# Systematic Uncertainties for Future Accelerator-Based Neutrino Oscillation Experiments

Mike Wilking
Stony Brook University
NNN 2018 (Halloween Edition)
November 3rd, 2018

# Oscillation Measurements

What we want to measure:

$$P_i^{osc}(\nu_\mu \rightarrow \nu_e) = \frac{\Phi_i^{\nu_e FD}}{\Phi_i^{\nu_\mu ND}}$$

True neutrino energy bin, *i*

What we actually measure:

Reconstructed neutrino energy bin, *j*

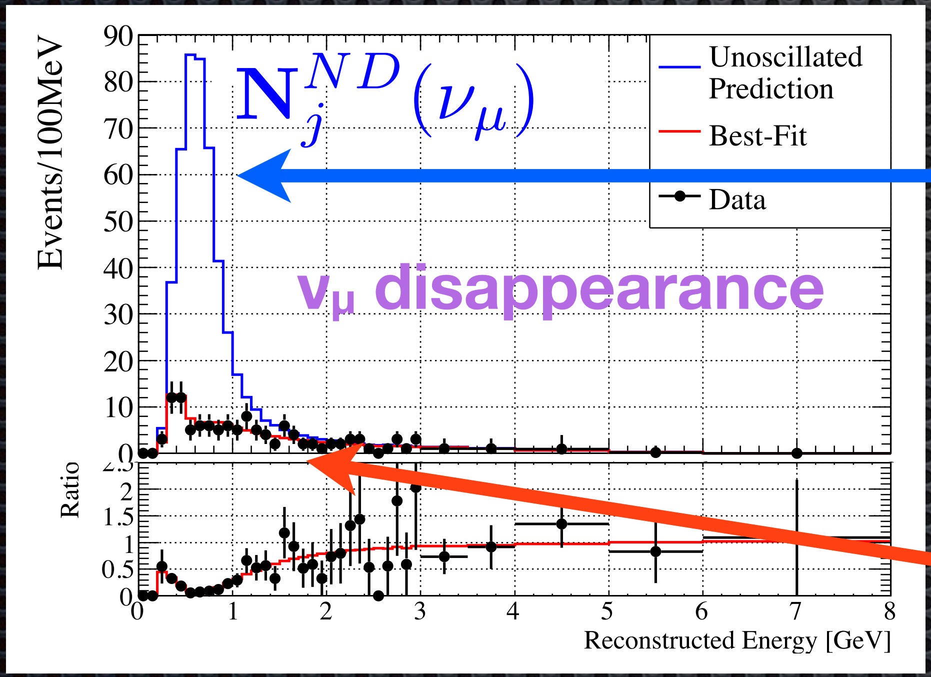
$$N_j^{FD}(\nu_e) = \sum_i^{E_\nu^{true}\ bins} \Phi_i^{\nu_e FD} \sigma_i^{\nu_e CC} \epsilon_i^{FD} M_{ij}^{FD}$$

**Can't extract with a simple ratio :(**

$$N_j^{ND}(\nu_\mu) = \sum_i^{E_\nu^{true}\ bins} \Phi_i^{\nu_\mu ND} \sigma_i^{\nu_\mu CC} \epsilon_i^{ND} M_{ij}^{ND}$$

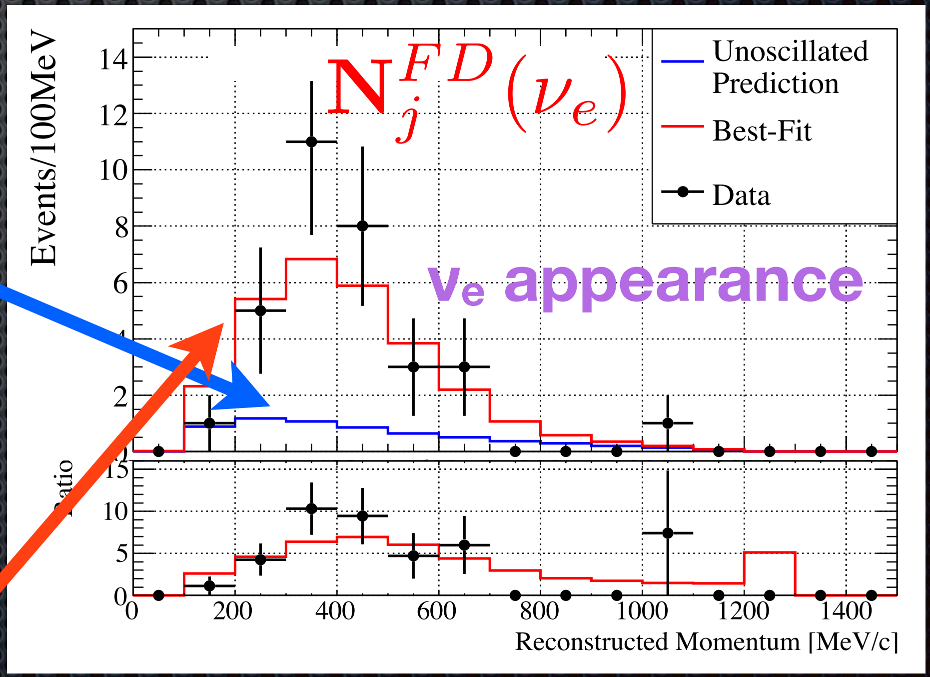
(excluding backgrounds for now)

- **$M_{ij}$ = $E_{\nu,true}$ to $E_{\nu,reco}$ smearing**
  - Caused by coupling between cross section and detector effects



**Near Detector Measures:**
- $\nu_\mu$ energy spectrum
- Small $\nu_e$ component

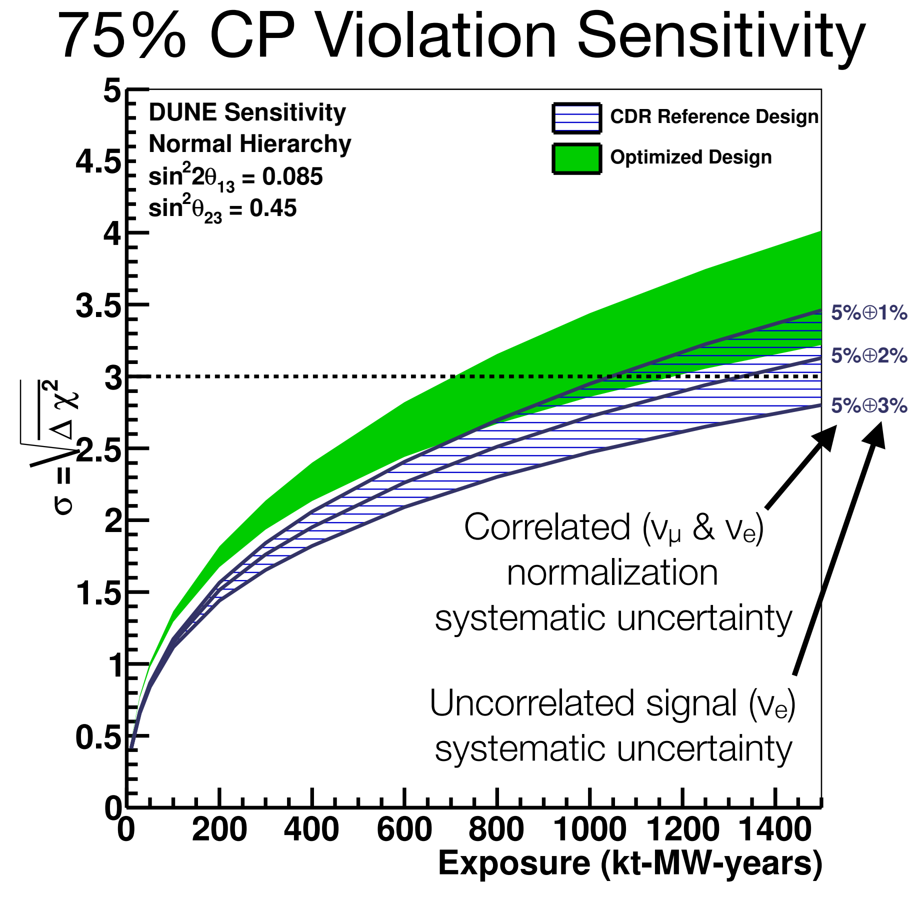
**Far Detector Measures:**
- Osc. $\nu_\mu$ energy spectrum
- Large $\nu_e$ appearance signal



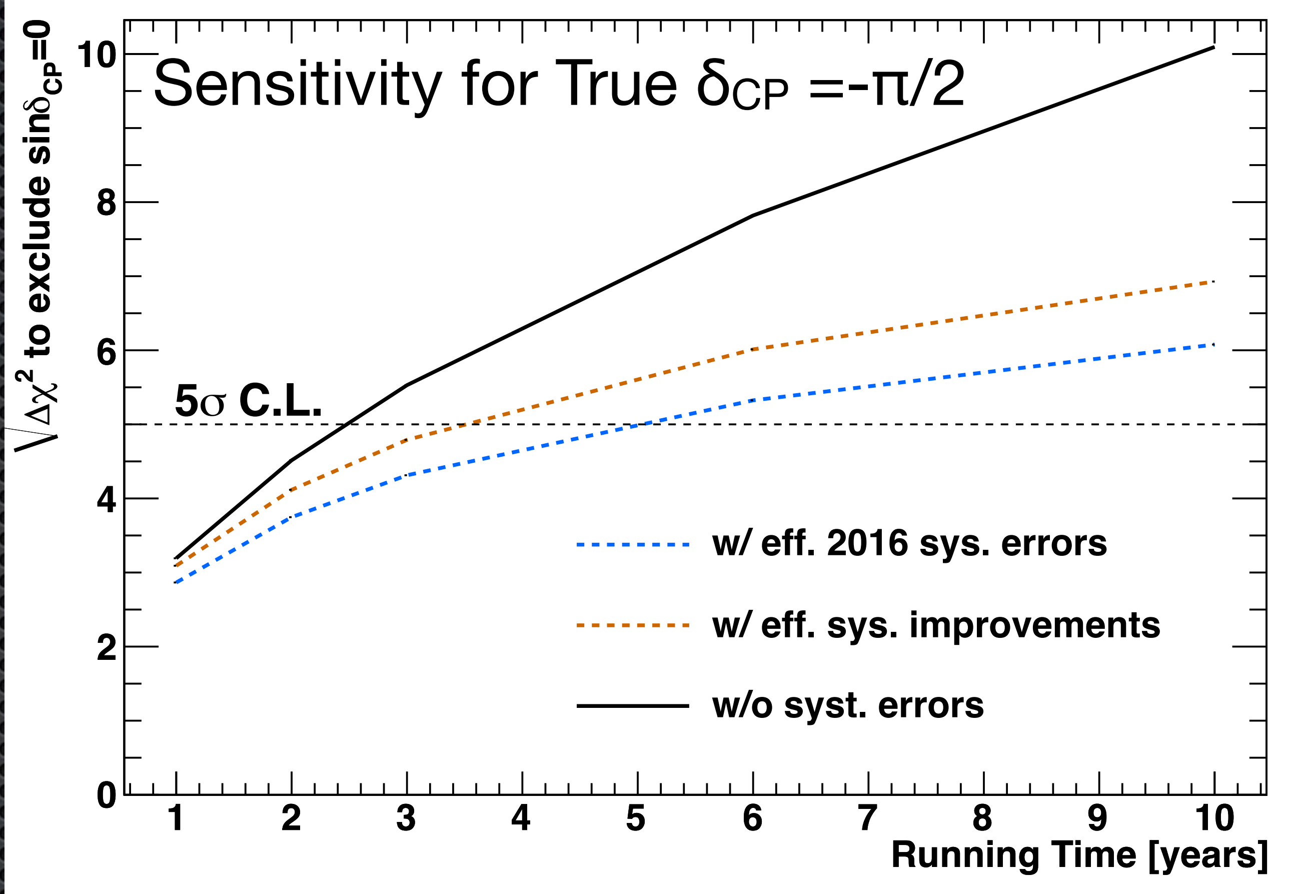
- **$\Phi_i$ is very different at ND & FD, primarily due to oscillations**
  - **$E_\nu$ smearing ($M_{ij}$) has a very different impact on $N^{ND}$ & $N^{FD}$**

# The Next Generation (CP Violation)

DUNE



Hyper-K



- Both DUNE and Hyper-K have made initial studies on the effects of systematic uncertainties on CP sensitivity
- Although more detailed studies are expected in the future, it is clear that ~few percent level uncertainties on the predicted event rates are needed

# Systematic Errors in Sensitivity Estimates

- Both Hyper-K and DUNE utilize T2K systematic error studies to project their future sensitivities

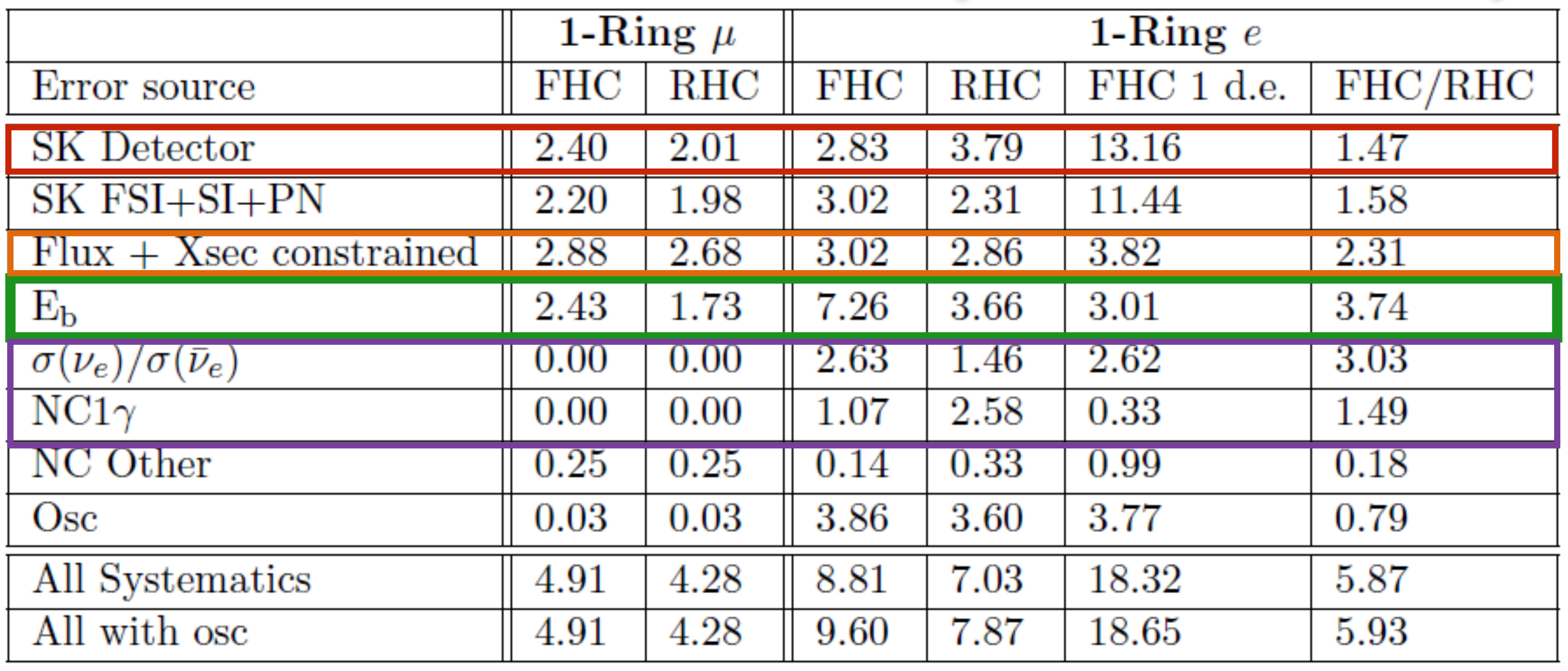
## T2K 2018 Uncertainties (% Event Rate )

