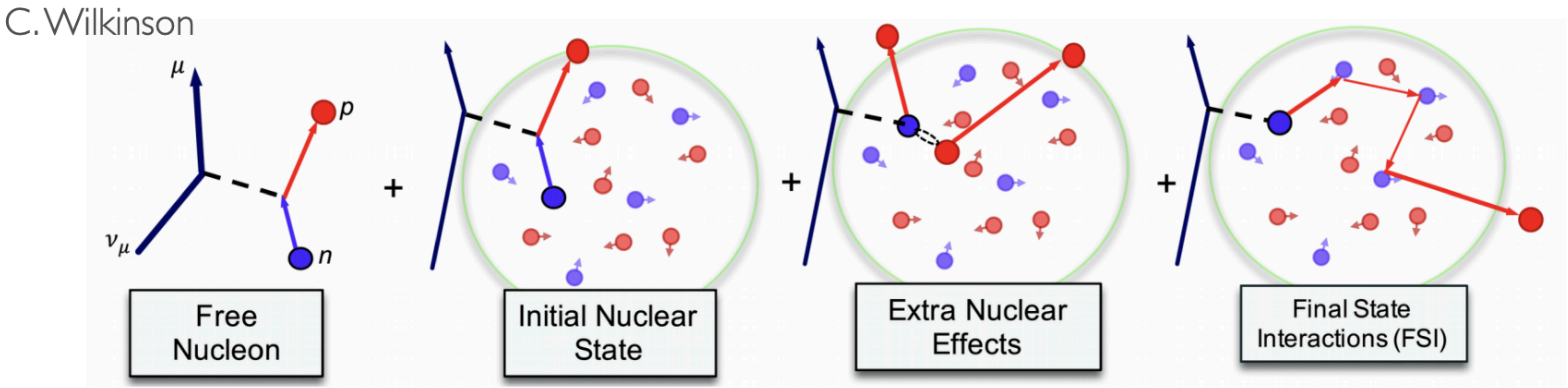
Model Validation Techniques for Cross-Section Extraction

Ben Bogart, London Cooper-Troendle, Wenqiang Gu, Lee Hagaman, Xin Qian, Nitish Nayak

4th October, 2023

NuXtract Conference

Introduction



- Many different effects, even for “simple” interactions :
- No first-principle based full description yet — different models superimposed
  - QCD nature of interactions means difficult to envision a non-perturbative approach
- Models necessarily approximations => CV + uncertainties (i.e need to evaluate “model space” using data)
- Real data can be very informative to test self-consistency and probe model bias

Approach to Cross-Section Extraction

For other MicroBooNE approaches, see Afro's talk earlier

Measured Distribution ↖ Response Matrix ↑

$$M = R \cdot S + B$$

→ Background Spectrum

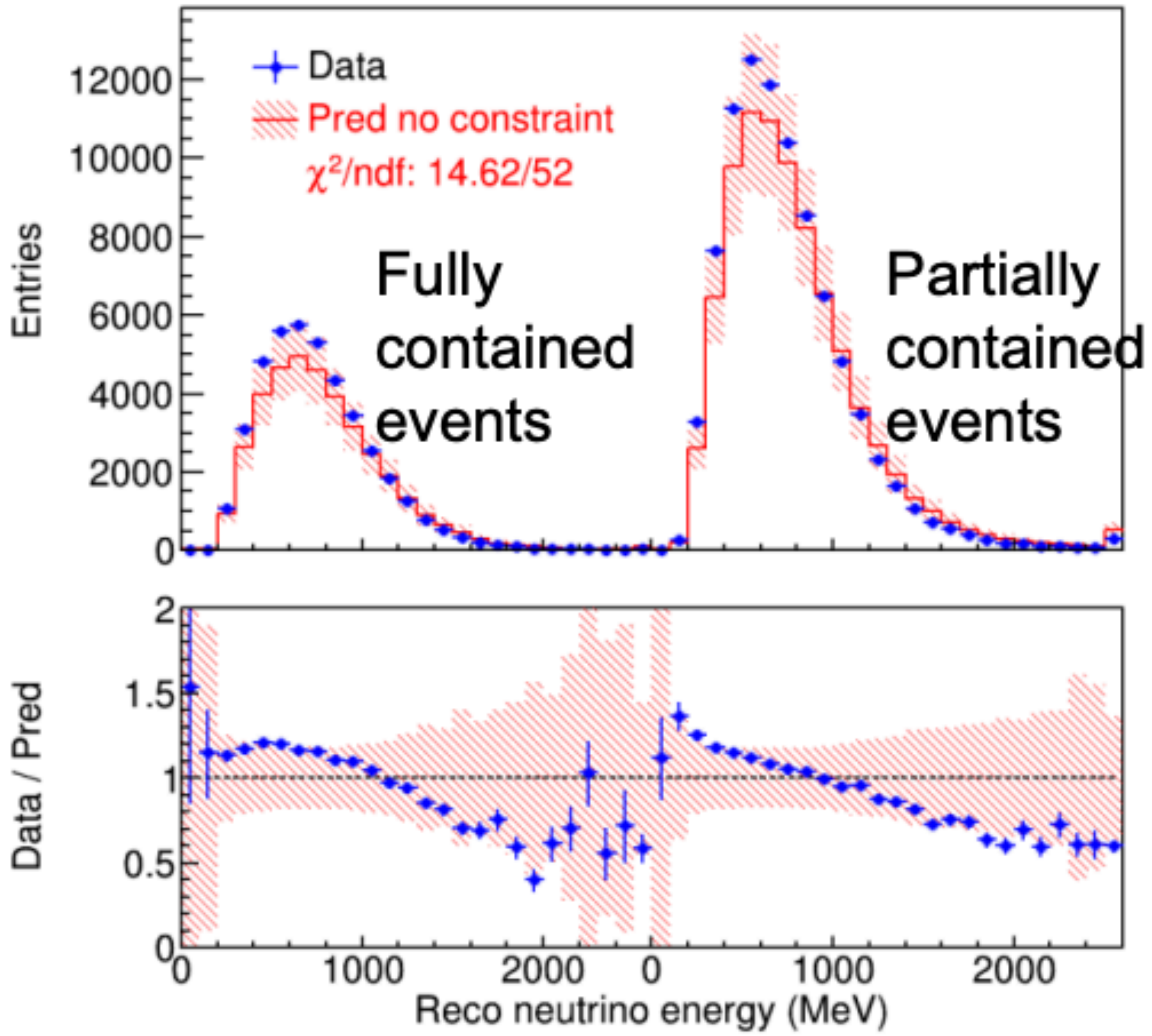
↓ Signal - Cross-Section

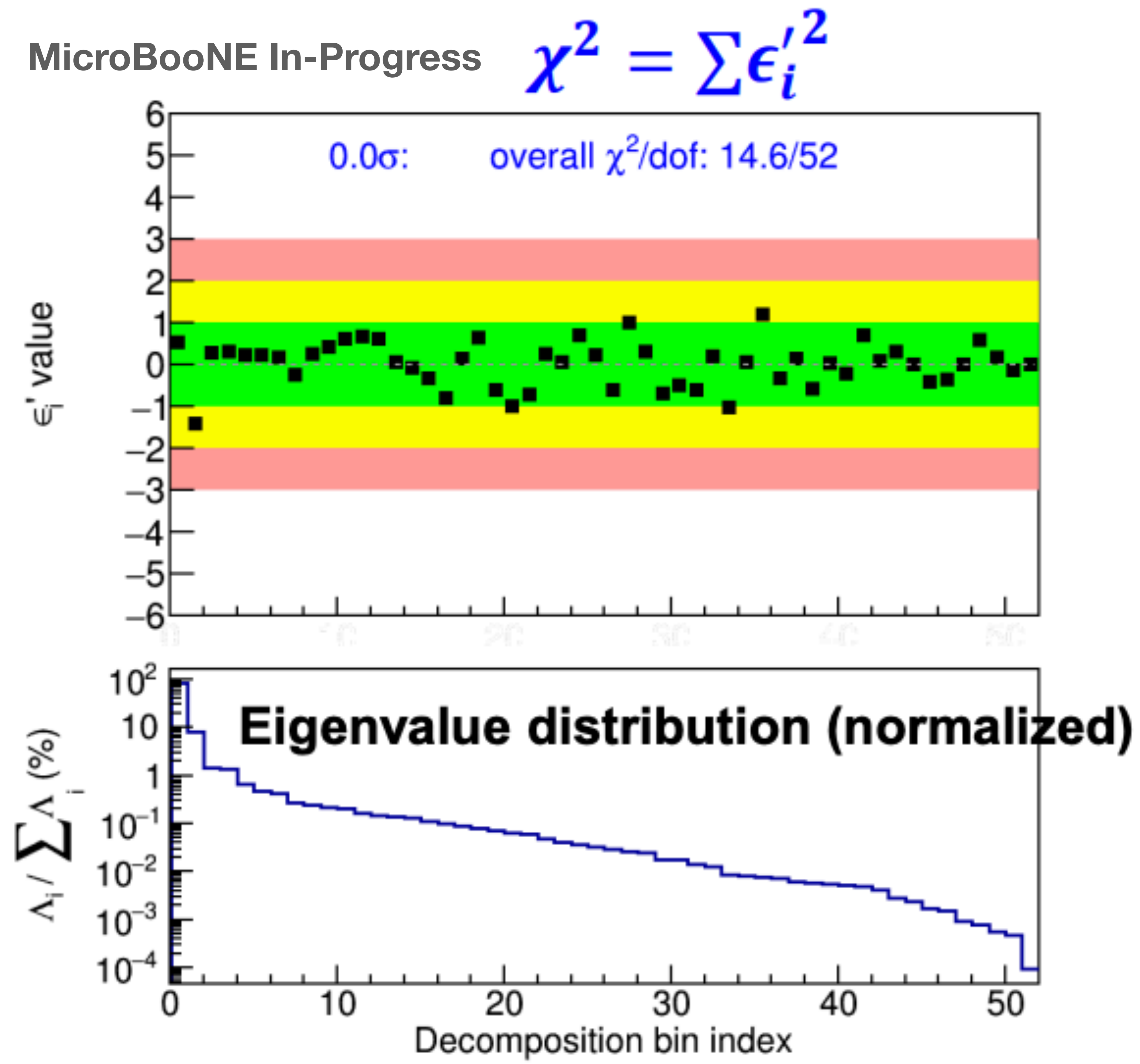
$$S_j = \frac{\int \bar{F}(E_{\nu j}) \cdot \sigma(E_{\nu j})\, dE_{\nu j}}{\int \bar{F}(E_{\nu j})\, dE_{\nu j}}$$

Explicitly nominal flux-averaged

- Very straightforward to compare to model predictions (just use nominal flux)
- To use nominal flux in comparisons, one must show small model bias for $D(E_\nu \rightarrow T_{reco})$ wrt quoted uncertainty budget (L. Koch, S. Dolan - Phys Rev D 102,113012)
- We check for model bias through a variety of tests. However, principal approach is to look at things "post-data"
  - Use real data to inform whether uncertainties are enough to cover observed discrepancies
- Fake data tests play an important role too :
  - Probe sensitivity of test to uncover injected model bias
- Unfold IFF a sufficiently sensitive test tells us the model bias is sub-dominant wrt total uncertainties, otherwise inflate them as necessary

Assessing Model using Data





- We can use standard goodness of fit tests (χ^2) to probe model performance
- But a single test isn't enough (bins are correlated significantly) :
  - Probe χ^2 along each eigenvector of covariance matrix -> check outliers (after correcting for look-elsewhere) : "differential goodness of fit"
- Moreover just goodness of fit doesn't validate $D(E_\nu \rightarrow T_{reco})$ — need more sensitive tests

Conditional Constraining Method

For some more discussion, see Xin's talk earlier

- Joint distribution of observables (X, Y) : assumed gaussian in each bin — for eg (E_μ, E_{had})

- We can use real data observed in one observable (E_μ) to inform how the model space performs in another observable (E_{had})

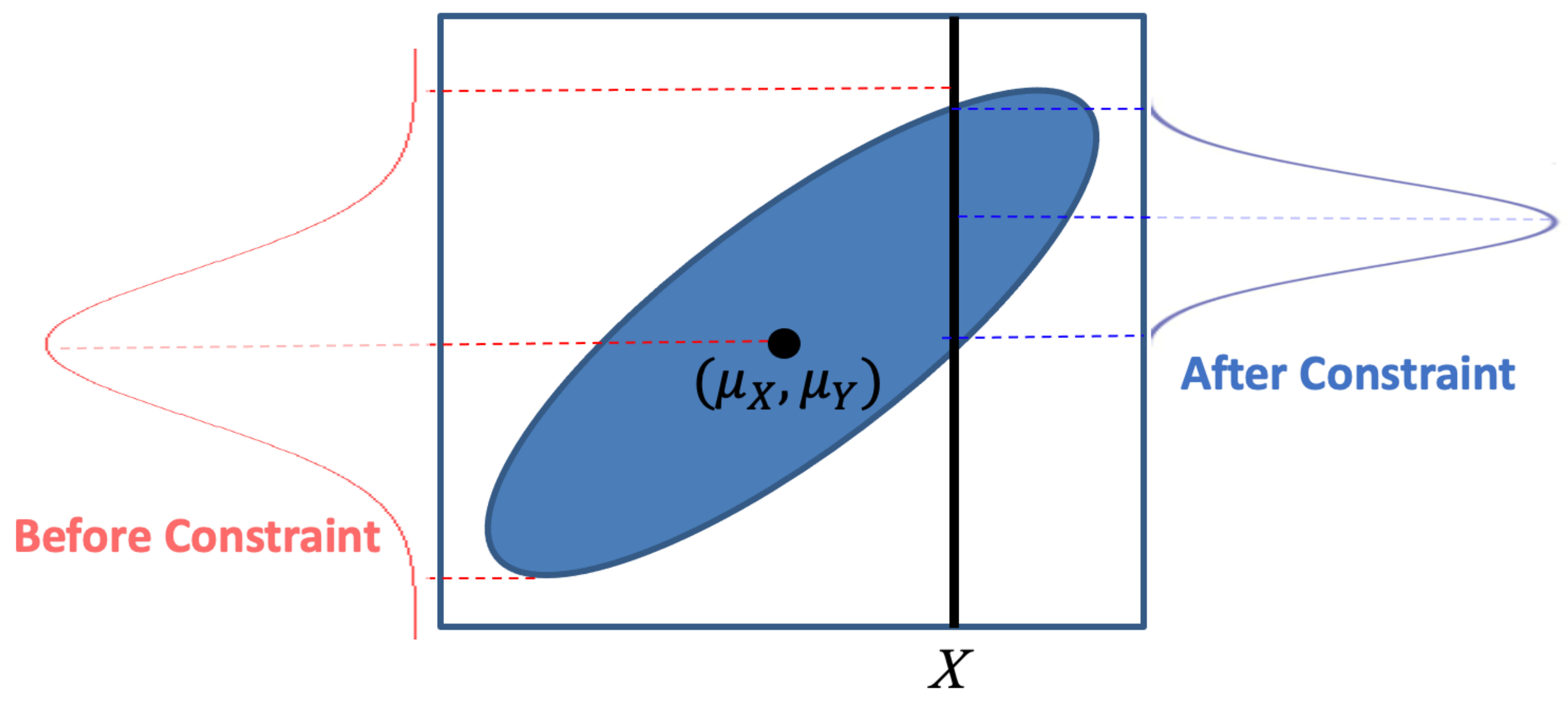
  - More stringent than just comparing E_{had} data and MC directly with uncertainties

Conditional expectation & covariance

$$\mu_{X,Y} = \begin{pmatrix} \mu_X \\ \mu_Y \end{pmatrix}, \qquad \Sigma_{X,Y} = \begin{pmatrix} \Sigma_{XX} & \Sigma_{XY} \\ \Sigma_{YX} & \Sigma_{YY} \end{pmatrix}$$

$$\mu_{Y|X} = \mu_Y + \Sigma_{YX}\Sigma_{XX}^{-1}(X - \mu_X)$$

$$\Sigma_{Y|X} = \Sigma_{YY} - \Sigma_{YX}\Sigma_{XX}^{-1}\Sigma_{XY}$$



$$\begin{pmatrix} \mu(E_{had}^{rec}) \\ \Sigma(E_{had}^{rec}) \end{pmatrix} + M(E_\mu^{rec}) = \begin{pmatrix} \mu(E_{had}^{rec} \mid E_\mu^{rec}, E_\nu) \\ \Sigma(E_{had}^{rec} \mid E_\mu^{rec}, E_\nu) \end{pmatrix}$$

Prior model — Sideband — Posterior model

- Uncertainties include the flux, detector response, cross-section model etc

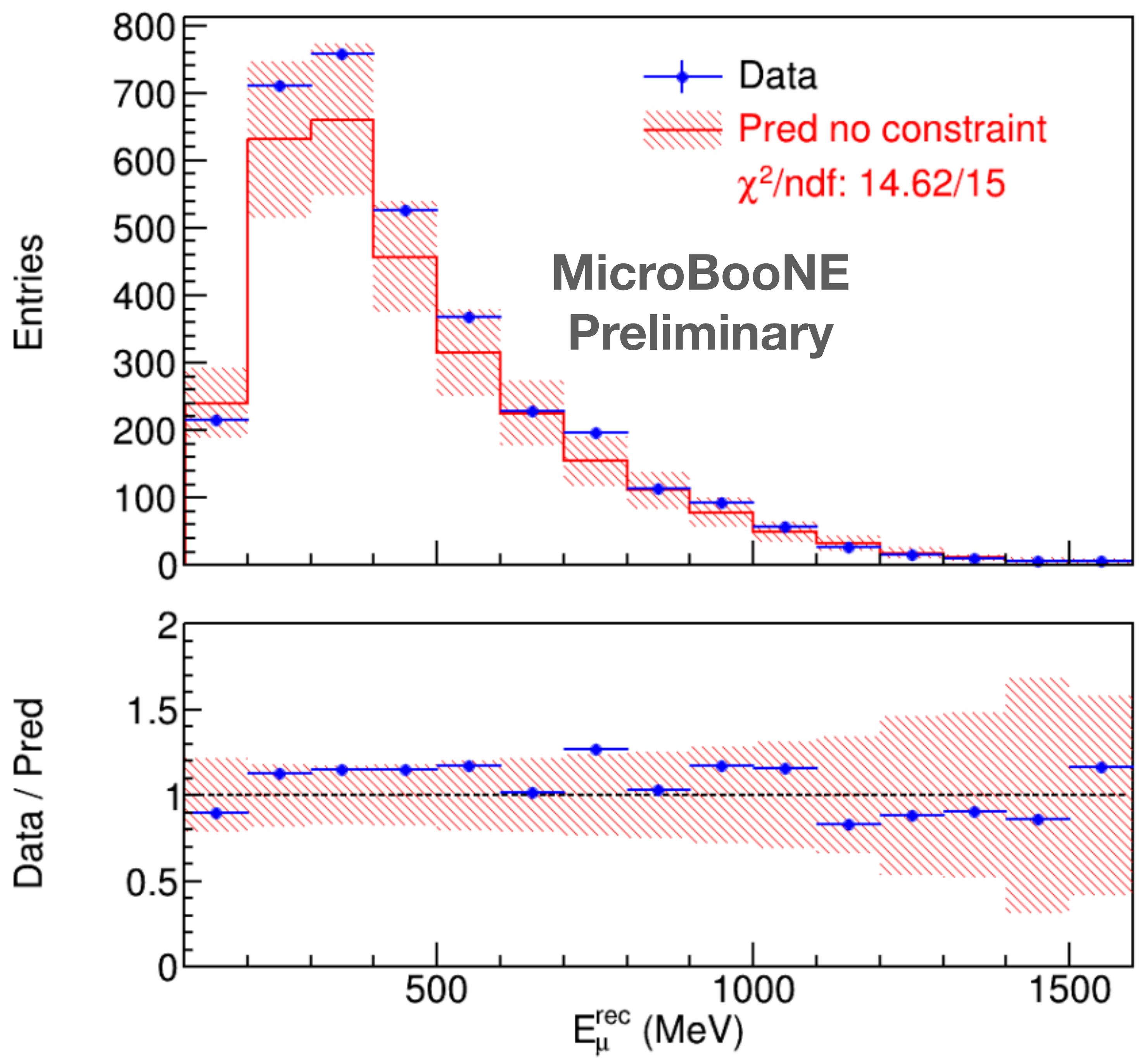
  - Post-constraint - only uncorrelated uncertainty on E_{had} remains

