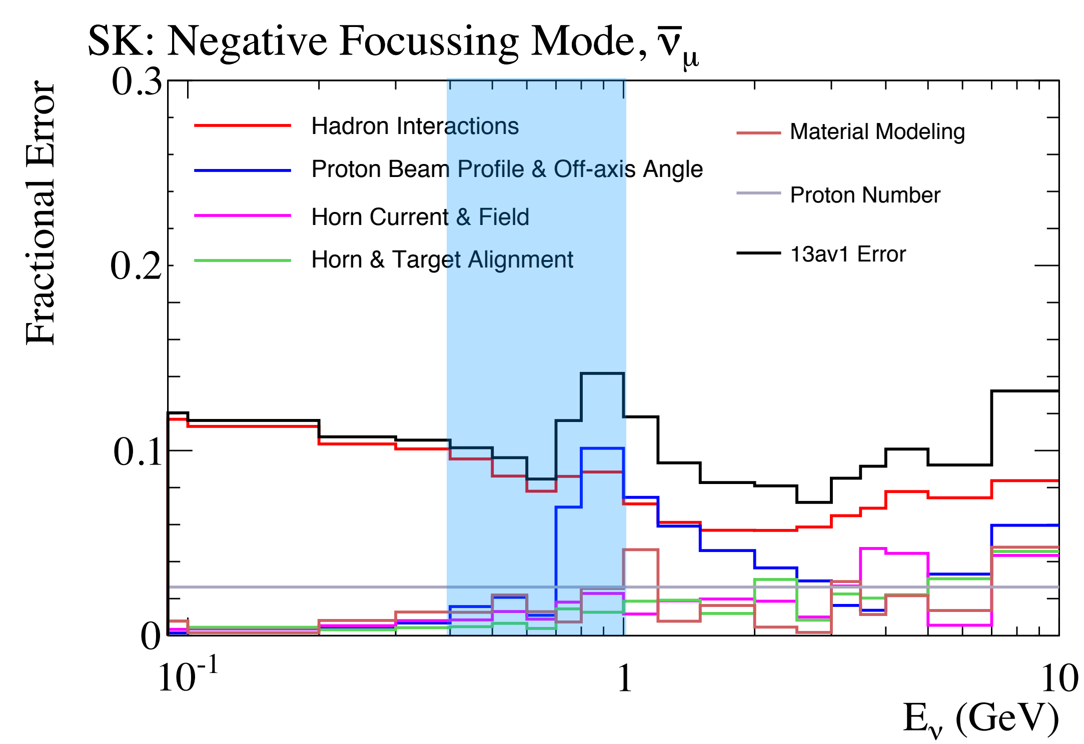

NA61/SHINE Data
INGRID/Beam Monitor Data
Flux Model
ND280 Detector Model
ND280 Data
External Cross Section Data
Cross Section Model
ν_l
l^-
n
p
OA 0.0°
OA 2.0°
OA 2.5°
E_ν (GeV)
41m
SK Data
Oscillation Fit
SK Detector Model
Oscillation Parameters

Beam Uncertainties

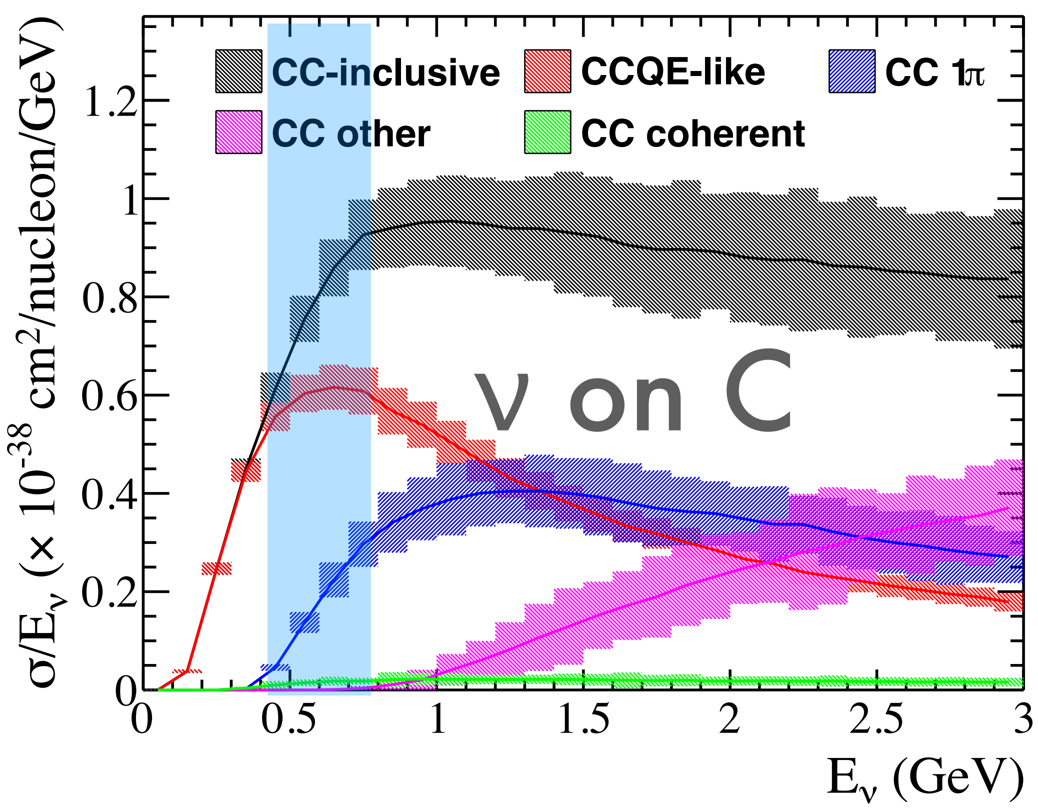


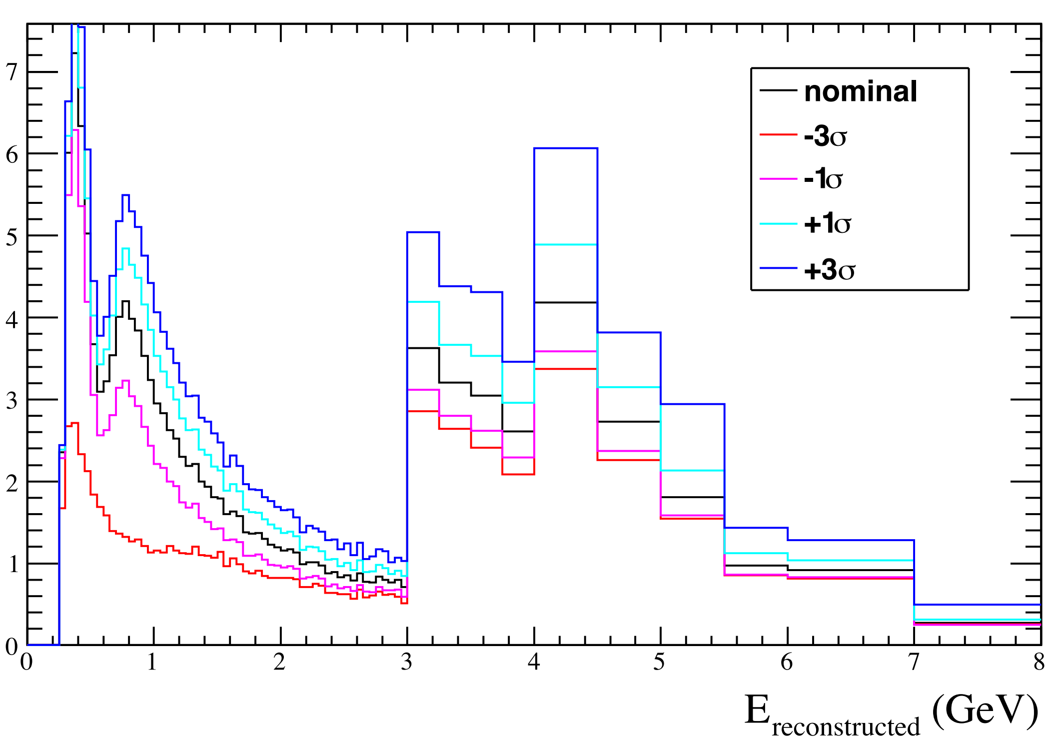






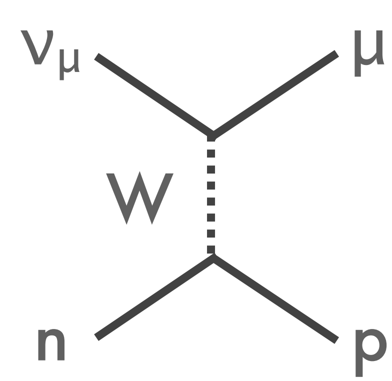
ν-N Cross Section Model





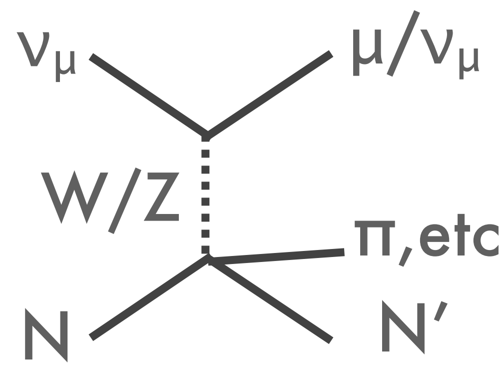
Uncertainties come from underlying model parameters and normalizations

Charged current quasi-elastic

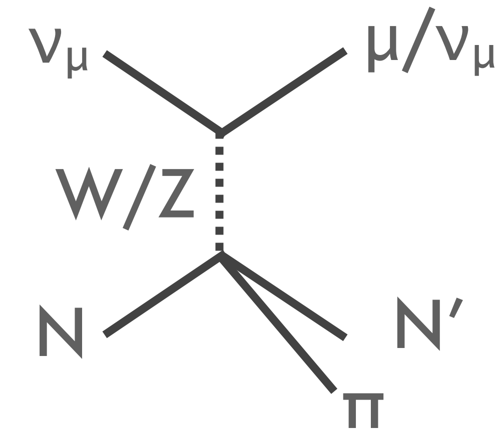


- MAQE
- MARES
- Fermi Mom.
- Binding Energy
- FSI
- 2p2h
- ...

Deep Inelastic Scattering

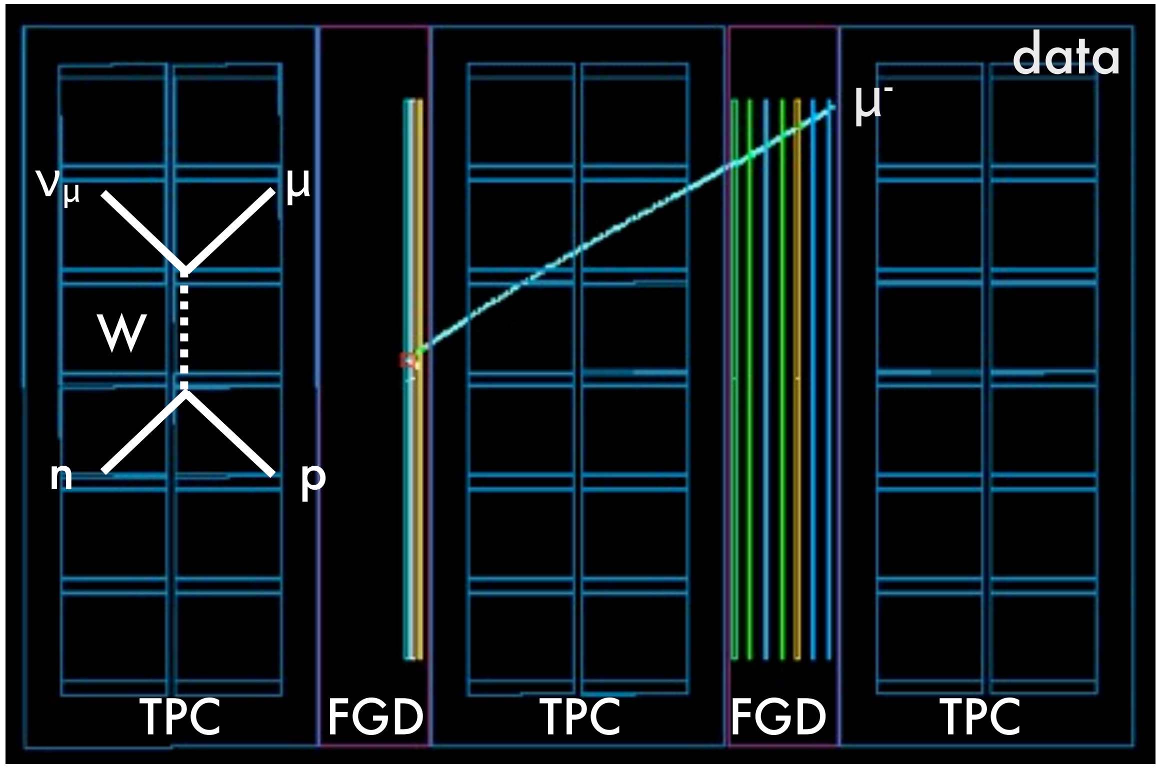


Charged Current 1π

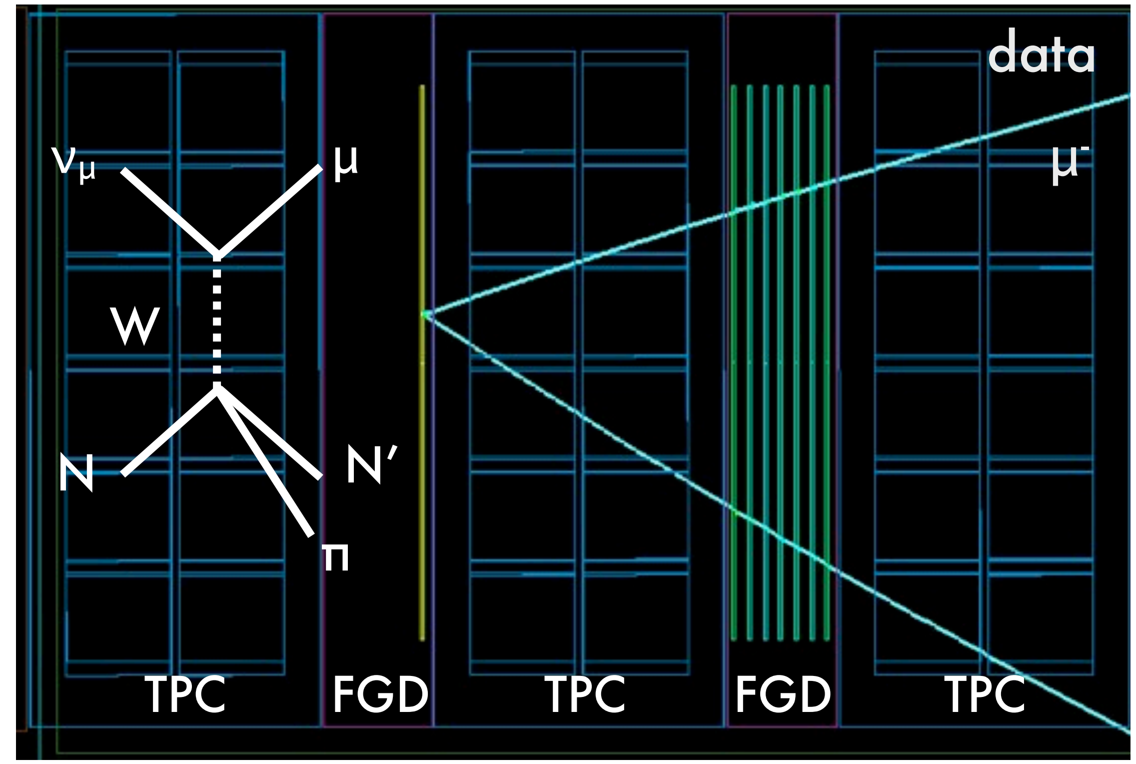


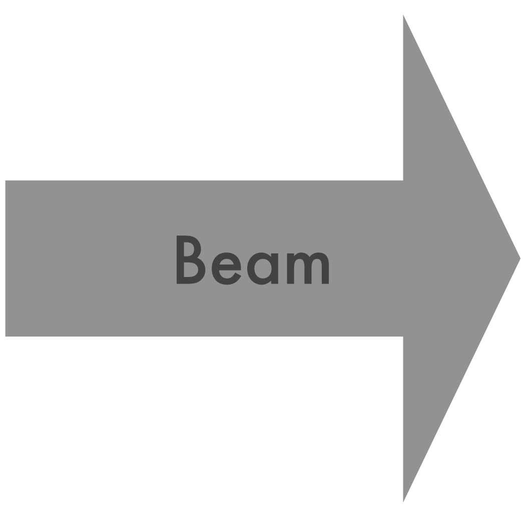
Talk by C. Wilkinson and poster by C. Wret on issues with external constraints