**Detector Error**

**Near/Far Fit**

**Fake data study (more later)**

**Theory Estimate**

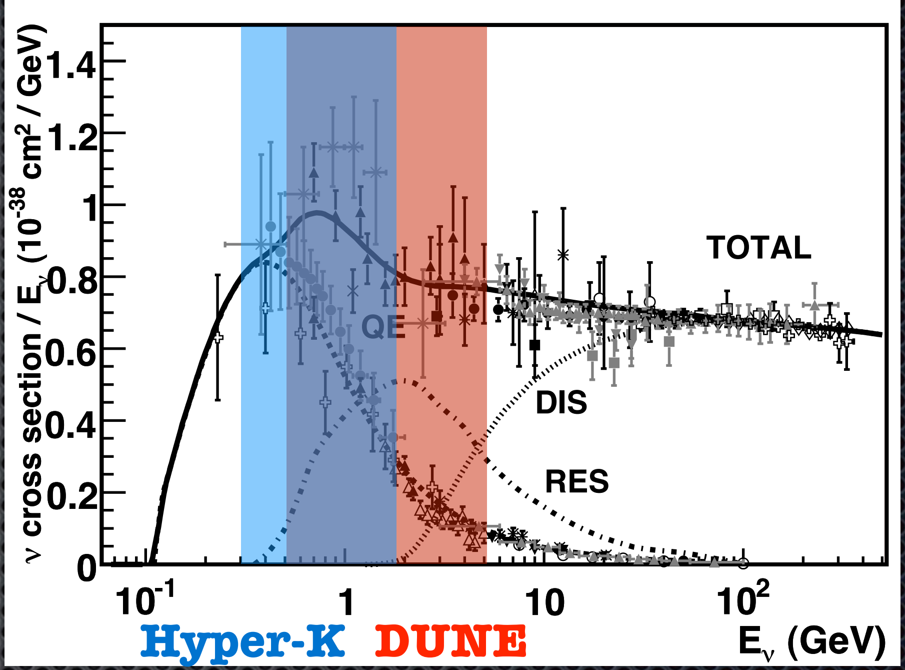


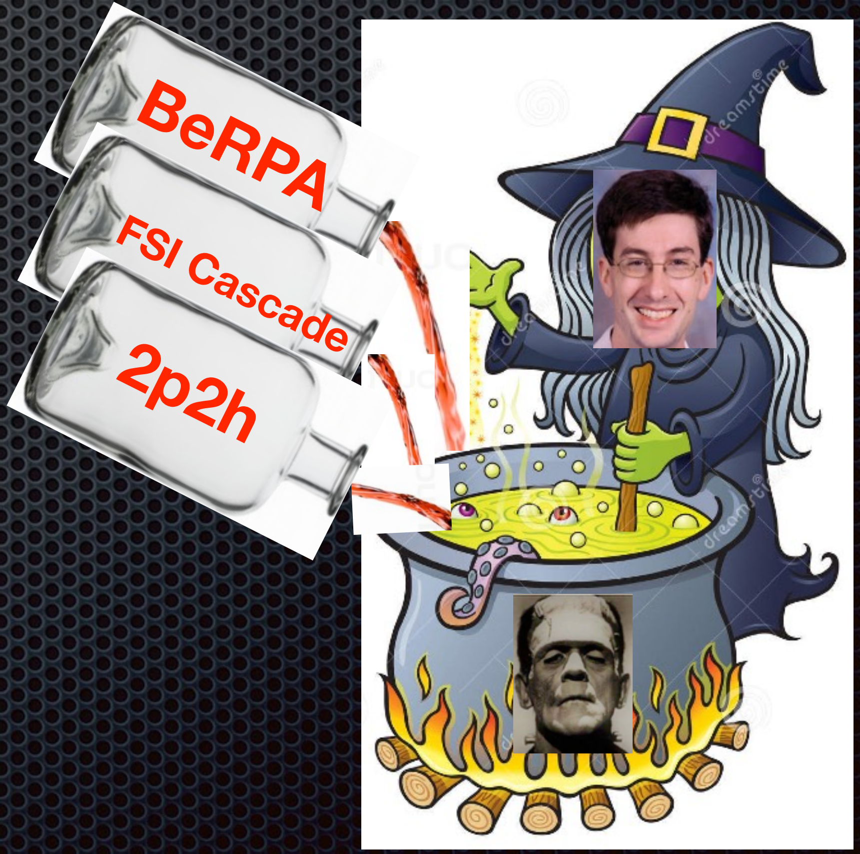
| | 1-Ring $\mu$ | | 1-Ring $e$ | | | |
|---|---|---|---|---|---|---|
| Error source | FHC | RHC | FHC | RHC | FHC 1 d.e. | FHC/RHC |
| SK Detector | 2.40 | 2.01 | 2.83 | 3.79 | 13.16 | 1.47 |
| SK FSI+SI+PN | 2.20 | 1.98 | 3.02 | 2.31 | 11.44 | 1.58 |
| Flux + Xsec constrained | 2.88 | 2.68 | 3.02 | 2.86 | 3.82 | 2.31 |
| $E_b$ | 2.43 | 1.73 | 7.26 | 3.66 | 3.01 | 3.74 |
| $\sigma(\nu_e)/\sigma(\bar{\nu}_e)$ | 0.00 | 0.00 | 2.63 | 1.46 | 2.62 | 3.03 |
| NC1$\gamma$ | 0.00 | 0.00 | 1.07 | 2.58 | 0.33 | 1.49 |
| NC Other | 0.25 | 0.25 | 0.14 | 0.33 | 0.99 | 0.18 |
| Osc | 0.03 | 0.03 | 3.86 | 3.60 | 3.77 | 0.79 |
| All Systematics | 4.91 | 4.28 | 8.81 | 7.03 | 18.32 | 5.87 |
| All with osc | 4.91 | 4.28 | 9.60 | 7.87 | 18.65 | 5.93 |

- Note that in some cases, the total error is larger than in 2016

# Cross Section Modeling

- Long baseline experiments have chosen a particularly difficult energy scale for nuclear physics (~GeV)
  - Lower energies: inverse beta decay
  - Higher energies: deep inelastic scattering
- At the GeV scale, we rely on effective theories that can get the qualitative features correct, but are not exact (i.e. no NNLO calculations from first principles)
- Theorists, model builders, and experimentalists then proceed to make model corrections and add tunable parameters until some level of agreement with data is achieved





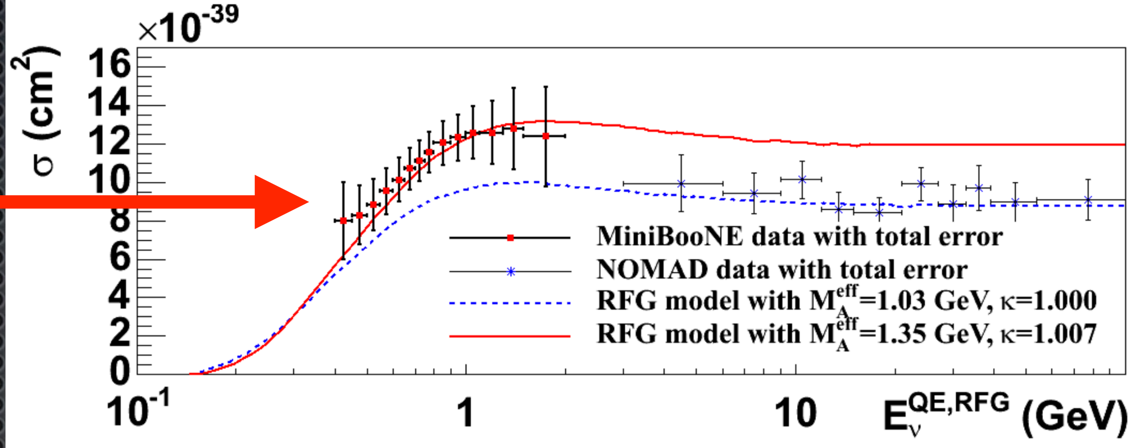
(See K. McFarland's talk from yesterday for a detailed accounting of T2K's ingredients)

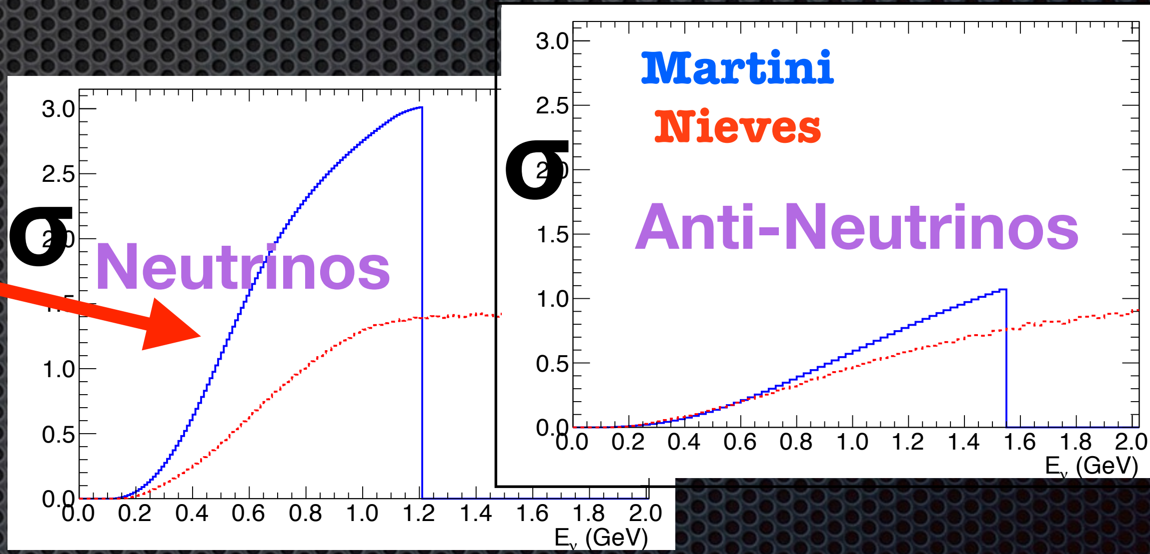
# Model Decisions

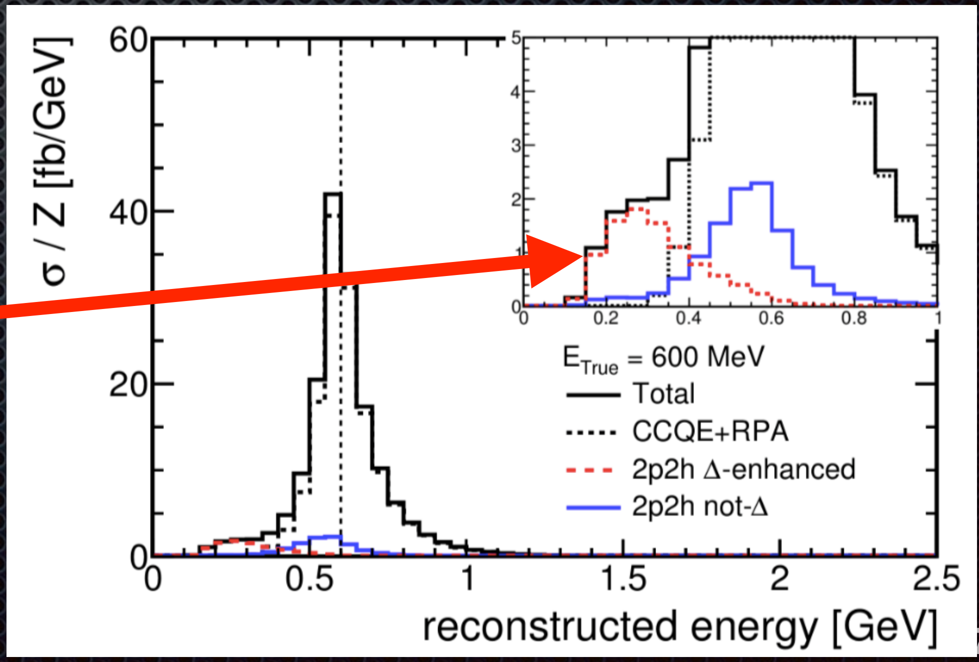
- Building a model requires decisions about which modifications are **1. "sufficiently physically motivated"** and/or **2. desperately needed** to achieve "sufficient agreement" with near detector data
  - Attempts are also made to simultaneously **compare** new model features across **multiple experiments** (not just neutrino scattering)
    - However, experimental differences often make such **comparisons very difficult** (neutrino fluxes, detector systematics, phase space / efficiencies, backgrounds, and especially different target nuclei)
- Note that this iterative process is very much **NOT a blind analysis**/procedure
  - i.e. each experiment will eventually match their model to ND data, but the model details may still be incorrect
- **Key question**: to what level can we trust such a model to extrapolate from ND to FD?

# Simplified Example (Part I)

- MiniBooNE: the $\nu_\mu$ cross section is higher than expected, so increase value of $M_A$
- Response from theory: introduce nucleon-nucleon correlations to grow the cross section
  - Some disagreements in calculations (Nieves vs Martini)
- Minerva & T2K: agreement with near detector data is qualitatively better, but still not good enough
- T2K: Introduce parameters for scaling 2 components of the model (Δ-enhanced & not-Δ), but otherwise do not change the predicted shapes of these components

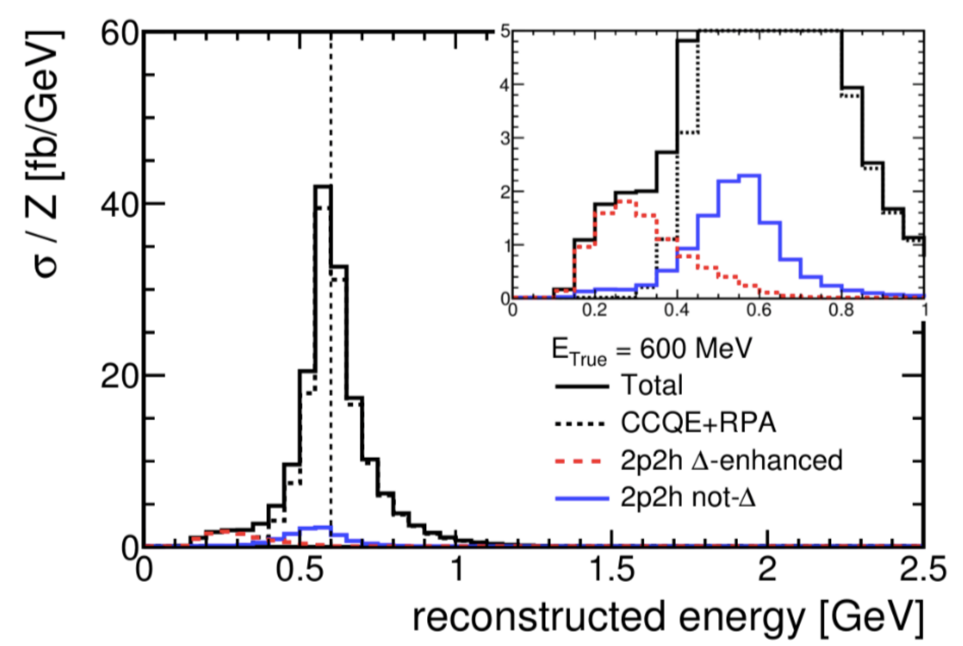


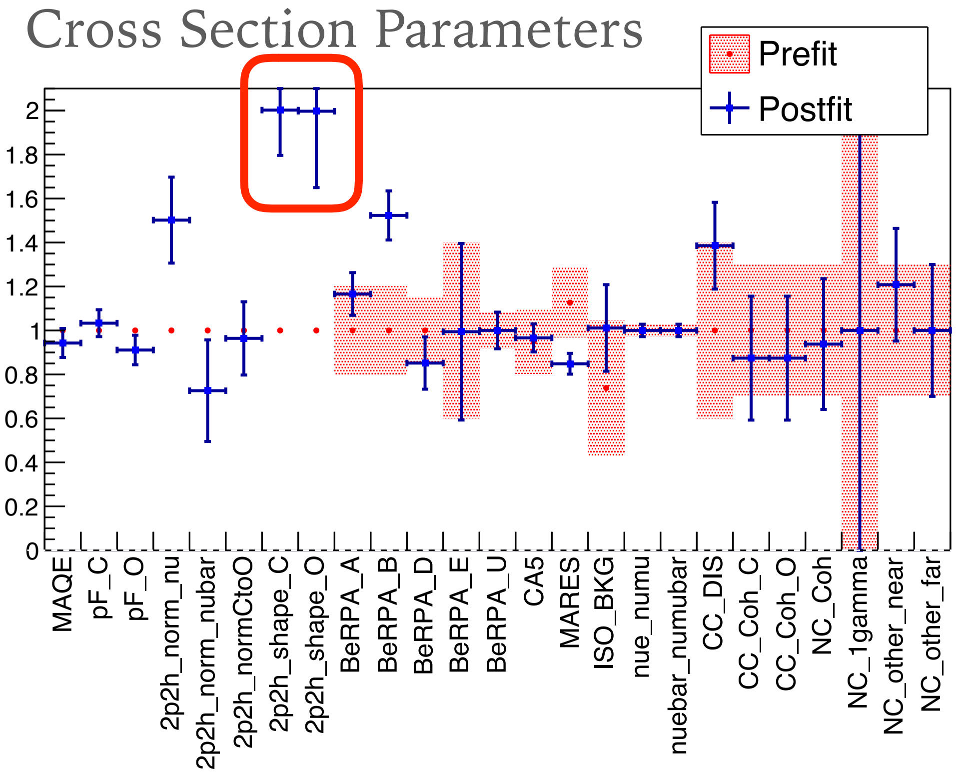




# Simplified Example (Part II)

- T2K fits near detector data to set values of these 2 new parameters
  - Parameters move substantially, but resulting agreement with near detector data is good
- Is the resulting model now sufficiently accurate to extrapolate to the far detector?
  - How many different model shape & parameter choices could have been made to achieve similar near detector / simulation agreement?
  - Most crucially: can different model choices produce different far detector predictions?