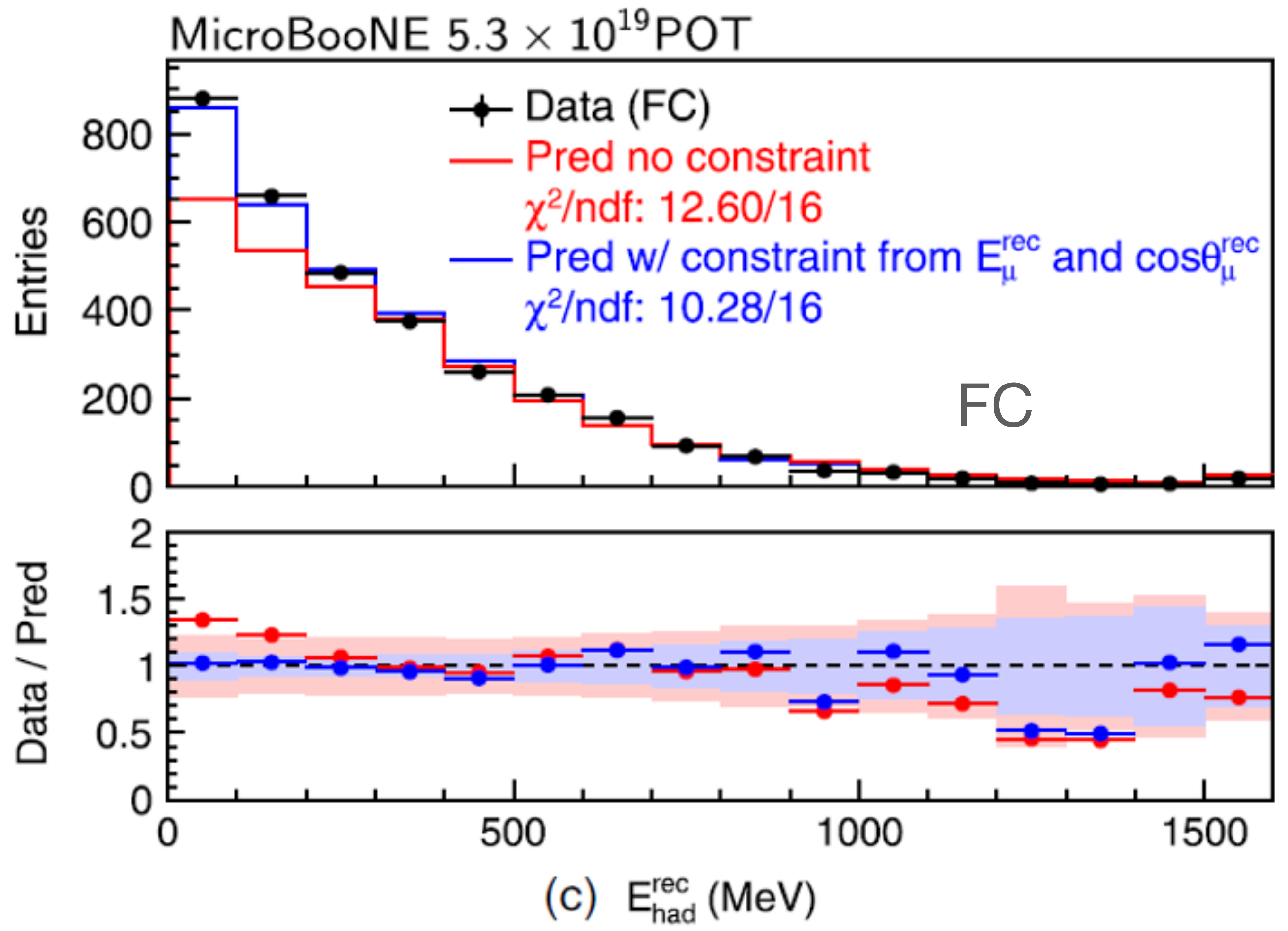
- Probes :

  - assumptions used by the model for $D(E_\nu \rightarrow E_\nu^{reco})$ + uncertainties taken on that assumption

Conditional Constraining Method

For some more discussion, see Xin's talk earlier



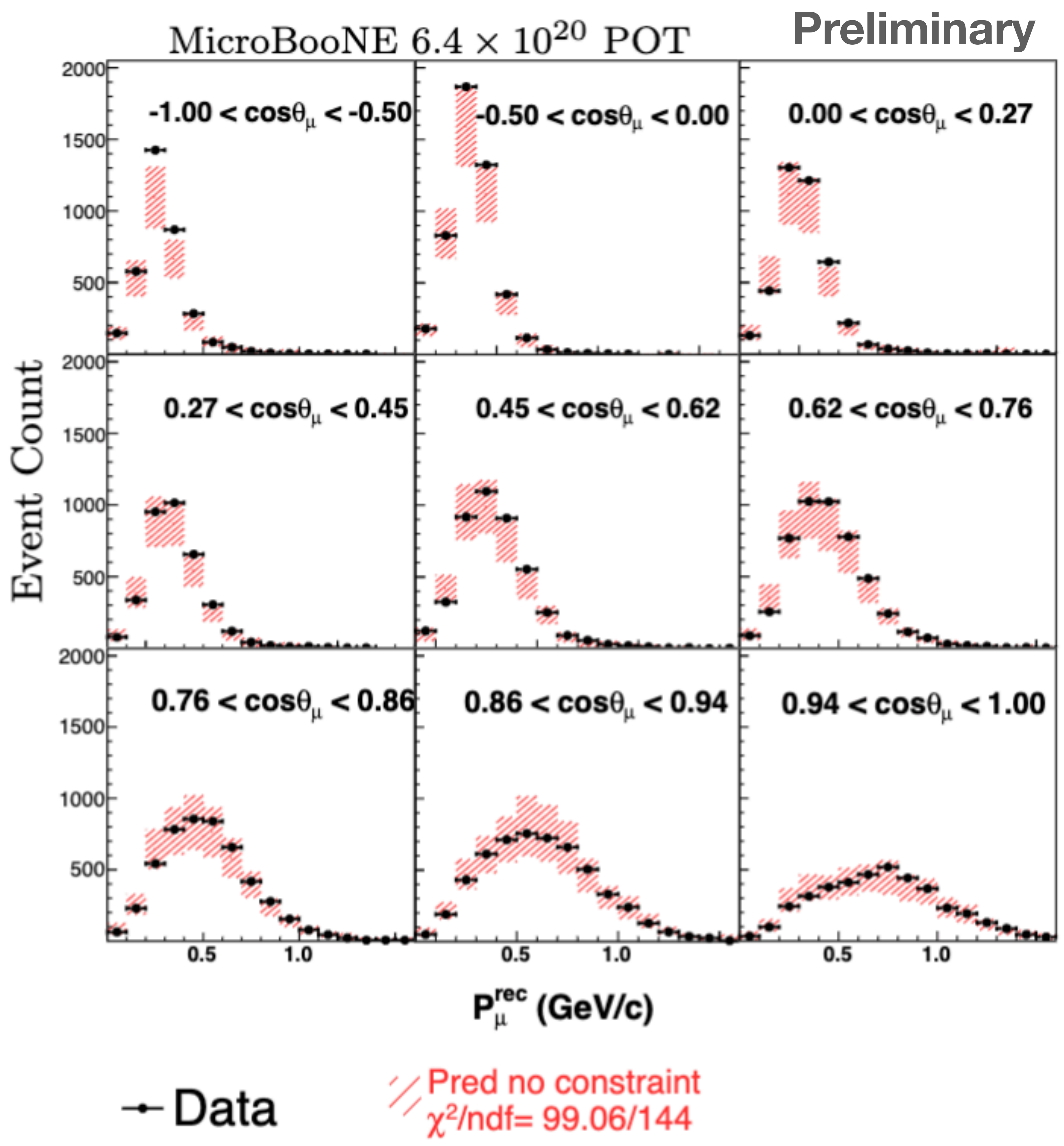


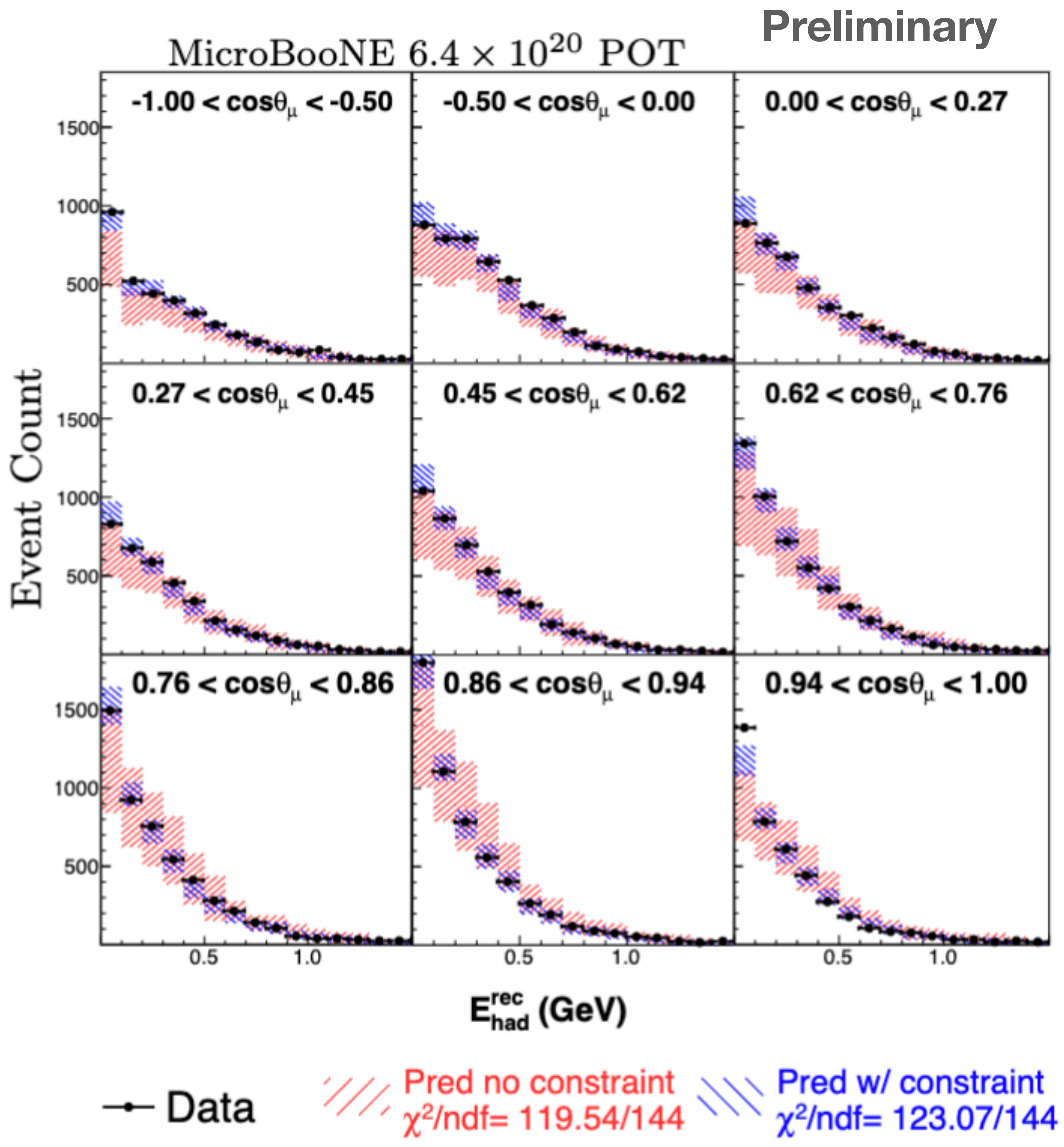
Phys. Rev. Lett. 128, 151801 (2022)

- See both - sizable drop in post-constraint error and good agreement with data
- Given data in $(E_\mu, \cos\theta_\mu)$ = within model possibilities in $(E_\mu, \cos\theta_\mu)$, we can say :
    - Data in E_{had} is also within same model possibilities => $D(E_\nu \to E_\nu^{Reco})$ is "validated" by data
- Sensitivity to the mapping ~ post-constraint error, (smaller error => more stringent test using data)
- Centerpiece of our approach, we use conditional constraint tests exhaustively to ensure robustness before unfolding
    - (NB : validation test, unfolding still uses unconstrained model)

Constraint tests in Multiple Dimensions

Triple-Differential Inclusive $\nu_\mu CC$ (submitted to PRL)





- Bin constraining $(P_\mu, \cos\theta_\mu)$ and constrained variables $(E_{had}, \cos\theta_\mu)$ in 2D
- Constrained prediction is compatible with data. Our $D(E_\nu \rightarrow E_\nu^{Reco})$ model still capable at this level
- Even more stringent $D(E_\nu \rightarrow E_\nu^{Reco})$ validation for inclusive channels

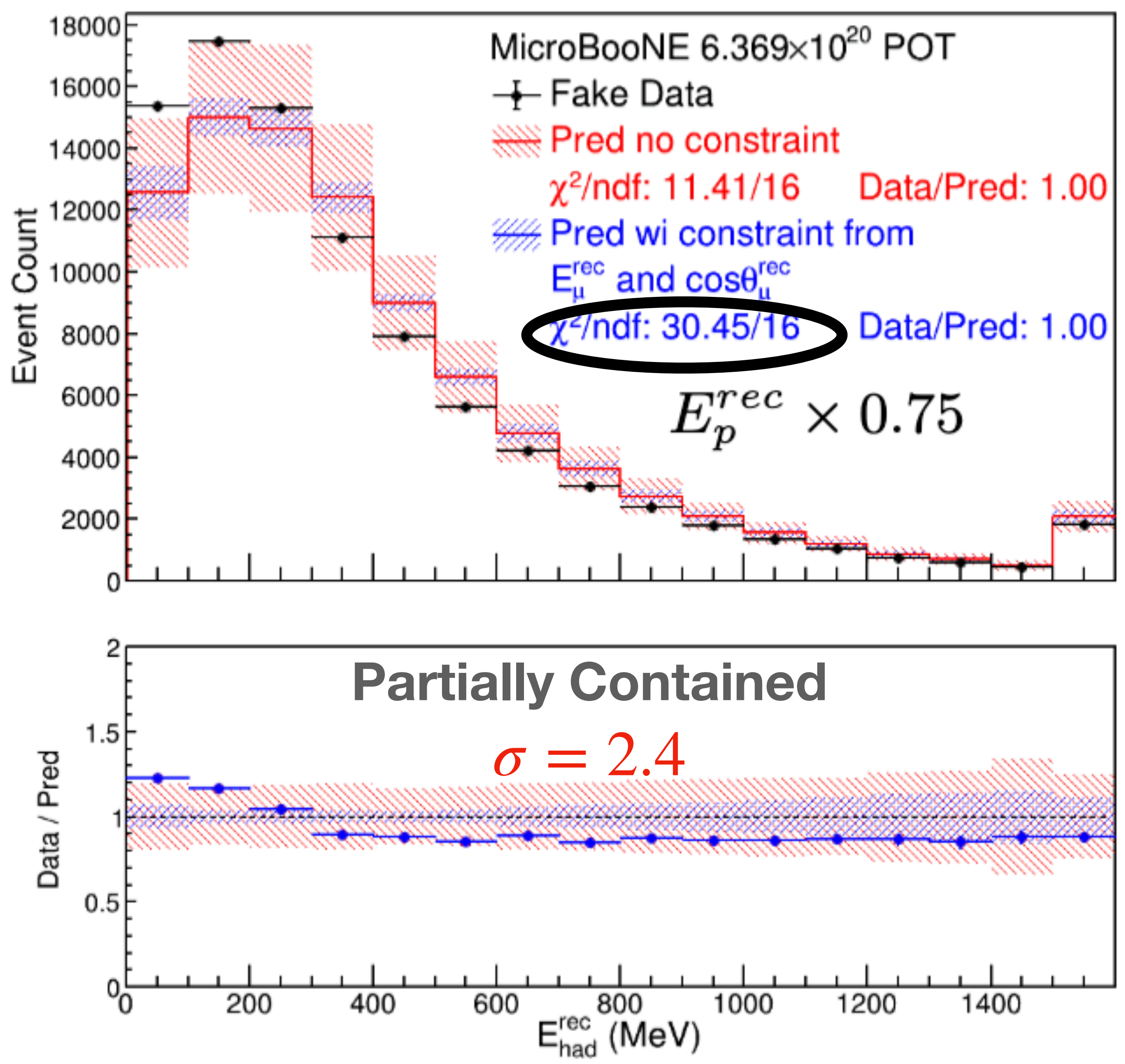
Fake Data Studies

Full Systematics

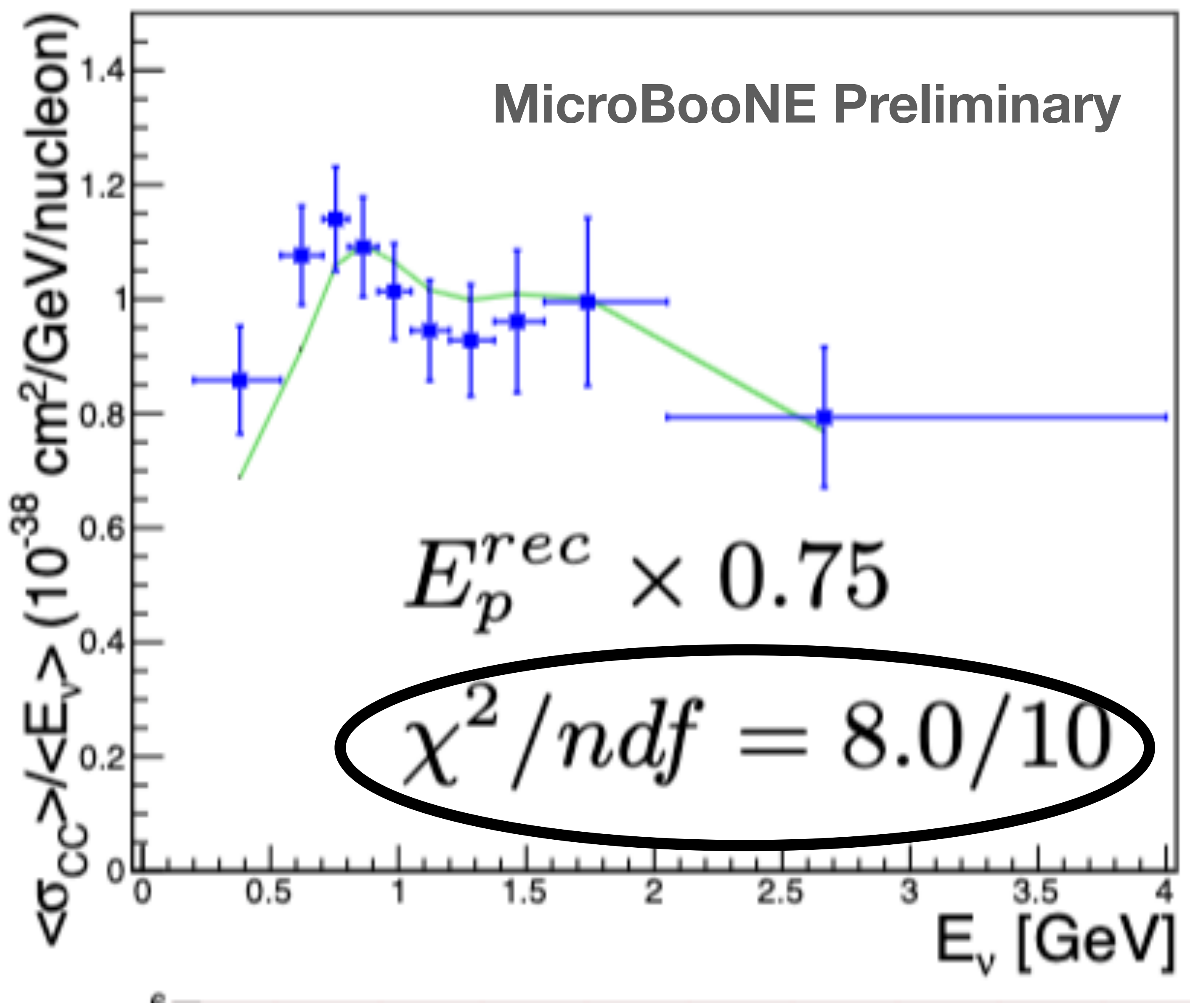
Proton energy scaling	Conditional Constraint Test Sensitivity (FC & PC) [Sigma Deviation]	Conditional Constraint Differential Test Sensitivity [Sigma Deviation]	E_ν Cross-section Bias [Sigma Deviation]
0.95	~0.0	~0.0	~0.0
0.85	~0.0	0.4	~0.0
0.75	1.3	1.5	0.5
0.65	4.2	5.7	1.9