CC0π

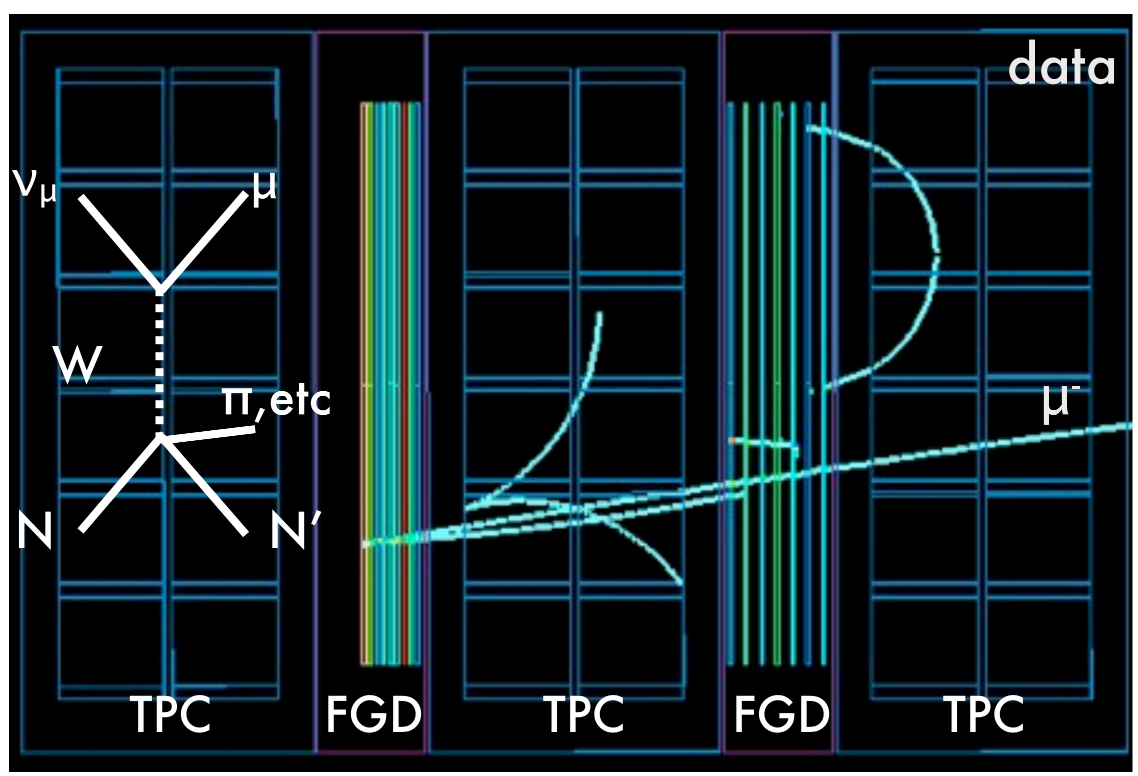


CC1π⁺



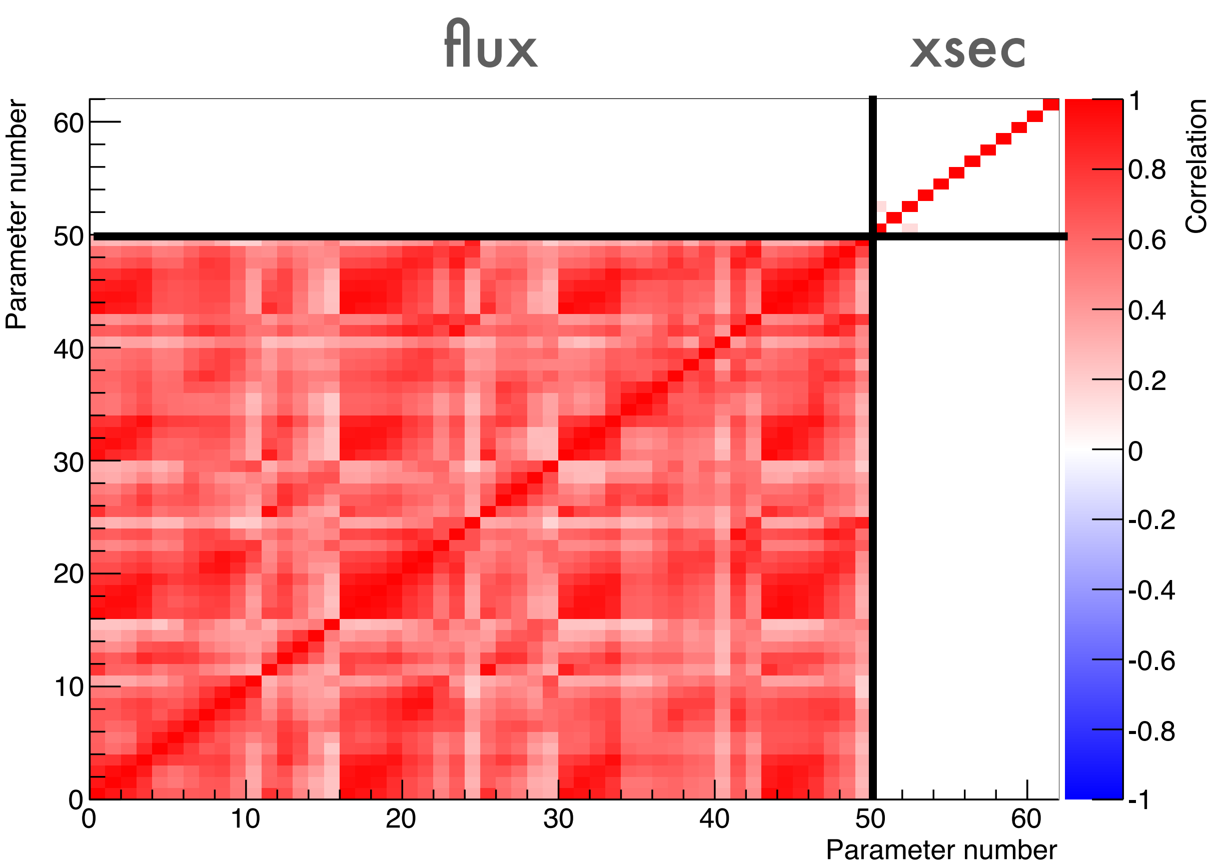


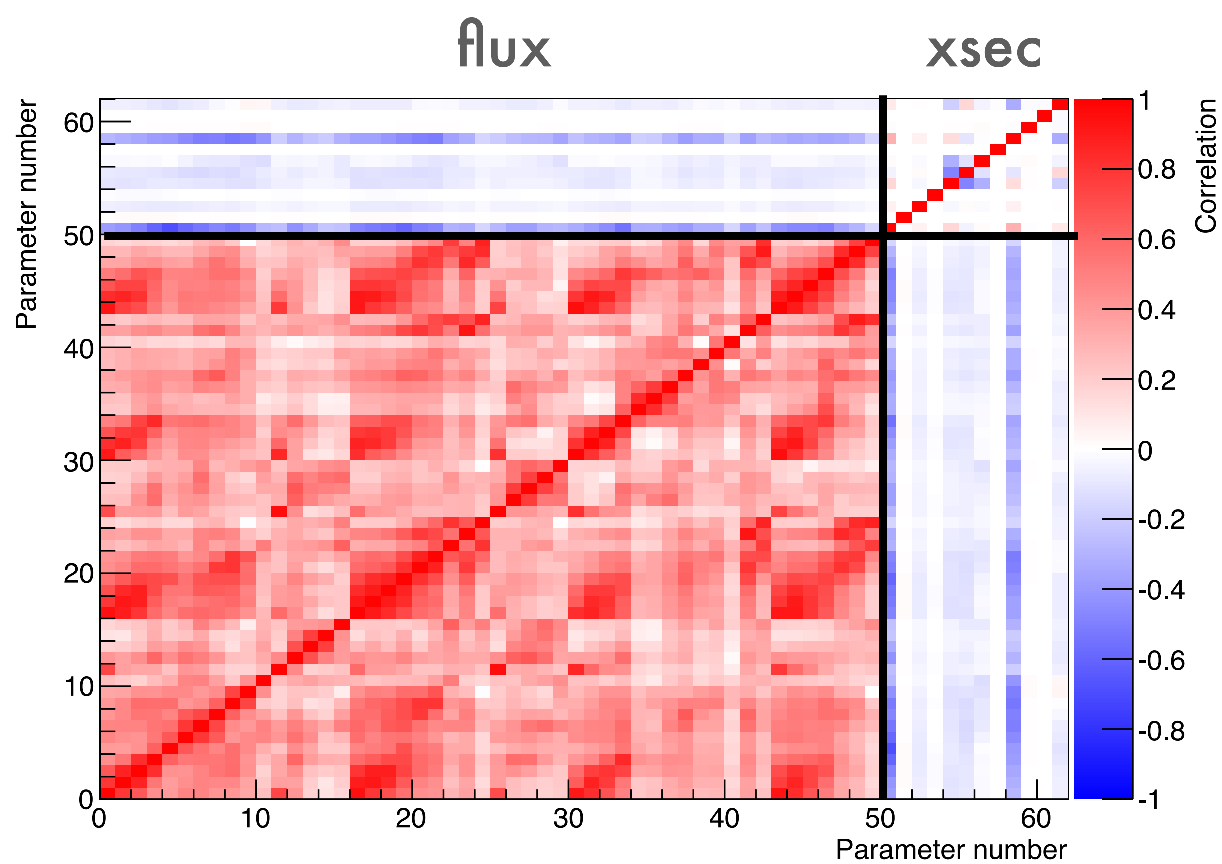
CC other



Near Detector Results

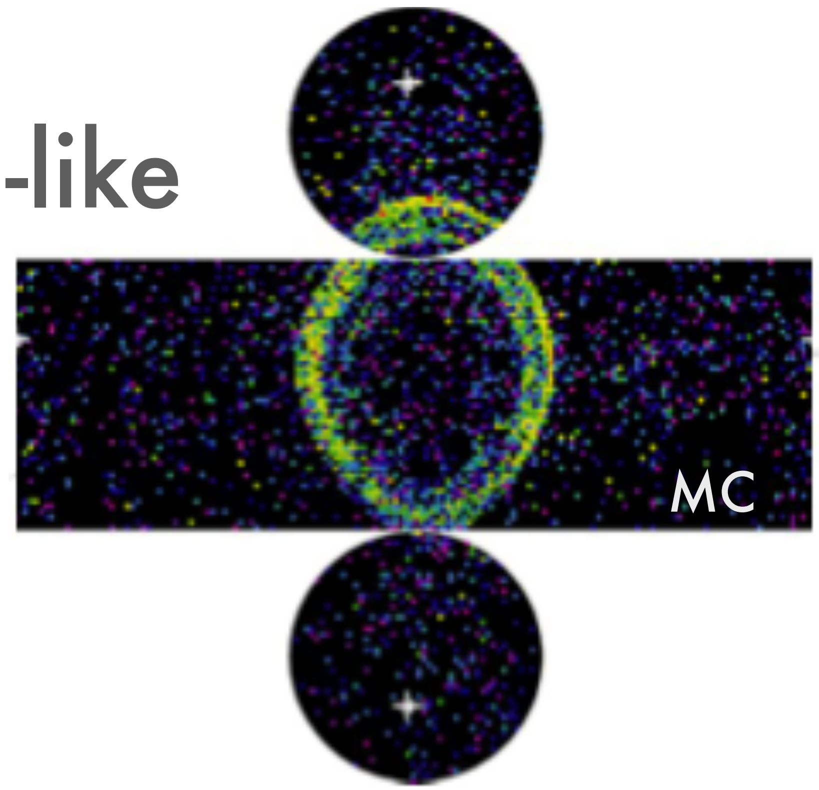
- Measure negative correlation between flux and cross section parameters
- Variety of analyses checks that gaussian assumptions are justified