# Fake Data Studies

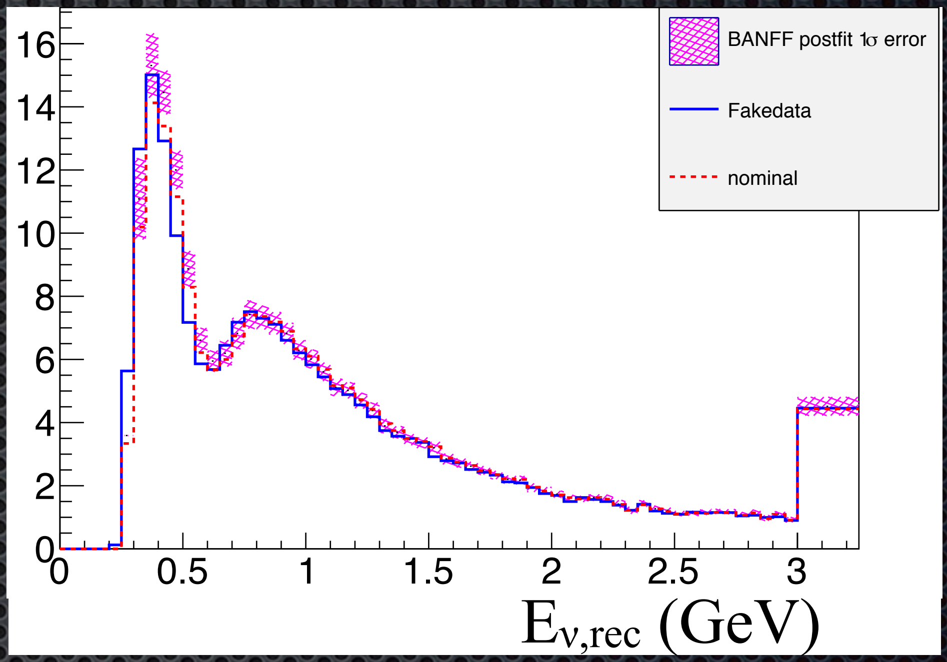
- How can we probe the consequences of having the wrong model?
- We can produce “fake data” that includes effects that are not included in the model used to fit the fake data
    - Hope to learn:
        1. How wrong could our measured oscillation parameters be? (bias)
        2. How does this bias compare to the systematic errors we calculated with our ND constrained model?
            - i.e. how much did we underestimate our systematic uncertainty due to our reliance on an imperfect model?

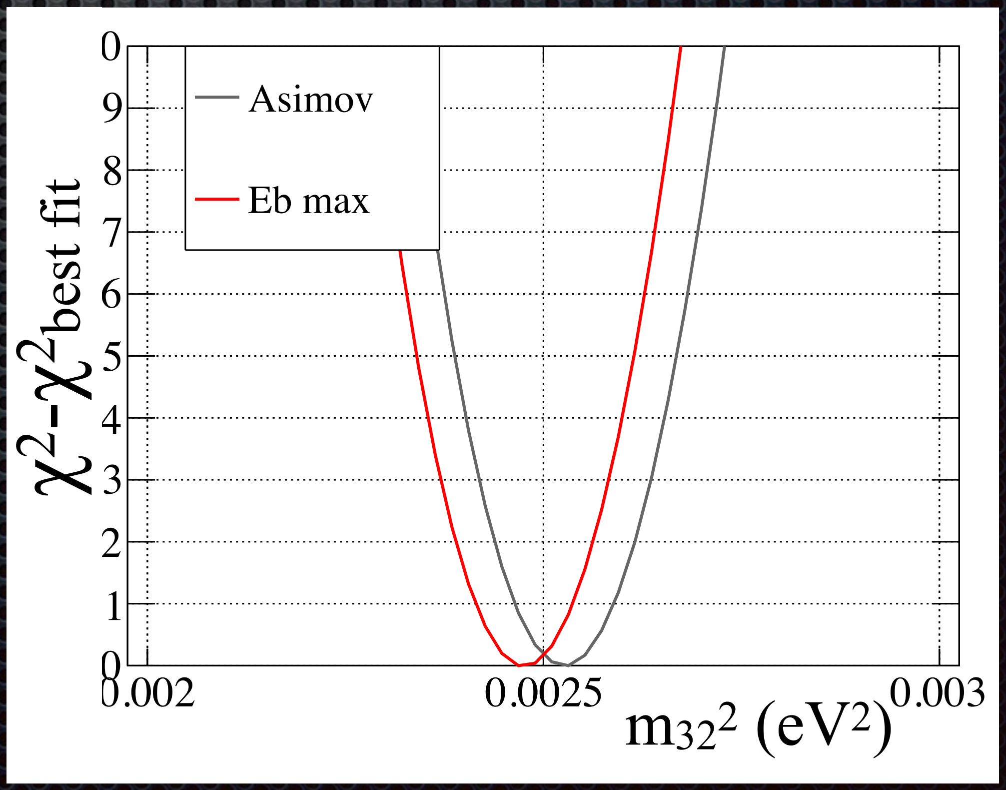
# Example: T2K $E_B$ Fake Data

- Implementing nucleon binding energy ($E_B$) variations is difficult (requires an energy shift, rather than a pure reweight), and was not available for this year's T2K analysis
- Fake data was generated by varying $E_B$ by 9 MeV, and then fit with full near/far framework
- Resulting bias in $\Delta m_{32}^2$ was above a predefined threshold, so an additional uncertainty was added to the fit contours in $\Delta m_{32}^2$

- Note: this uncertainty is expected to be reduced in the future by:
    - Implementing $E_B$ reweighing in the model (so fake data studies are no longer relevant)
    - Reducing the allowed variation of $E_B$ to **±3** MeV (based on arXiv:1801.07975)

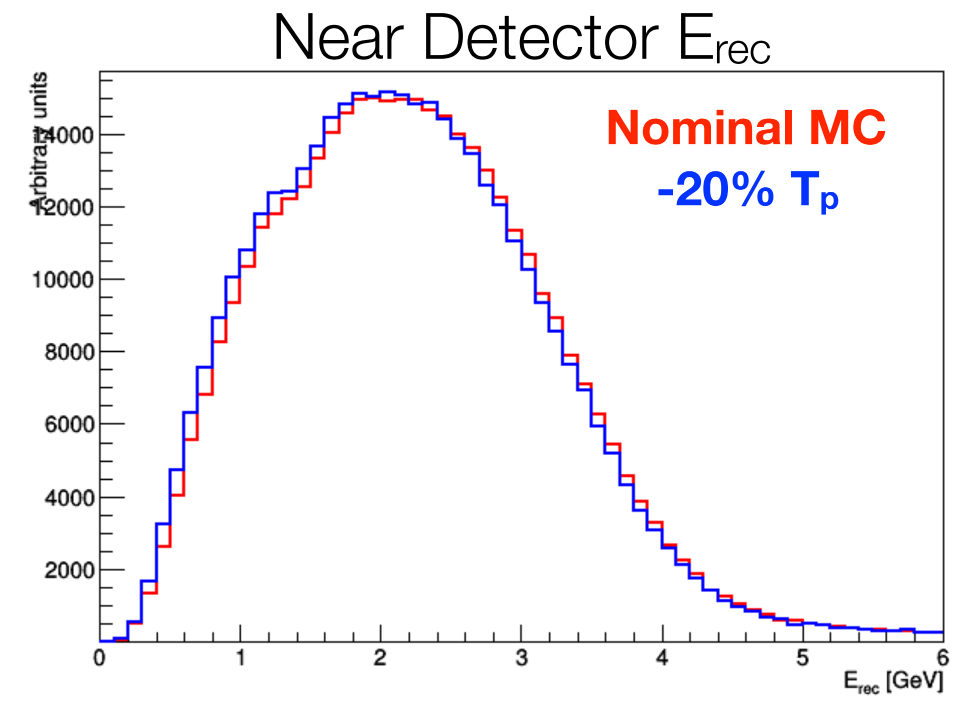
Far Detector $\nu_\mu$ $E_{rec}$



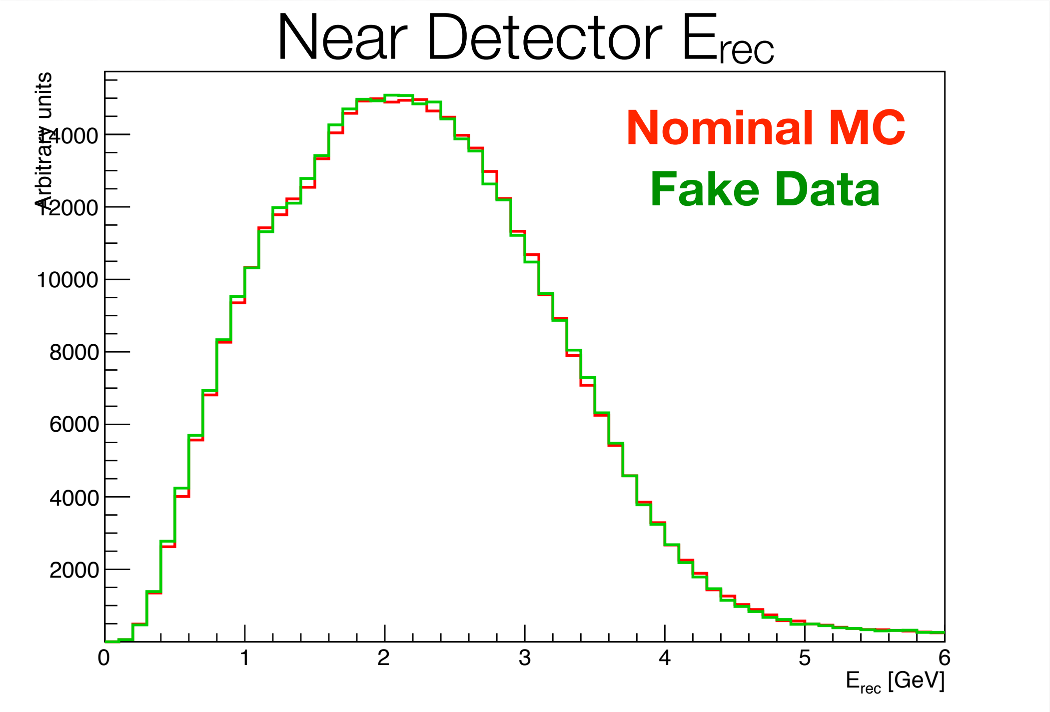


# DUNE Fake Data

- The details of the emitted hadronic state in ν-nucleus interactions have been less studied than outgoing lepton kinematics
- What if the energy sharing between outgoing protons and neutrons was incorrectly modeled?
- Study with a fake dataset:
  - Step 1: Transfer 20% of proton kinetic energy, $\mathbf{T_p}$, to (unseen) neutrons, $\mathbf{T_n}$
  - Step 2: Adjust model (mostly $d\sigma/dT_p$) to reproduced observed spectra

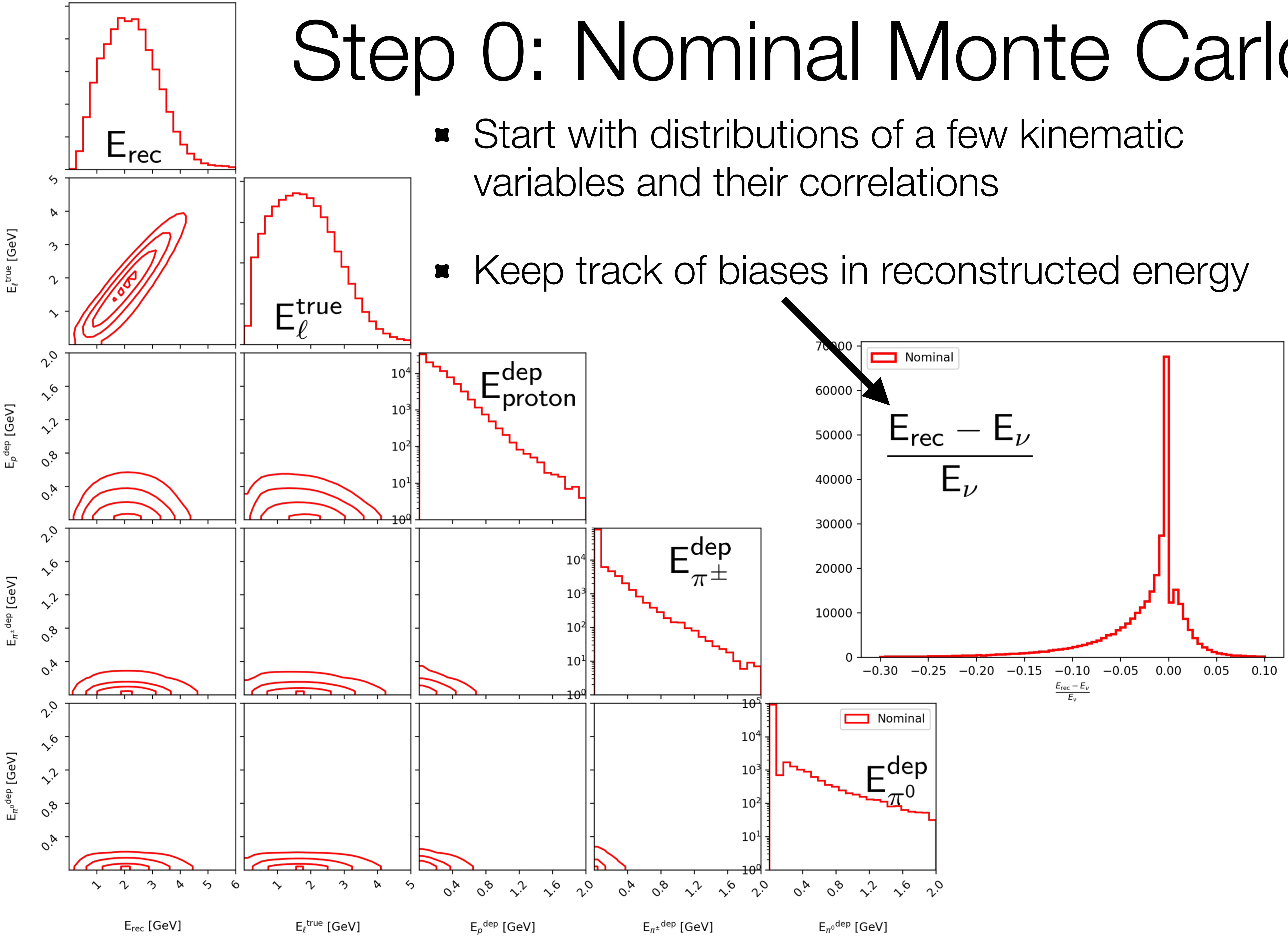


Fake Data Model Adjustment



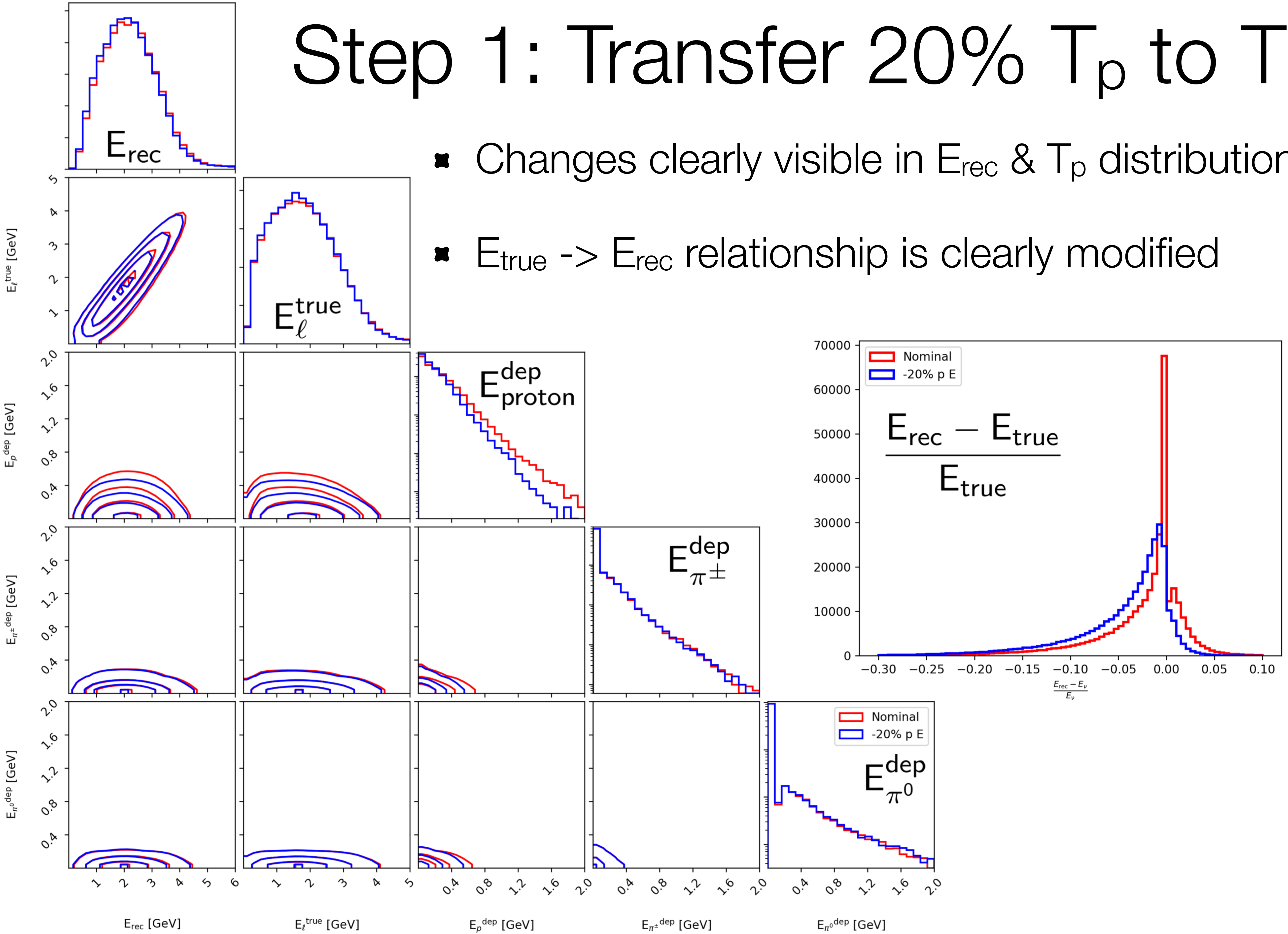
# Step 0: Nominal Monte Carlo

- Start with distributions of a few kinematic variables and their correlations
- Keep track of biases in reconstructed energy