More missing energy injected

MicroBooNE In-Progress

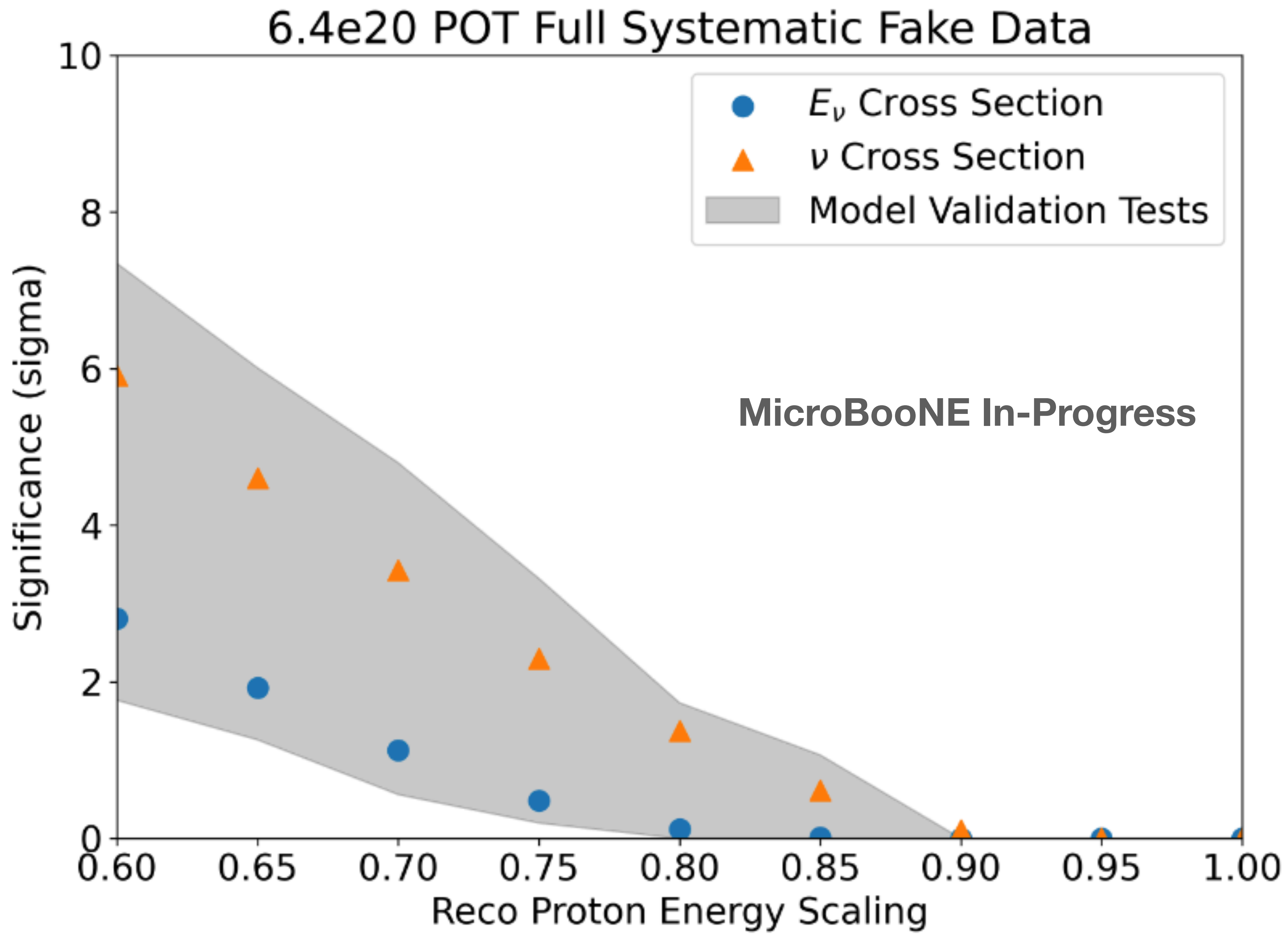


MicroBooNE In-Progress



- Fake data tests help us visualize sensitivity of tests for uncovering model bias
- Down-scale proton energy ("energy goes missing") and perform conditional constraint tests and compare to χ^2 after extracting the cross-section
- Shows model validation tests for our uncertainty budget is more sensitive to the bias than cross-section extraction — especially for E_ν

Sensitivity of Constraint Tests vs Unfolded Cross-Sections

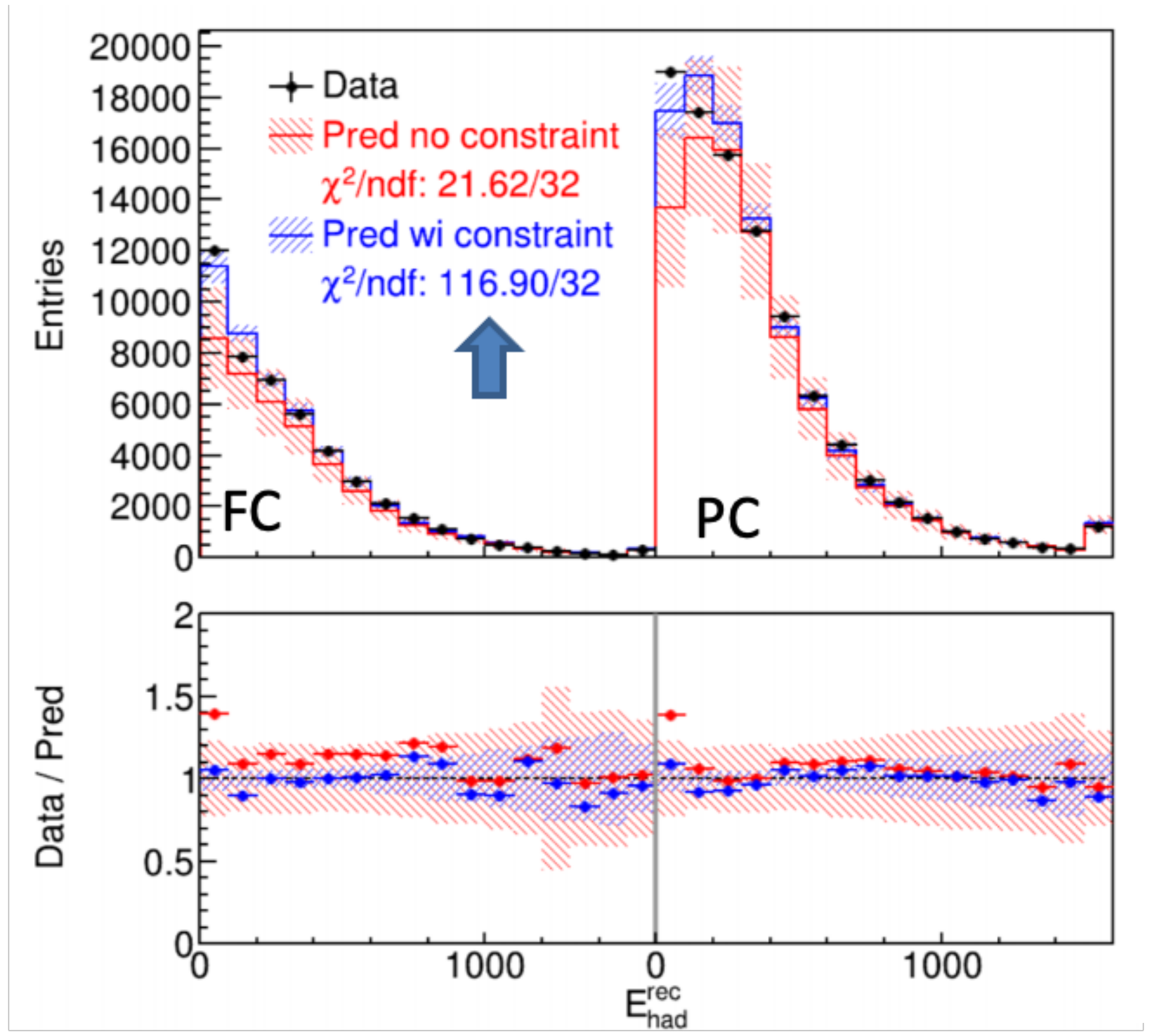


- Perform many model validation tests for eg individually in each channel (FC, PC, FC&PC, differential etc)
- Comparing to χ^2 of extracted cross-section using fake data with downscaled proton energy

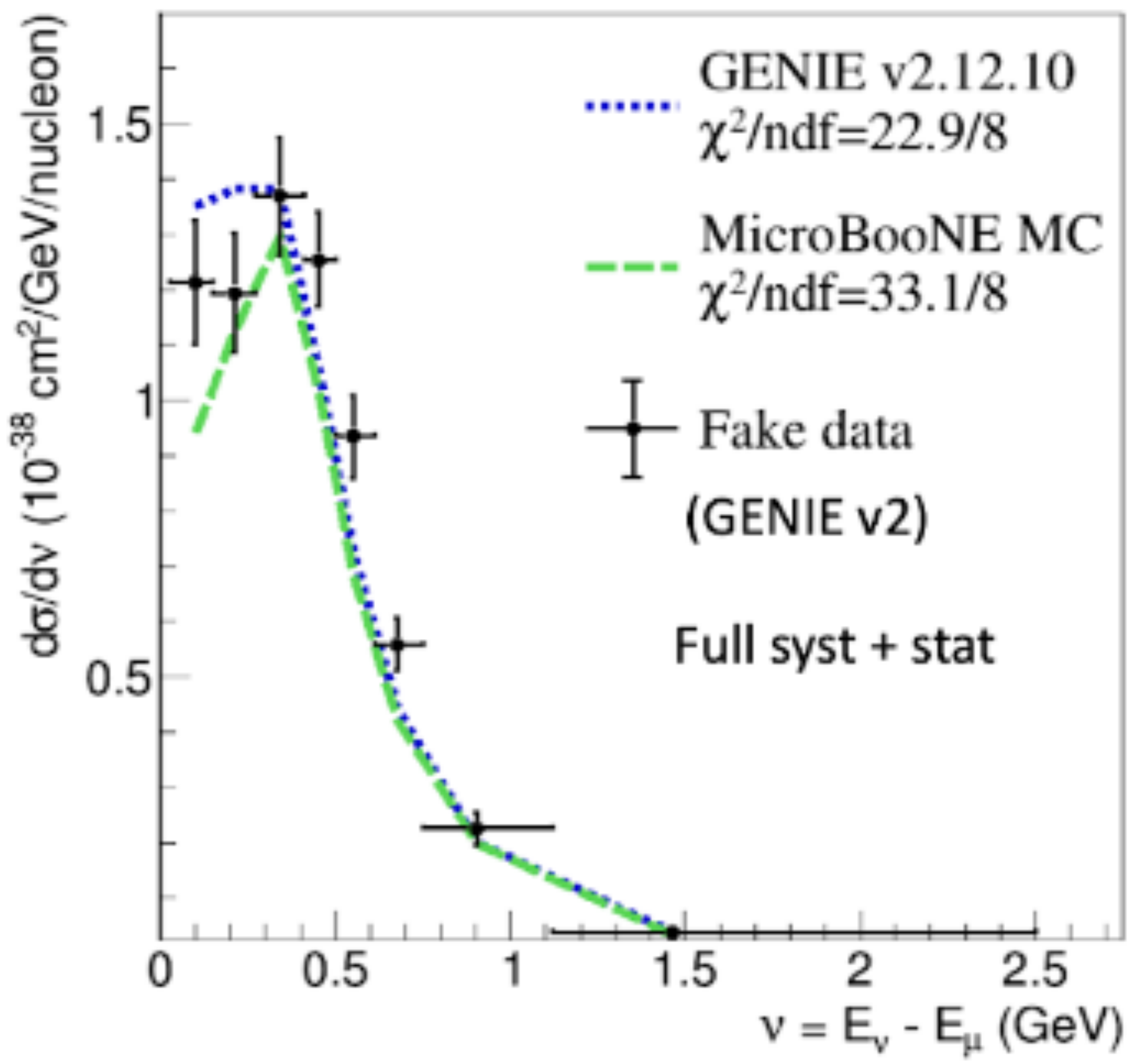
- We see conditional constraint tests are more sensitive than inherent bias in cross-section
  - We will be able to uncover this bias in real data before extracting the cross-section
  - Successful validation of real data => bias is small wrt uncertainties

FDS using different generators

MicroBooNE Preliminary



MicroBooNE Preliminary



Full Systematics

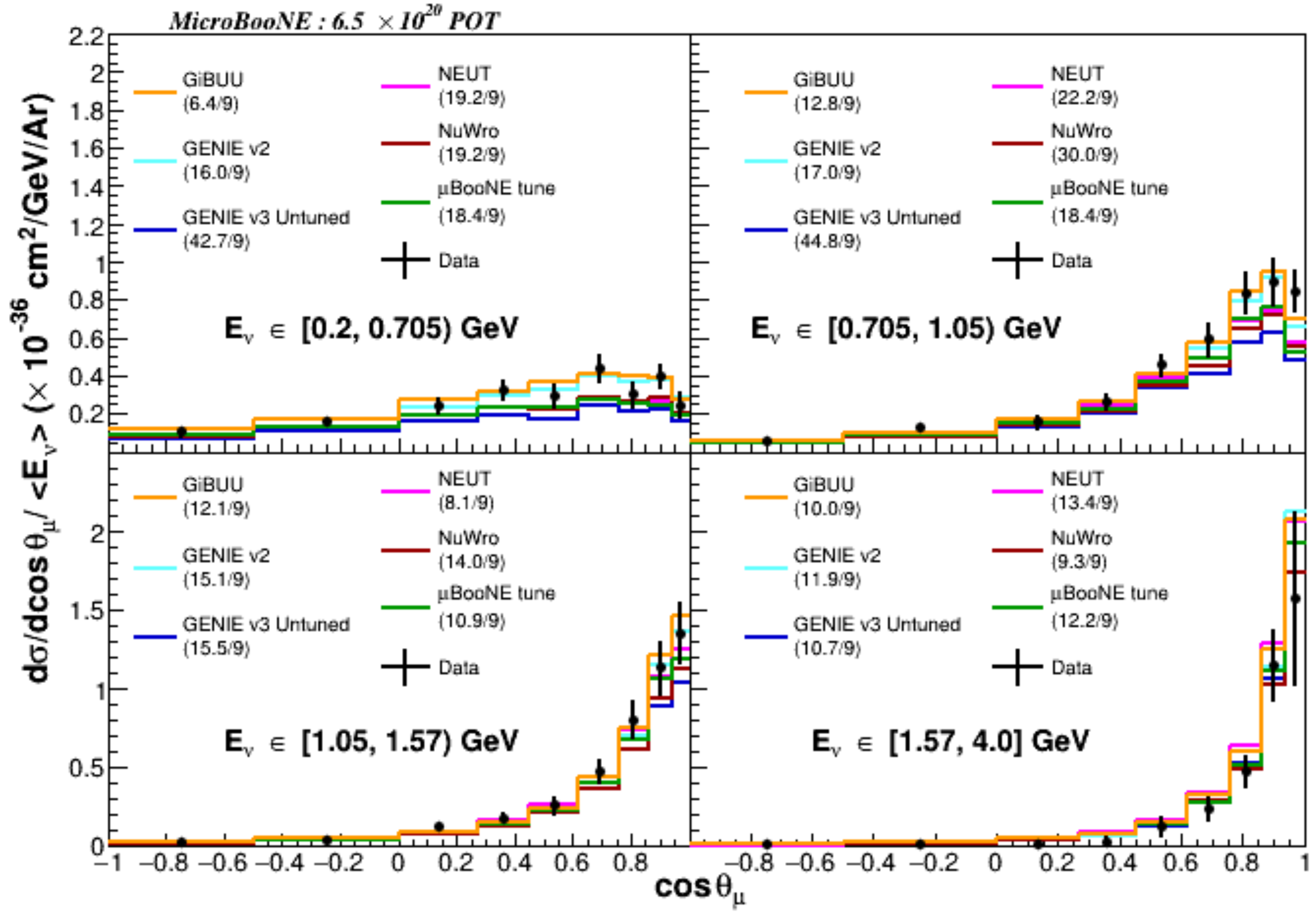
Generator Used for FDS	E_ν Cross-section Bias [Sigma Deviation]	ν Cross-section Bias [Sigma Deviation]	Model Validation [Sigma]
GENIEv2 (7.24E20)	0.2	2.9	6.8
GENIEv3 (5.33E19)	~0.0	~0.0	~0.0
NuWro (6.11E20)	0.01	0.1	0.01

- Check conditional constraint tests against different generators for eg, GENIEv2, NuWro
- Behavior as expected from before — model validation using conditional constraint is (more) sensitive
- Successful validation of real data => bias is small wrt uncertainties

Summary and Next Steps

- We propose a new approach to validate model esp $D(E_\nu \rightarrow T_{reco})$ in order to extract robust cross-sections to compare to various predictions
- Allows unfolding to indirect/direct observables including neutrino energy, energy transfer etc
- Enables easy comparisons of extracted cross-section across different models

MicroBooNE Preliminary



- So far focused on ν_μ-CC inclusive results
- Future results will focus on 0pNp, π^0 final states and other exclusive measurements as well => more model validation needed as well

Thank you!

201 collaborators, 36 institutions, 5 countries

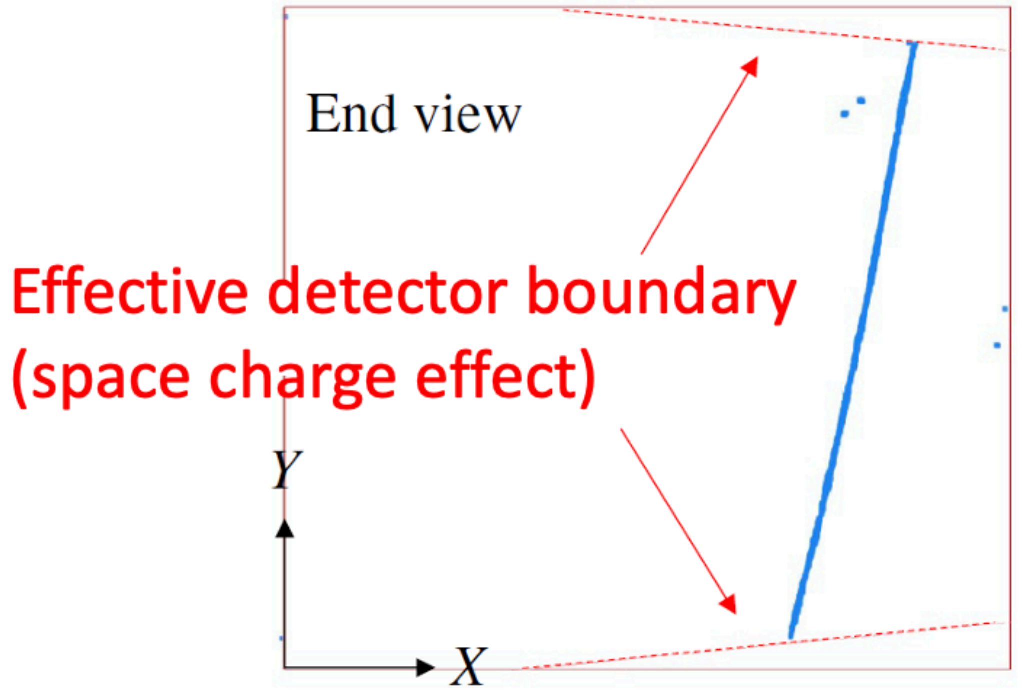
MicroBooNE collaboration @ 2022

Backup

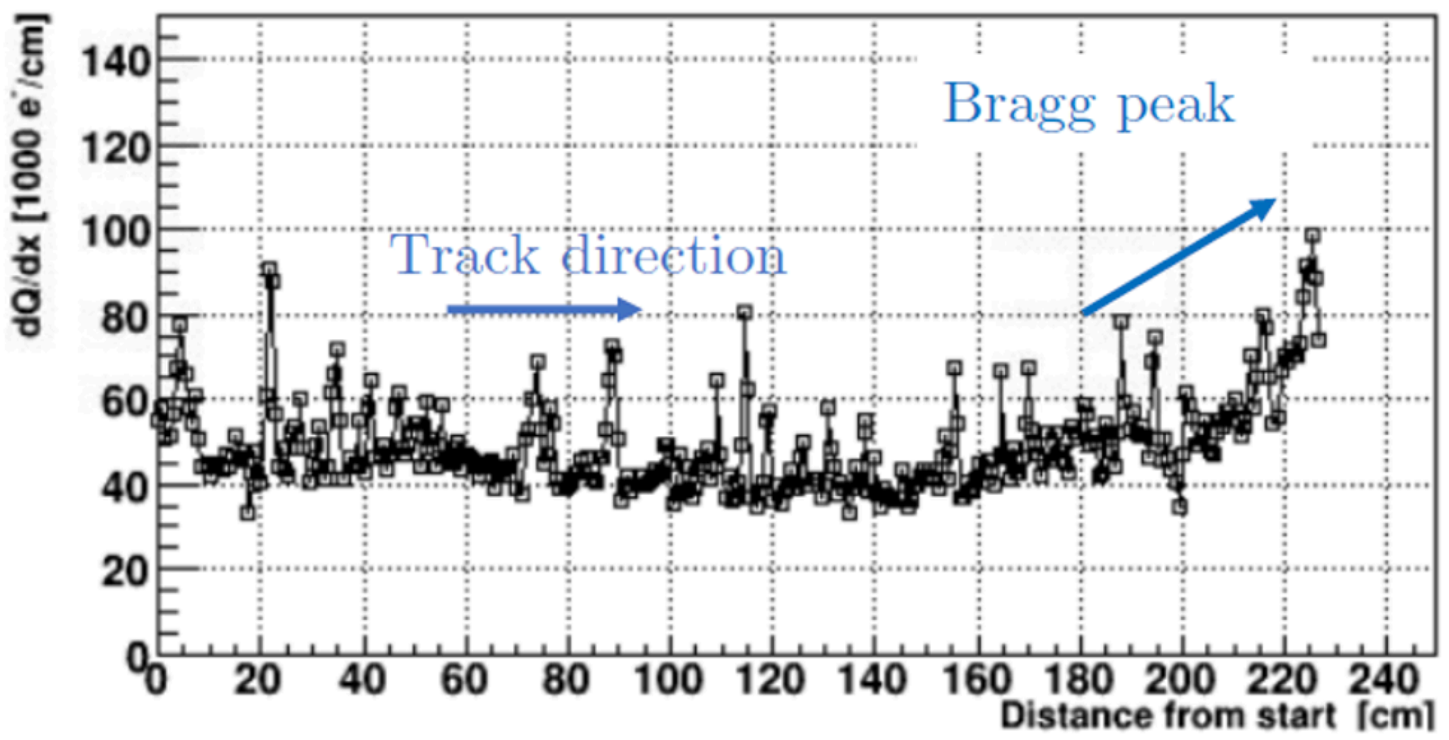
Neutrino Identification

Phys. Rev. Applied 15, 064071 (2021)