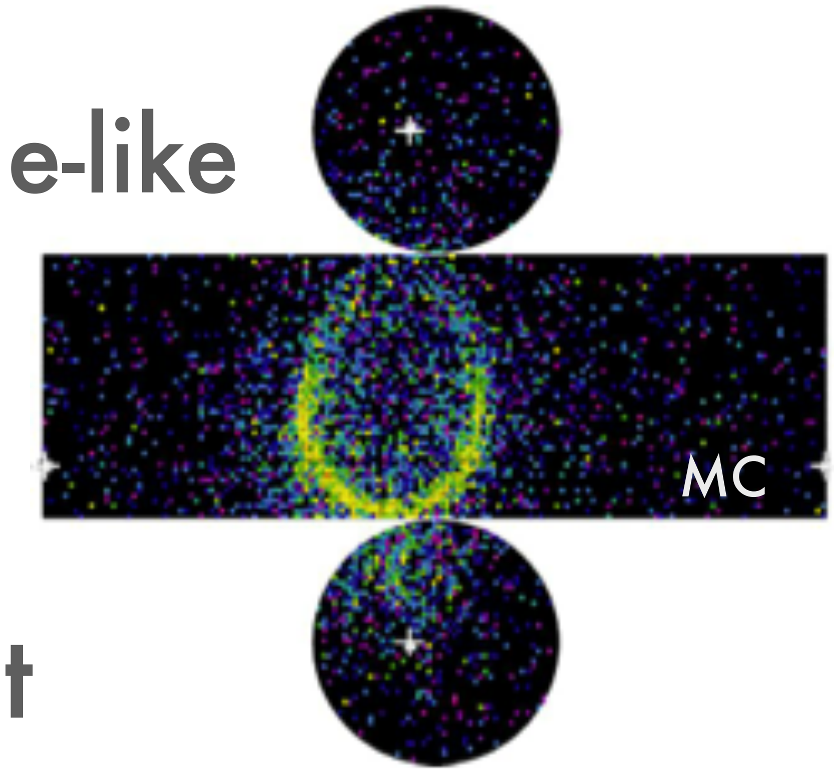


SK Data Samples

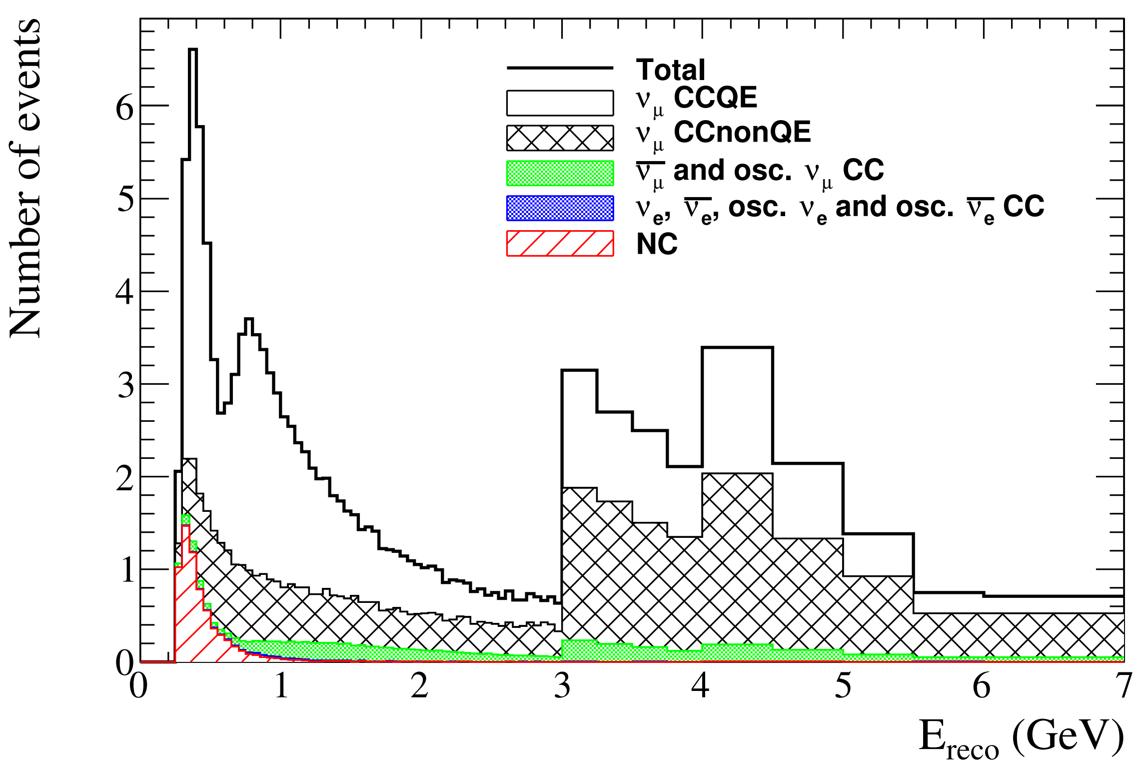
μ-like

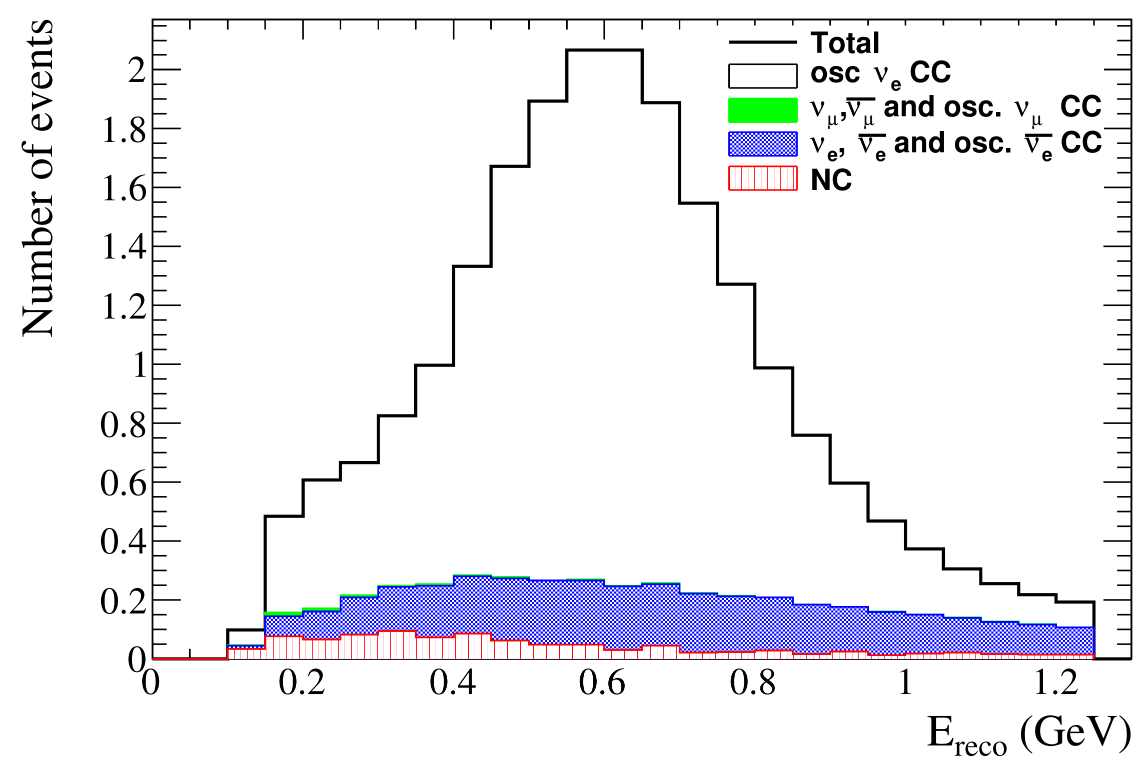


e-like



Binned ML Fit



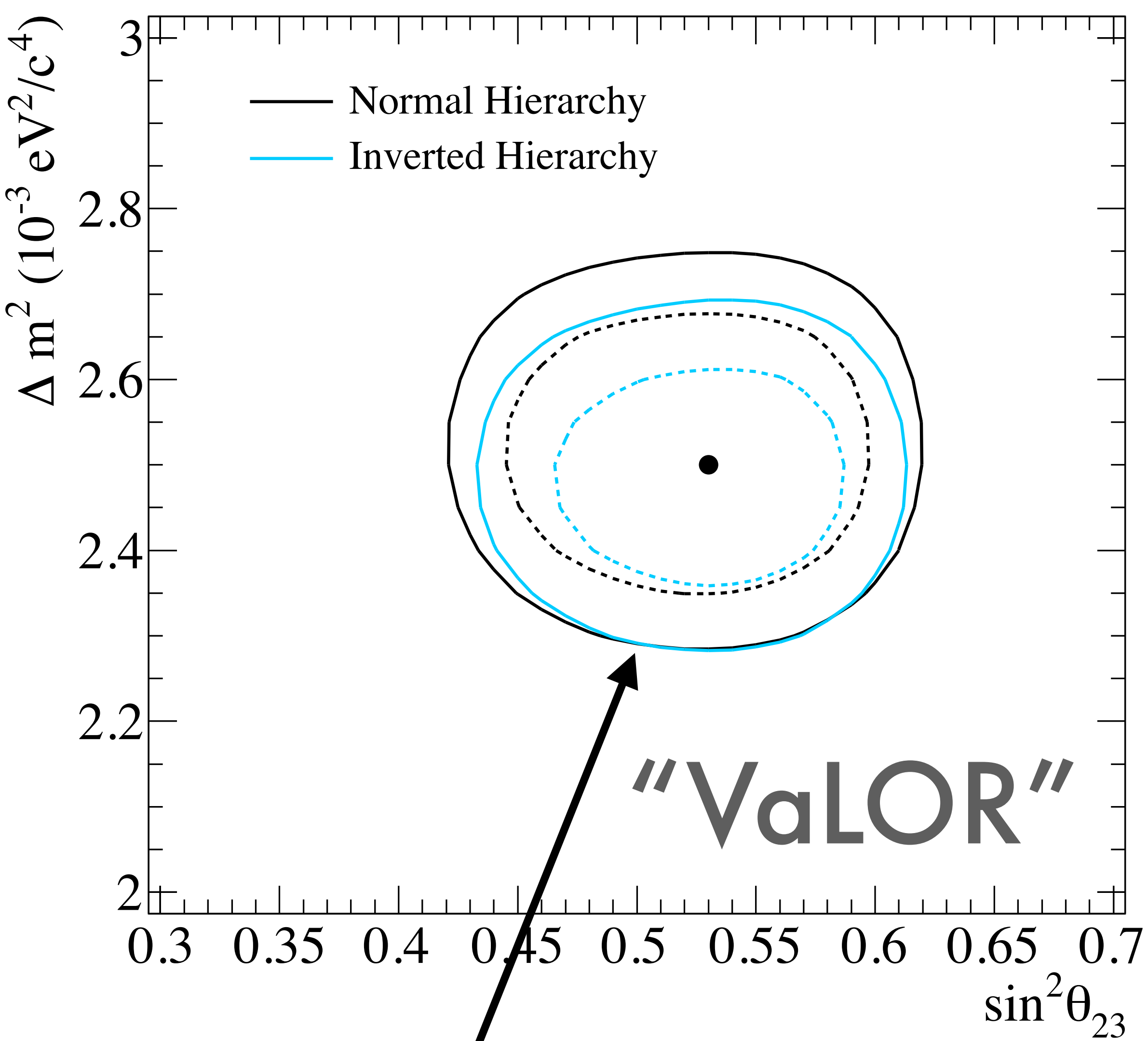


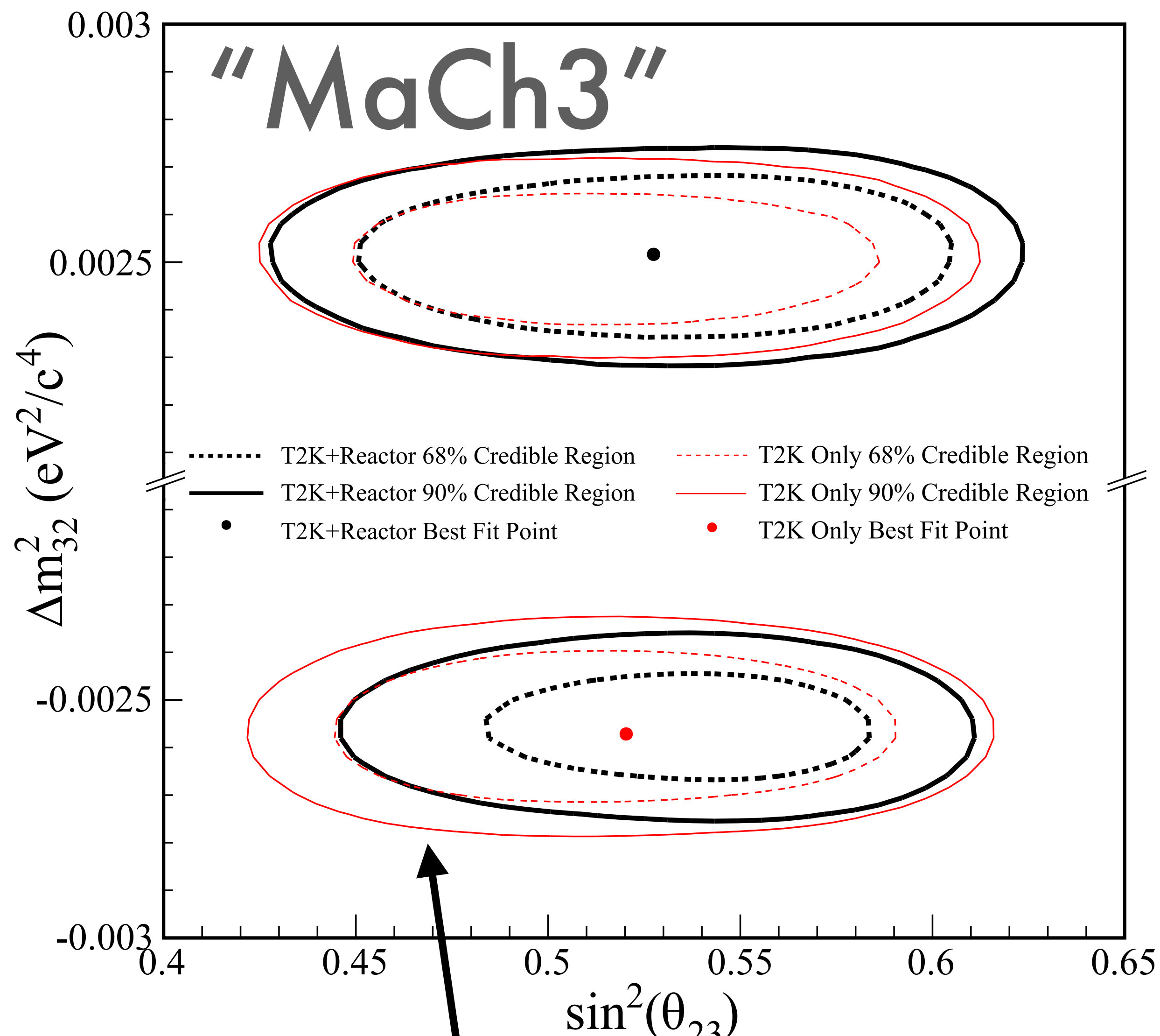
Abe, K., et al. "Measurements of neutrino oscillation in appearance and disappearance channels by the T2K experiment with 6.6× 10^{20} protons on target." *Physical Review D* 91.7 (2015): 072010.

Results

See posters by: L. Haegel, D. Sgalaberna, and R. Shah

T2K loves all statistical methods!



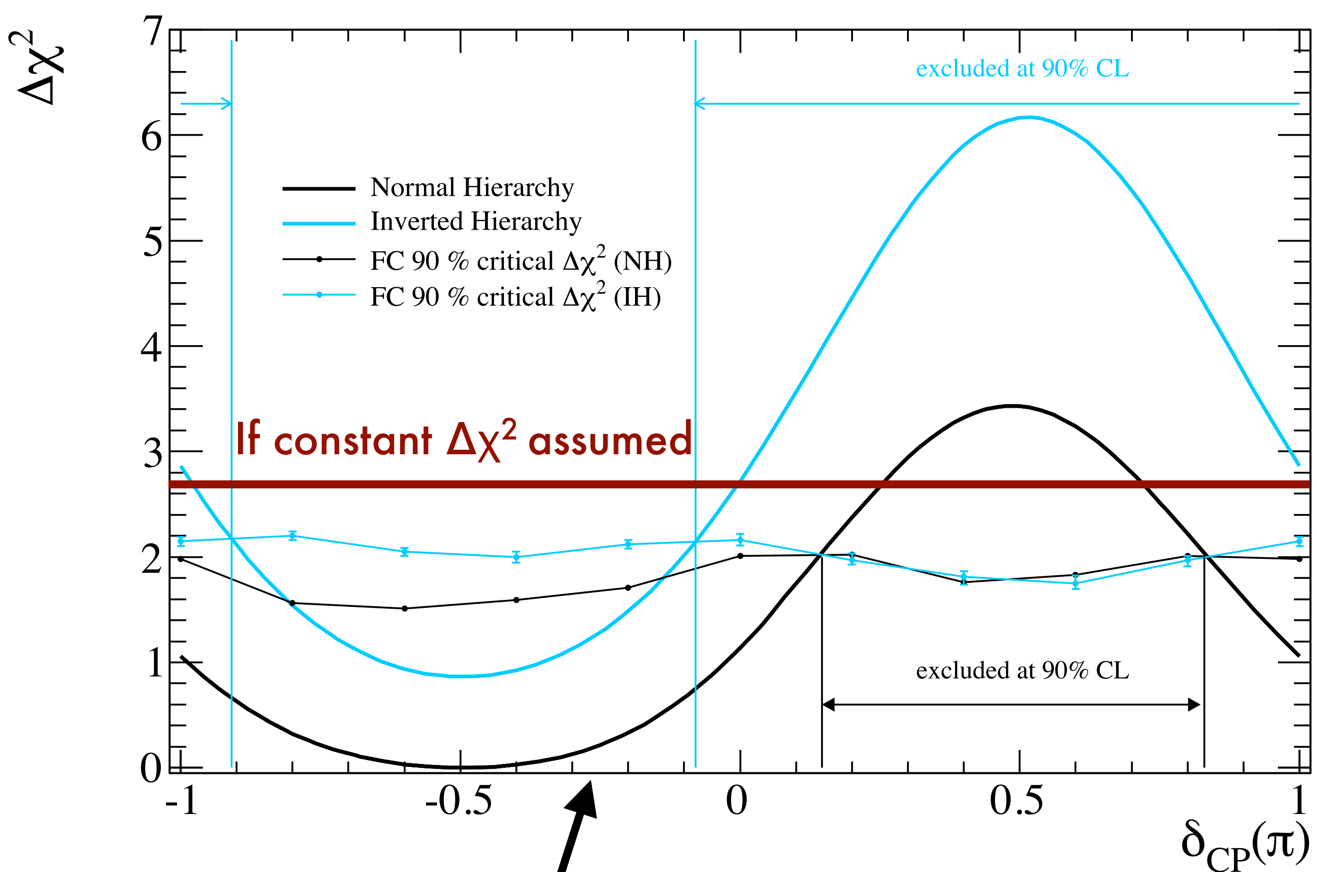


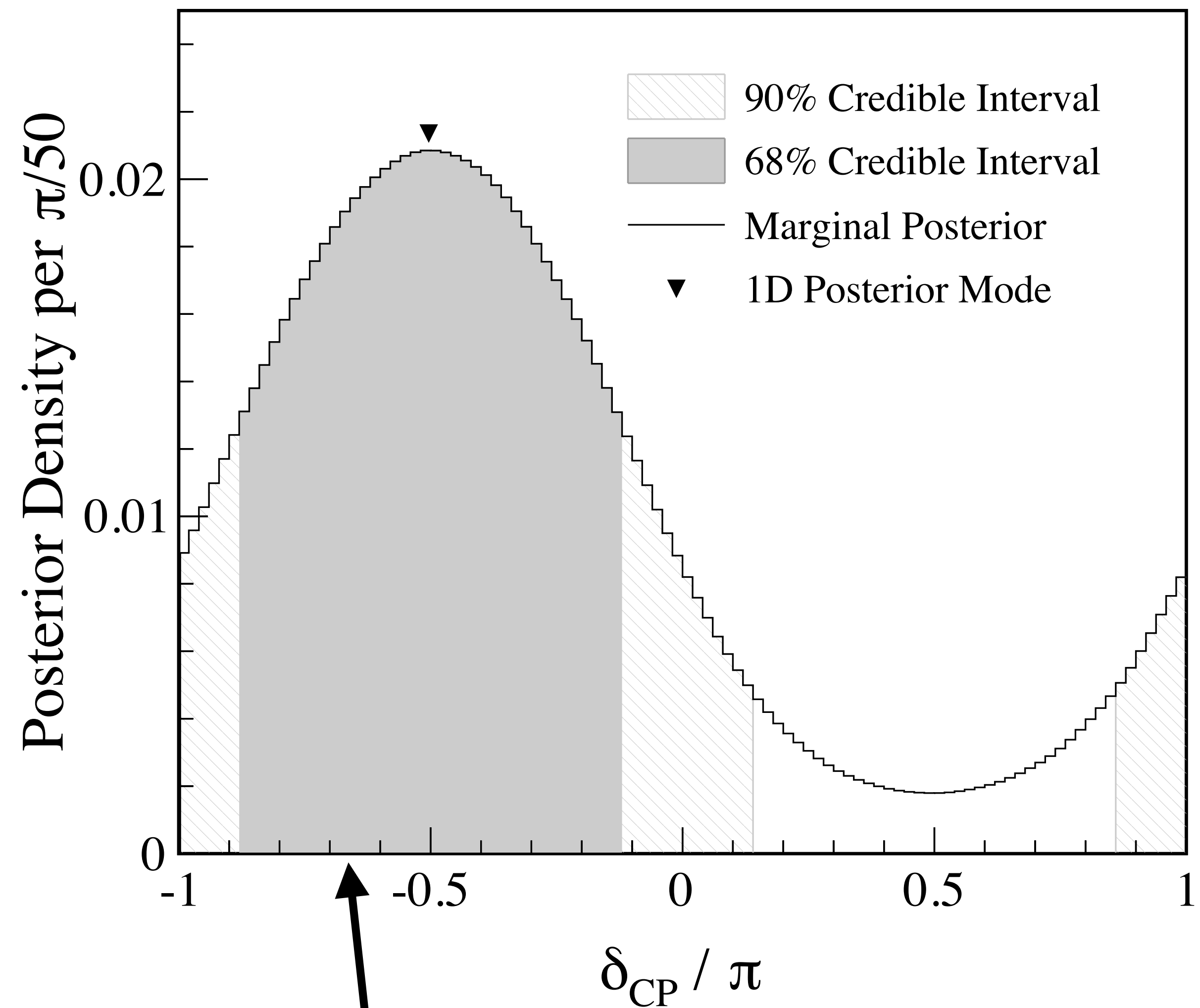
confidence level contours
constant $\Delta\chi^2$
Profiling treatment of systematics
Gradient descent

credible interval contours
highest posterior density method
Marginalization treatment of systematics
Markov Chain MC (Metropolis-Hastings)