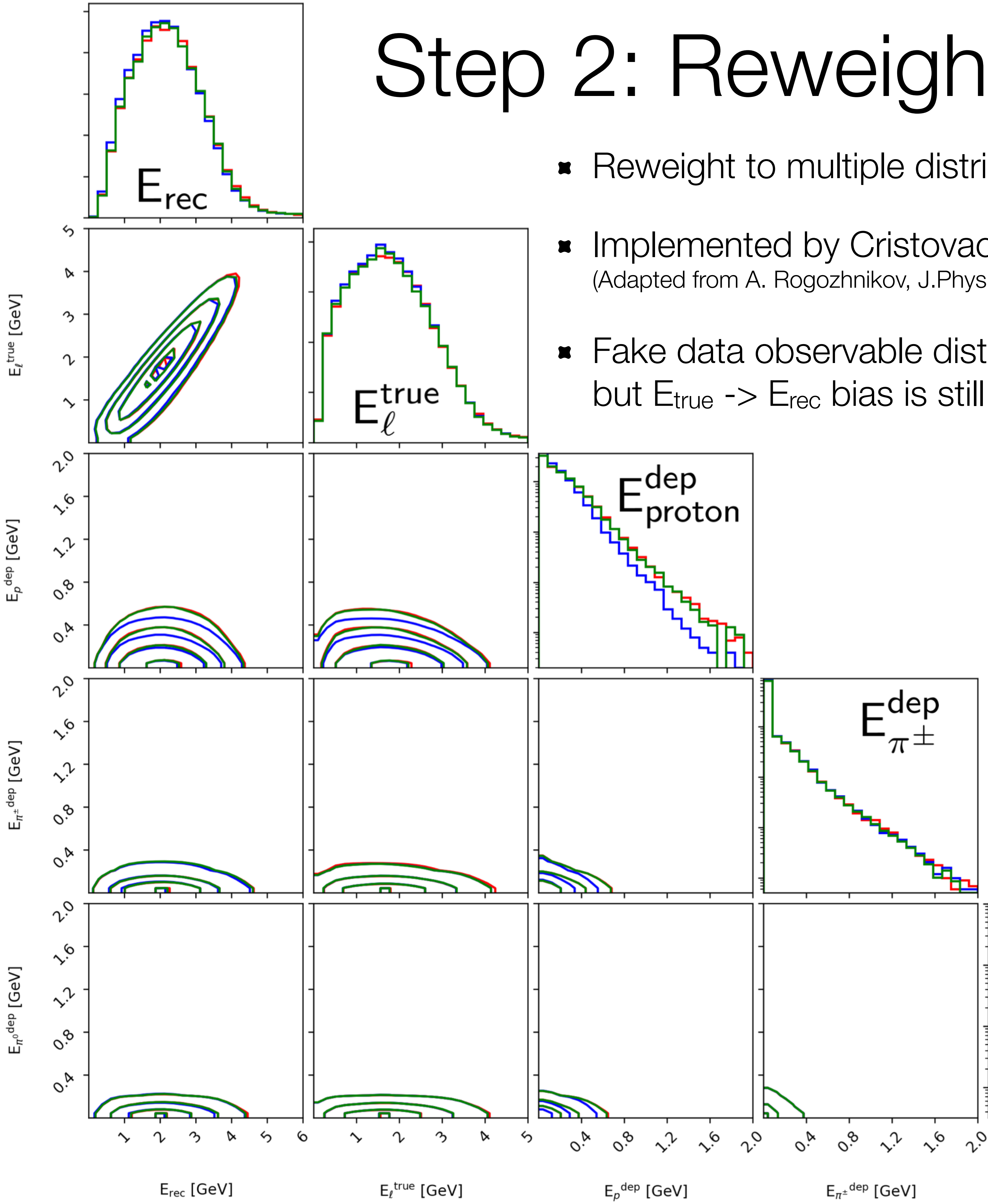
# Step 1: Transfer 20% $T_p$ to $T_n$

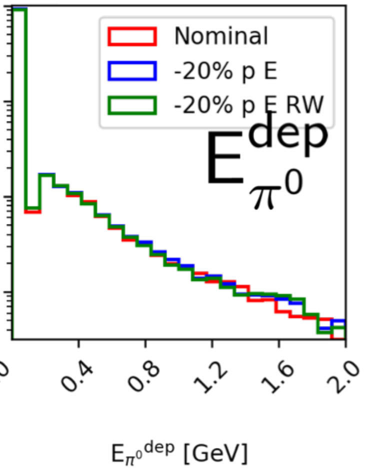
- Changes clearly visible in $E_{rec}$ & $T_p$ distributions
- $E_{true}$ -> $E_{rec}$ relationship is clearly modified

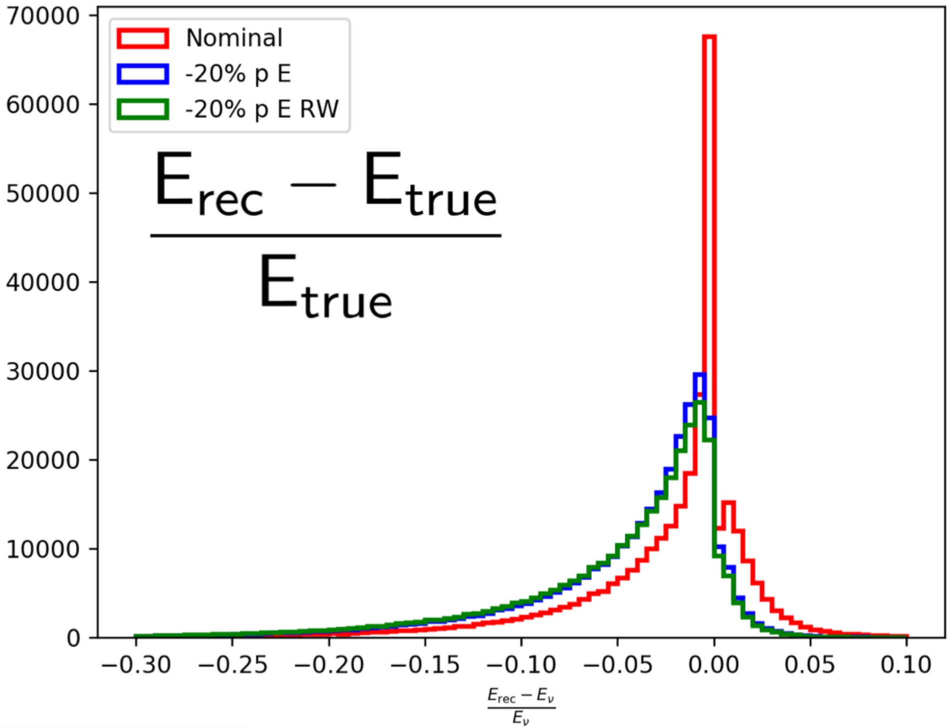


# Step 2: Reweight to Match Nominal

- Reweight to multiple distributions simultaneously, including correlations
- Implemented by Cristovao Vilela, Stony Brook University
  (Adapted from A. Rogozhnikov, J.Phys.Conf.Ser. 762 (2016) no.1, 012036 [arXiv:1608.05806])
- Fake data observable distributions now perfectly match nominal MC, but $E_{true}$ -> $E_{rec}$ bias is still present

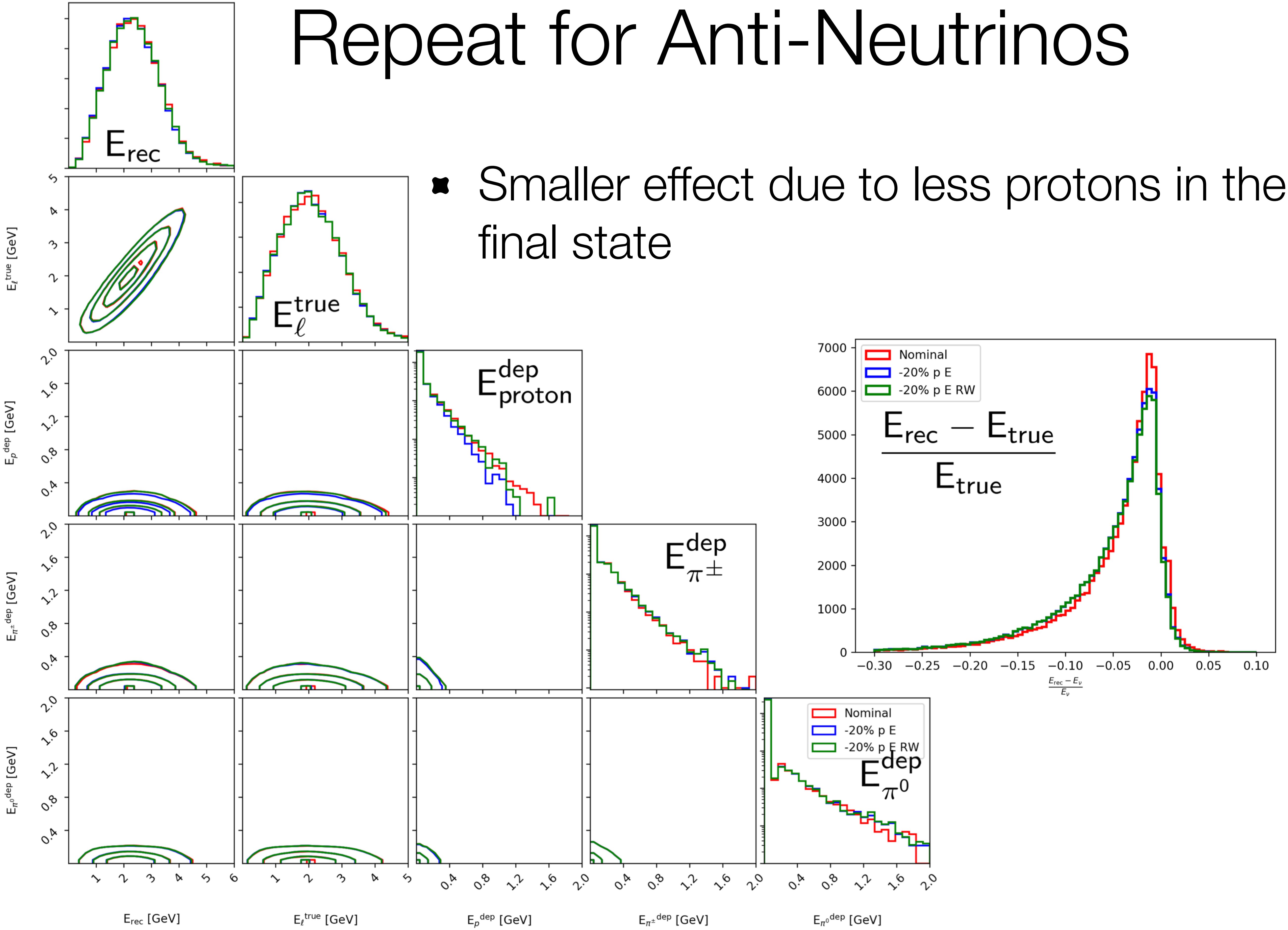






# Repeat for Anti-Neutrinos

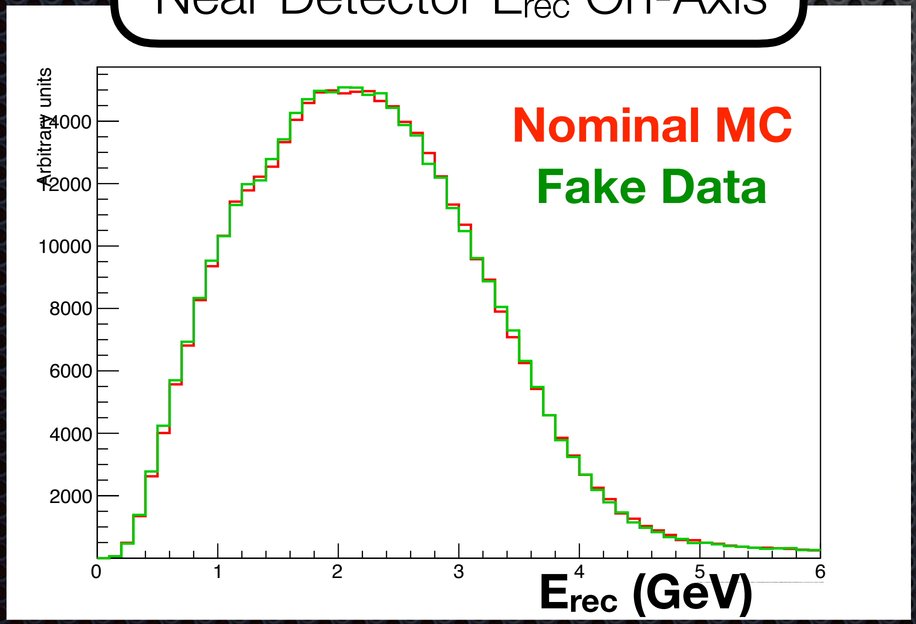
- Smaller effect due to less protons in the final state



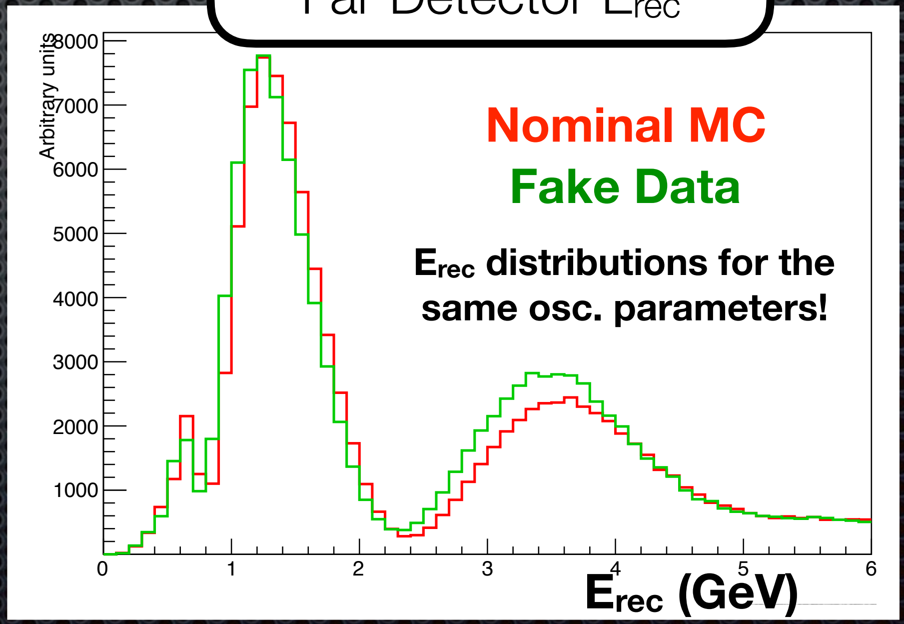
# Impact on Oscillation Measurements

- The far detector prediction does not match the far detector fake data (for the same oscillation parameters)
  - Therefore, the near/far fit gets the wrong answer for $\delta_{CP}$ (and $\theta_{23}$, $\Delta m_{32}^2$)
  - In fact, the correct answer is excluded at many sigma
- **<u>Moral</u>**: long-baseline experiments rely heavily on the underlying details of the cross section model
  - …even (especially?) the details that don't have a tunable dial