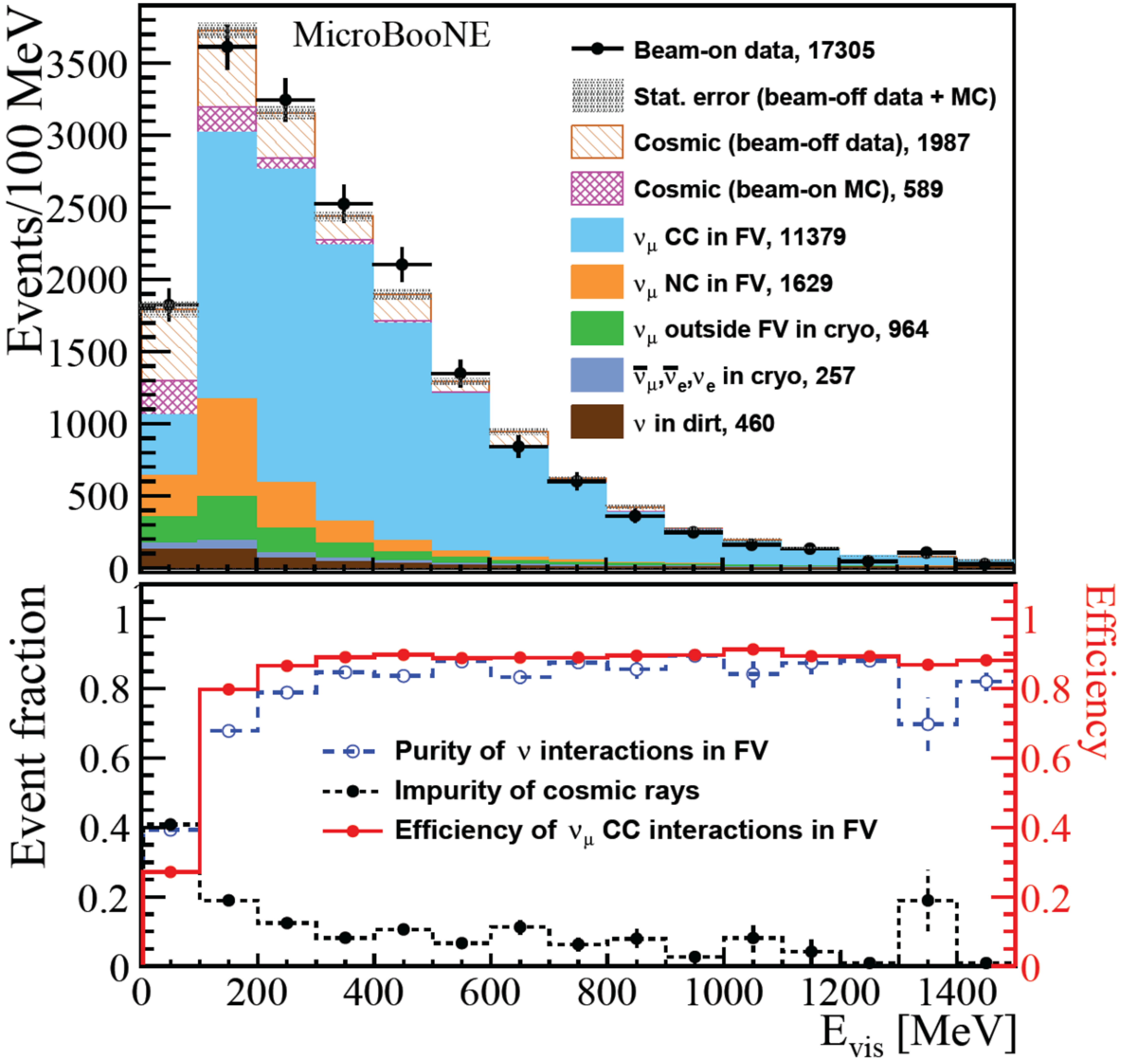
Through-going muon (TGM)



Stopping muon (STM)

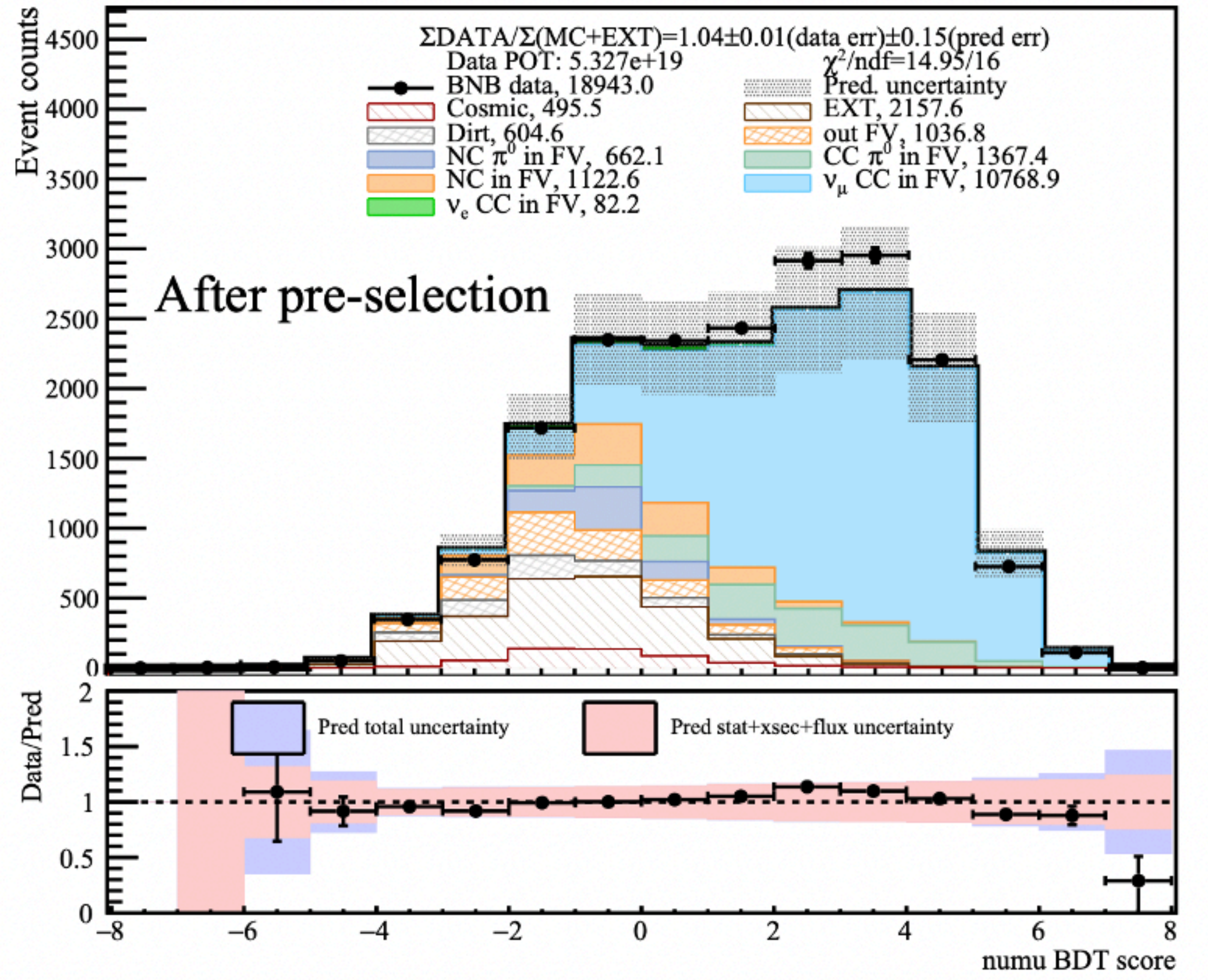


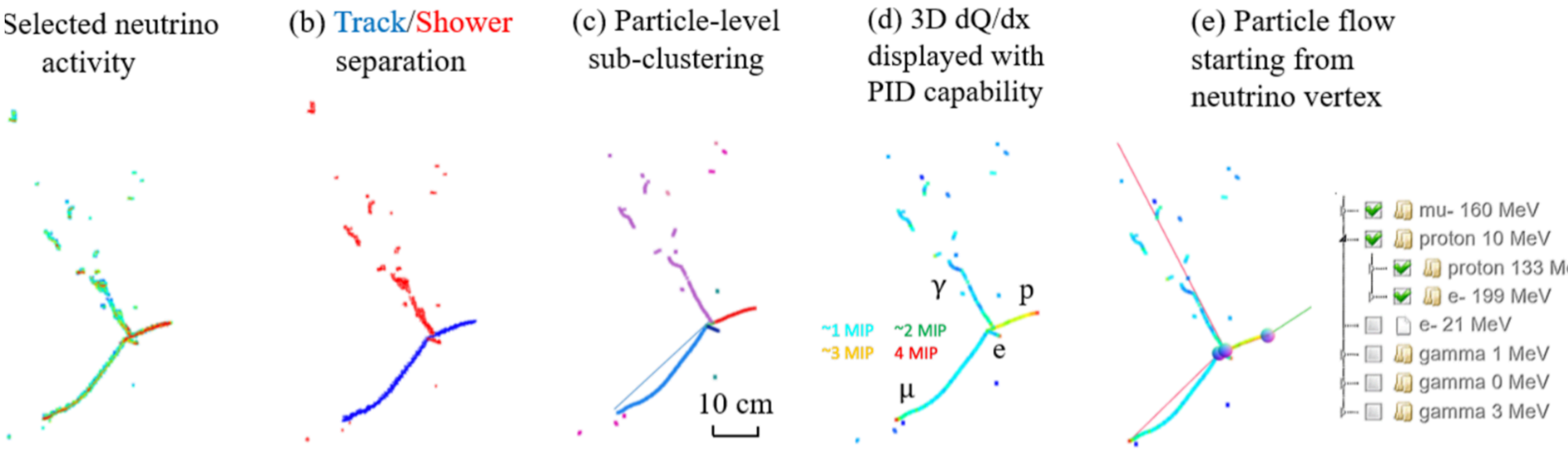
	Neutrino:Cosmic-ray	
Charge-light matching	1 : 6.4	Improved by factor of >6
TGM rejection	1 : 0.91	
STM rejection	1 : 0.36	Improved by factor of ~3
Additional Cuts	1 : 0.20	



- Topology agnostic reconstruction => generic neutrino selection
- Use effective detector boundaries and directionality from trajectory and dQ/dx fitting to reject cosmic muons
  - Through-going and stopping
- 80% efficiency for selection, similar purities
  - Downstream selections/PID for various topologies
- Reject 99.999% of cosmic activity!

ν_μ-CC Selection

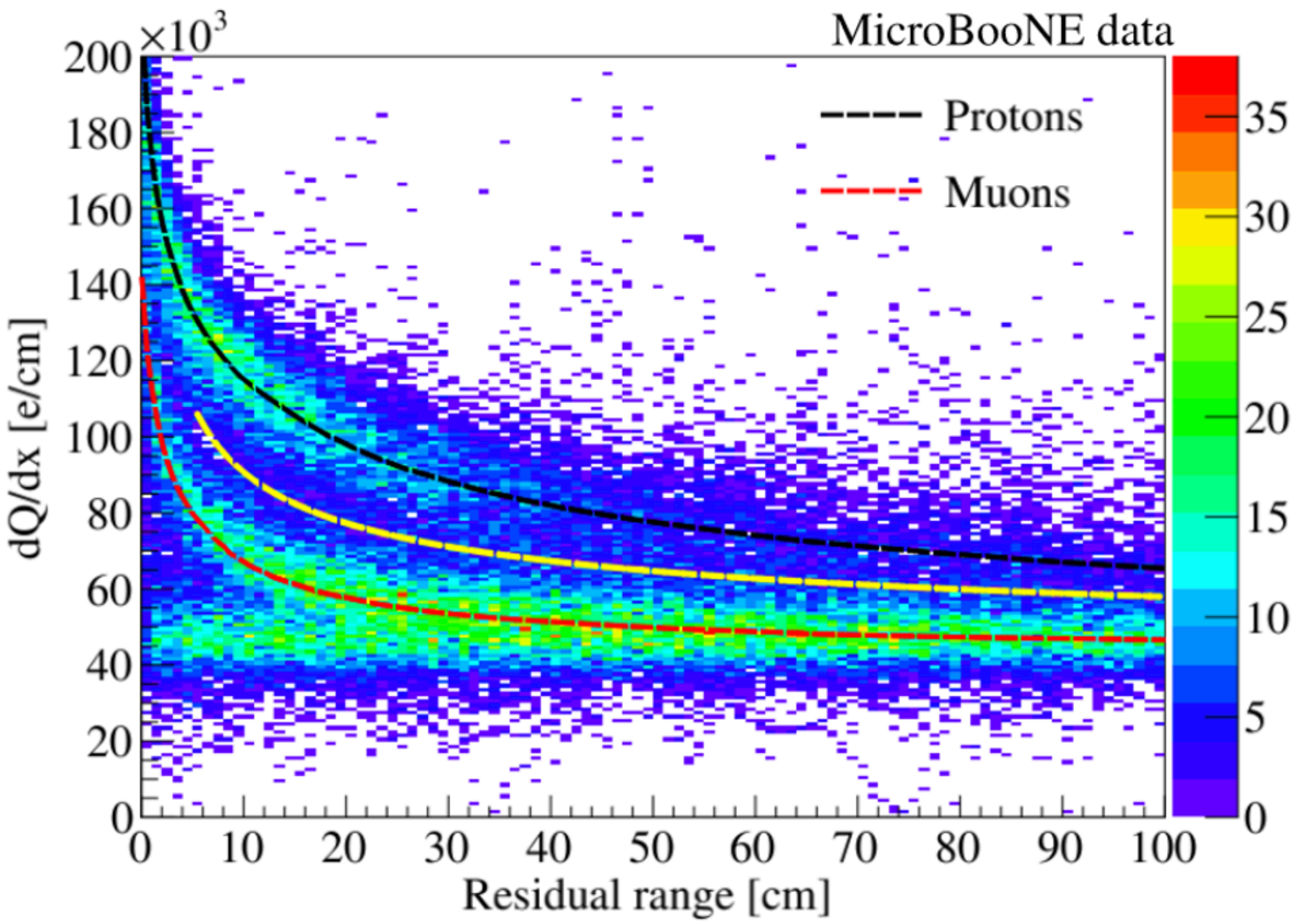




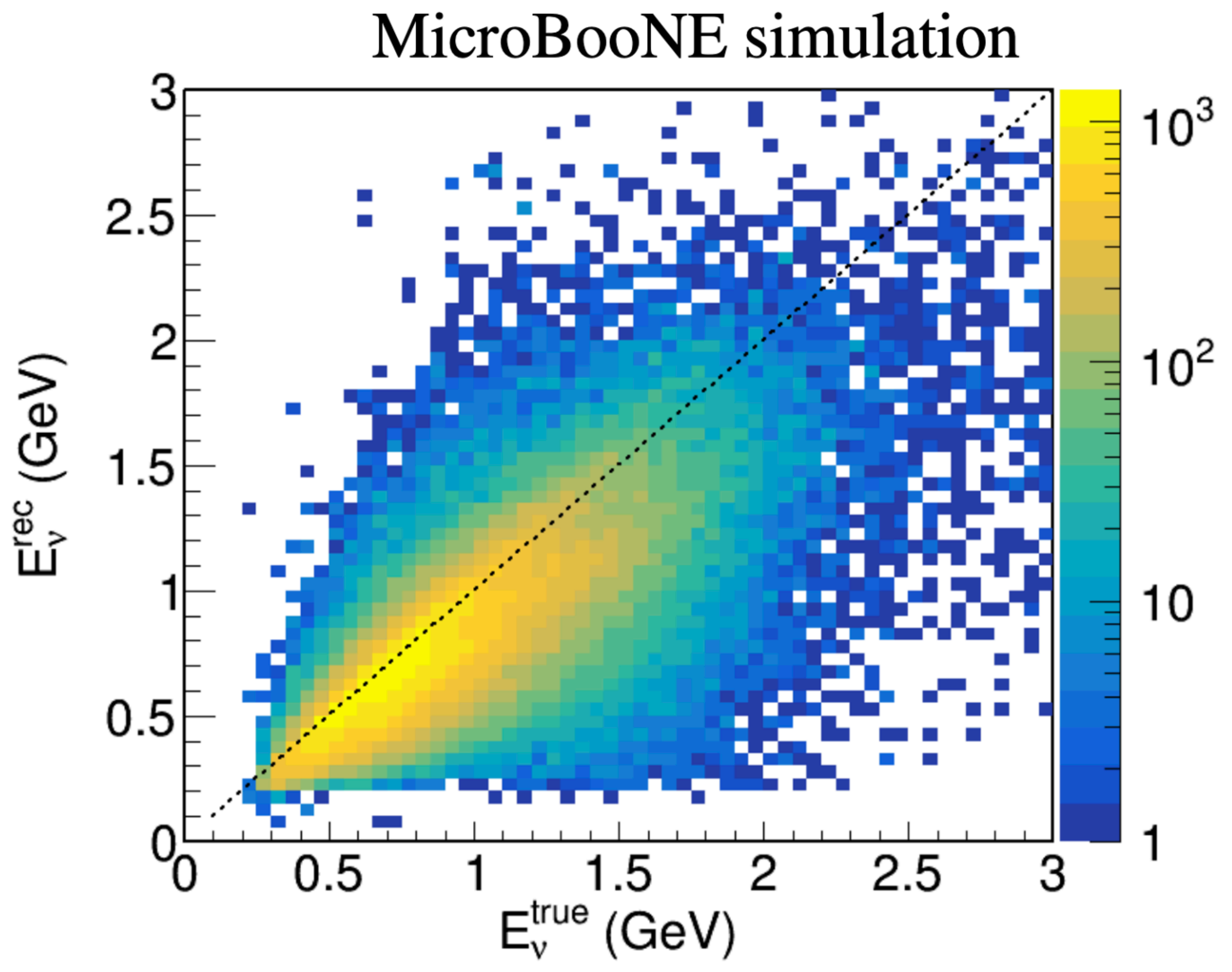
- XGBoost-based BDT based on human-engineered inputs
- Achieves >90% purity

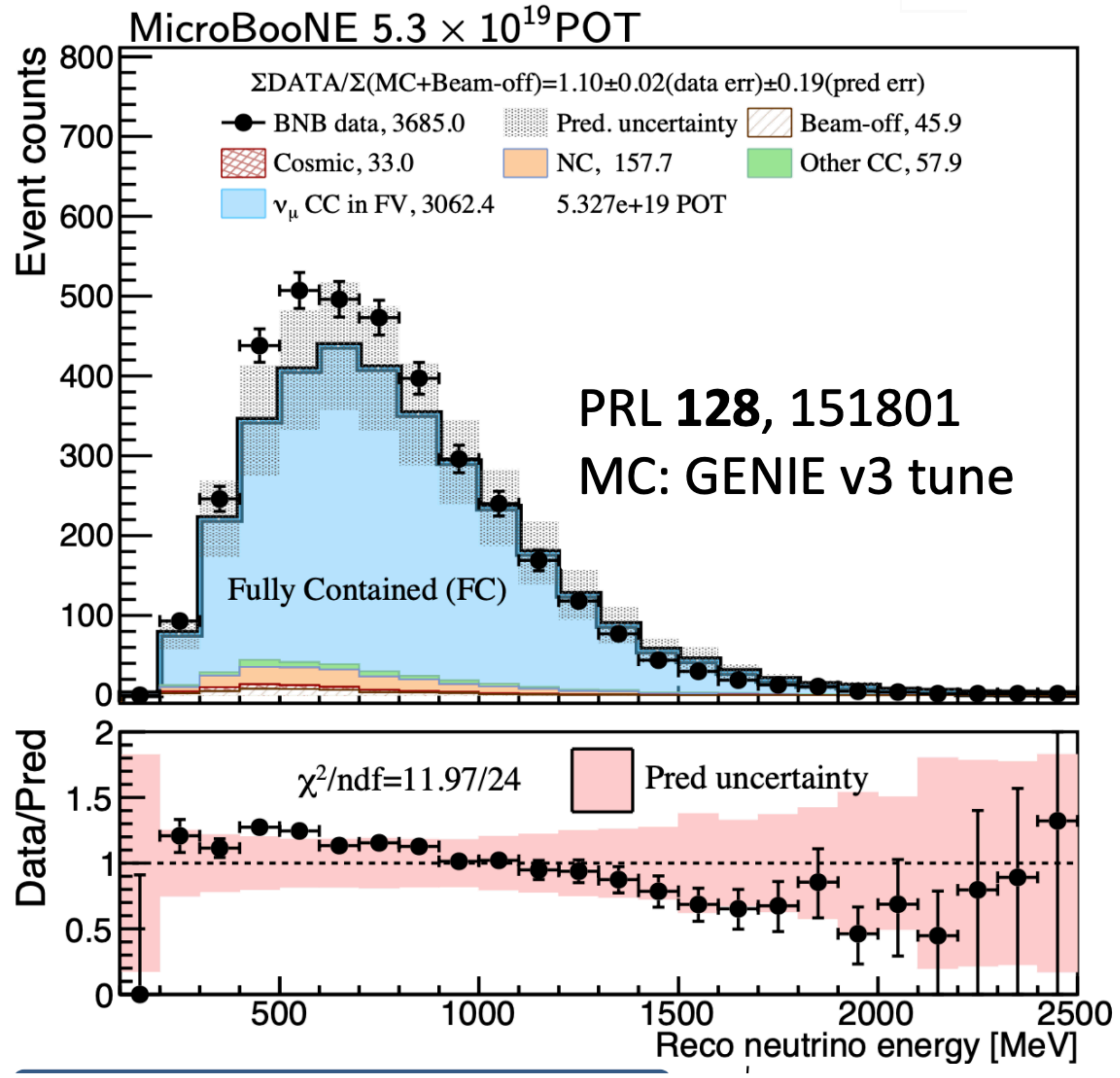
	Efficiency	Purity	Cosmic-μ rejection
Trigger	1	5e-5	1
Cosmic-ray rejection	80%	65%	7e-6
ν_μCC with pattern recognition **(Fully & Partially Contained)**	**68%**	**92%**	7e-7