Results

See posters by: L. Haegel, D. Sgalaberna, and R. Shah





confidence level interval
Feldman-Cousins-Highland
separate for NH/IH

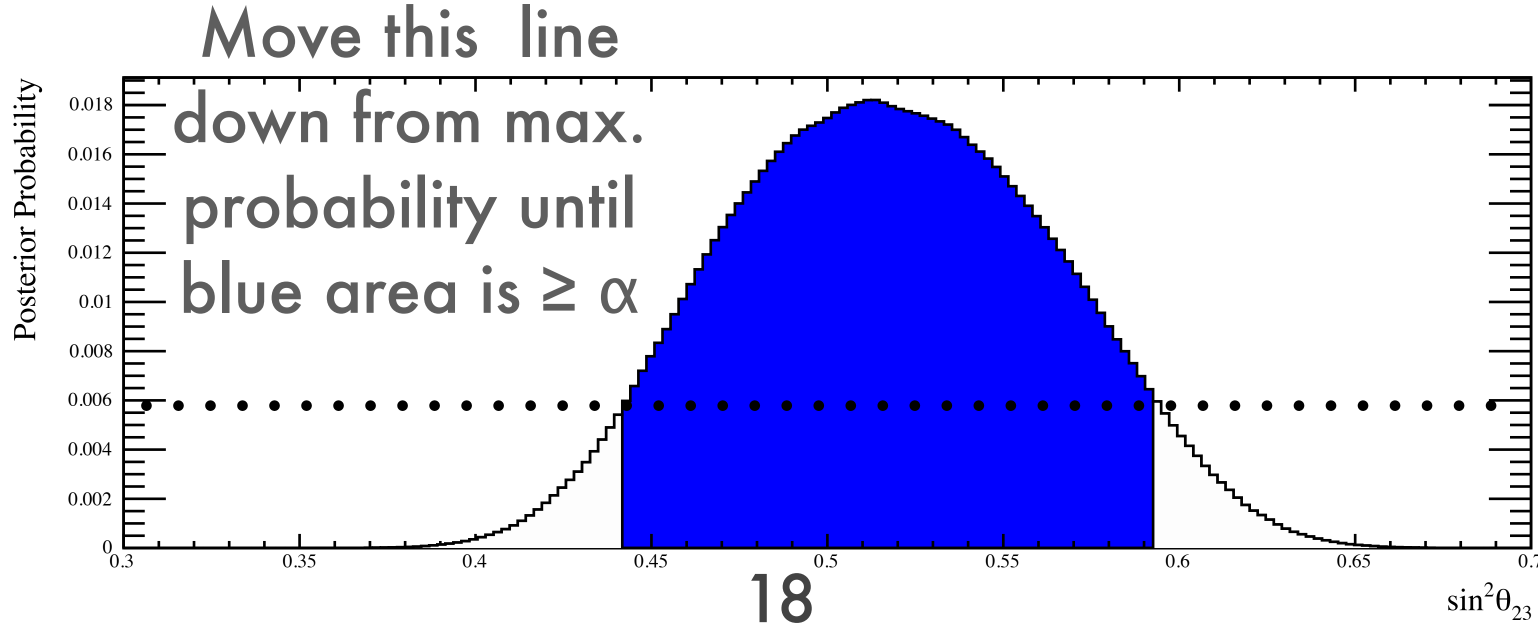
credible interval interval
highest posterior density method
marginalized over hierarchy

A Brief Review of F-C-H

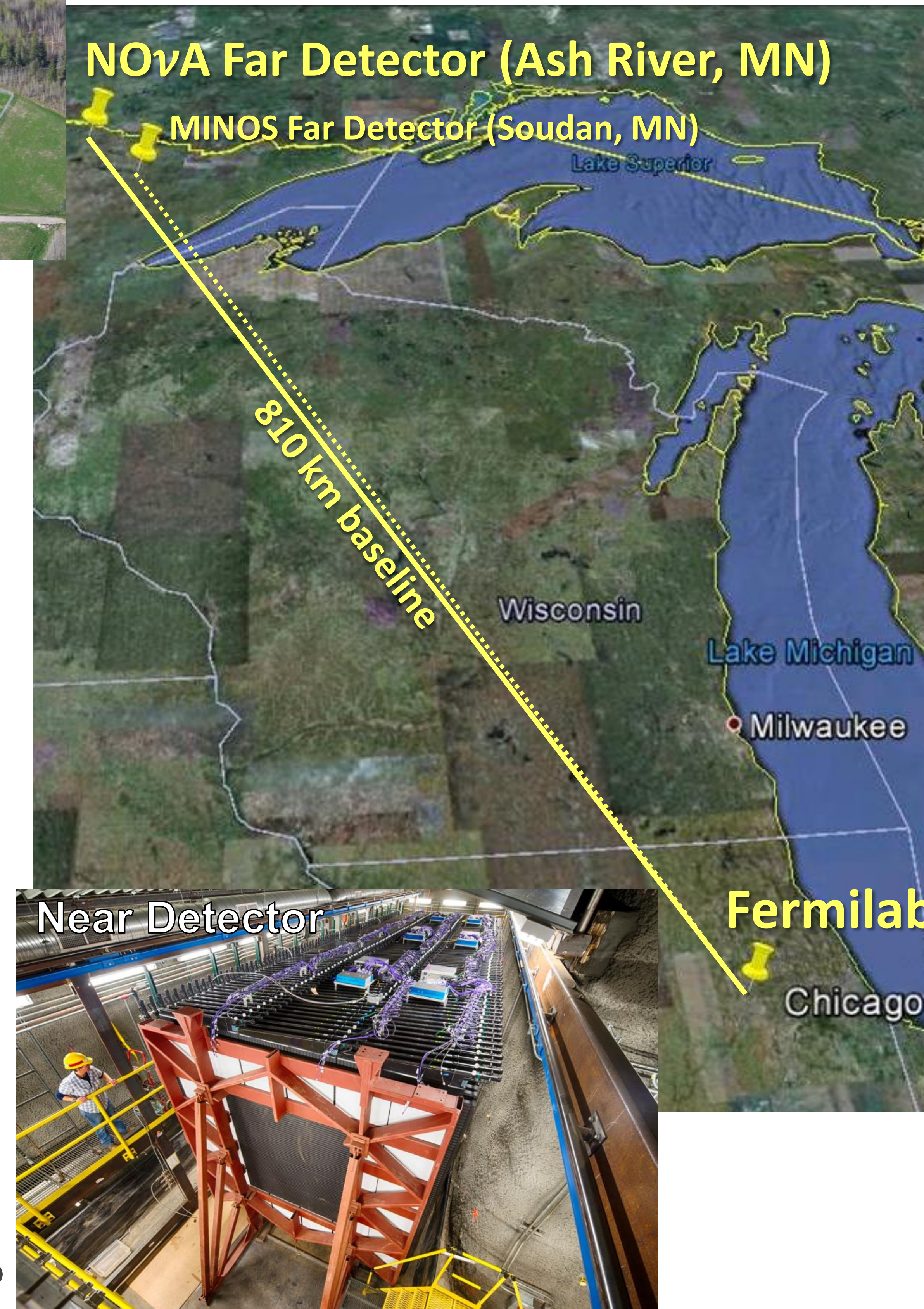
- Scan over your parameter space
- At every point, generate many toy MC experiments, including systematic parameter variations
- Fit your toys and generate the distribution of the test statistic
- Find the value of the test statistic where α% are excluded; this is the critical value of the test statistic at that point in the parameter space
- For the distribution of your data test statistic over your parameter space, draw a contour to exclude those where the data test statistic is greater than the critical value

A Brief Review of HPD Credible Intervals

- Build your posterior in your parameter space of interest
- Draw your contour such that the integral of the probability inside your contour is α%, and such that the posterior density at any point inside the contour is higher than any point outside the contour
- This minimizes the size of the contour

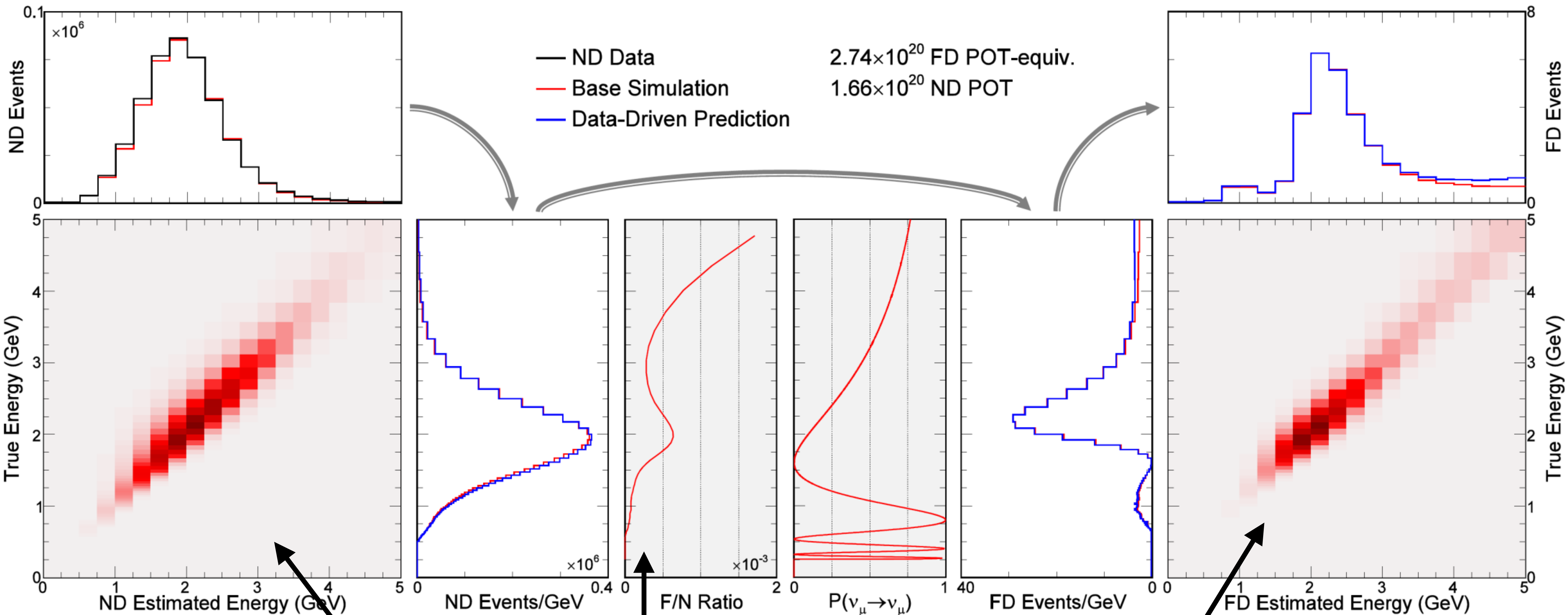


NOvA


NOνA Far Detector (Ash River, MN)
MINOS Far Detector (Soudan, MN)
Lake Superior
810 km baseline
Wisconsin
Lake Michigan
Milwaukee
Fermilab
Chicago
Near Detector

Select CC inclusive

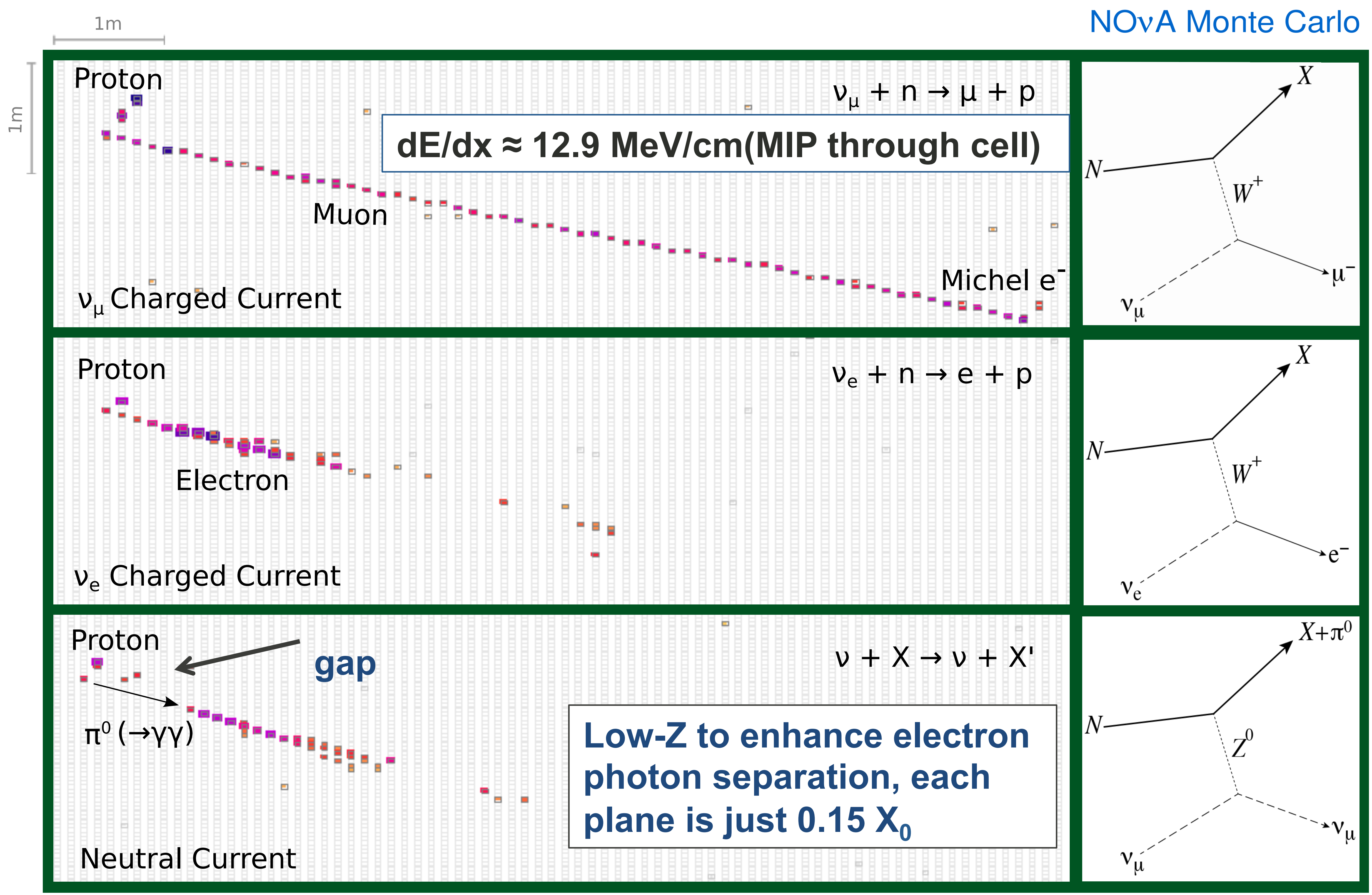
Select CC inclusive



Estimated from simulation

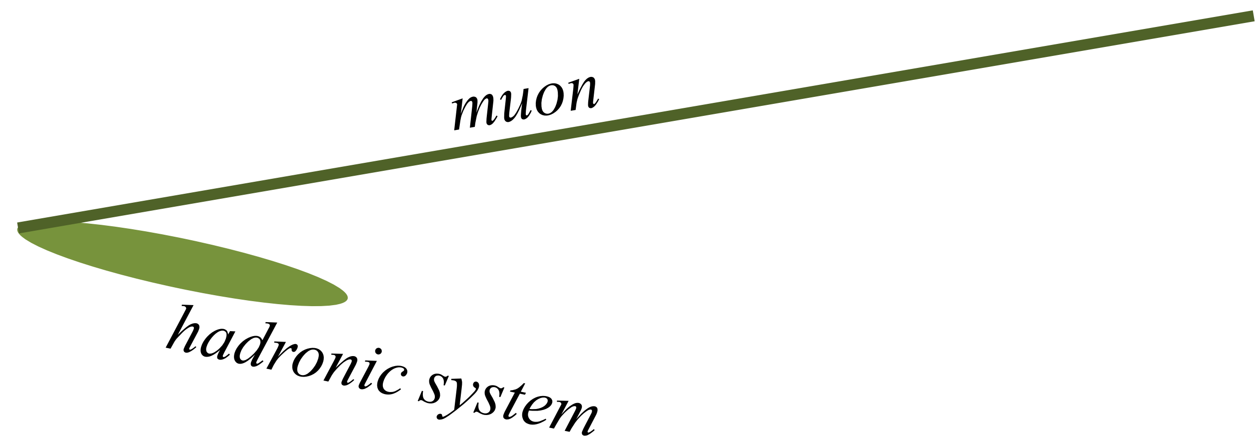
Use a binned ML method

R. Patterson


NOvA Monte Carlo
1m
1m
Proton
Muon
Michel e-
ν_μ Charged Current
$\nu_\mu + n \rightarrow \mu + p$
dE/dx ≈ 12.9 MeV/cm(MIP through cell)
N
X
W^+
ν_μ
μ^-
Proton
Electron
ν_e Charged Current
$\nu_e + n \rightarrow e + p$
N
X
W^+
ν_e
e^-
Proton
gap
π^0 ($\rightarrow\gamma\gamma$)
Neutral Current
$\nu + X \rightarrow \nu + X'$
Low-Z to enhance electron photon separation, each plane is just 0.15 X_0
N
$X+\pi^0$
Z^0
ν_μ
ν_μ

R. Patterson

Energy estimation



Reconstructed muon track:

length $\Rightarrow E_\mu$

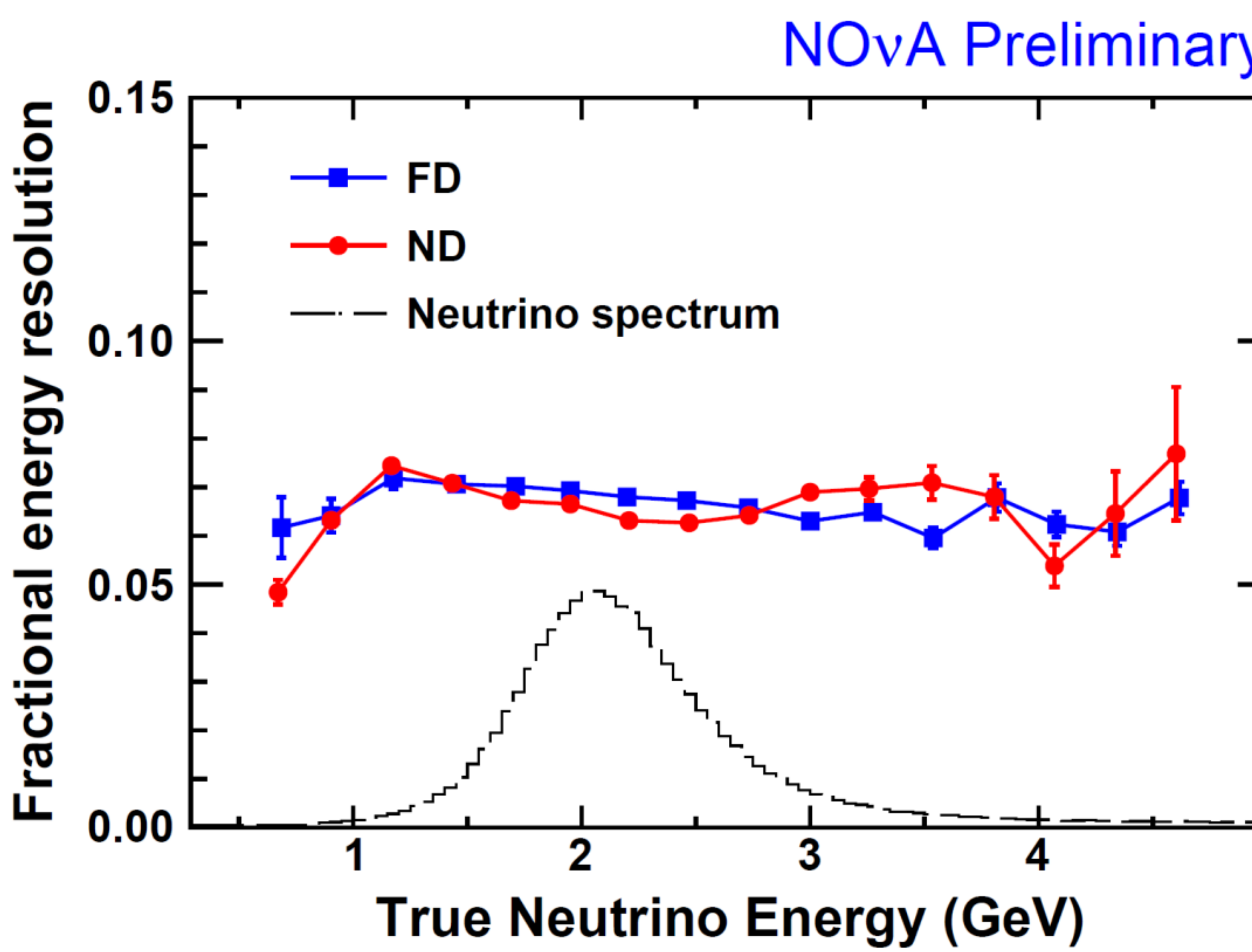
Hadronic system:

$\sum_{\text{cells}} E_{\text{visible}} \Rightarrow E_{\text{had}}$

Reconstructed ν_μ energy is the sum of these two:

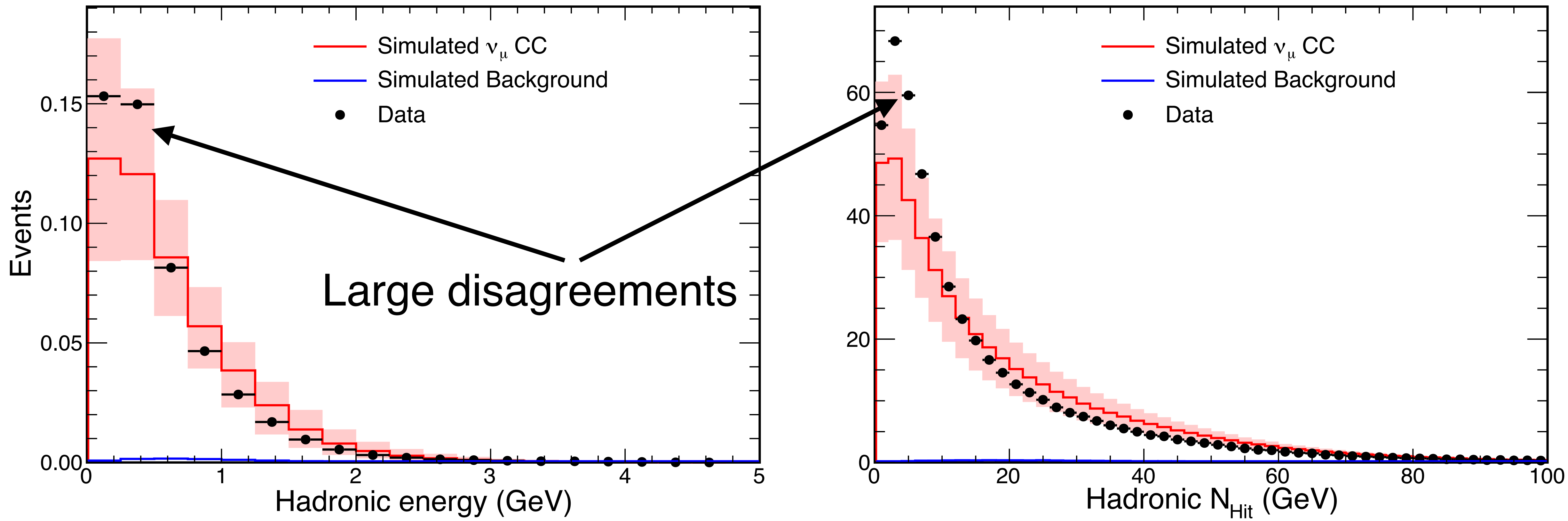
$$E_\nu = E_\mu + E_{\text{had}}$$

Energy resolution at beam peak ~7%



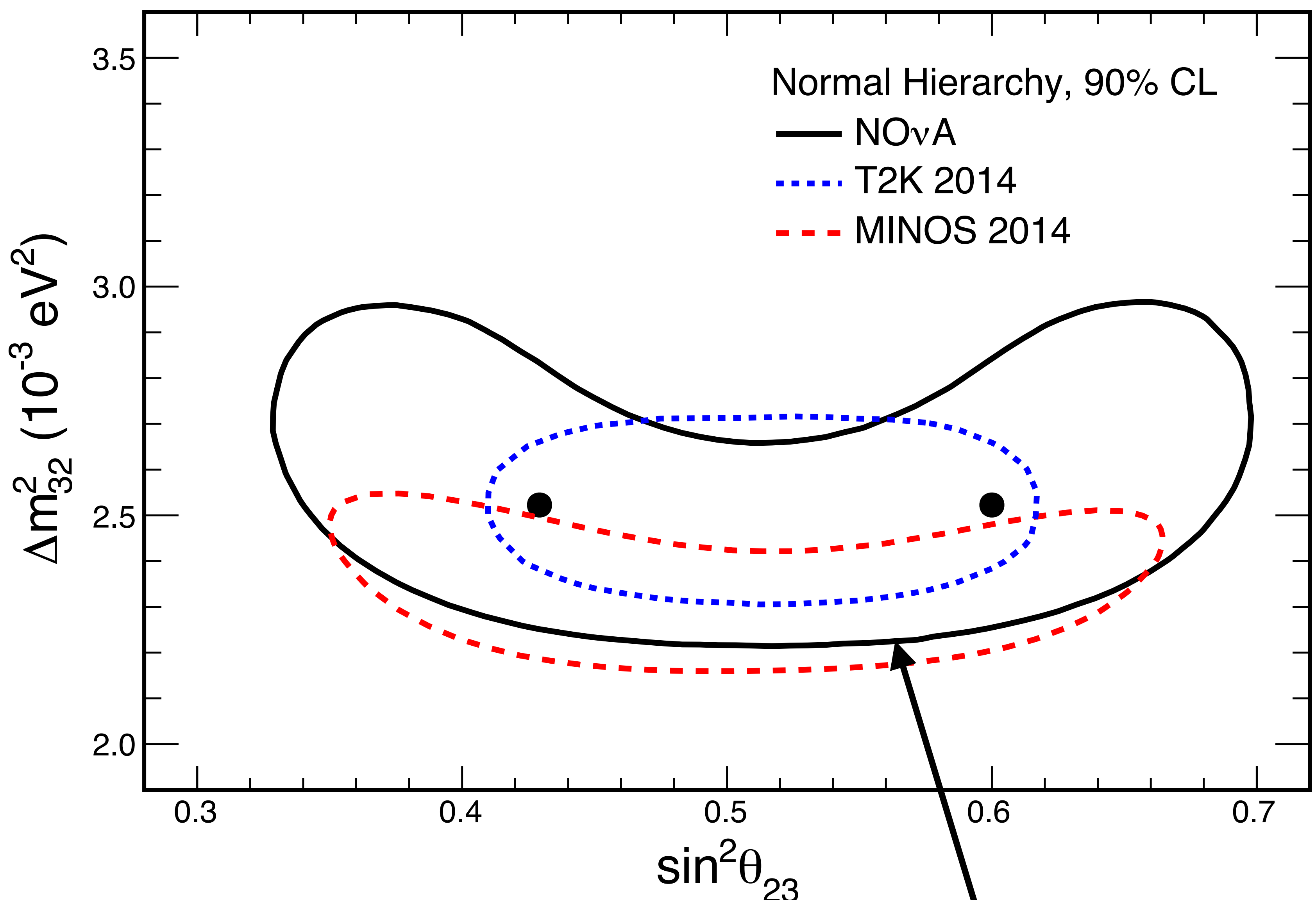
R. Patterson

Cross Section Uncertainties



- Change underlying parameters in neutrino model, similar to T2K
- NOvA uses a CC-inclusive sample, so individual modes are less important
- However, see a large uncertainty in the hadronic energy component which is currently a major systematics driver

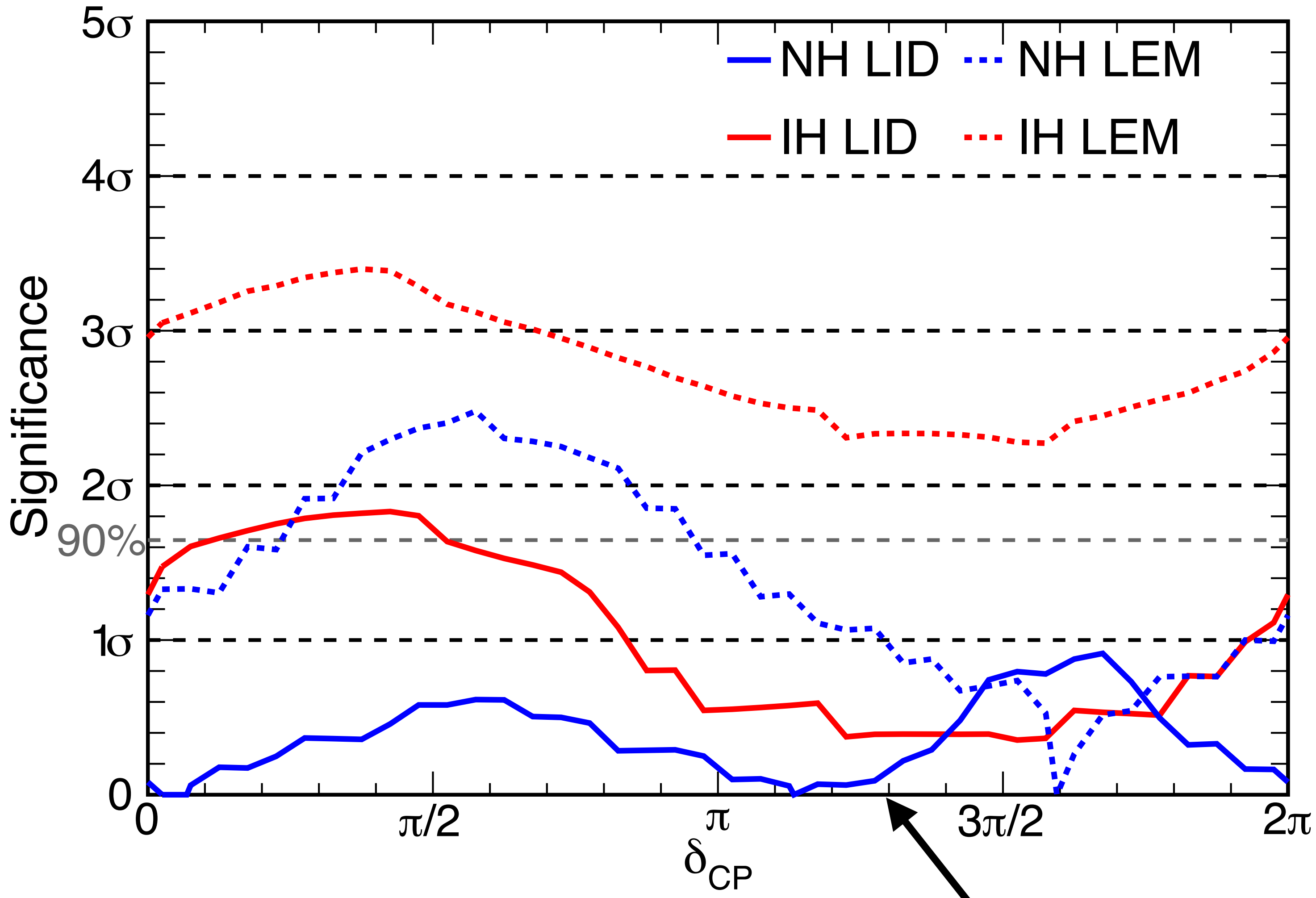
Results



confidence level contours
Feldman-Cousins
Profiling treatment of systematics
Gradient descent

Adamson, P., et al. "First measurement of muon-neutrino disappearance in NOvA." *arXiv preprint arXiv:1601.05037* (2016).

Results



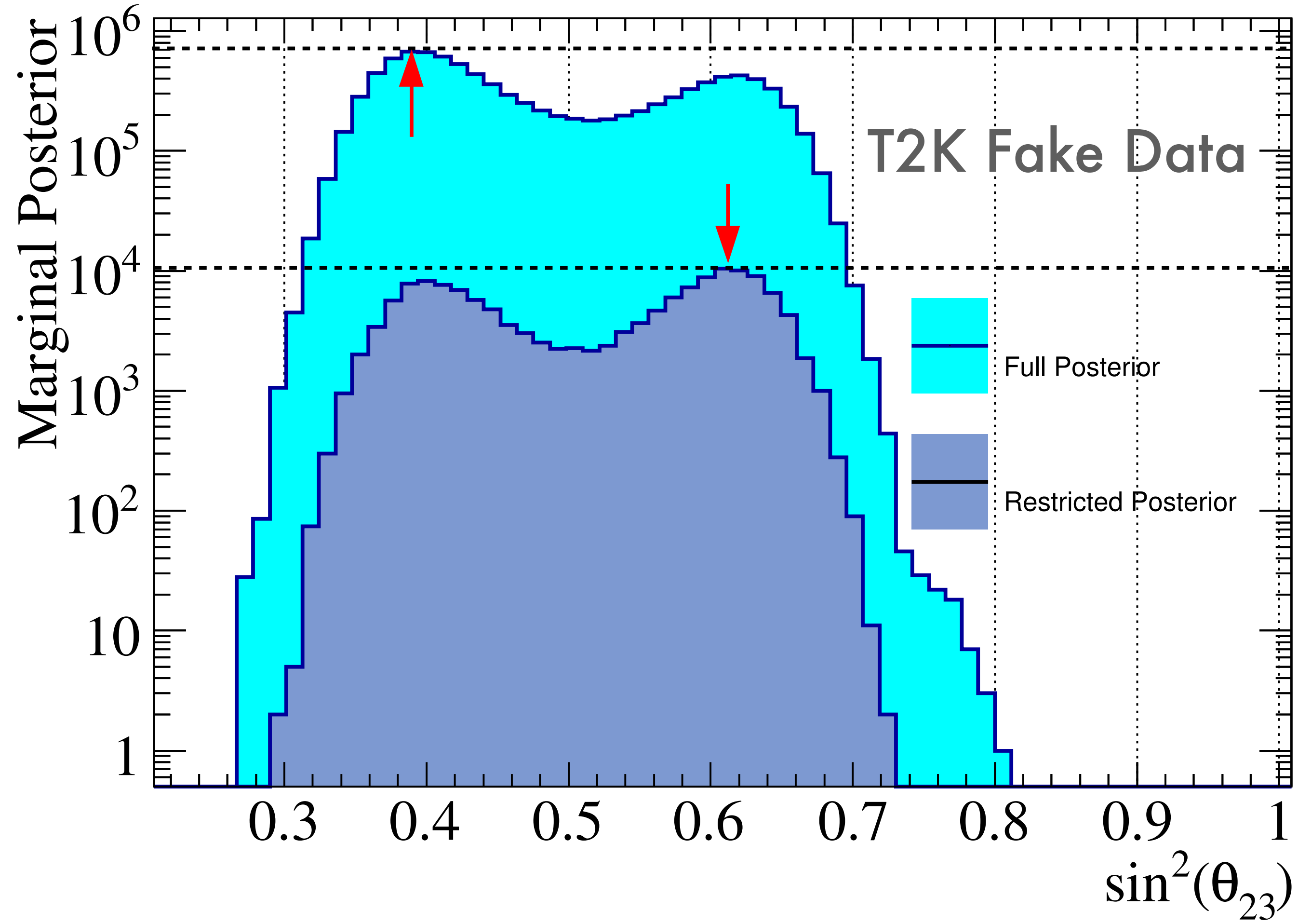
confidence levels
Feldman-Cousins
separate for NH/IH

Adamson, P., et al. "First measurement of electron neutrino appearance in NOvA." *arXiv preprint arXiv:1601.05022* (2016).