Near Detector $E_{rec}$ On-Axis

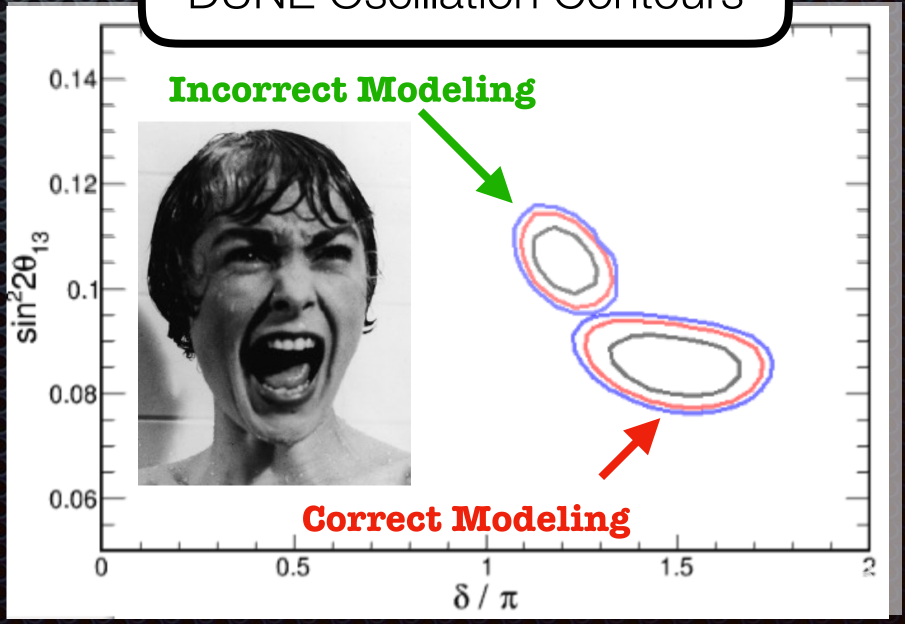


Far Detector $E_{rec}$



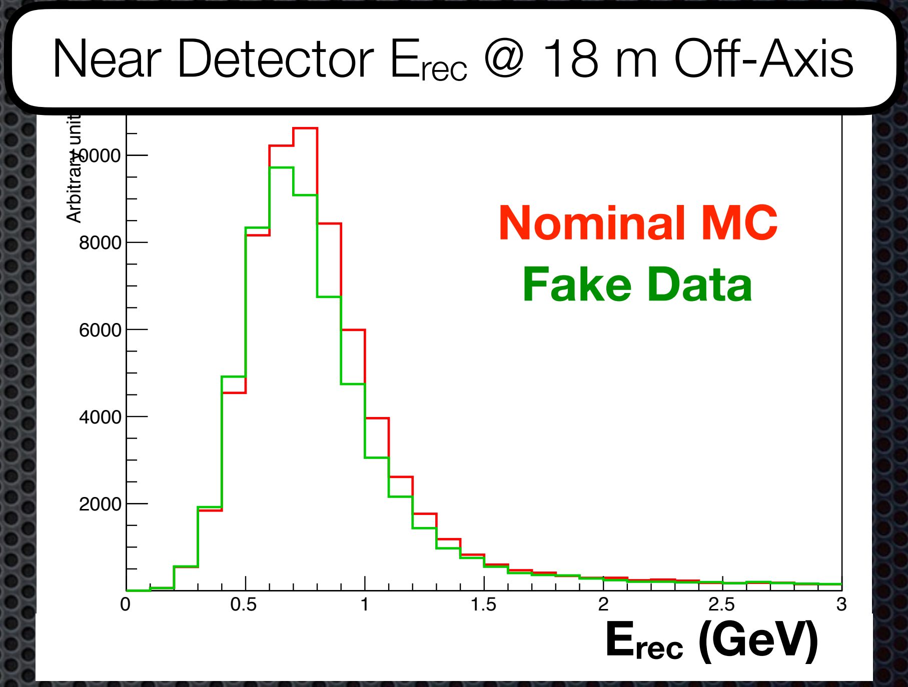
$E_{rec}$ distributions for the same osc. parameters!

DUNE Oscillation Contours



# A Path Forward (Cross Sections)

- **Redundancy!**
  - Measurements across a variety of different energy spectra can be a very powerful tool for identifying model deficiencies that are important to oscillation experiments
    - The main goal of an oscillation analysis is to extrapolate event rates from a ND flux to a FD flux
    - If a model can successfully move between many different energy spectra, its credibility for an ND to FD extrapolation will be greatly improved
  - The can be accomplished by making measurements off-axis to the beam direction



Near Detector $E_{rec}$ @ 18 m Off-Axis

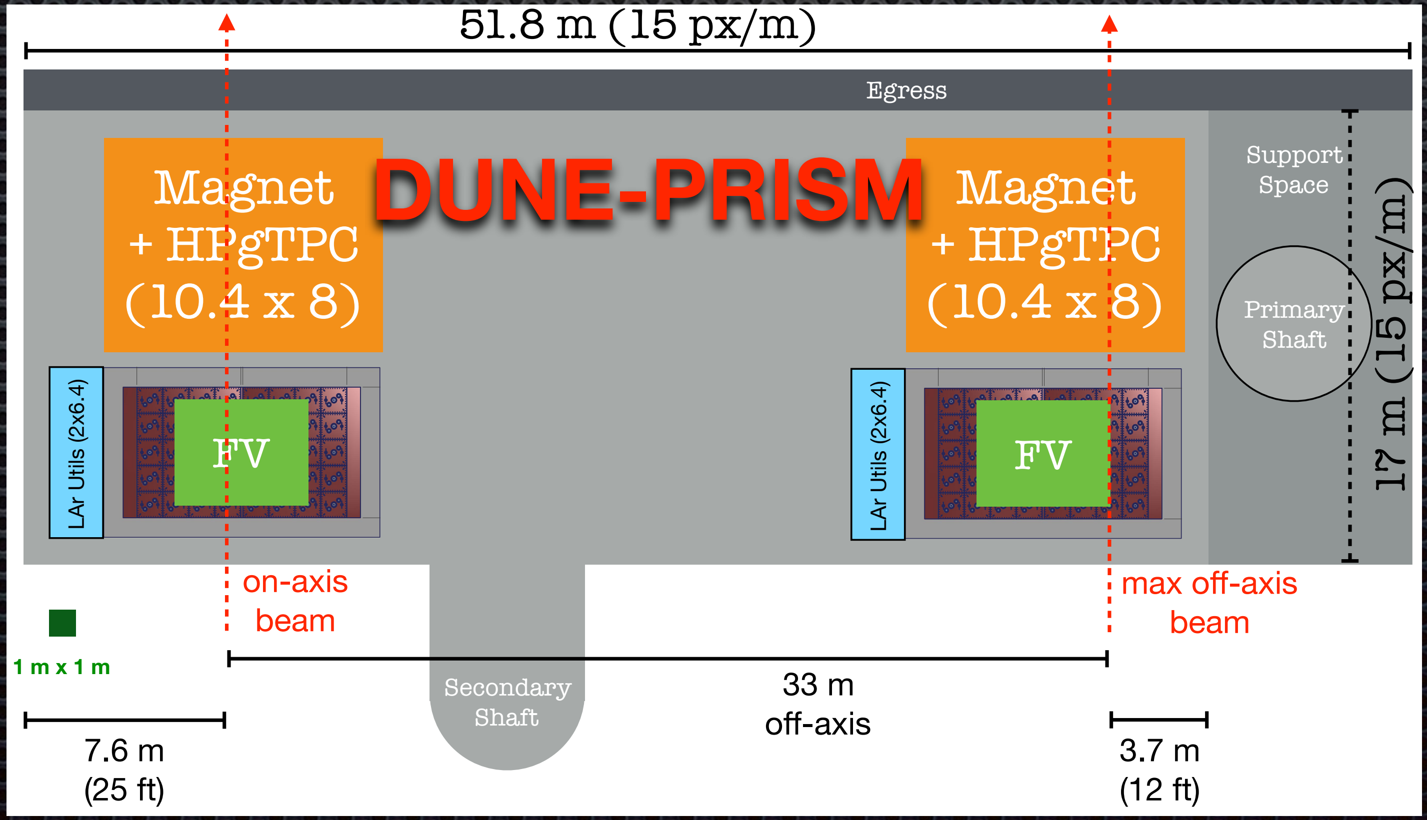
**The modeling problems that produced biased $\delta_{CP}$ measurement are easily identified off-axis**

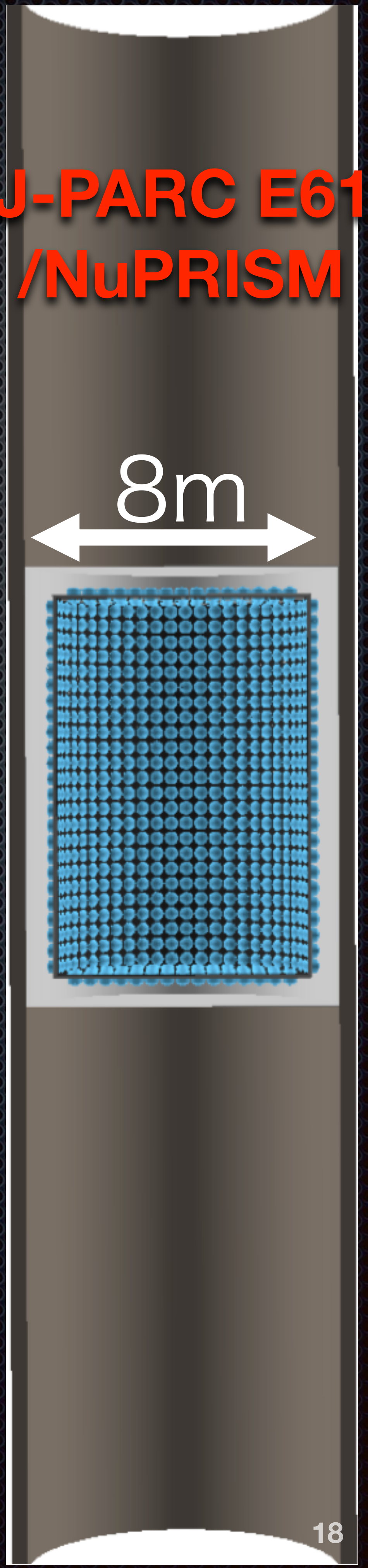
**It is extremely useful if such redundant measurements occur in the same beamline so that many important flux uncertainties will cancel**

recision
eaction
ndependent
pectrum
easurement

# PRISM Detectors

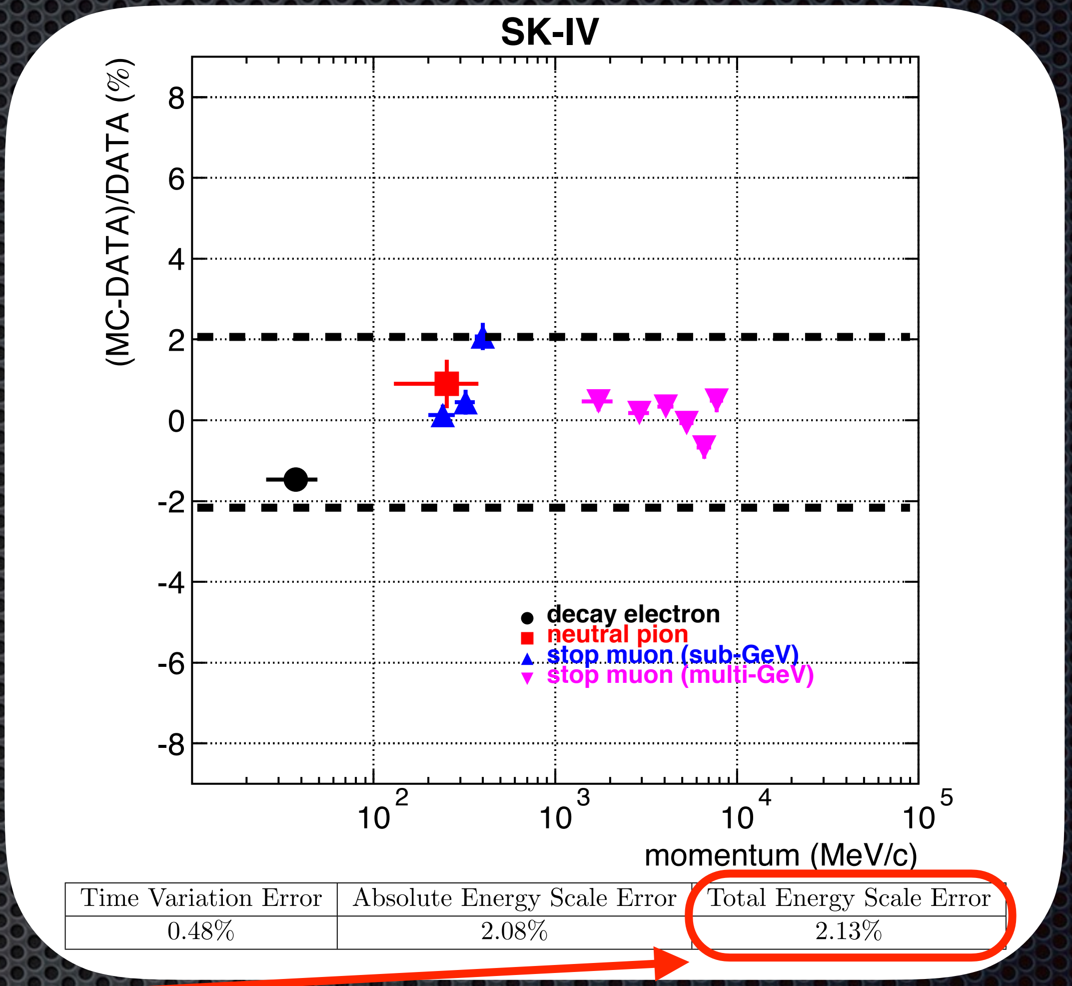
- Detectors that can move off-axis to sample different neutrino energies provide essential measurement redundancy
  - This information can also be used to make far detector predictions via linear combinations of near detector data
    - This would bypass model-based extrapolations to “first order” (see next talks)
- Both DUNE & Hyper-K are now planning to implement such detectors (see next talks)



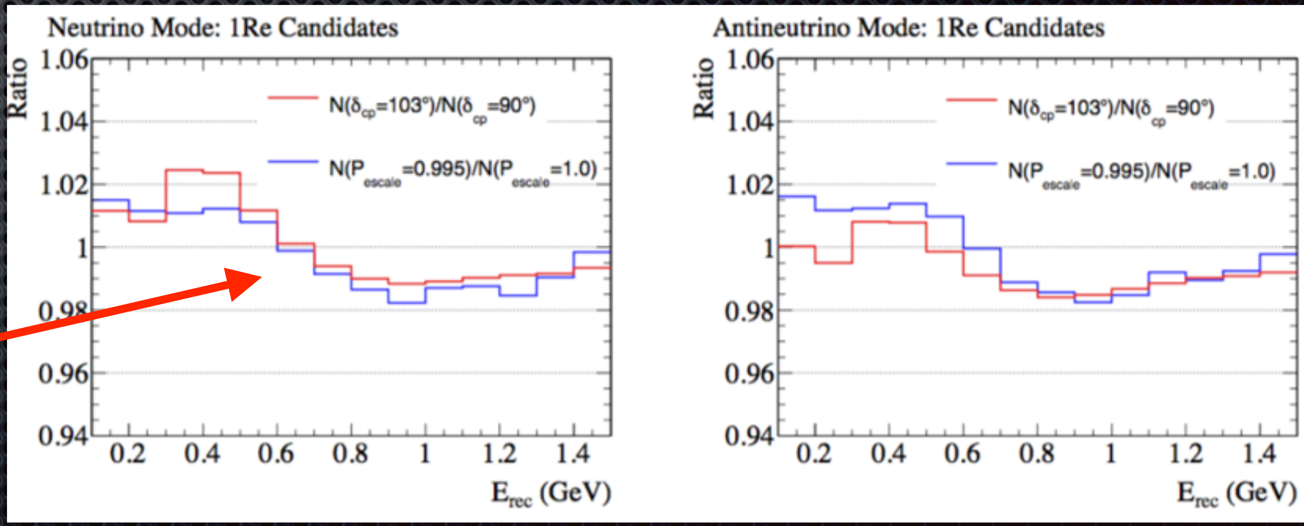


# Detector Systematic Uncertainties

- Detector systematic uncertainties must also be substantially improved for the next generation of experiments
- Example:
    - In Super-K, energy scale is considered known to 2.1%
    - In Hyper-K, a 0.5% energy scale uncertainty is equivalent to a 13° uncertainty in δCP