ν_μ-CC Energy Reconstruction



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



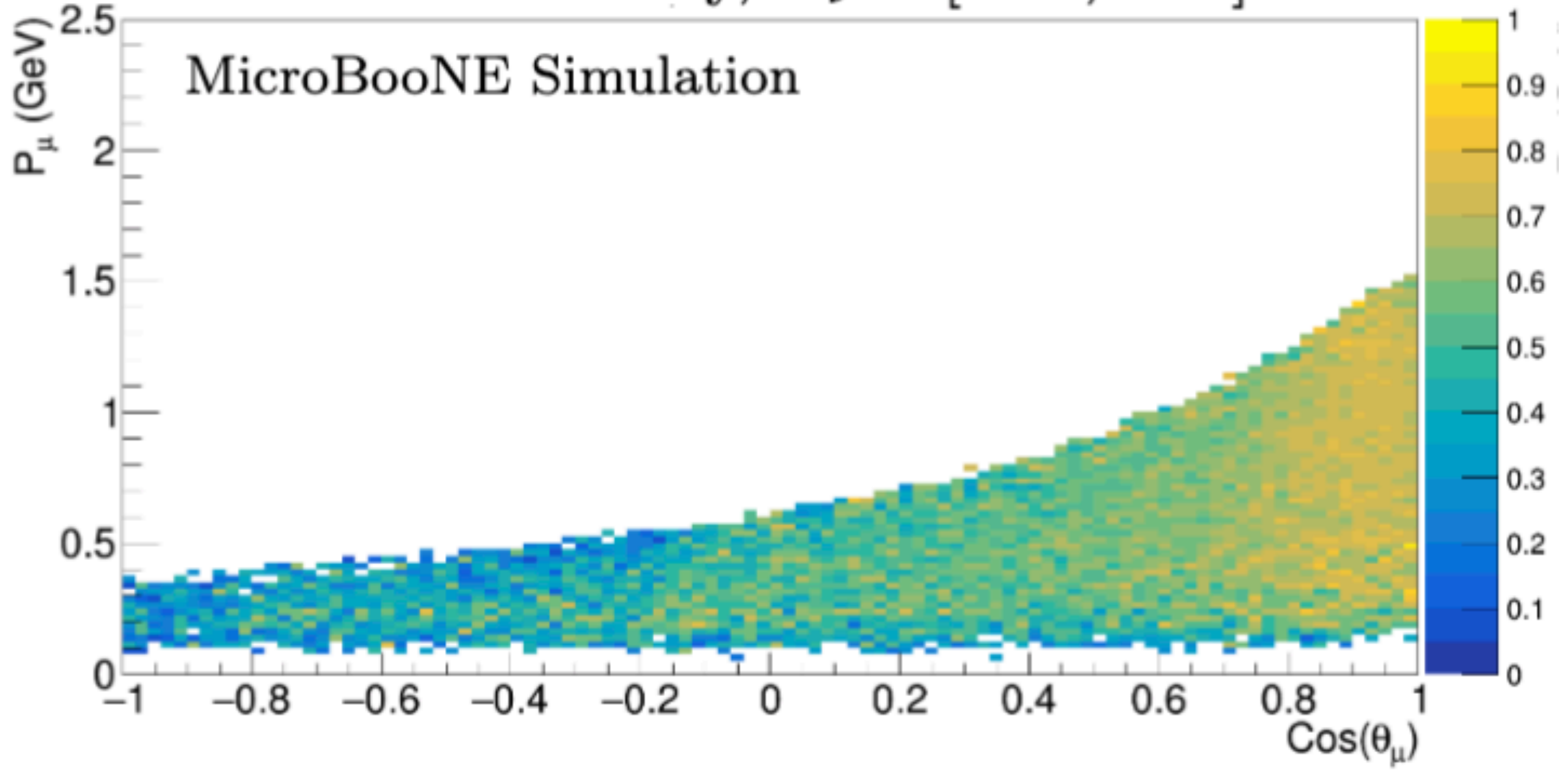


- Range based estimator for μ
- For rest : dQ/dx -> dE/dx energy scale calibration using stopping muon/ proton samples
- Additional scaling for EM showers to match π^0 mass peak
- Achieve 15-20% resolution for ν_μ-CCs w/ ~10% bias

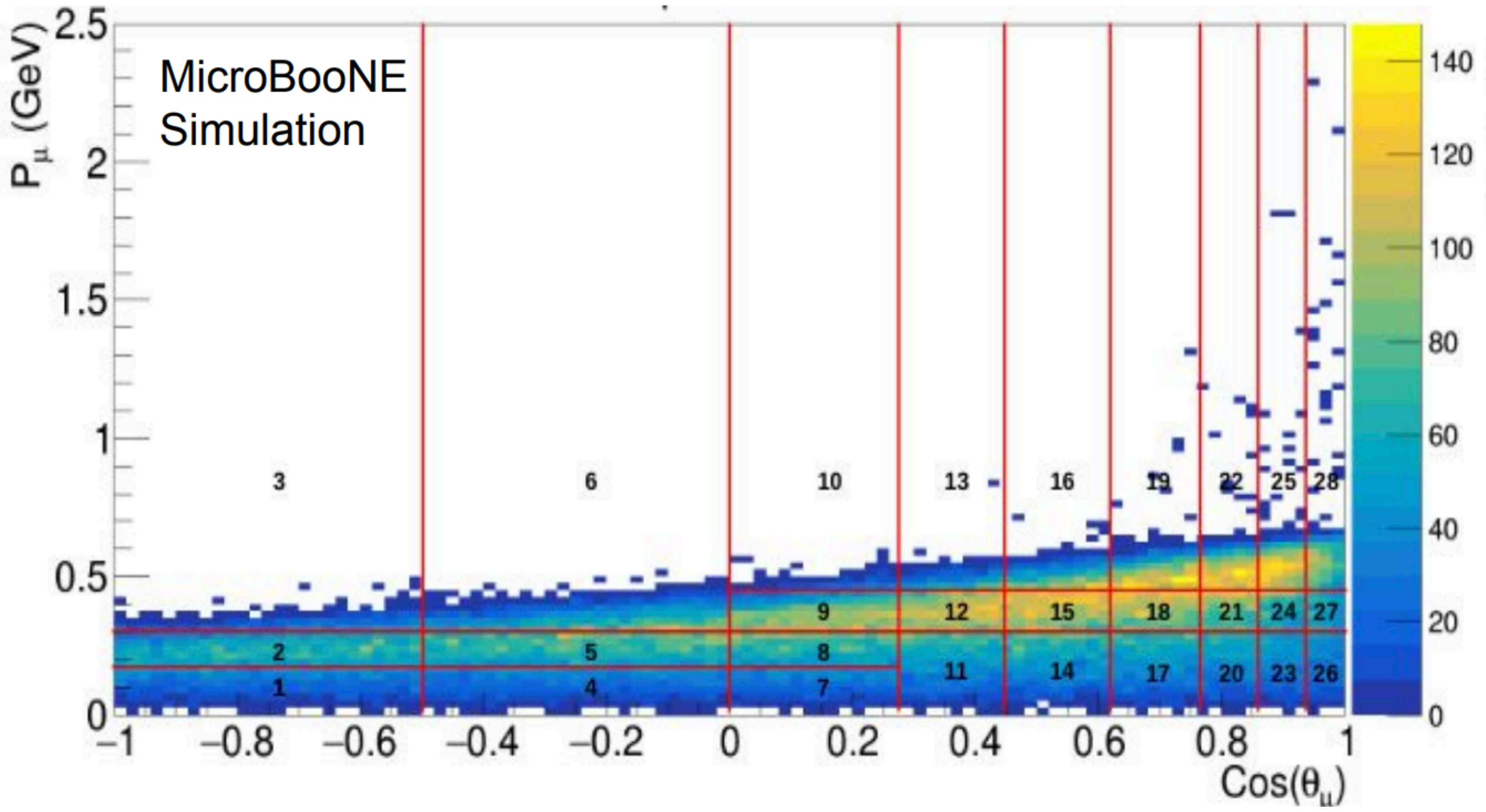
Analysis Binning

- 138 analysis bins in total in $(E_\nu, P_\mu, \cos\theta_\mu)$
  - ~10% resolution on P_μ
  - $\delta\theta_\mu$ ~ 5° for forward angles
- Binning chosen so we have good stats everywhere
  - Efficiency is pretty good across phase-space

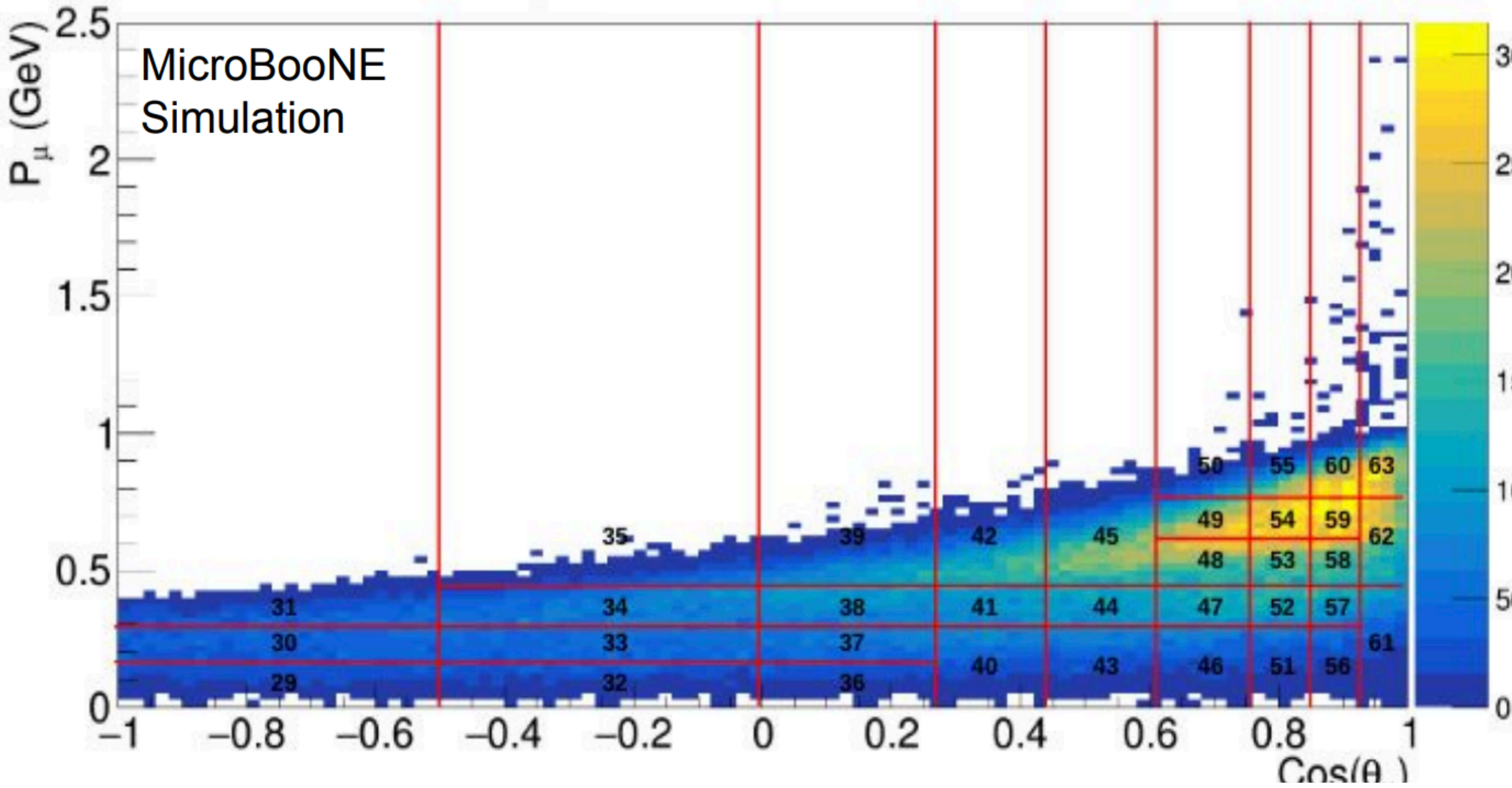
Selection Efficiency, $E_\nu \in [1.05, 1.57]$ GeV