Special Issues

Marginalization vs Projection

- T2K found that treatment of systematics can have noticeable effects on analyses, especially when treating oscillation parameters, which are very non-gaussian
- T2K has moved to marginalization for all oscillation analyses, using both MCMC and likelihood averaging over toys

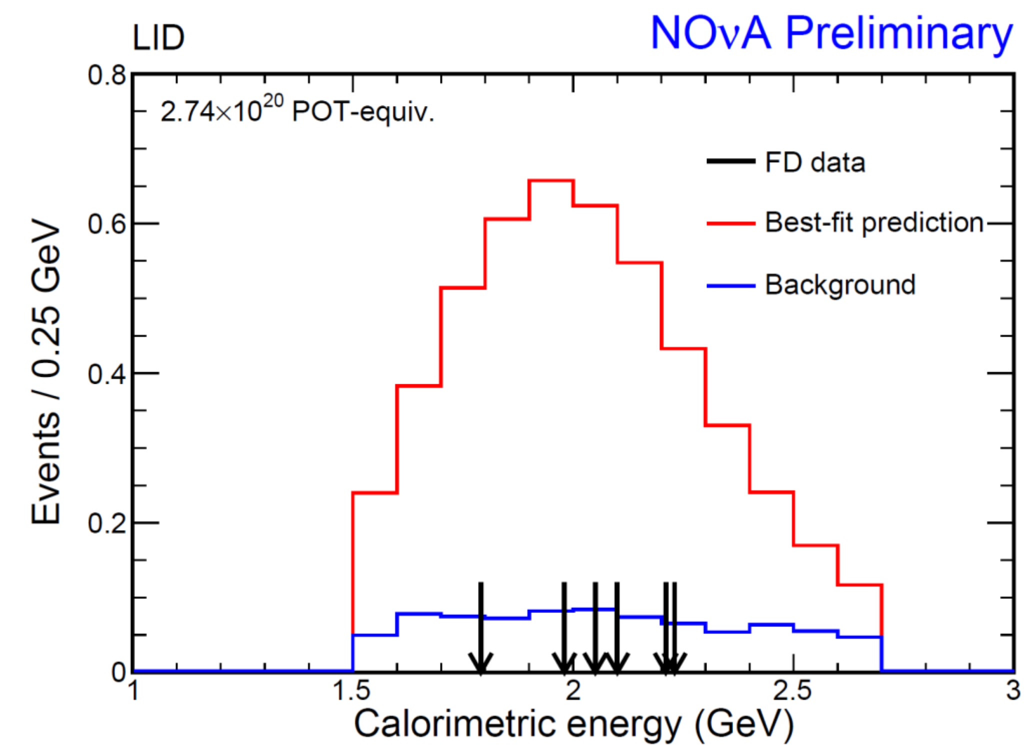


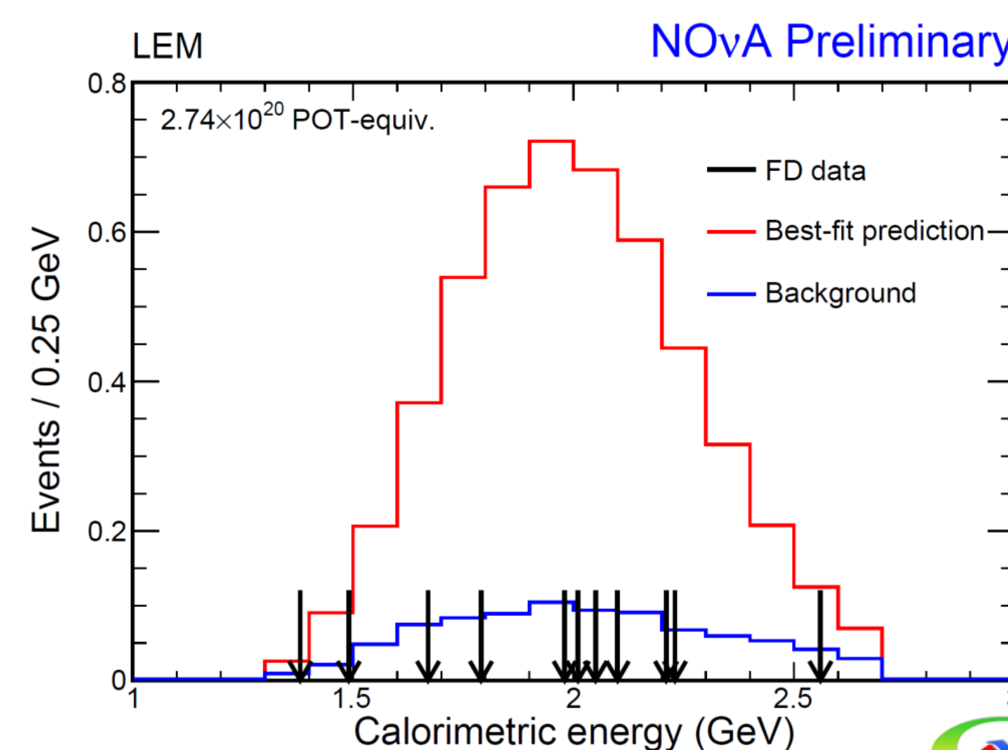
$$f_{marg}(x) = \int dy\, f(x, y)$$

$$f_{proj}(x) = \max_y f(x, y)$$

Selections with Tension

- NOvA has two event selections:
  - Likelihood Identification (LID) with 6 events
  - Library Event Matching (LEM) with 11 events
- All 6 LID events are in the LEM selection
- The predicted number of events for the two selections was roughly equal (5.6 vs 5.9)
- Should we be worried?





Selections with Tension

- Use MC to estimate p_1, p_2, and p_3 for signal and background separately, and calculate P for N_b=0,1,2,...
- What is the total probability for combinations which are less probable than the data combination?

Probability of LEM only

$n_2=5$

Probability of LID only

$n_1=0$

$$P = \frac{n!}{(n_1!)(n_2!)(n_3!)} p_1^{n_1} p_2^{n_2} p_3^{n_3}$$

Probability of LEM and LID

$n_3=6$

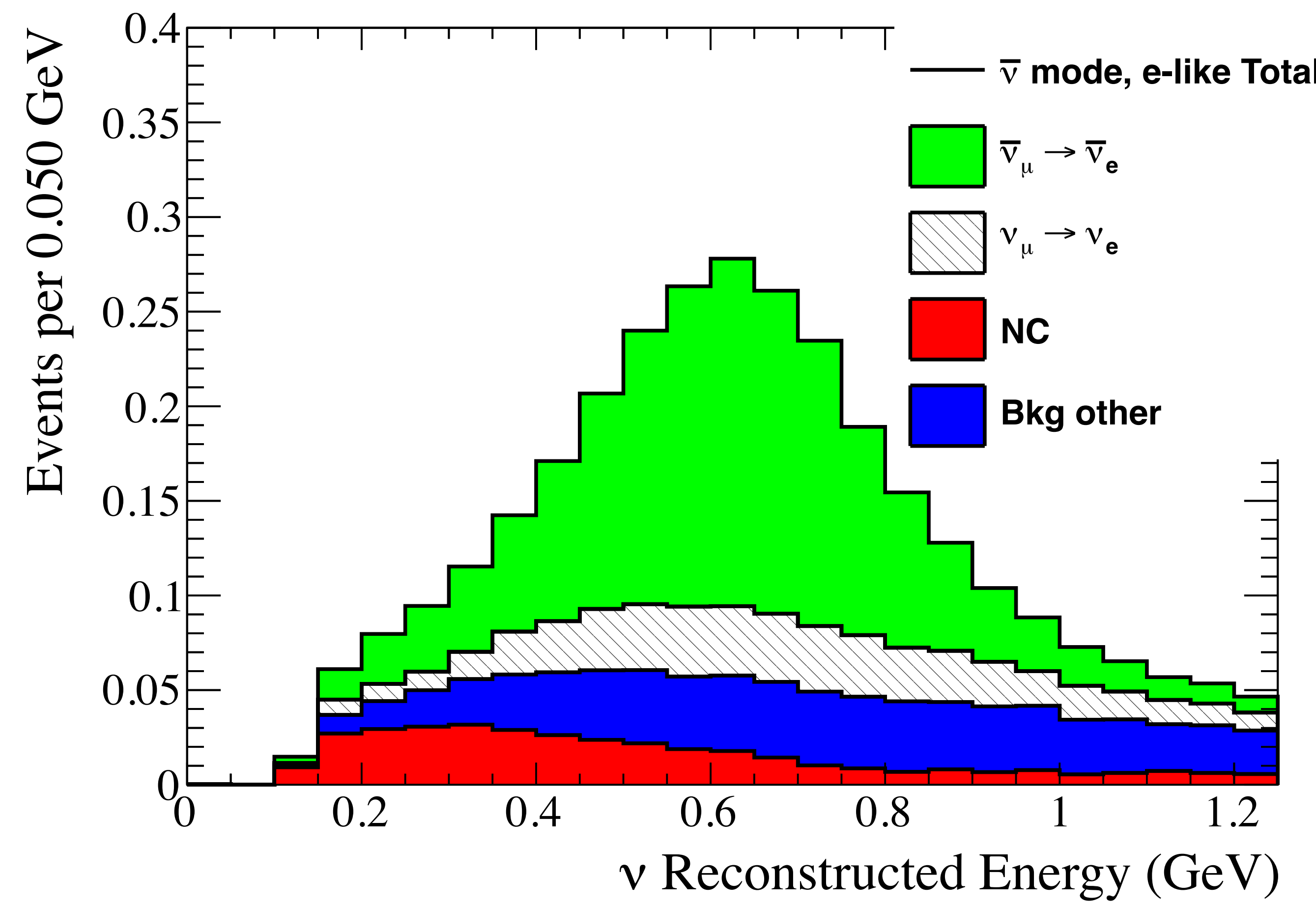
Selections with Tension

$$\begin{aligned}
B = 0, S = 11 &\quad \rightarrow \quad p_{1,B=0}, p_{2,B=0}, p_{3,B=0} \rightarrow P(0/5/6+|B=0) = 6\% \\
B = 1, S = 10 &\quad \rightarrow \quad p_{1,B=1}, p_{2,B=1}, p_{3,B=1} \rightarrow P(0/5/6+|B=1) = 7.1\% \\
B = 2, S = 9 &\quad \rightarrow \quad p_{1,B=2}, p_{2,B=2}, p_{3,B=2} \rightarrow P(0/5/6+|B=2) = 8.3\% \\
&\quad \ldots \\
&\quad \text{where} \\
p_{1,B=N_b} &\quad = \quad p_1^b N_b/11 + p_1^s(11-N_b)/11 \\
&\quad \text{and} \\
P(B=0) &\quad = \quad 20.9\% \\
P(B=1) &\quad = \quad 35.2\% \\
P(B=2) &\quad = \quad 26.9\% \\
&\quad \ldots \\
&\quad \text{so} \\
P &\quad = \quad \sum_B P(0/5/6 \text{ or worse}|B)P(B) = \boxed{7.8\%}
\end{aligned}$$

Uses central value of background prediction

Parameterizing $\nu/\bar{\nu}$ Differences

- After establishing ν_e appearance, T2K wanted to analyze $\bar{\nu}_e$ appearance
- How to build a model that is "minimally" CP-violating?
- Decided to introduce a parameter β that multiplies $P(\bar{\nu}_\mu \rightarrow \bar{\nu}_e)$ while keeping all other oscillation parameters the same between ν and $\bar{\nu}$
- Allow β to only take the values 0 (no $\bar{\nu}_\mu \rightarrow \bar{\nu}_e$ oscillation) and 1 ($\bar{\nu}_\mu \rightarrow \bar{\nu}_e$ oscillates as $\nu_\mu \rightarrow \nu_e$)



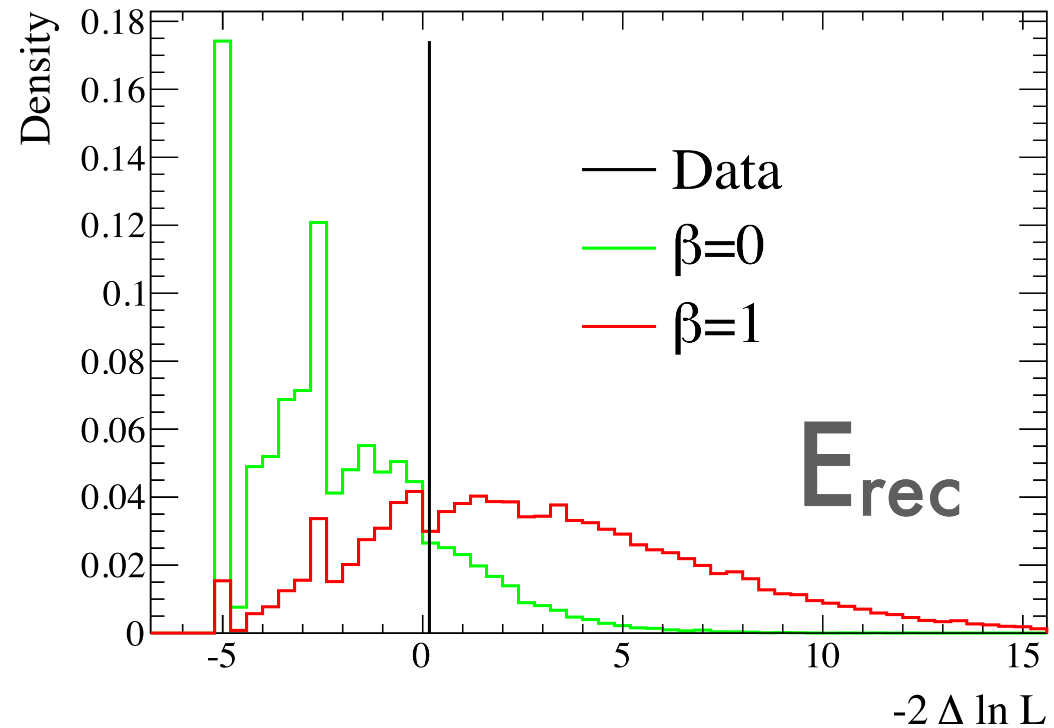
Parameterizing $\nu/\bar{\nu}$ Differences