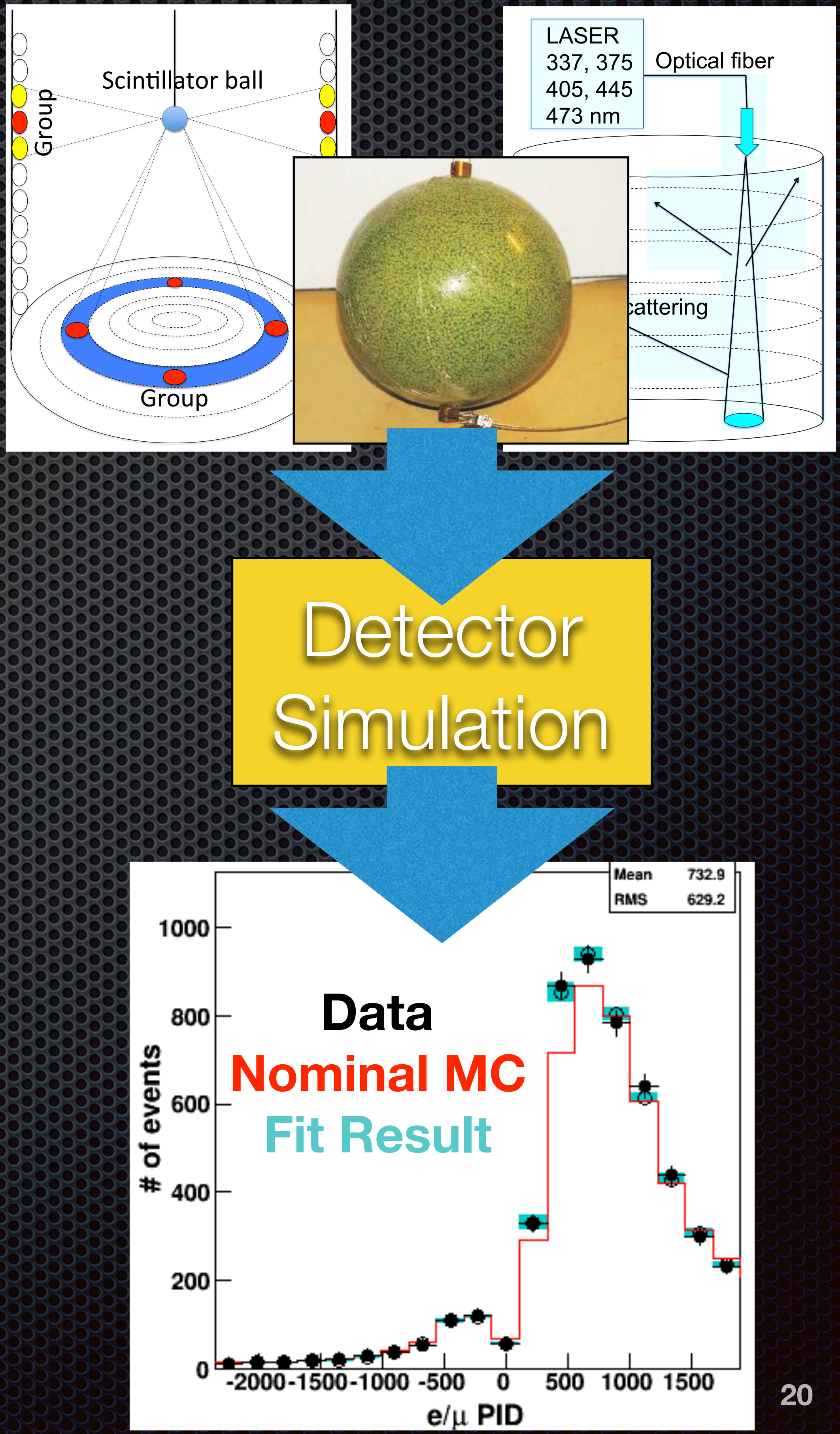


| Time Variation Error | Absolute Energy Scale Error | Total Energy Scale Error |
|---|---|---|
| 0.48% | 2.08% | 2.13% |

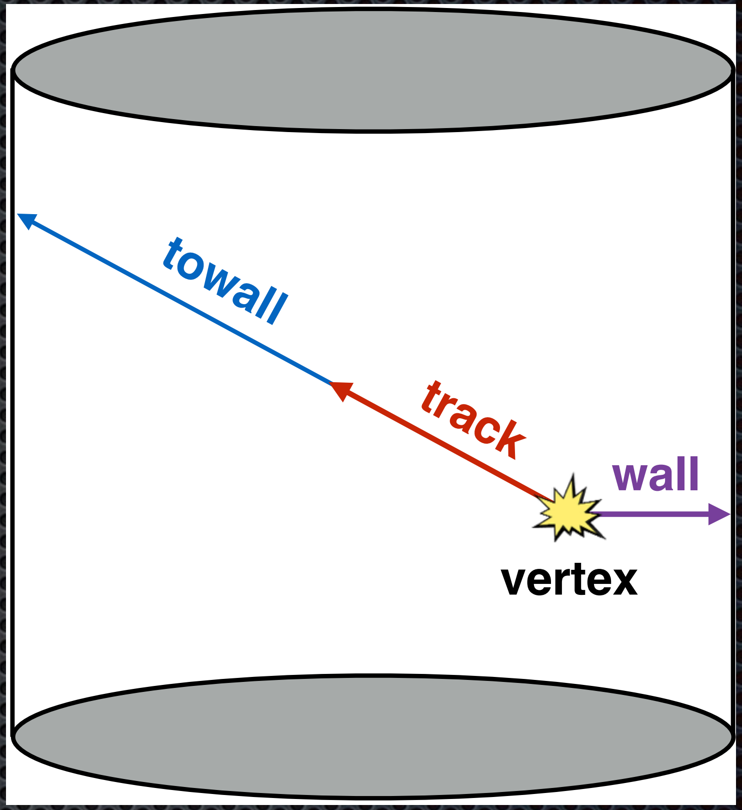


# Detector Errors ⇔ Calibration

- Detector uncertainties are intimately linked to calibration
    - A perfectly understood (and simulated) detector incurs no detector uncertainty
- However, there are a variety of ways to link calibration data to detector errors
- One method (T2K, Super-K):
    - Implement all detector calibrations into the detector simulation
    - Shift and smear simulated high level distributions (e.g. particle ID) until MC matches the data
        - Uncertainties on these shifts and smears are translated into detector uncertainty



# T2K/Super-K Detector Systematics

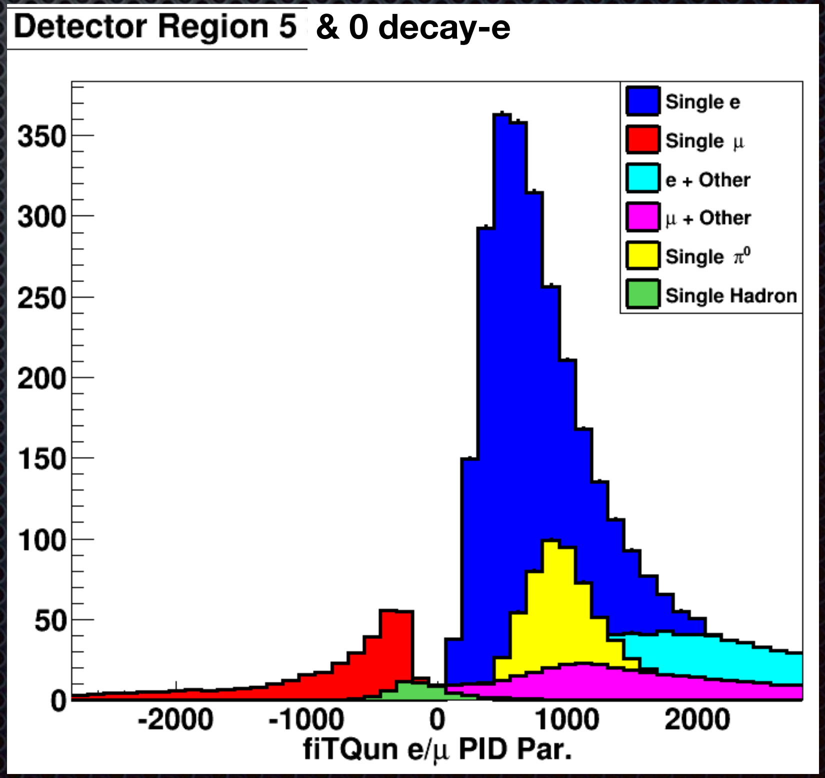


1. Divide atmospheric neutrino events into $N_{\text{decay-e}}$, wall/towall bins, Evisible, etc.

2. For each sample, plot the high level variables to be constrained ($N_{\text{rings}}$, e/μ PID, e/$\pi^0$ PID, etc.)

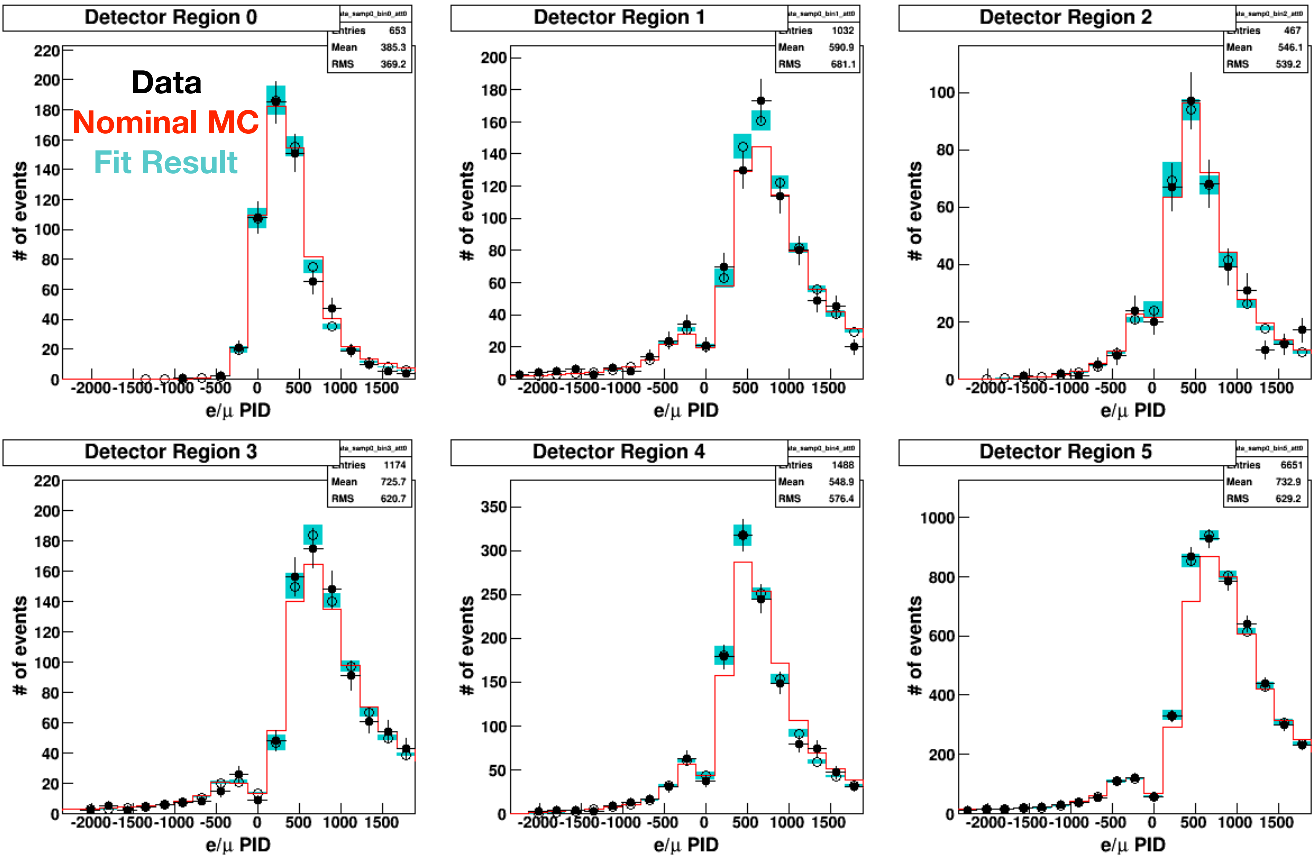
3. Allow each MC component to be smeared and shifted: $X' = \alpha X + \beta$

4. Constrain all nuisance parameters ($\alpha$ & $\beta$) in a fit to the atmospheric neutrino data

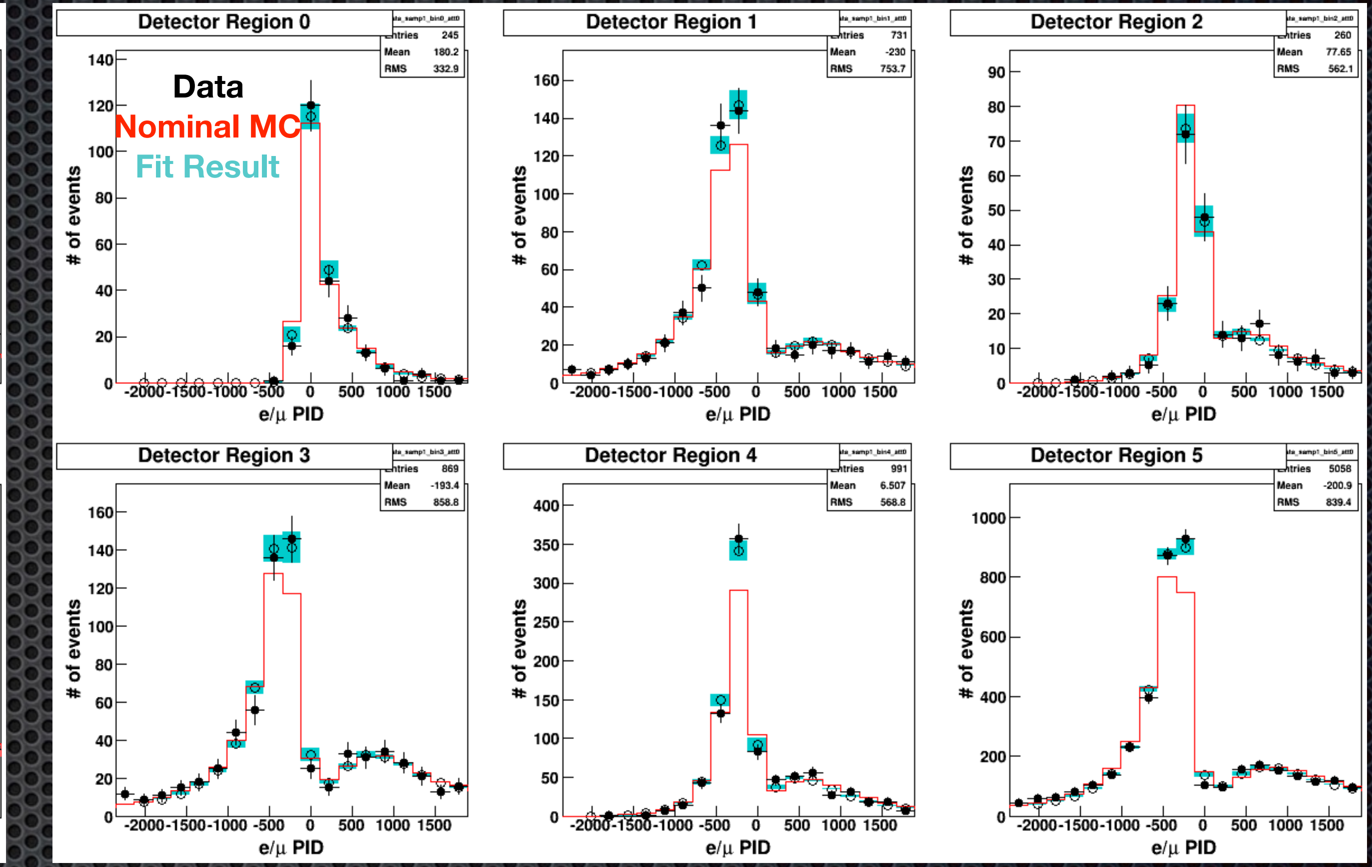


# (A Few) Atmospheric Fit Results

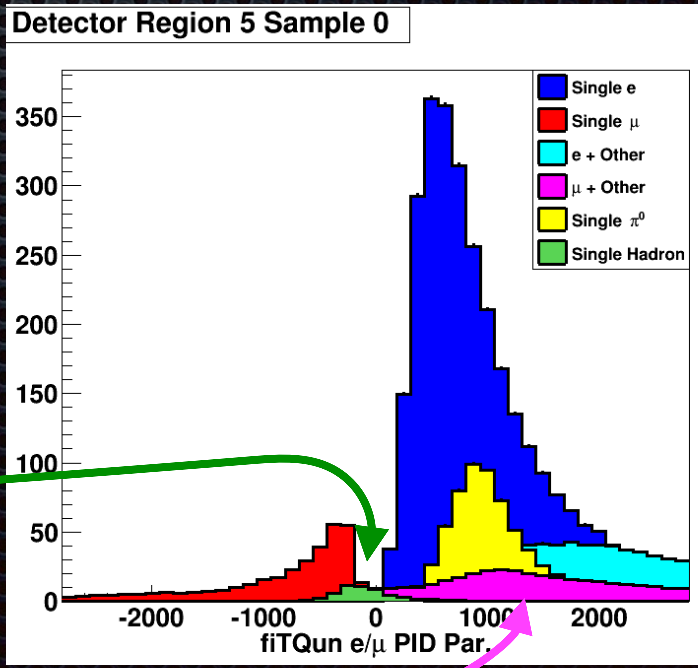
## 0 Decay-e



## 1 Decay-e



- Generally good fit agreement, but we used a lot of nuisance parameters within an unlikely-to-be-correct model to get there
- And, ultimately, some of the most problematic samples are poorly constrained:
  - Single $\pi^+$ backgrounds
  - CC$\pi^+$ events with a $\pi^+$ near the Cherenkov threshold