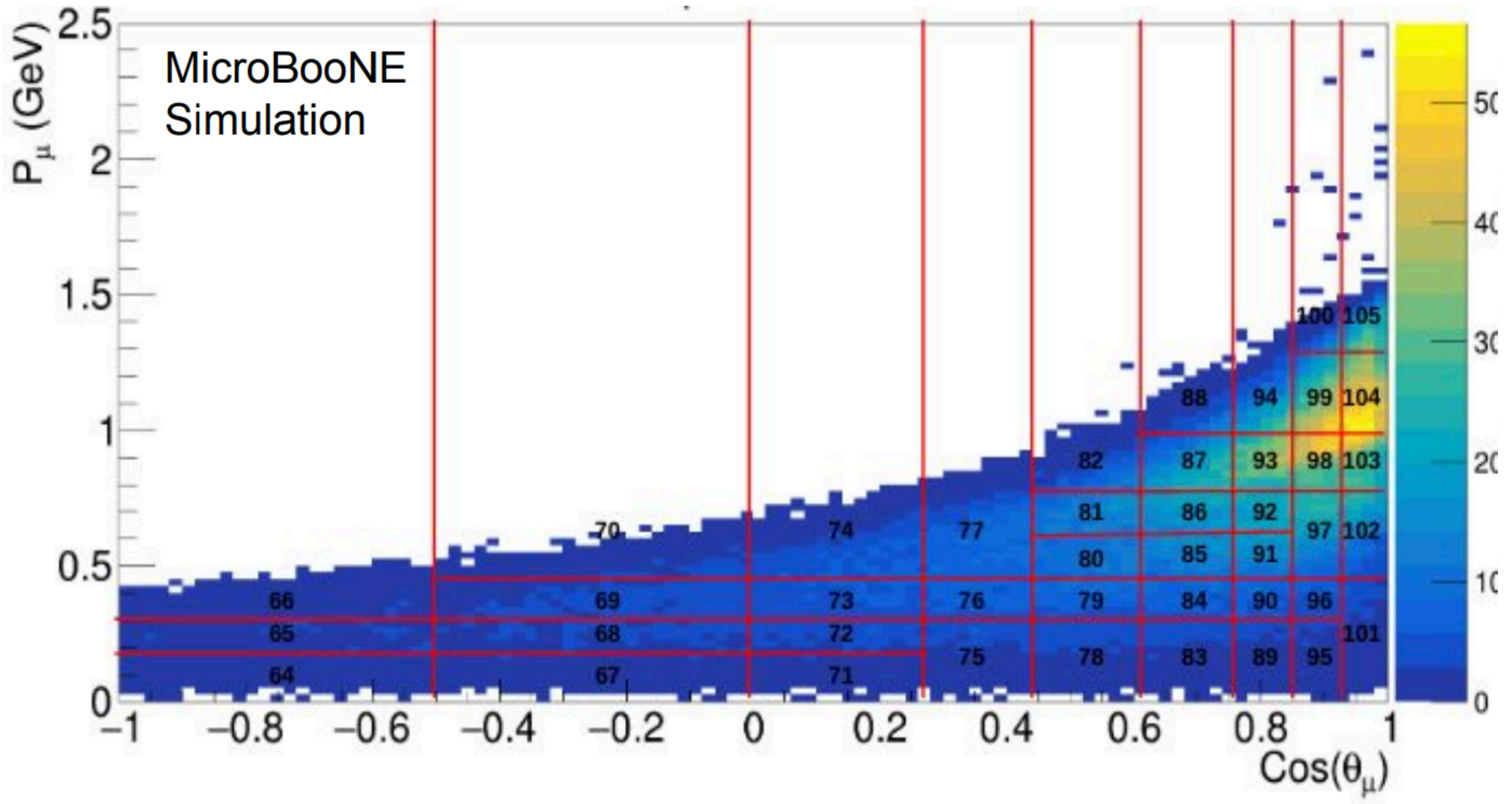
Binning for E_ν in [0.2,0.705] GeV



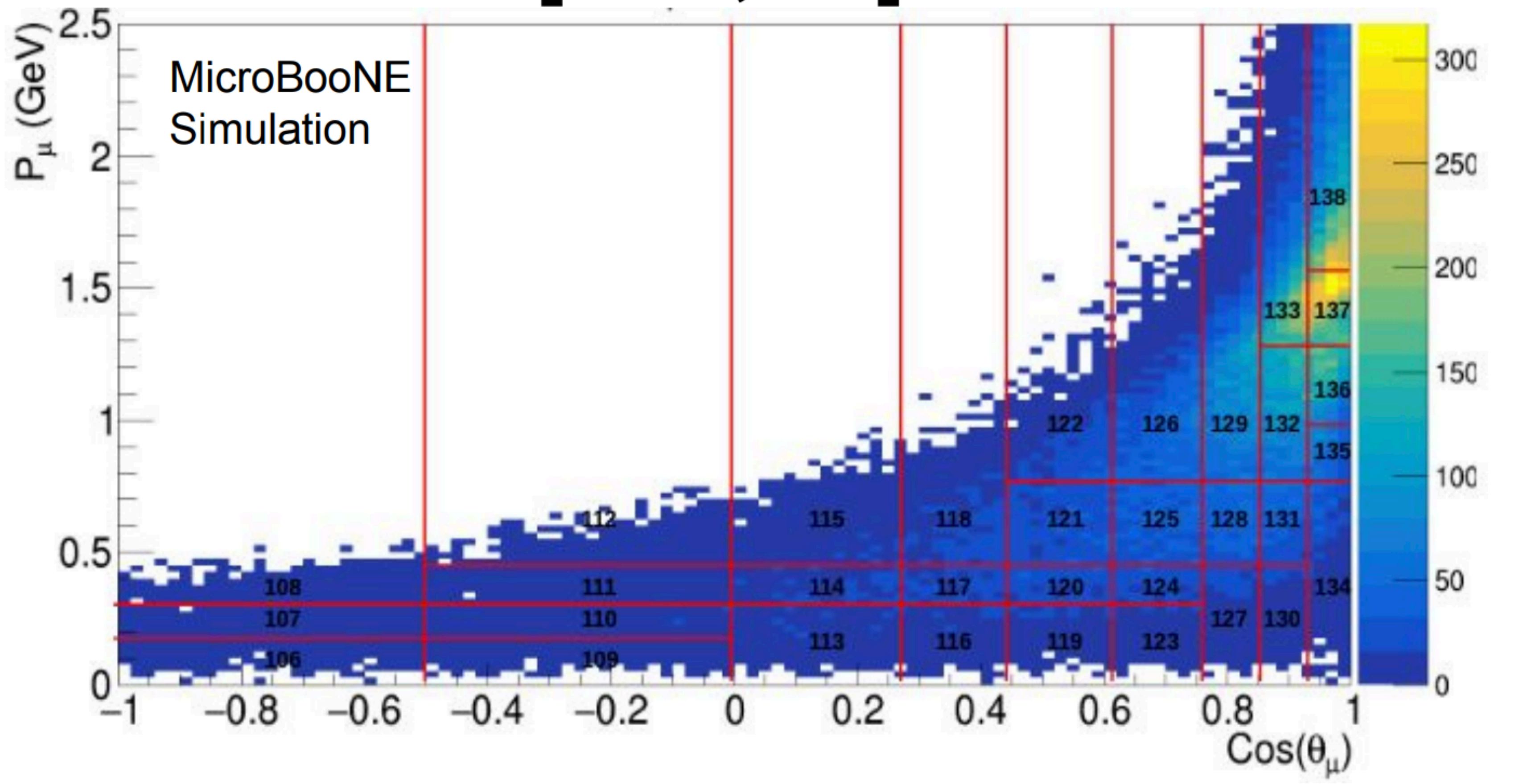
Binning for E_ν in [0.705,1.05] GeV



Binning for E_ν in [1.05,1.57] GeV



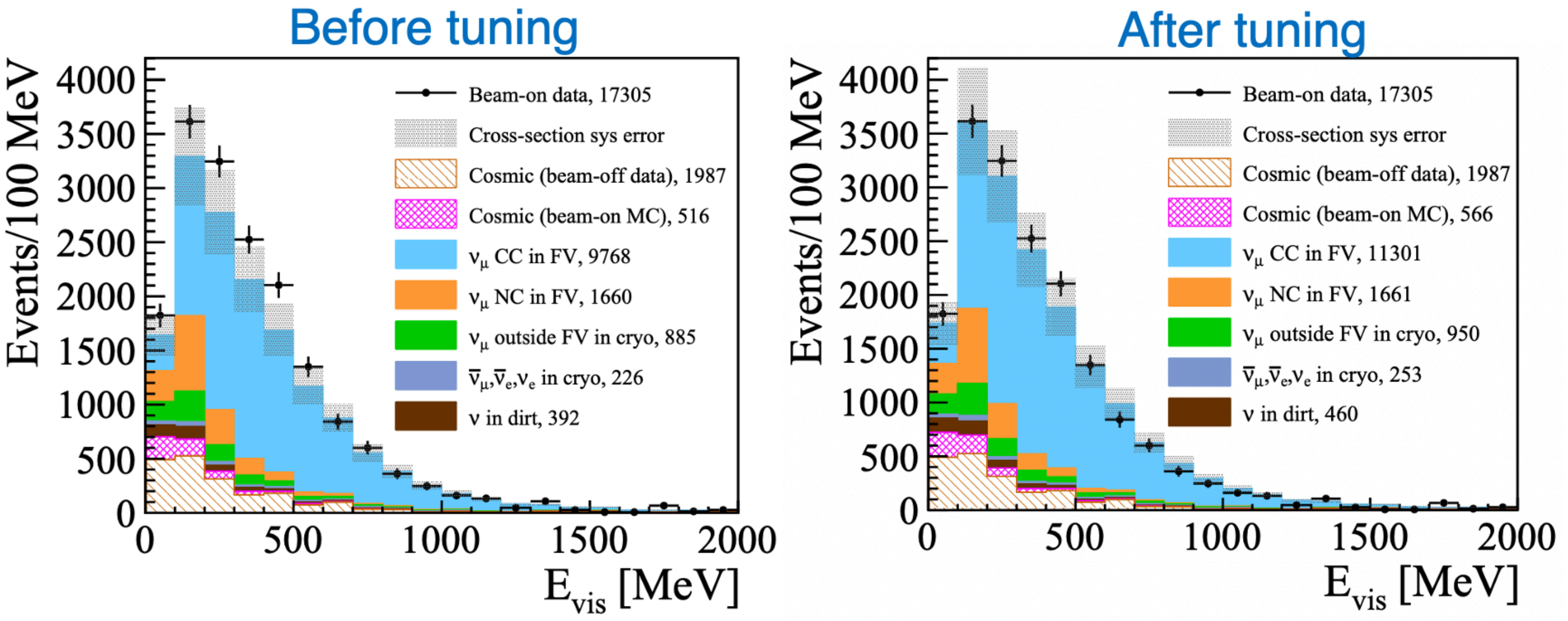
Binning for E_ν in [1.57,4.0] GeV

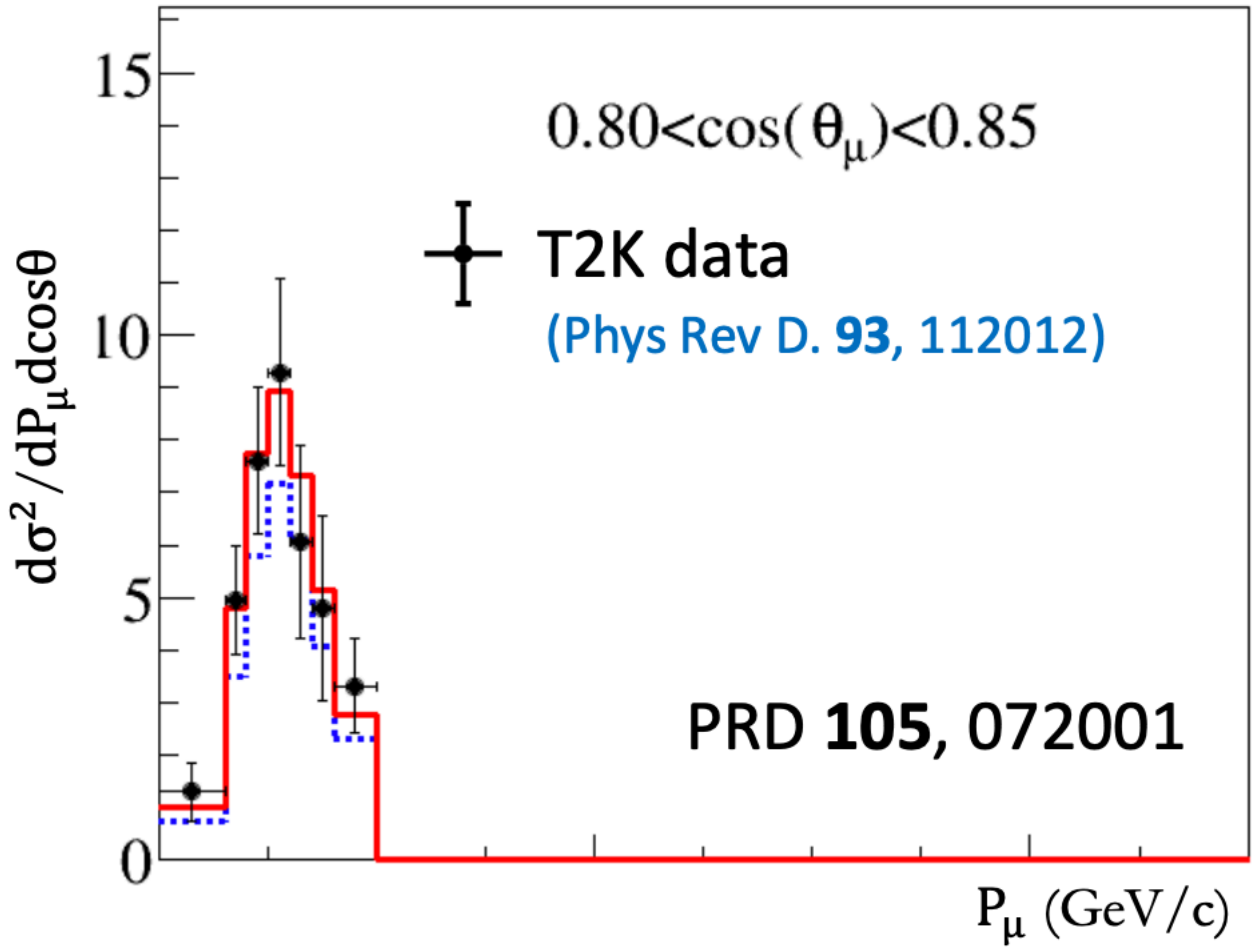


Neutrino Interaction Model - "μB tune"

GENIE 3.0.6

Generic neutrino preselection



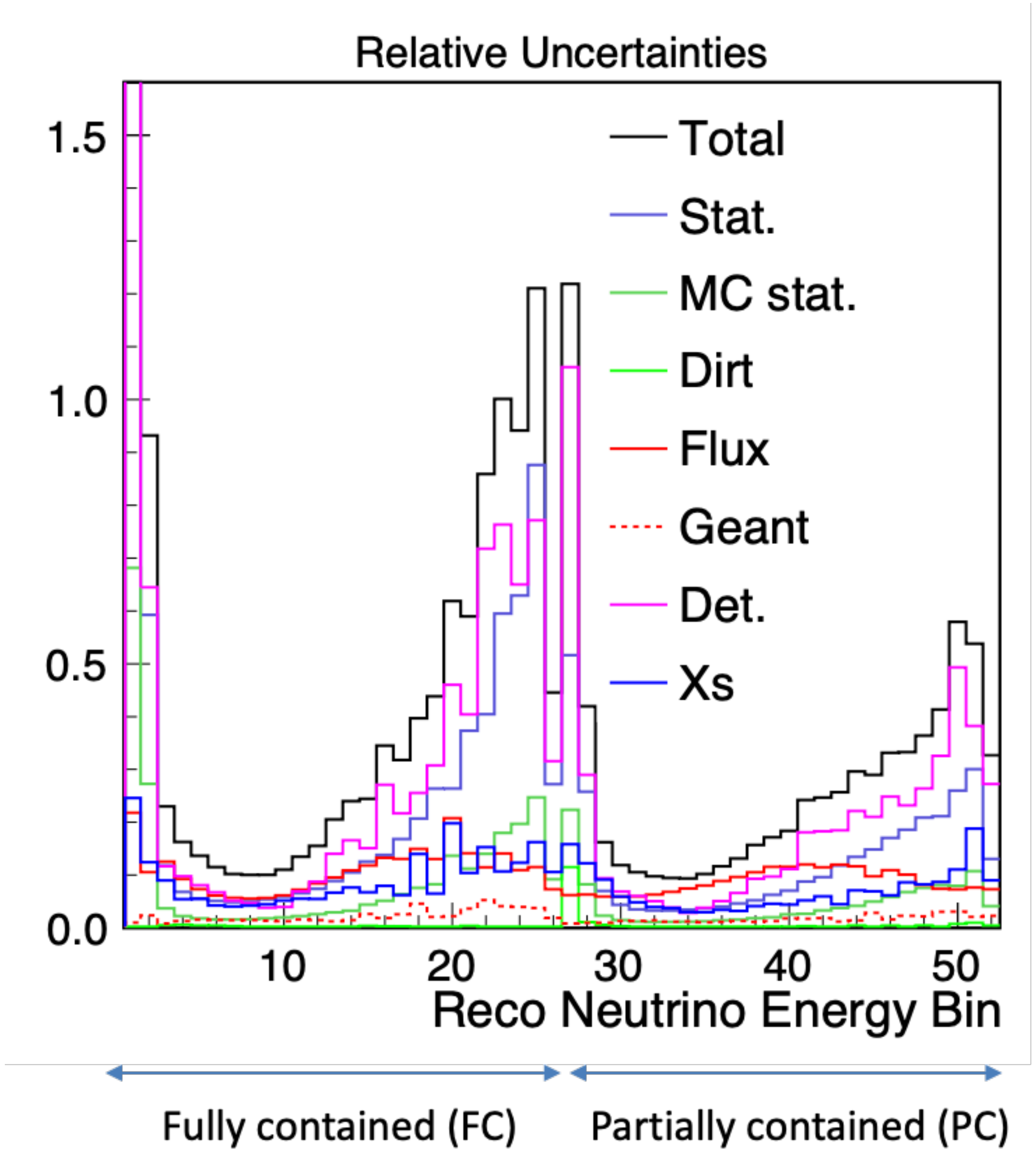


	MaCCQE (GeV)	CC2p2h Norm.	CCQE RPA Strength	CC2p2h Shape
Untuned	0.961242	1	100%	0
Tuned	1.10 ± 0.07	1.66 ± 0.19	(85 ± 20) %	$1^{+0}_{-0.74}$

- Updated to GENIEv3
  - "μB-tune" tuned to T2K CC0π data (no actual μB data used)
  - Consider mainly CCQE (RPA, Ma) and 2p2h parameters for CV + uncertainties
  - Additional GENIE + custom knobs (> 50) considered for uncertainties

Systematic Uncertainties

- BNB Flux
  - MiniBooNE (hadron production, beam transport, POT counting)
- Neutrino Cross-Sections
  - Apart from GENIE v3.0.6 knobs for model set G18_10a_02_11a
  - Custom knobs from "μB-tune" to T2K CC0π data : MaCCQE, RPA, 2p2h norm & shape
  - Others (Coulomb corrections, Δ-contributions to 2p2h etc)
- Detector systematics
  - TPC, Light yield, SCE, Recombination
- Hadron-Argon re-interactions :
  - Geant4Reweight-based
- MC Statistical
- Dirt Systematics
  - Material outside cryostat



Traditional approach to Cross-Section Analyses

$$\frac{d^2\sigma}{dT_l cos\theta} = \frac{\sum_j U_{ij}(d_j - b_j)}{\Phi \cdot T \cdot \varepsilon_i \cdot (\Delta T_l, \Delta cos\theta)_i}$$

- Typically unfolded to unknown true flux, not nominal flux-averaged
- In principle, robust to true neutrino energy -> reco observable map, $D(E_\nu \to T_{reco})$ especially when looking at direct observables
- However, still need reference flux to compare to model predictions
    - Can be quite complicated in practice (L. Koch, S. Dolan - PhysRevD.102.113012)
- Model comparisons usually done using nominal flux as a result since its simple

Our approach to Cross-Section Analyses

$$M_i = R_{ij} S_j + B_i$$

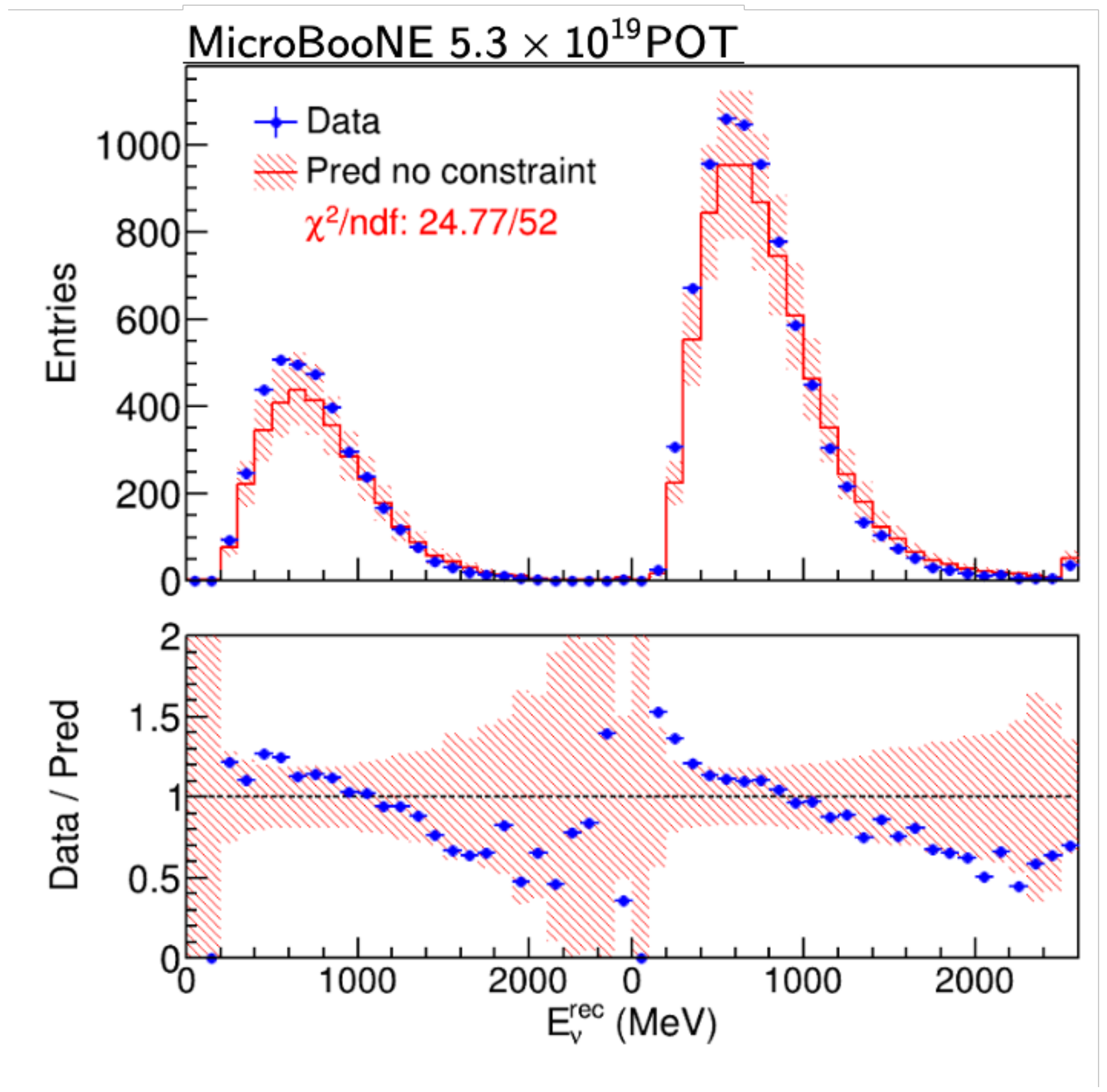
- M_i : Measurement in reco bin i of some observable
- R_{ij} : mapping between true bin j and reco bin i constructed using nominal flux
- B_i : Background in reco bin i
- S_j : Extracted cross-section result, given by
  $S_j = \frac{\int \bar{F}(E_{\nu j}) \cdot \sigma(E_{\nu j})\, dE_{\nu j}}{\int \bar{F}(E_{\nu j})\, dE_{\nu j}}$ (explicitly nominal flux-averaged xsec)
- Easier to compare to model predictions (just use nominal flux)

- Model validation allows us to probe $D(E_\nu \to T_{reco})$ and use nominal flux-weighted xsec for easy model comparisons