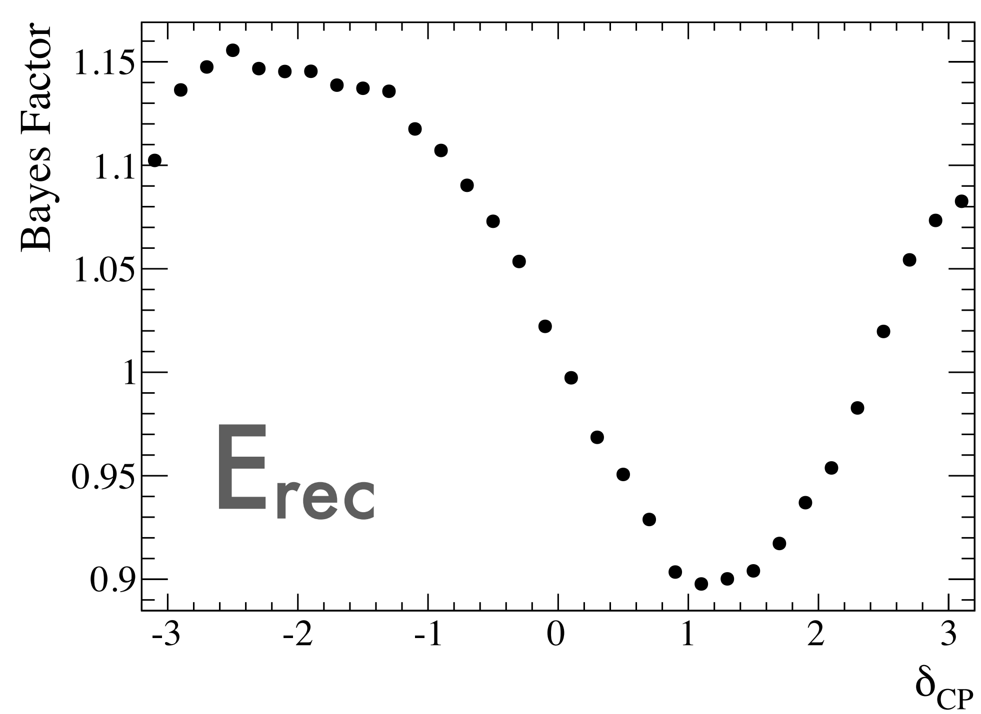
- Use a test statistic of

$$-2\left(\ln \mathcal{L}(\beta = 1) - \ln \mathcal{L}(\beta = 0)\right)$$

- Throw toy data sets for the cases β=0 and β=1 and compare what data looks like

- Also calculate a Bayes factor

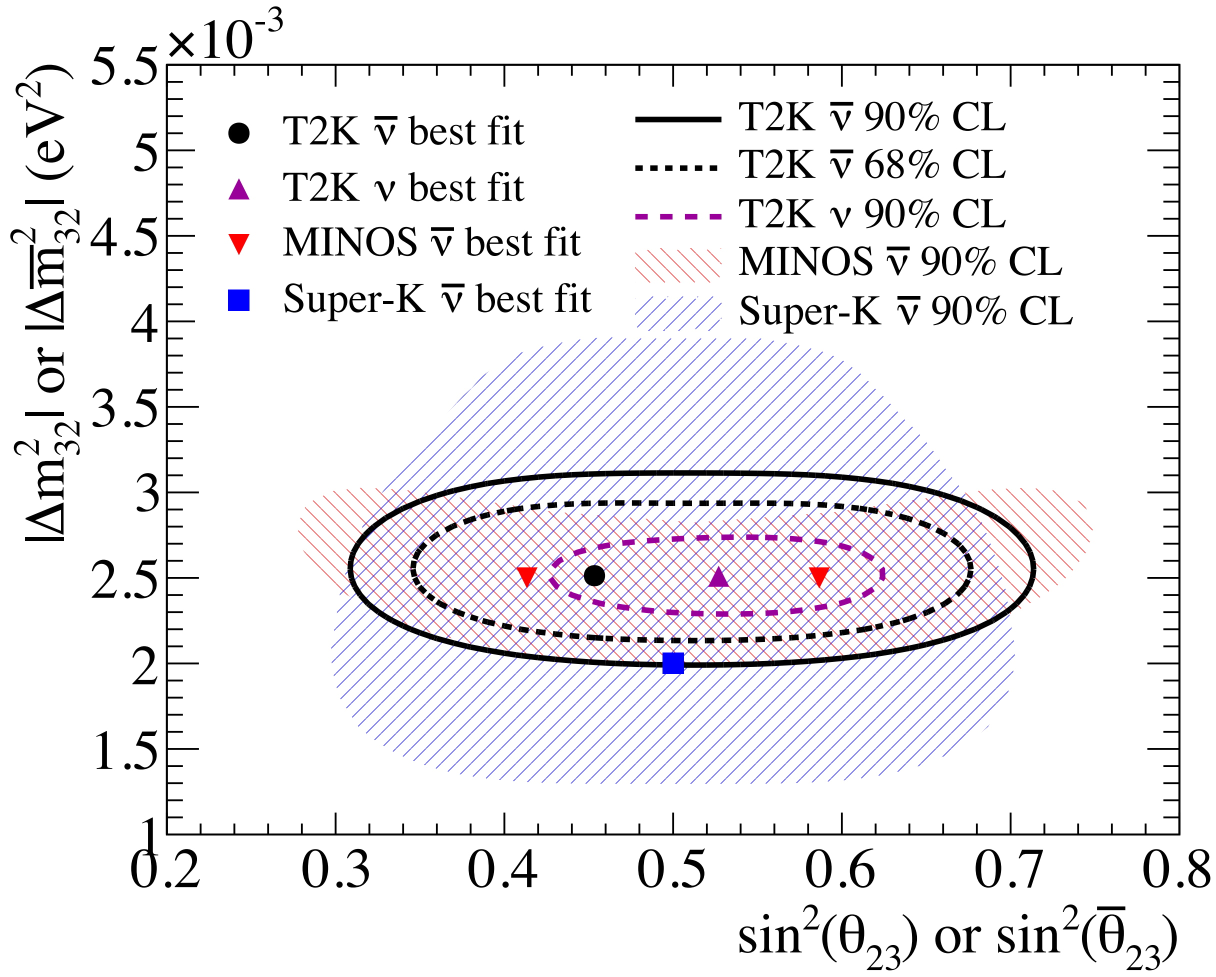




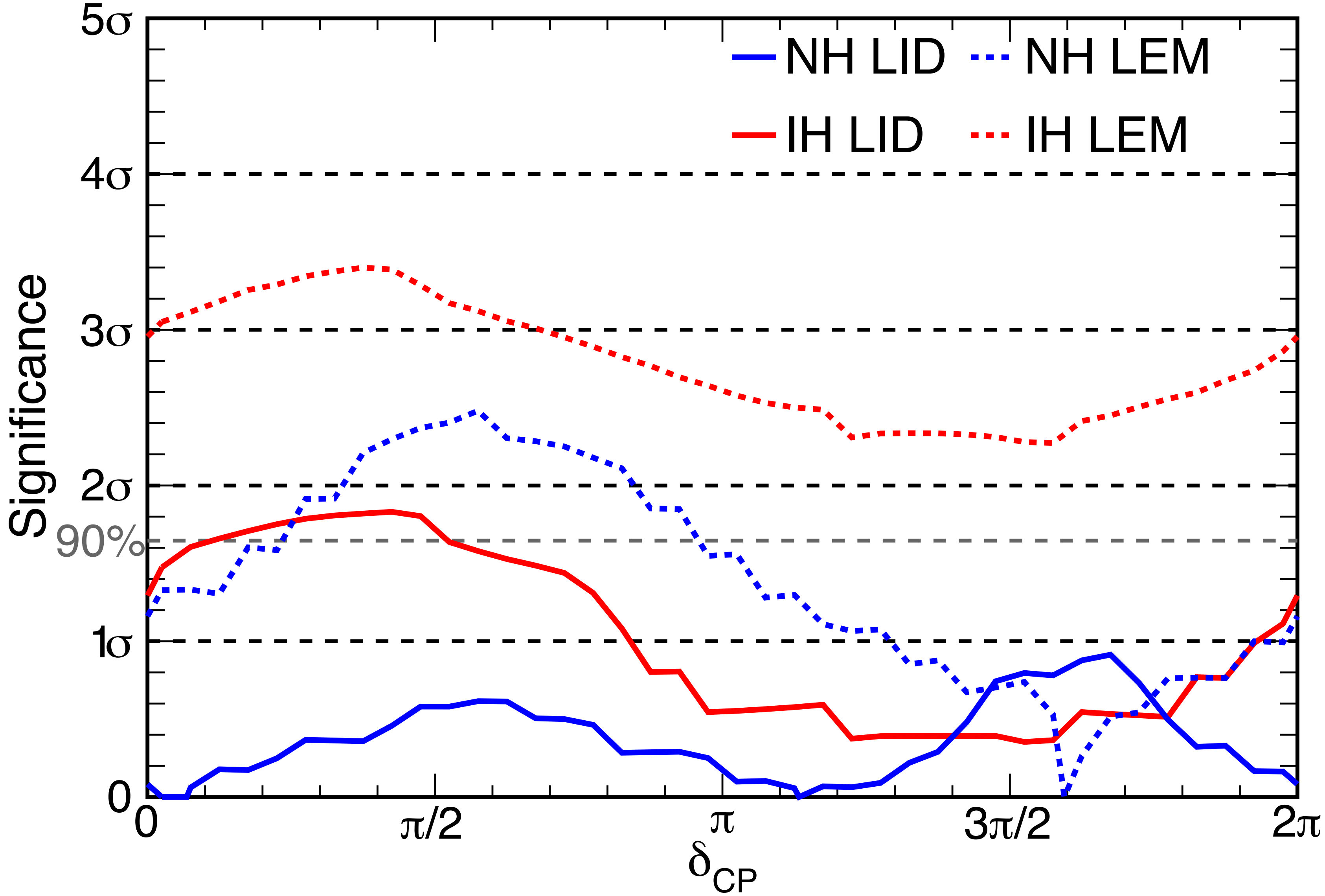
	E_{rec}	p-theta
p-value	0.16	0.34
BF	1.1	0.6

Parameterizing $\nu/\bar{\nu}$ Differences

- For $\bar{\nu}$ disappearance measurements, T2K chose to duplicate the full set of oscillation parameters (θ_{23}, $\bar{\theta}_{23}$, etc.)
- Because of a low statistics sample, all parameters except the 23 parameters were fixed as equal between ν and $\bar{\nu}$; only $\bar{\theta}_{23}$ and $\Delta\bar{m}^2_{23}$ are variable



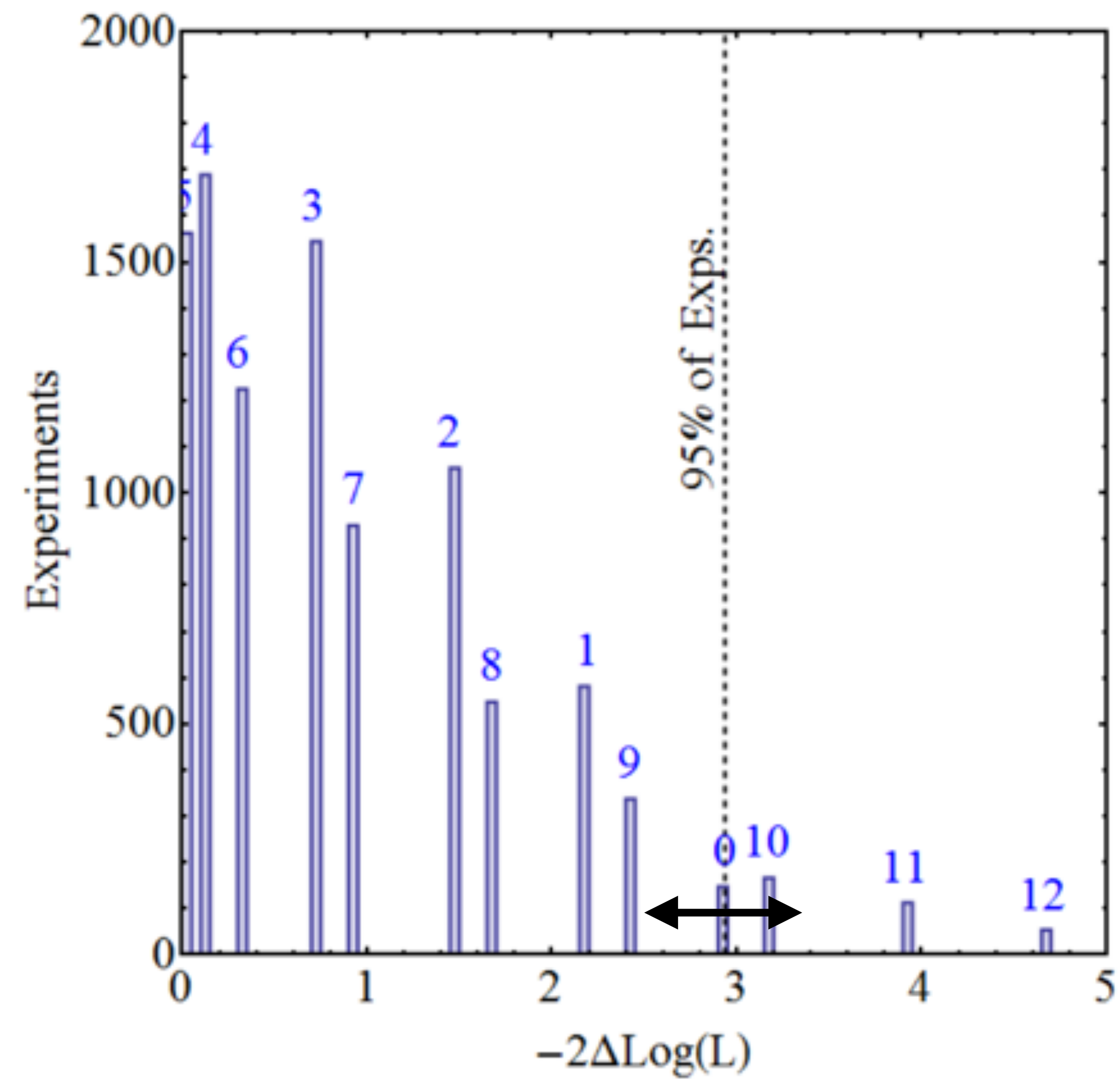
P-Values (1)



Why are these so jagged?

P-Values (1)

- These p-values are generated as a counting rate
- When the assumptions about the oscillation parameters are varied, these spikes can cross the critical value, which causes jaggedness
- Things like shape information and more data will smooth this out



P-Values (2)

- How do we treat p-values in the presence of systematic uncertainties that may be constrained by the analysis?
- This is an issue for δ_{CP} and mass hierarchy measurements
- T2K has been discussing which of several methods to use: http://physics.rockefeller.edu/luc/proceedings/phystat2007.pdf (Thank you to PhyStats of years gone by!)

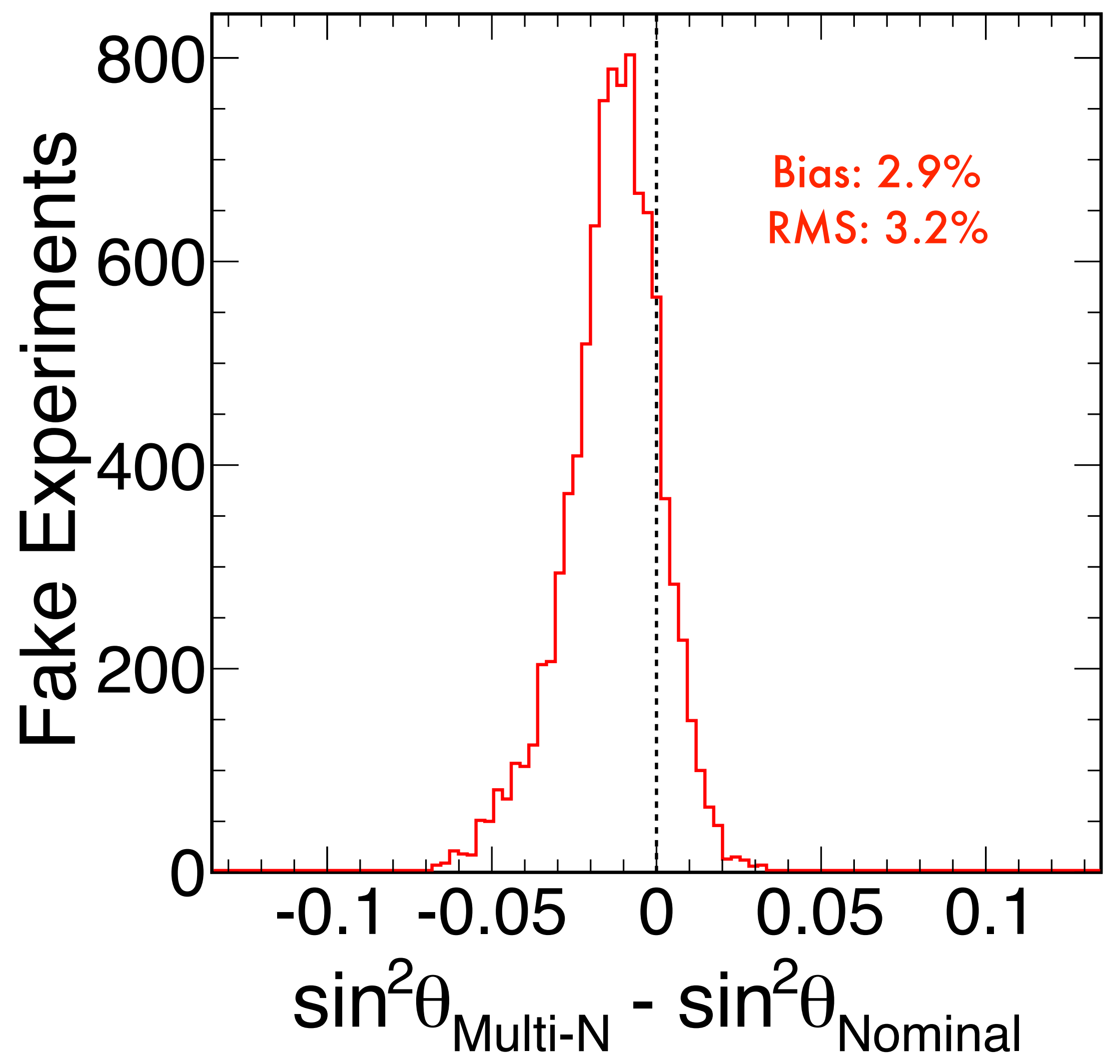
P-Values (2)

- Two options, both which integrate the p-value over the systematics
  - Prior-predictive
    - $$p_{prior} = \int d\nu \, \pi(\nu) \, p(\nu).$$
    - "A small value of p_{prior} is therefore evidence against the overall model used to describe the data, and could in principle be caused by a badly elicited prior as well as by an invalid likelihood (or unlikely data)"
  - Posterior-predictive
    - $$p_{post} = \int d\nu \, \pi(\nu \,|\, n_0) \, p(\nu)$$
    - "Estimates the probability that a future observation will be at least as extreme as the current observation if the null hypothesis is true"
- Ongoing discussion!

Discrete Model Issues

- We do our best to have the best model of cross section, but this is hard
- Test impact using fake data studies—can see bias in parameters
- We may not have complete implementation of a particular model, which makes it hard to use that model to fit data
- How can we include the error from these discrete model differences on our final parameters?

Martini et al, Phys. Rev. C, 81:045502, Apr 2010

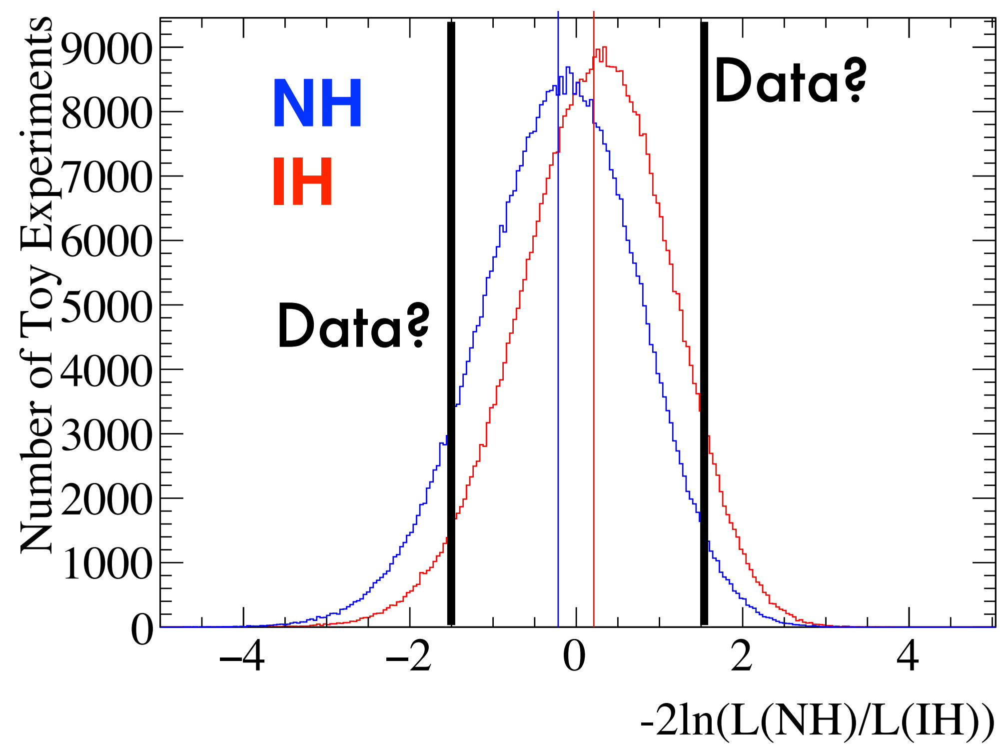


Mass Hierarchy

- What should we report?
- Answer: report it all!
  - Show distributions for test statistic
  - Show Bayes Factor
  - Show sensitivity (and assumptions)
- Let the consumer make choice—somewhat help them along!

BF=2.2

For P(NH)=P(IH)=0.5	NH	IH	Total
$\sin^2\theta_{23} \leq 0.5$	17.9%	7.8%	25.7%
$\sin^2\theta_{23} > 0.5$	50.5%	23.8%	74.3%
Total	68.4%	31.6%	100%



Questions for Discussion

- Are we doing anything that's flat out wrong?
- Do we need to do anything more?
- How do we best communicate our choices?

Thank You!