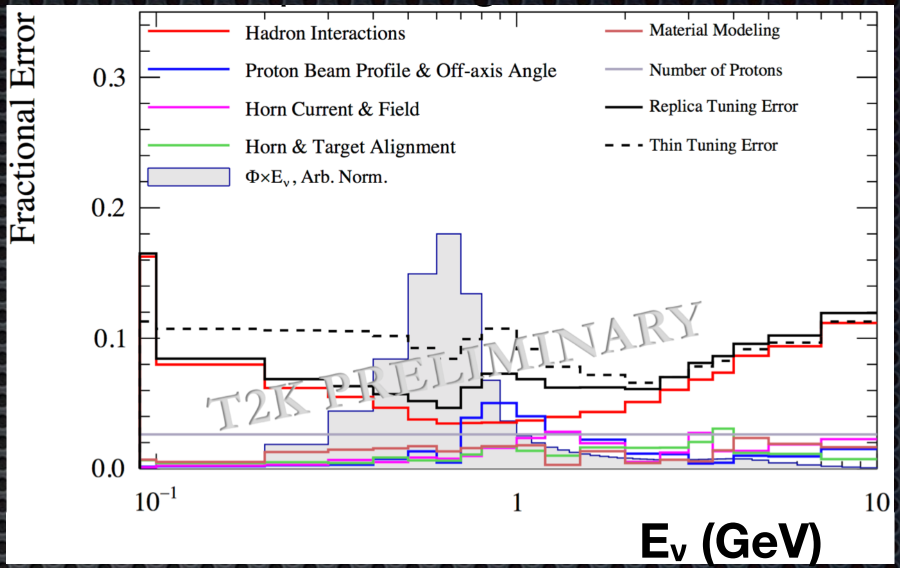
# A Path Forward (Detector Errors)

- A more accurate model for detector uncertainties would be a model based directly on the **underlying physics of the detector**
    - PMT QE, discriminator thresholds, reflectivity of PMTs and detector walls, water attenuation, Rayleigh scattering, Mie scattering, etc.
    - While not perfect (or complete), such a model would directly translate physical uncertainties in the detector to the high level variables used in event selections
    - This model can then constrain all types of interactions, even those that are rare in control samples (or, e.g. atmospheric neutrino data)
- These models are often **computationally difficult**, since detector parameters are **not easily reweightable**, but several approaches are possible
    - Brute force simulations of key calibration/control samples can be used to reduce the allowed detector parameter space (e.g. MiniBooNE, NOvA)
    - Several interesting ideas exist to translate detector parameter PDFs to high-level variable PDFs (e.g. using a BDT)
- Ultimately, regardless of the model, control of detector systematics relies on a **high-precision and redundant calibration program**

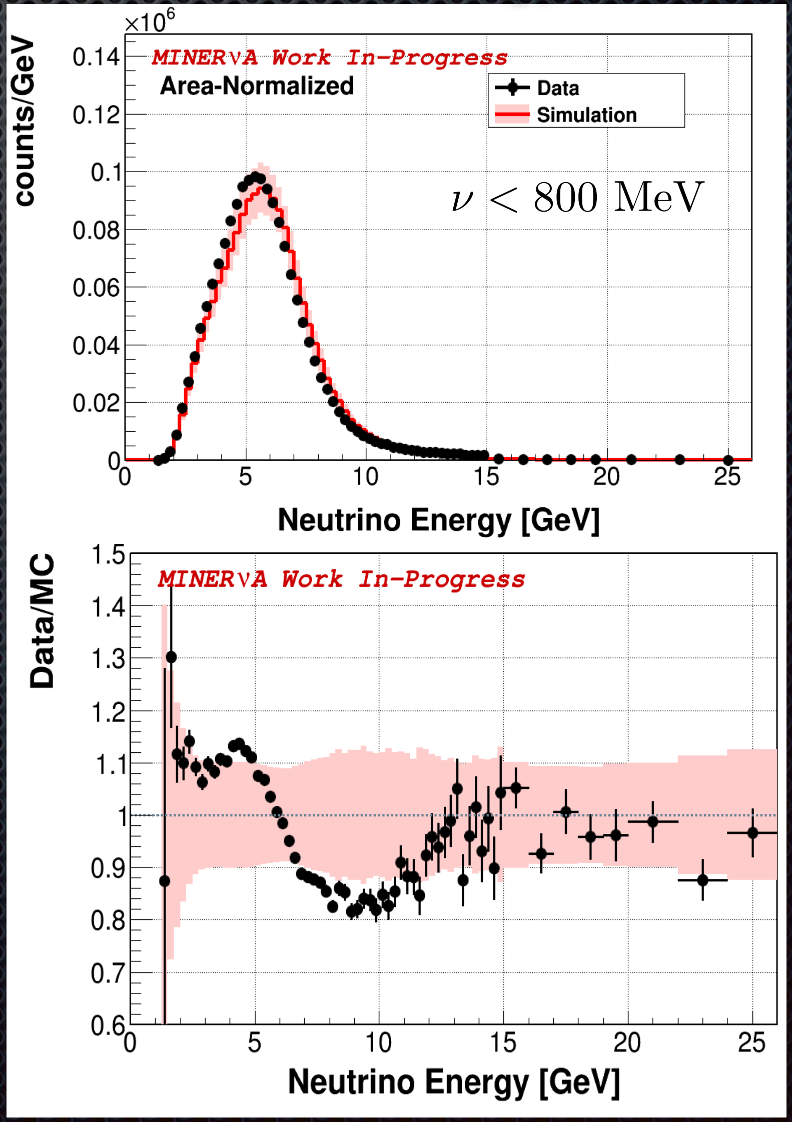
# Flux Uncertainties

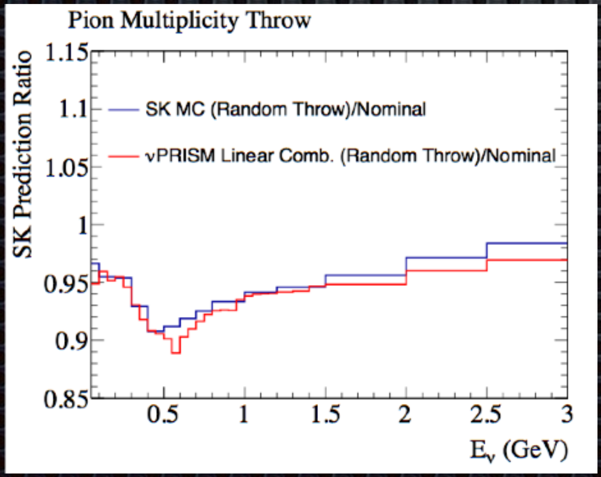
- Substantial progress has been made on reducing flux uncertainties
  - Replica target data has recently reduced hadron production uncertainties at J-PARC to the level of beam direction and horn focusing uncertainties
  - Secondary (and tertiary) interactions on other material outside the target must also be constrained (see T. Yoshida talk yesterday)
- Detailed characterization and monitoring of the beam are essential, particularly to search for deficiencies in the flux modeling
  - Large discrepancies seen in Minerva data for NuMI medium energy flux, but not seen in horn-off data (suggests beam optics)
- For analysis methods using off-axis ND measurements, beam focusing uncertainties can be more important than hadron production uncertainties

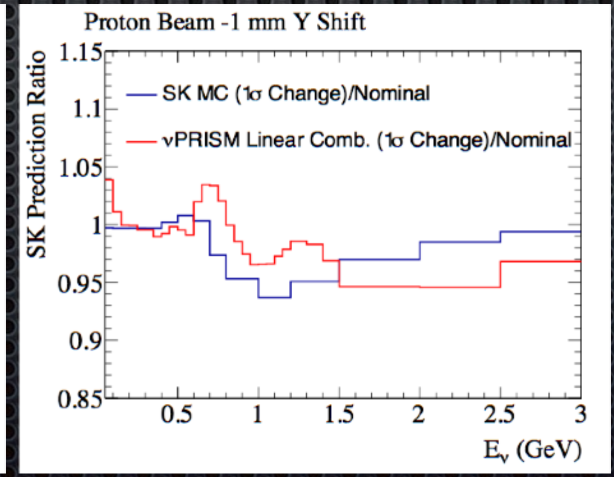
J-PARC Neutrino Beam

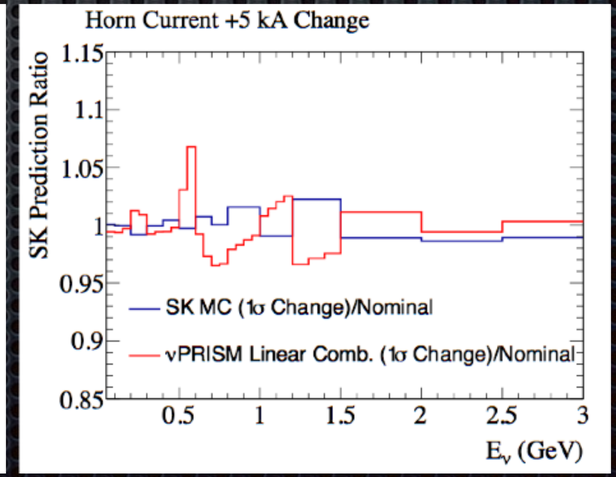


Minerva data with NuMI Medium Energy Flux









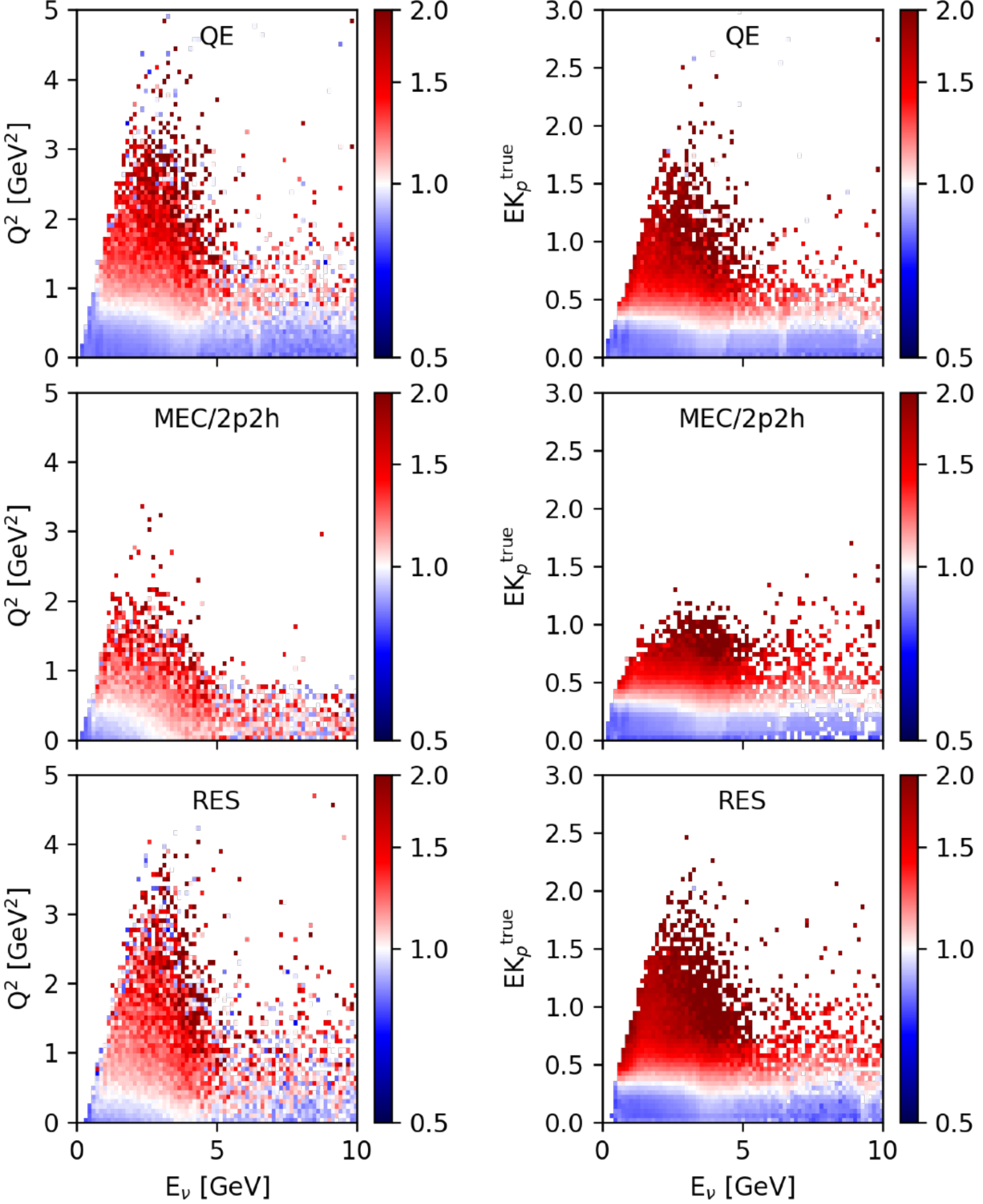
# Summary

- The next generation of accelerator-based oscillation experiments will require few percent uncertainties (on far detector event rate and shape predictions)
- Neutrino-nucleus interaction modeling is difficult at the GeV scale, and existing models are unlikely to provide sufficient precision for future experiments
    - T2K is already starting to run into model limitations
    - With incorrect cross section modeling, it is possible to get the wrong answer for $\theta_{23}$, $\delta_{CP}$, etc., even with good agreement in near detector samples
    - DUNE and Hyper-K plan address this problem with measurements at a variety of off-axis angles
- Detector systematic errors are already contribute a substantial fraction of the total systematic error
    - New methods to more-closely link detector uncertainties to low-level detector response variables are under development, and may be essential for future experiments
- Neutrino flux uncertainties must also be precisely controlled, and sufficient monitoring redundancy to search for beam modeling problems are essential