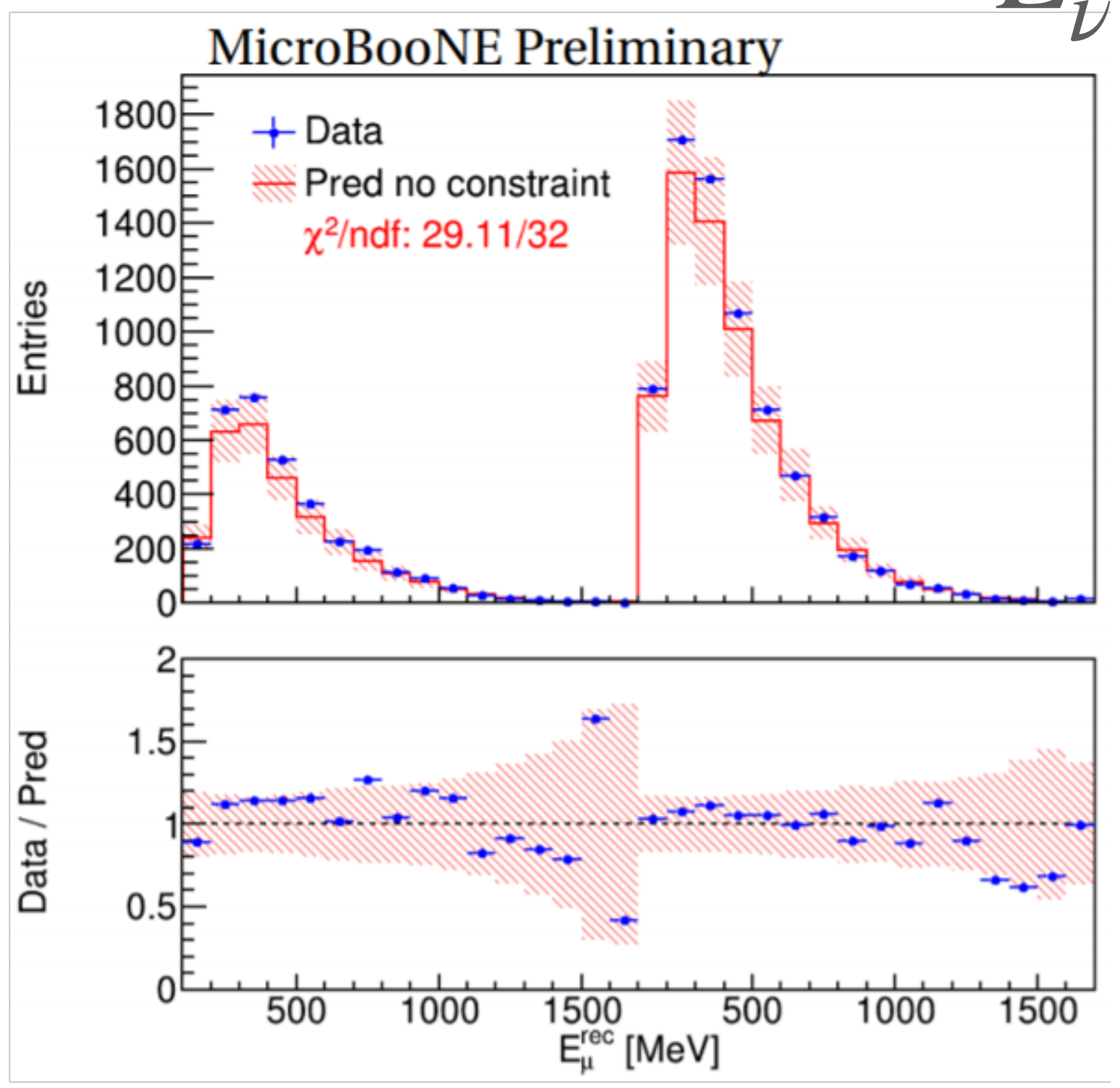
Model Validation Principle

$$\chi^2 = (M-P)^T \times Cov^{-1}_{full}(M,P) \times (M-P)$$

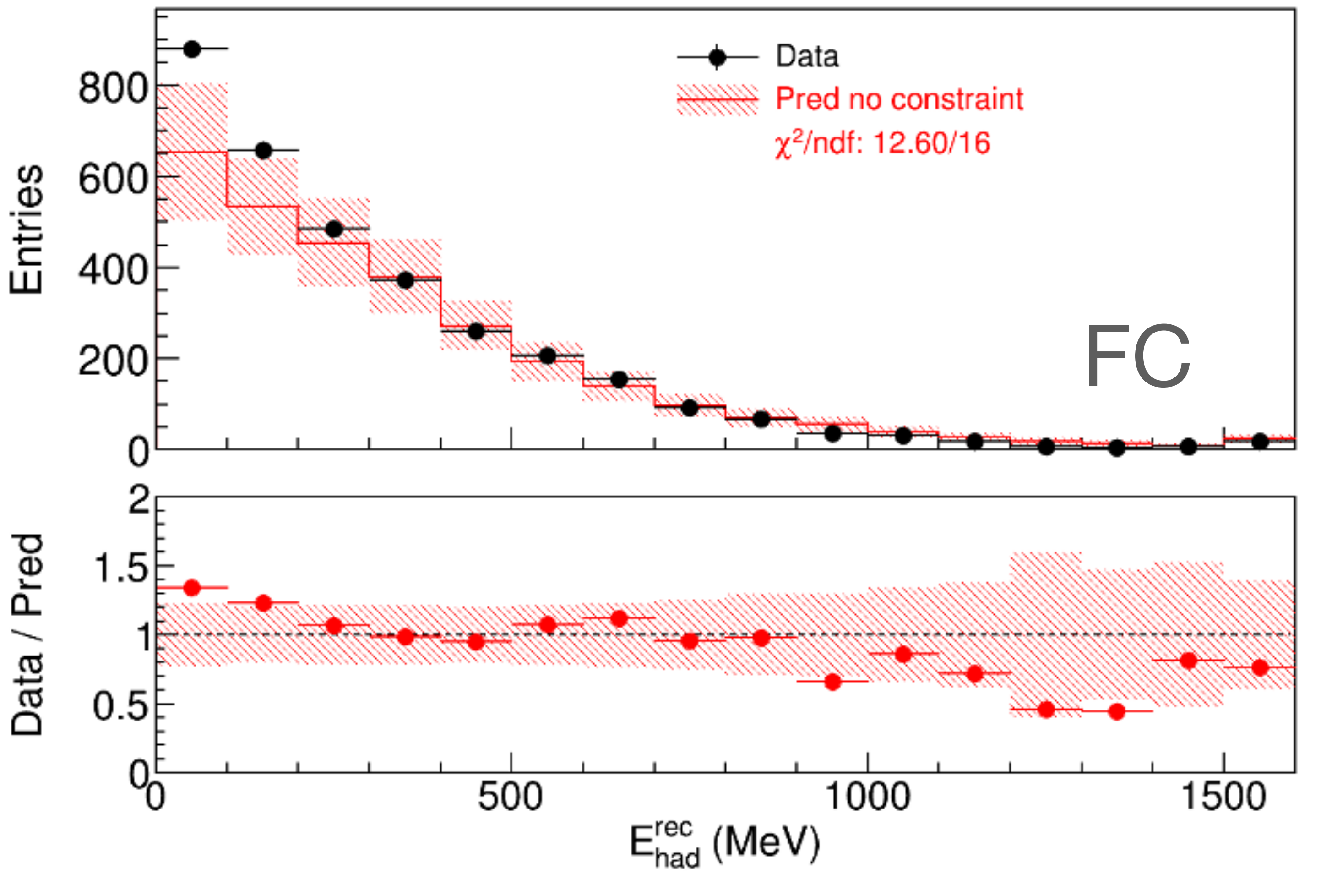
$$E_\nu^{rec} = E_\mu^{rec} + E_{had}^{rec}$$

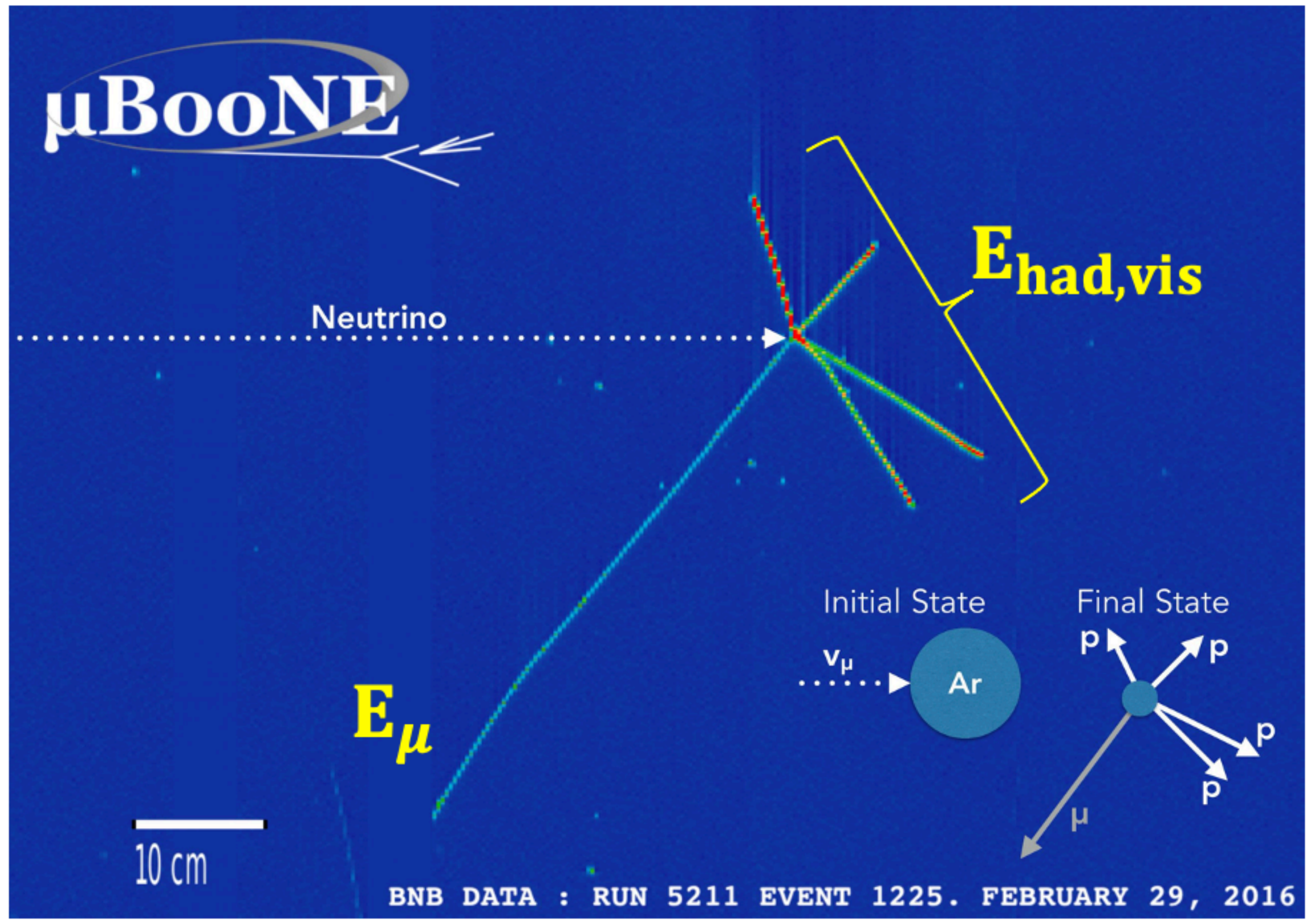


=



+





- Unfolding to E_ν we want to be able to validate

$$E_\nu^{true} = E_\mu + E_{had}^{vis} + E_{had}^{missing} \quad \Rightarrow \quad D(E_\nu \rightarrow E_\nu^{Reco})$$

- But we're not running completely blind (energy is conserved)
- Its not event by event but we have information about distributions of
  - E_ν (from flux), E_μ and E_{had}^{vis} (from data)

Equation For Unfolding

$$M_i - B_i = \sum_j R_{ij} \cdot S_j$$

$$\chi^2 = (\boldsymbol{M} - \boldsymbol{B} - \boldsymbol{R} \cdot \boldsymbol{S})^T \cdot \boldsymbol{V}^{-1} \cdot (\boldsymbol{M} - \boldsymbol{B} - \boldsymbol{R} \cdot \boldsymbol{S})$$

$$R_{ij} = \widetilde{\Delta}_{ij} \cdot \tilde{F}_j$$

$$\widetilde{\Delta}_{ij} = \frac{POT \cdot T \cdot \int_j F\left(E_{\nu\,\mathrm{j}}\right) \cdot \sigma\left(E_{\nu\,\mathrm{j}}\right) \cdot D\left(E_{\nu\,\mathrm{j}}, E_{rec\,i}\right) \cdot \varepsilon\left(E_{\nu\,\mathrm{j}}, E_{rec\,i}\right) \cdot dE_{\nu\,\mathrm{j}}}{POT \cdot T \cdot \int_j \overline{F}\left(E_{\nu\,\mathrm{j}}\right) \cdot \sigma\left(E_{\nu\,\mathrm{j}}\right) \cdot dE_{\nu\,\mathrm{j}}}$$

➥ a MC ratio, less sensitive to Xs uncertainty

$$\tilde{F}_j = POT \cdot T \cdot \int_j \overline{F}\left(E_{\nu\,\mathrm{j}}\right) \cdot dE_{\nu\,\mathrm{j}}$$

$$S_j = \frac{\int_j \overline{F}\left(E_{\nu\,\mathrm{j}}\right) \cdot \sigma\left(E_{\nu\,\mathrm{j}}\right) \cdot dE_{\nu\,\mathrm{j}}}{\int_j \overline{F}\left(E_{\nu\,\mathrm{j}}\right) \cdot dE_{\nu\,\mathrm{j}}}$$

Not subject to prior knowledge of the Xs uncertainty

- ***V*** is the covariance matrix encoding:
  - Data statistical uncertainty: **M**
  - Flux uncertainty: **B, R (F)**
  - Cross-section (Xs) uncertainty: **B, R ($\boldsymbol{\sigma}$)**
  - GEANT4 hadron interaction uncertainty: **B, R (D, ε)**
  - Detector-model uncertainty: **B, R (D, ε)**
  - “Dirt” uncertainty: **B**
  - POT uncertainty (2%): **M**
  - MC statistical uncertainty: **M**
- The unfolded cross section S_j is defined based on the nominal flux $\overline{F}$
  - Easy for model comparisons
  - Simple for uncertainty calculation

PRD 102 (2020) 113012

Benefit Of the S_j Definition

- Define the flux-averaged cross section using the nominal flux $\overline{F}$, thus can be easily compared with any model prediction based on the nominal flux

$$S_j = \frac{\int_j \overline{F}\left(E_{\nu\,\mathrm{j}}\right) \cdot \sigma\left(E_{\nu\,\mathrm{j}}\right) \cdot dE_{\nu\,\mathrm{j}}}{\int_j \overline{F}\left(E_{\nu\,\mathrm{j}}\right) \cdot dE_{\nu\,\mathrm{j}}}$$

- Simplify the uncertainty calculation
  - Switch $\overline{F}$ *to F* would bring up complicated systematic correlation
  - Proper treatment of flux shape uncertainty: PRD **102** 113012

$$M_i - B_i = \sum_j R_{ij} \cdot S_j$$

⟺

$$\chi^2 = (\boldsymbol{M} - \boldsymbol{B} - \boldsymbol{R} \cdot \boldsymbol{S})^T \cdot \boldsymbol{V}^{-1} \cdot (\boldsymbol{M} - \boldsymbol{B} - \boldsymbol{R} \cdot \boldsymbol{S})$$

V is the covariance matrix encoding:

- Data statistical uncertainty: **M**
- Flux uncertainty: **B, R (F)**
- Cross-section (Xs) uncertainty: **B, R (*σ*)**
- GEANT4 hadron interaction uncertainty: **B, R (D, ε)**
- Detector-model uncertainty: **B, R (D, ε)**
- “Dirt” uncertainty: **B**
- POT uncertainty (2%): **M**
- MC statistical uncertainty: **M**

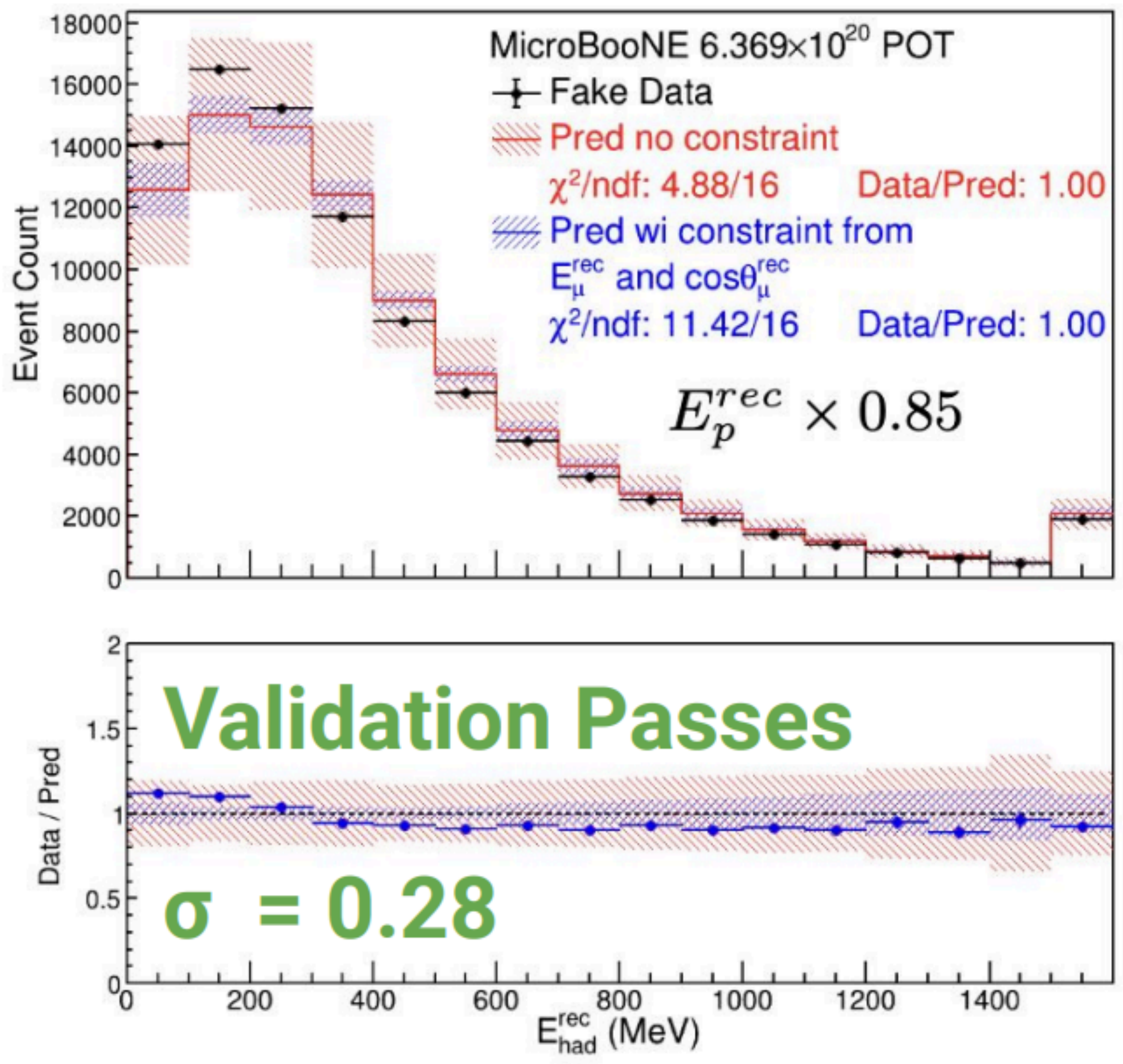
Graphic from Matt Toups

Outcome of “pre-data” interaction model fake data tests

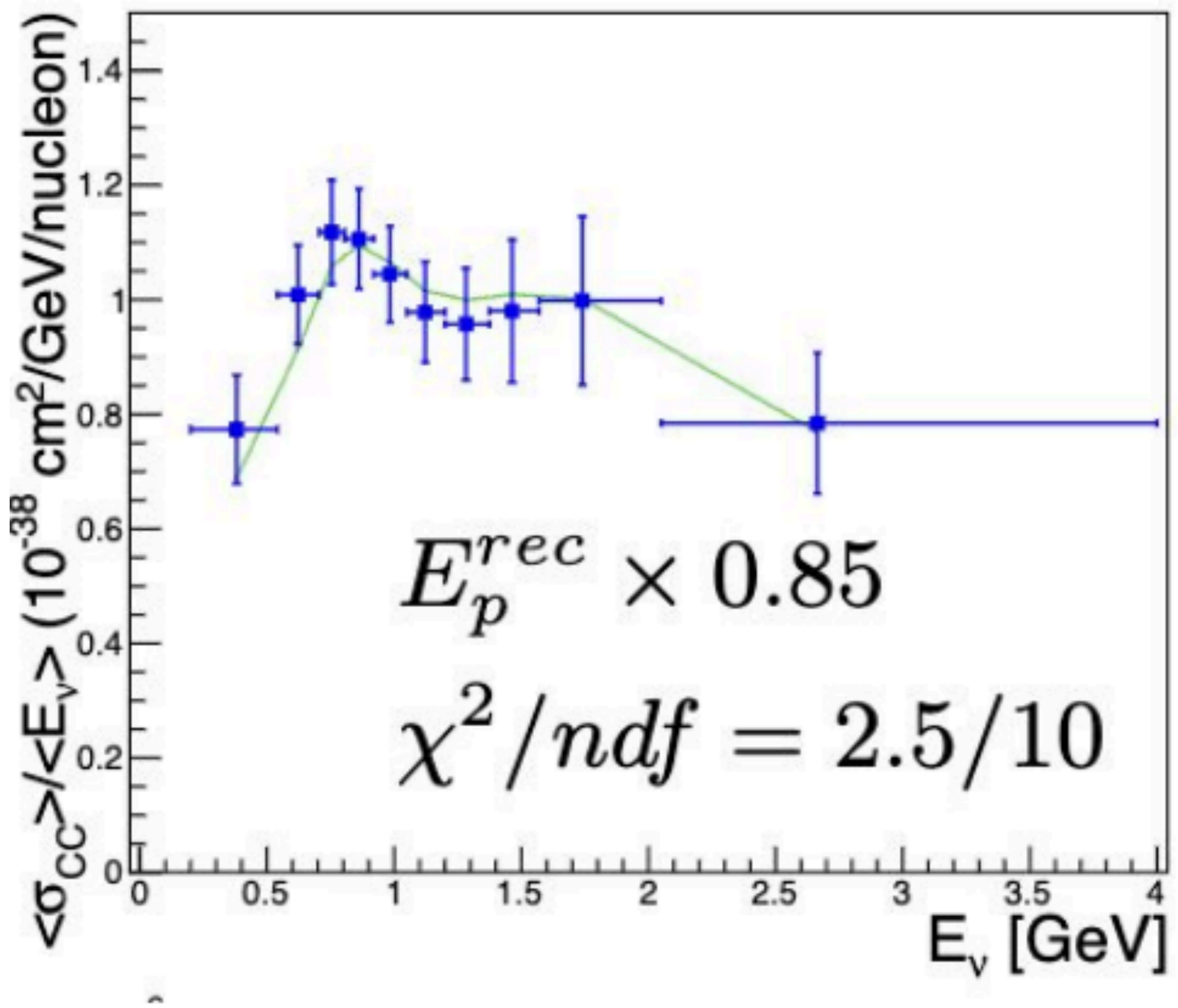
Outcome of “post-data” model validation tests