Supplementary

Posterior Probability

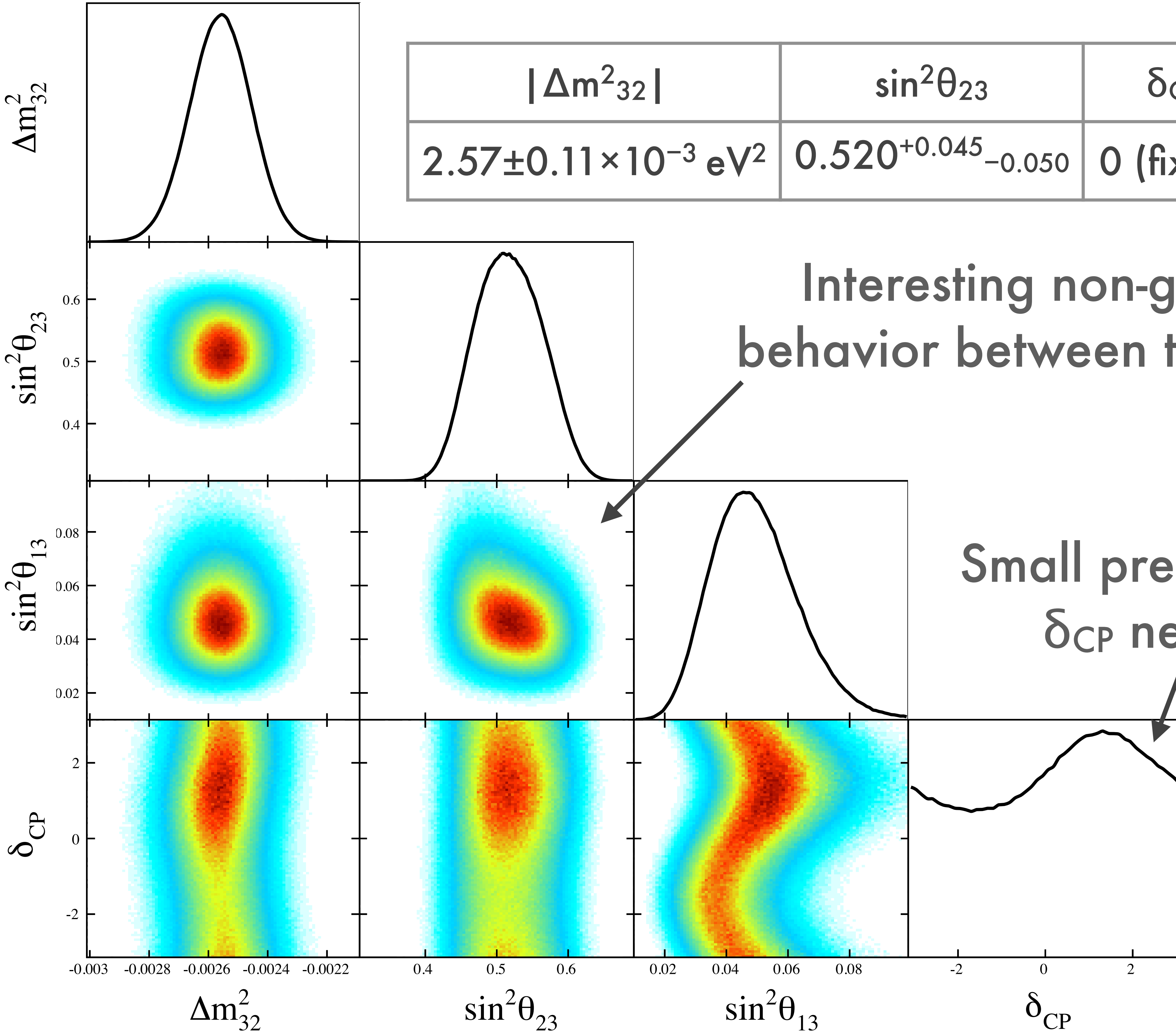
$$P(\vec{z}|\text{data}) \propto \boxed{P(\text{data}|\vec{z})}\boxed{P(\vec{z})}$$

$$
\begin{aligned}
-\ln(P) = c & + \sum_{i}^{ND280bins} N_i^p(\vec{b},\vec{x},\vec{d}) - N_i^d \ln N_i^p(\vec{b},\vec{x},\vec{d}) \\
& + \sum_{i}^{N_\mu \text{ bins}} N_{\mu,i}^p(\vec{\theta},\vec{b},\vec{x},\vec{s}) - N_{\mu,i}^d \ln N_{\mu,i}^p(\vec{\theta},\vec{b},\vec{x},\vec{s}) \\
& + \sum_{i}^{N_e \text{ bins}} N_{e,i}^p(\vec{\theta},\vec{b},\vec{x},\vec{s}) - N_{e,i}^d \ln N_{e,i}^p(\vec{\theta},\vec{b},\vec{x},\vec{s}) \\
& + \tfrac{1}{2}\,\Delta\vec{b}^T V_b^{-1}\Delta\vec{b} + \tfrac{1}{2}\,\Delta\vec{x}^T V_x^{-1}\Delta\vec{x} + \tfrac{1}{2}\,\Delta\vec{d}^T V_d^{-1}\Delta\vec{d} \\
& + \tfrac{1}{2}\,\Delta\vec{s}^T V_s^{-1}\Delta\vec{s} + \tfrac{1}{2}\,\Delta\vec{\theta}_{sr}^T V_{\theta sr}^{-1}\Delta\vec{\theta}_{sr}
\end{aligned}
$$

How well does the data agree with our prediction?

What prior understanding did we put into our prediction?

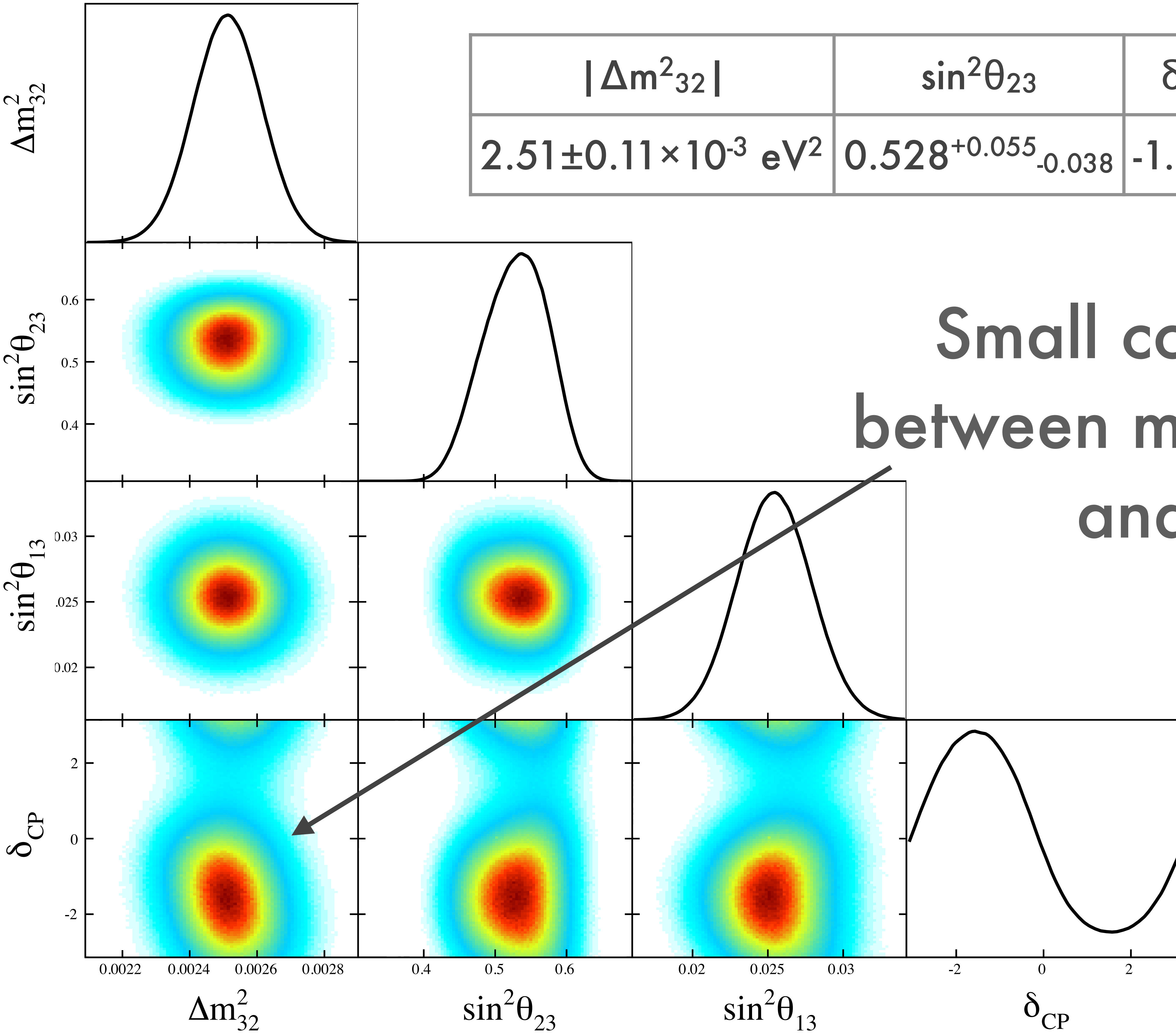
$\|\Delta m^2_{32}\|$	$\sin^2\theta_{23}$	δ_{CP}	Hierarchy
$2.57 \pm 0.11 \times 10^{-3}$ eV^2	$0.520^{+0.045}_{-0.050}$	0 (fixed)	IH



Interesting non-gaussian behavior between two angles

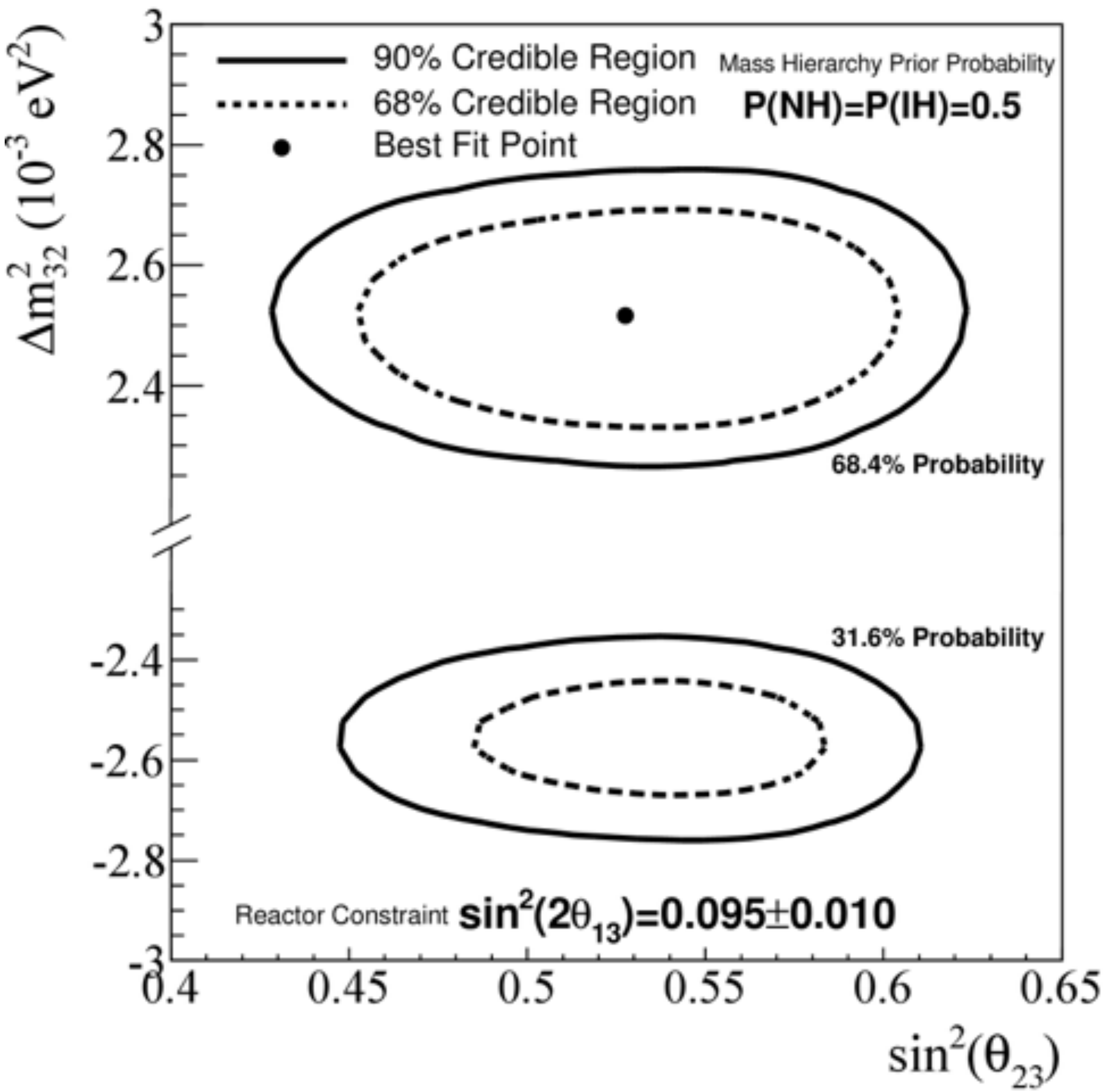
Small preference for δ_{CP} near $\pi/2$

$\lvert \Delta m^2_{32} \rvert$	$\sin^2\theta_{23}$	δ_{CP}	Hierarchy
$2.51 \pm 0.11 \times 10^{-3}$ eV^2	$0.528^{+0.055}_{-0.038}$	-1.601	NH



Small correlation between mass splitting and δ_{CP}

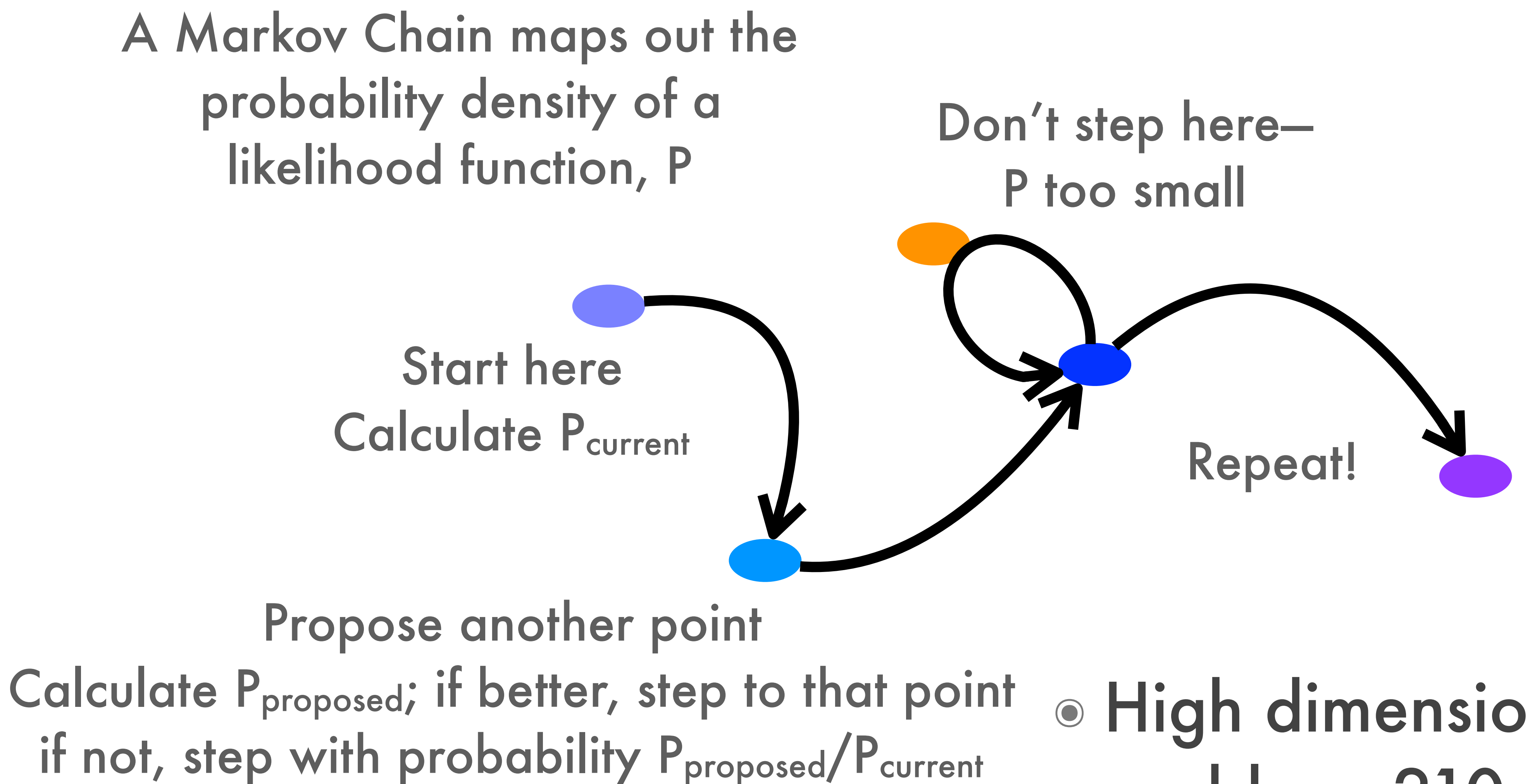
Mass Hierarchy



- Markov Chain steps between the two hierarchies
- Relative probability shows any preference for hierarchy

	NH	IH	Total
$\sin^2\theta_{23}\leq0.5$	17.9%	7.8%	25.7%
$\sin^2\theta_{23}>0.5$	50.5%	23.8%	74.3%
Total	68.4%	31.6%	100%

Markov Chain Monte Carlo and GPUs



- High dimensionality problem: 210 parameters
- Use Metropolis-Hastings algorithm with MCMC; doesn’t require calculating likelihood derivatives