# Backup

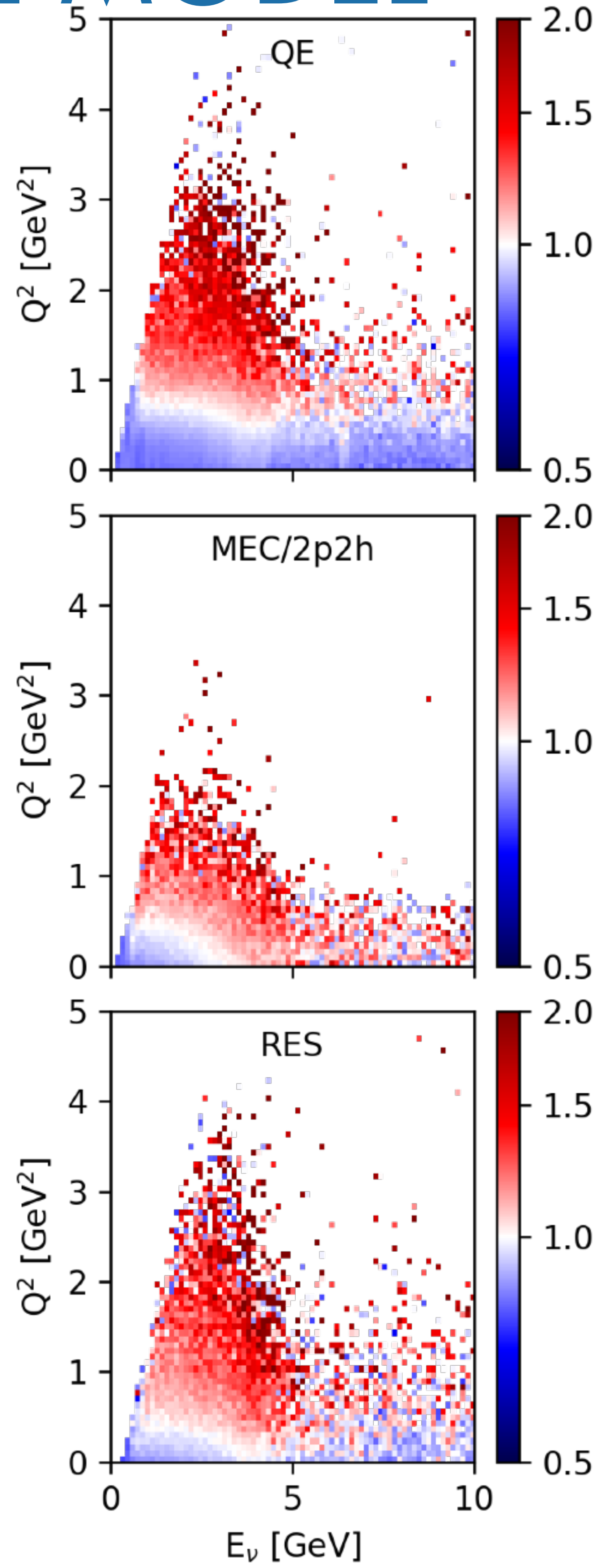
# PROPAGATING THE MODEL

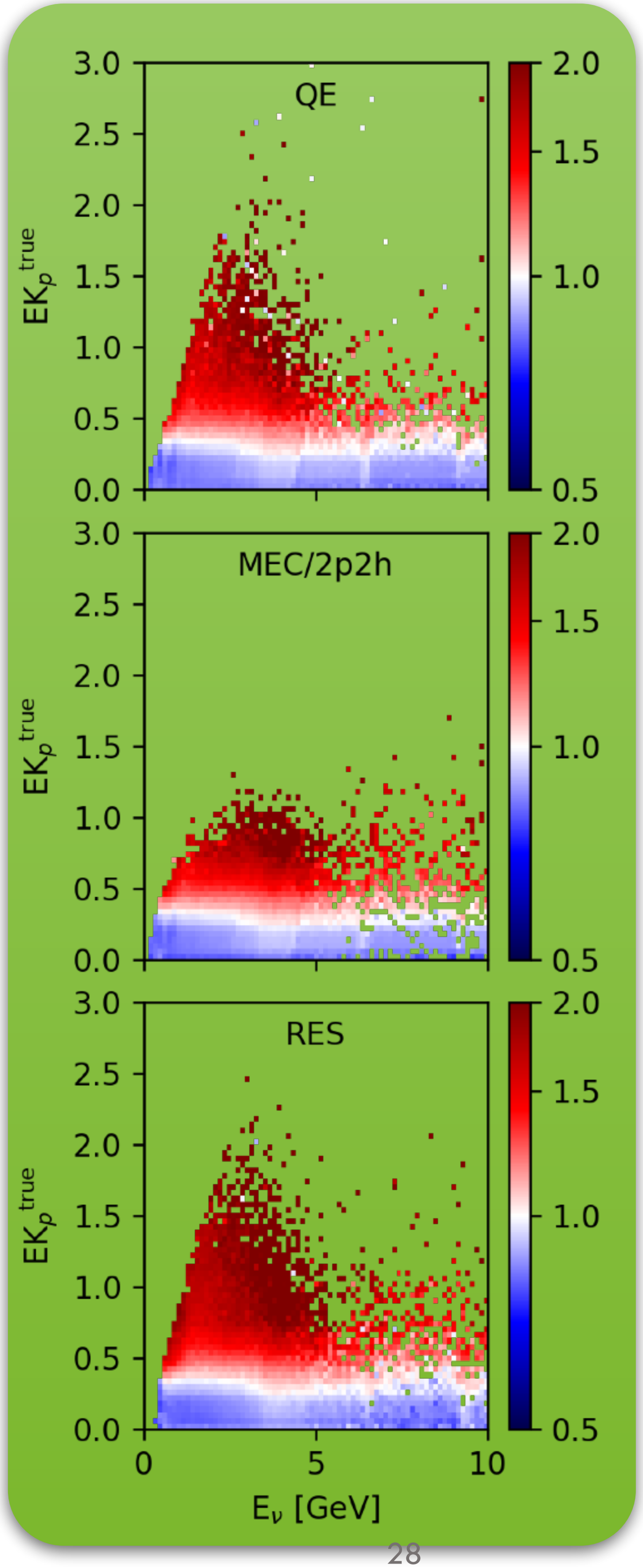
- To study the effect on oscillation fits, we need to propagate this model to far detector.
  - Also to off-axis near detector stops, to demonstrate the PRISM technique.
- Bin event weights in true variables typically used to describe interaction models.
  - Get smoothly varying functions!
  - MVA treats interaction modes differently.
    - Even though it doesn't "know" about them!



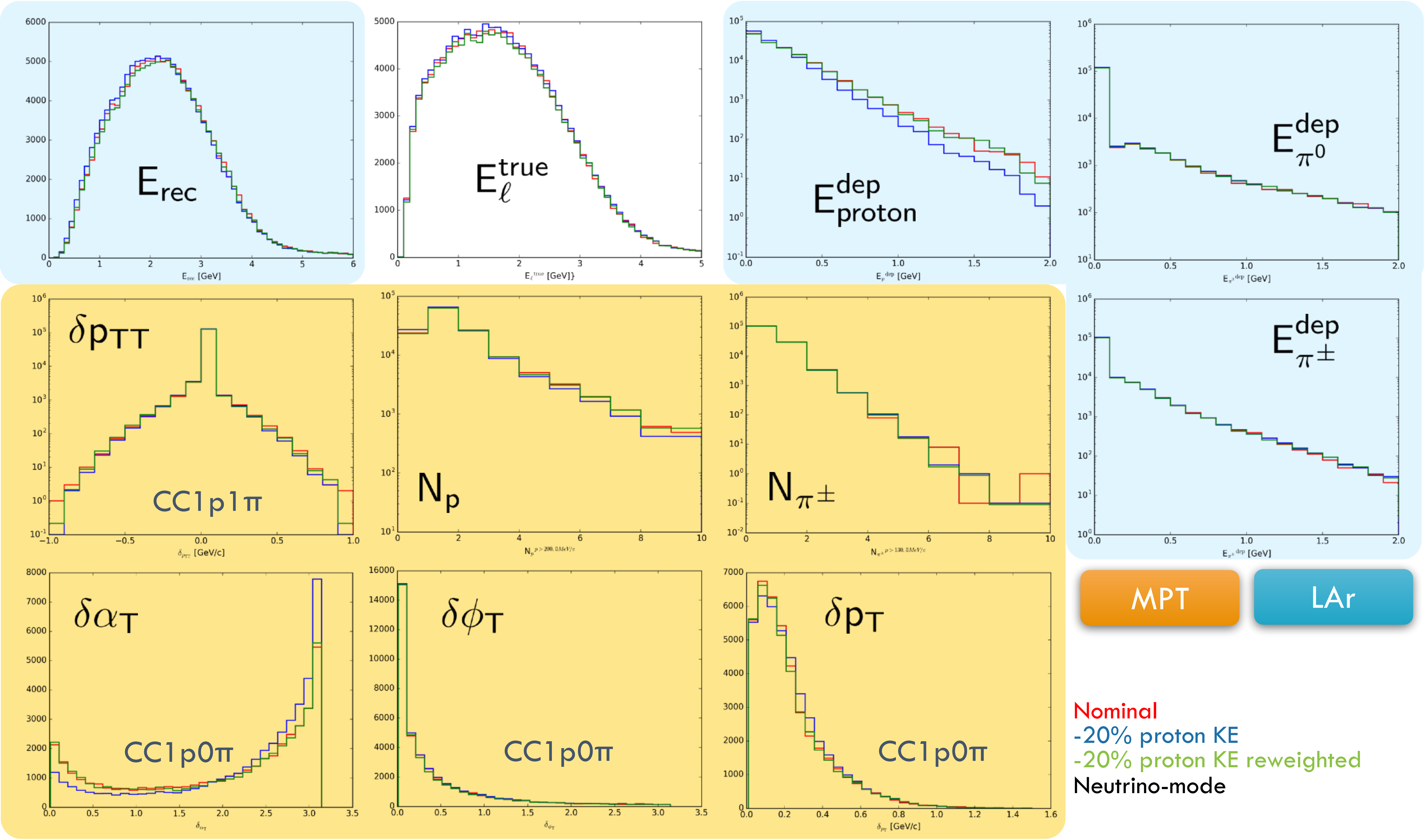
# PROPAGATING THE MODEL

- For this data set, use $E_\nu$ vs true proton kinetic energy.
- Extract weights separately for ν and anti-ν using FHC and RHC on-axis near detector data.
  - Assume perfect charge separation.
- Do not reweight regions of the space that fall outside of the ND acceptance.
  - These events get weight = 1, but 20% proton deposited energy removed.



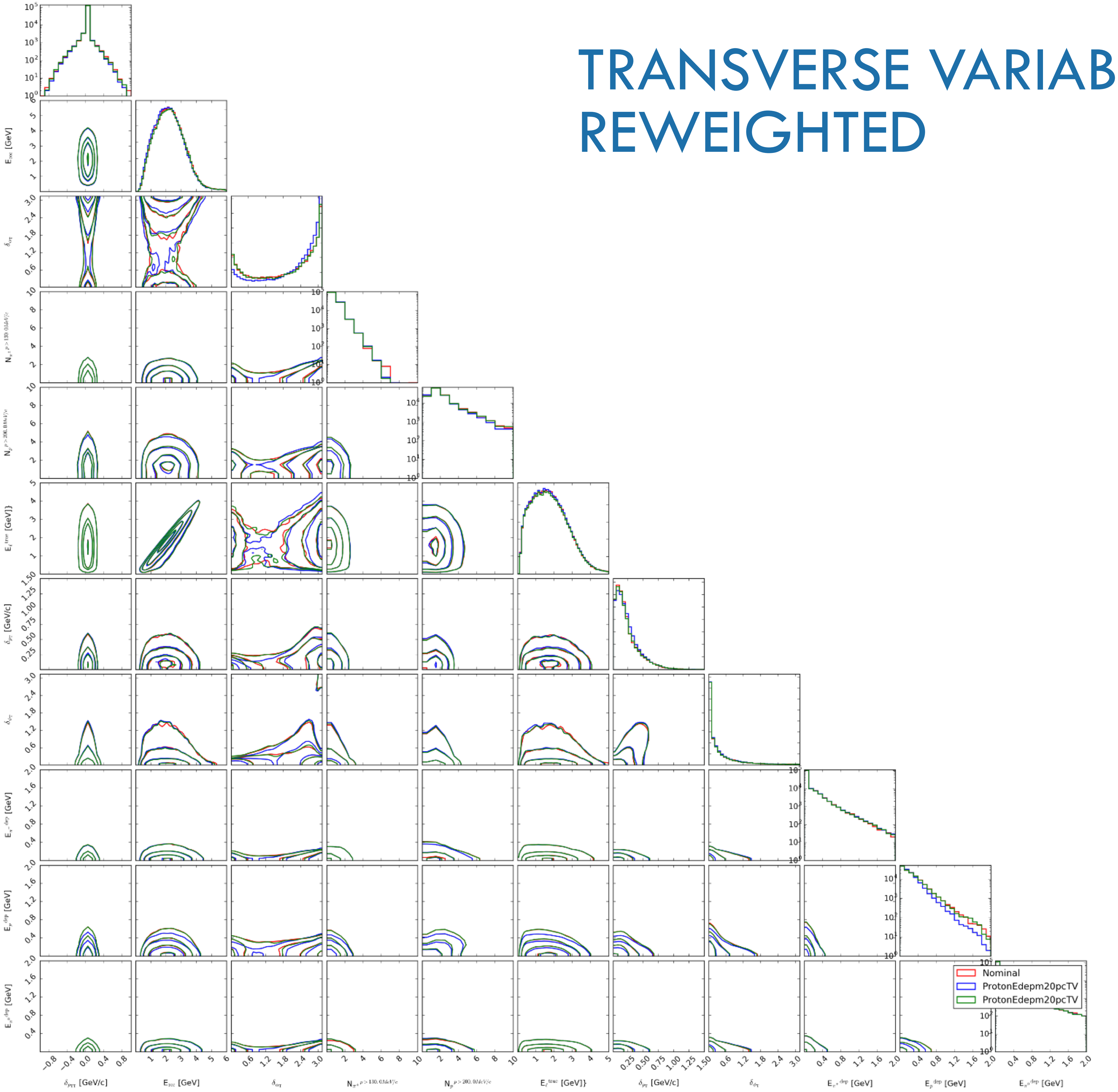


# TRANSVERSE VARIABLES, REWEIGHTED

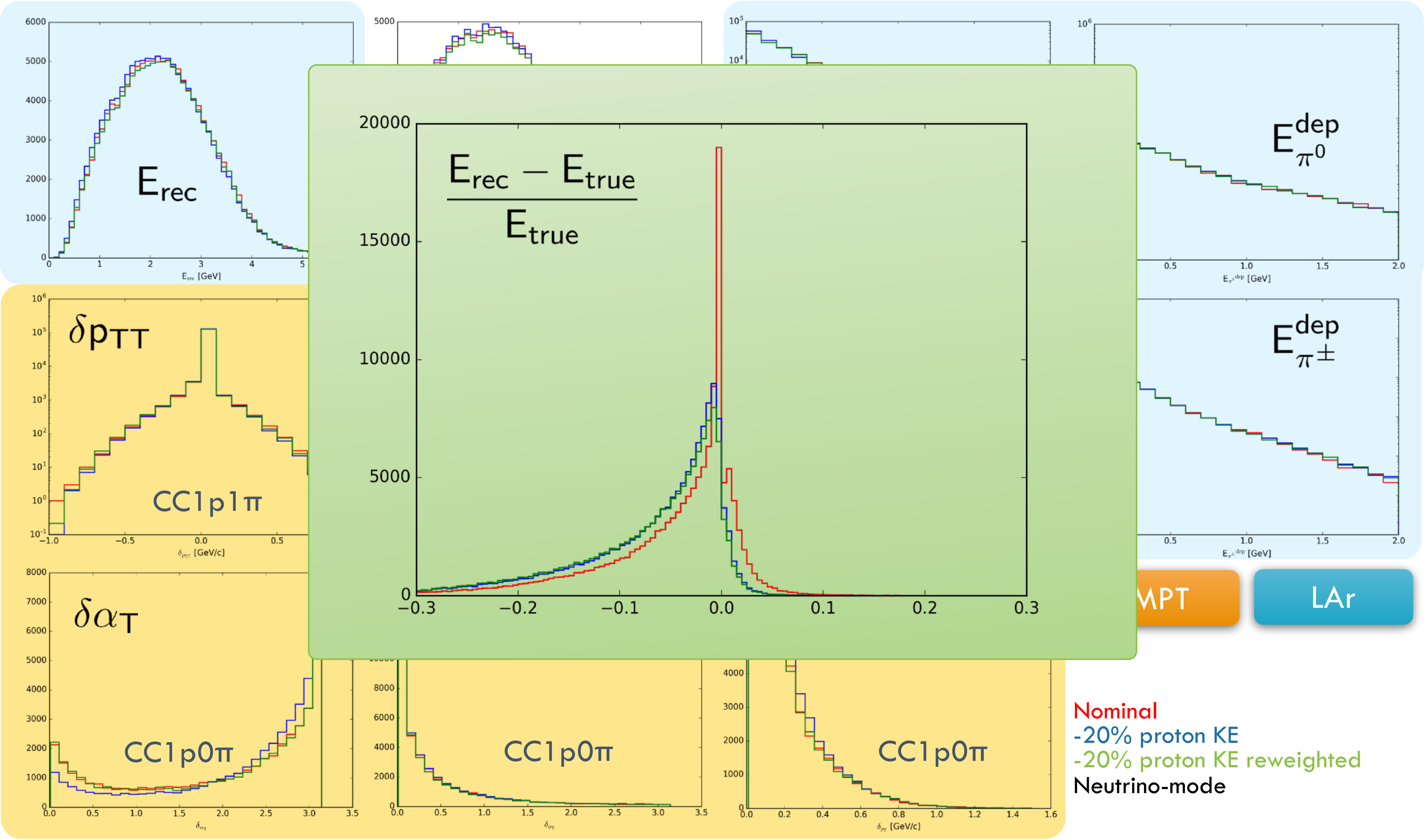


Nominal
-20% proton KE
-20% proton KE reweighted
Neutrino-mode

# TRANSVERSE VARIABLES, REWEIGHTED



# TRANSVERSE VARIABLES, REWEIGHTED



Nominal
-20% proton KE
-20% proton KE reweighted
Neutrino-mode