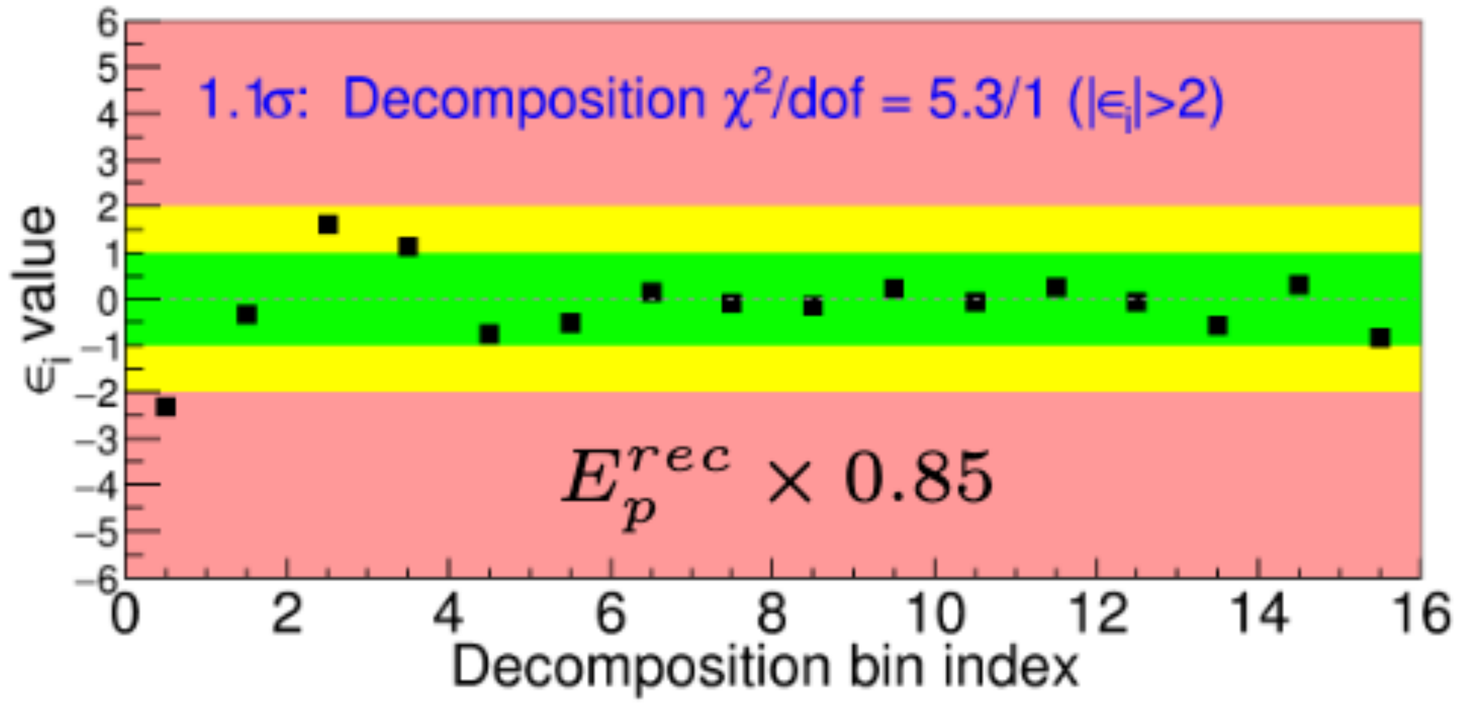
Outcome of “post-data” model validation tests	Interaction model systematics are inadequate	Interaction model systematics are adequate
Full systematics are inadequate	Both methods would have specific concerns about bias in the cross section extraction	Indication of a potential for bias in the cross section extraction from an unknown source, not necessarily due to the interaction model (possible Type-I error, mitigated if extracting cross sections a function of “directly observable” quantities)
Full systematics are adequate	Indication of a potential for bias due to the interaction model that the data itself indicates is only realized as a subdominant effect in the cross section extraction (possible Type-II error)	Both methods would not have specific concerns about bias in the cross section extraction

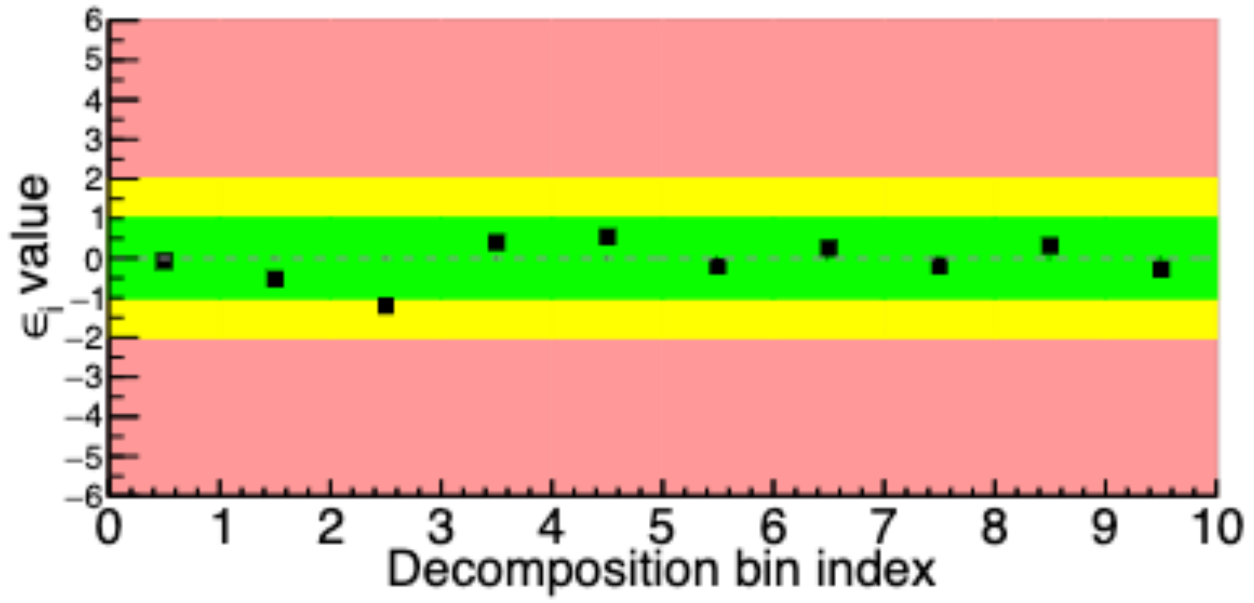
MicroBooNE In-Progress

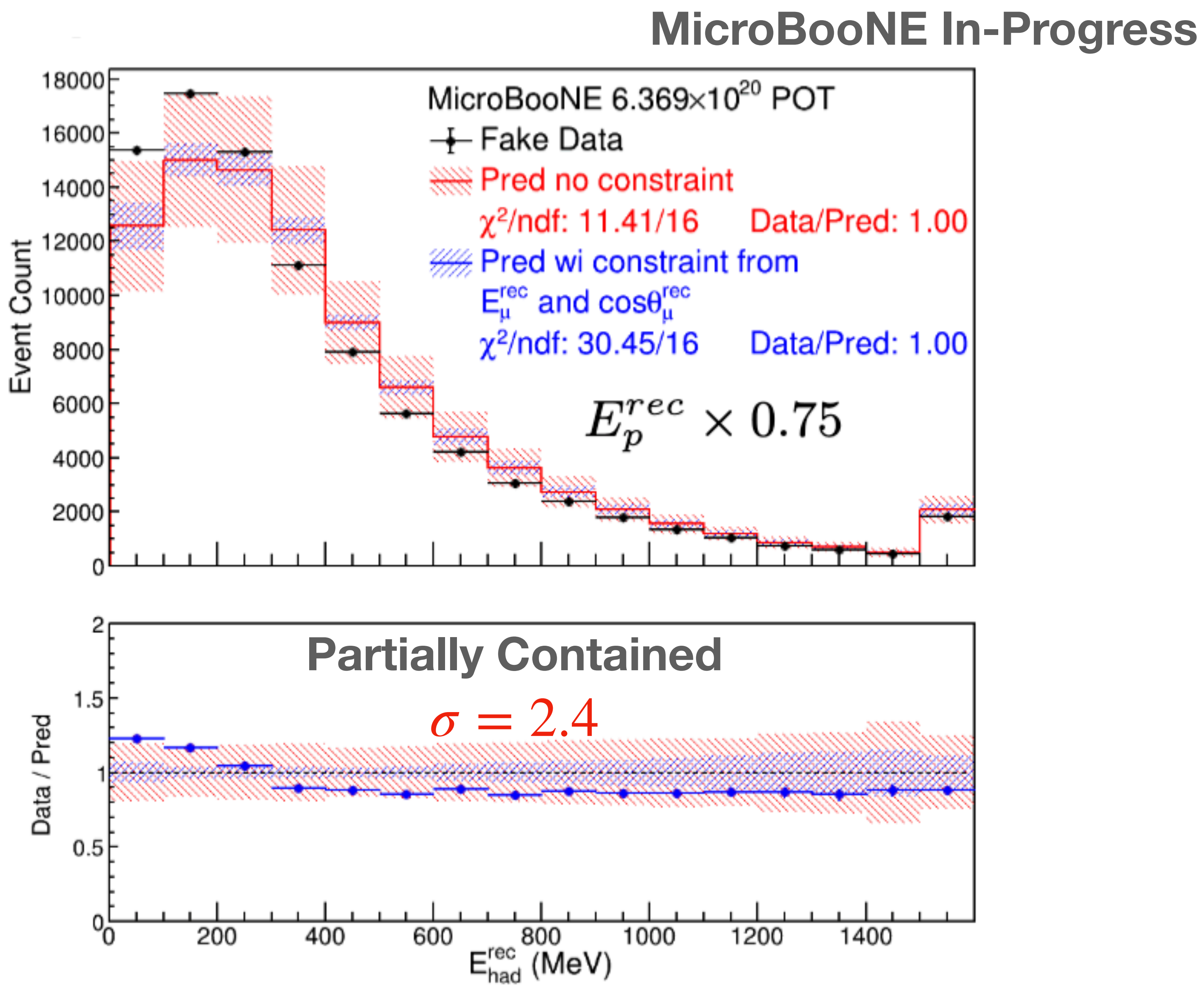


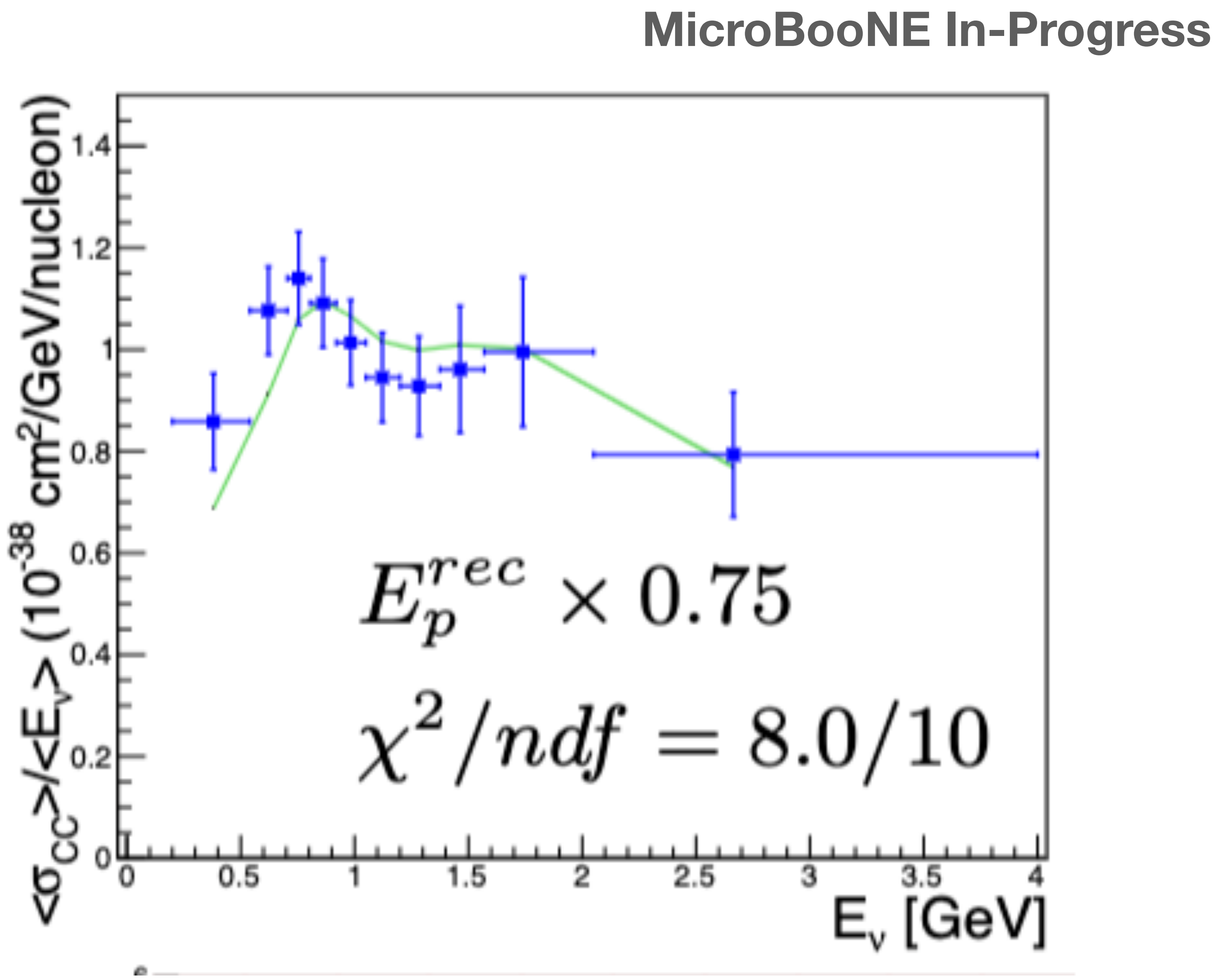
MicroBooNE In-Progress

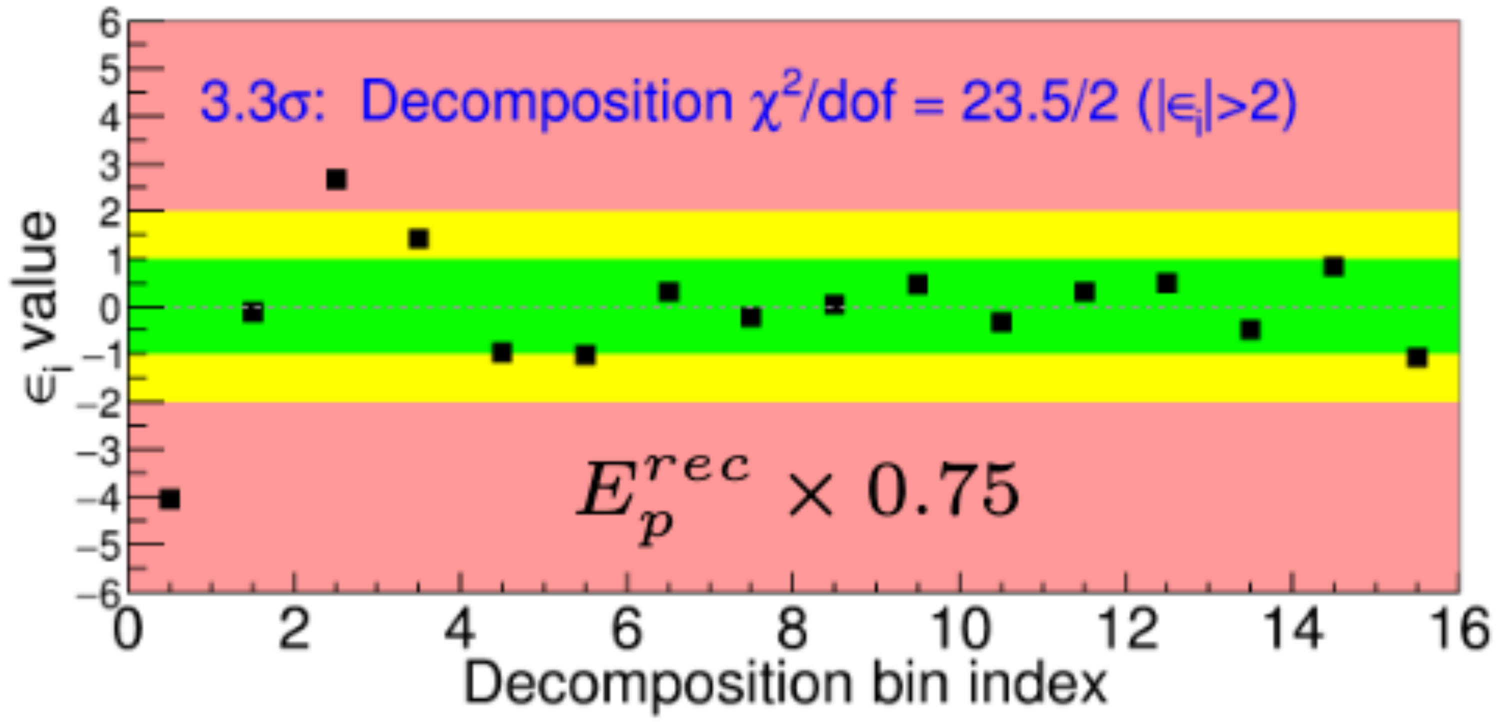


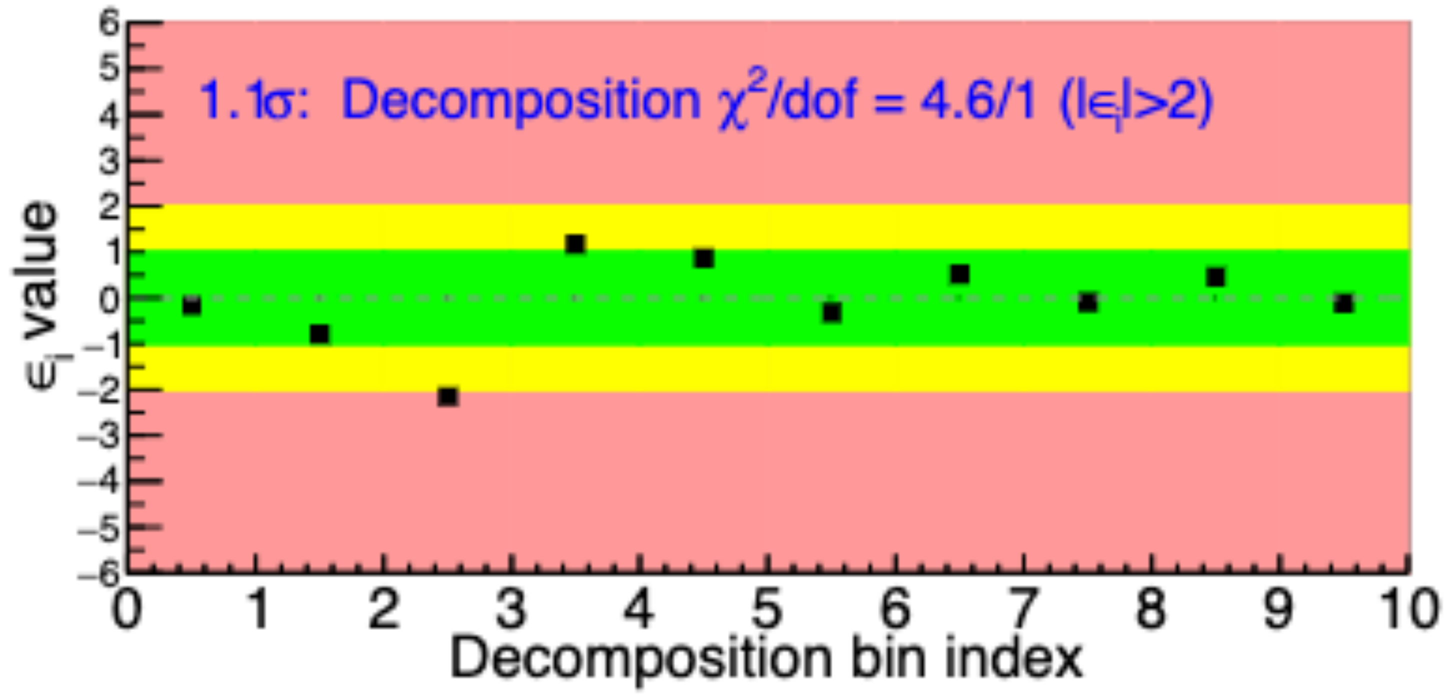





MicroBooNE In-Progress
MicroBooNE 6.369×10^20 POT
Fake Data
Pred no constraint
χ²/ndf: 11.41/16 Data/Pred: 1.00
Pred wi constraint from
E_μ^rec and cosθ_μ^rec
χ²/ndf: 30.45/16 Data/Pred: 1.00
$E_p^{rec} \times 0.75$
Event Count
Partially Contained
$\sigma = 2.4$
Data / Pred
E_had^rec (MeV)


MicroBooNE In-Progress
$E_p^{rec} \times 0.75$
$\chi^2/ndf = 8.0/10$
<σ_CC>/<E_ν> (10^-38 cm²/GeV/nucleon)
E_ν [GeV]


3.3σ: Decomposition χ²/dof = 23.5/2 (|ϵ_i|>2)
$E_p^{rec} \times 0.75$
ϵ_i value
Decomposition bin index


1.1σ: Decomposition χ²/dof = 4.6/1 (|ϵ_i|>2)
ϵ_i value
Decomposition bin index