METHODOLOGY OF THE EXTRACTION OF MULTI-DIFFERENTIAL CROSS SECTIONS OF CHARGED-CURRENT ν_μ-ARGON INTERACTION IN MICROBOONE USING WIRE-CELL ν_μCC SELECTION

May 31, 2022
MICROBOONE_INFO@FNAL.GOV
MICROBOONE-NOTE-1110-PUB

The MicroBooNE Collaboration

Abstract

Neutrino physics experiments rely on accurate nuclear-interaction models that are in part guided by experimental observations. Multi-differential cross section measurements are particularly valuable for exploring the underlying physics described in these models, making them a significant step forward in the field. This note presents methodology in preparation to unfold neutrino flux-averaged double and triple-differential cross sections of the inclusive muon neutrino charged-current interaction on Argon. This work builds upon the the existing framework and methodology presented in recent energy-dependent cross section measurements at MicroBooNE using the Wire-Cell tomographic event reconstruction. The signal definition, choice of binning, and handling of estimated detector uncertainties are discussed, and the MicroBooNE simulation model is validated over the multi-dimensional phase space of muon momentum, muon polar angle, and visible hadronic energy.

Contents

1 Introduction

The MicroBooNE experiment [1] consists of a Liquid Argon Time Projection Chamber (LArTPC) detector along the Booster Neutrino Beam (BNB) at Fermilab. LArTPC physics, and ν-Ar interactions in general, are of particular interest for the field of neutrino physics such as for DUNE. The measurement of multi-differential cross sections at MicroBooNE allows ν-Ar models to be refined for more precise estimations in future experiments, and thus more accurate measurements of physics quantities such as θ_{13}, θ_{23}, Δm^2_{32}, and δ_{CP}.

MicroBooNE is well suited for differential cross section measurements due to the fully active calorimeter LArTPC detector. This detector allows for accurately reconstructed measurements of muon kinematics as well as the visible hadronic energy in the TPC, E_{had}. These quantities can be used to estimate the energy transfer to the argon system ν and the total neutrino energy E_ν, as well as to produce accurate differential and multi-differential cross section measurements, particularly for the relatively-easily-selected charged-current (CC) inclusive ν_μ-argon interactions. For example, the MicroBooNE collaboration has published a double-differential cross section measurement over muon polar angle and muon momentum [2], and a set of energy-dependent (differential) cross section measurements [3]. This note describes the methodology that will be used to produce a novel triple-differential cross section measurement of ν_μ-Ar CC inclusive interactions at MicroBooNE using the Wire-Cell tomographic event reconstruction [4, 5]. Additionally, the proper modeling of the undetectable hadronic energy from MicroBooNE's tuned GENIE model [6] is stringently validated.

2 Event Selection

Based on 6.5×10^{20} protons on target (POT) of data, this analysis involves a series of algorithms from the Wire-Cell event reconstruction that detect and remove specific cosmic-ray signature types, as well as Boosted Decision Trees (BDTs) which remove likely cosmic rays based on a large number of kinematic variables.

A triple-differential cross section can be measured using three reconstructed kinematic variables: muon momentum P_μ, muon polar angle $cos(\theta_\mu)$, and neutrino energy E_ν. P_μ and $cos(\theta_\mu)$ are directly reconstructed from the detector measurements, while E_ν is estimated through the muon energy E_μ and the visible hadronic energy E_{had}. Through the use of MicroBooNE's simulation model [6, 2], reconstruction biases due to undetectable energy such as energy loss through neutrons can be estimated and accounted for when extracting cross sections.

The ν_μ CC signal is defined to require a neutrino interaction vertex inside the fiducial volume of the detector, have a true neutrino energy between [0.2, 4.0] GeV, and have a muon momentum below 2.5 GeV/c. Selected event candidates are binned along the three dimensions of the reconstructed kinematic variables used. There are four E_ν bins defined by the bin edges $\{0.2, 0.705, 1.050, 1.570, 4.0\}$ GeV, nine $cos(\theta_\mu)$ bins defined by the bin edges $\{-1, -0.5, 0, 0.27, 0.45, 0.62, 0.76, 0.86, 0.94, 1\}$, and fifteen P_μ bins 0.1 GeV wide from 0 to 1.5 GeV/c plus an overflow bin up to 2.5 GeV/c. For the purposes of validation events are also binned over E_{had} following the same 15+1 bin structure up to 2.5 GeV. Additionally, events are separated into two categories based on whether they are fully contained (FC), if their main cluster activities reside within the fiducial volume, or partially contained (PC) otherwise.

Through the MicroBooNE simulation, reconstructed kinematics can be compared to the true quantities used to generate them. Fig. 1 shows the efficiency for selecting simulated events over the distribution of their true neutrino energy. Figs. 2 and 3 shows the distribution of reconstructed and truth events across the 2D $\{P_\mu, cos(\theta)\}$ phase space within each of the four E_ν slices. Fig 4 shows the reconstruction efficiency across the 2D $\{P_\mu, cos(\theta)\}$ phase space within each of the four E_ν slices.

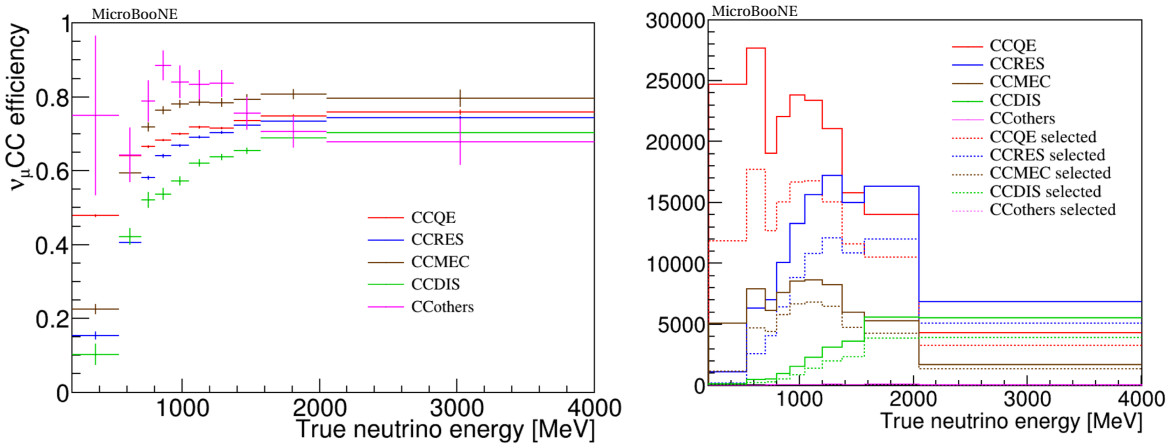

Figure 1: (Left) Efficiencies for ν_μCC selection as a function of true neutrino energy for different interaction types. The error bar includes statistical uncertainty. (Right) Event distributions before and after selection.

3 Systematic Uncertainty Treatment

Many details about systematic uncertainty treatment in this work are in common with those in MicroBooNE's recent neutrino-energy-dependent cross section measurement [3], including the neutrino flux model uncertainty [2], the neutrino-argon interaction cross

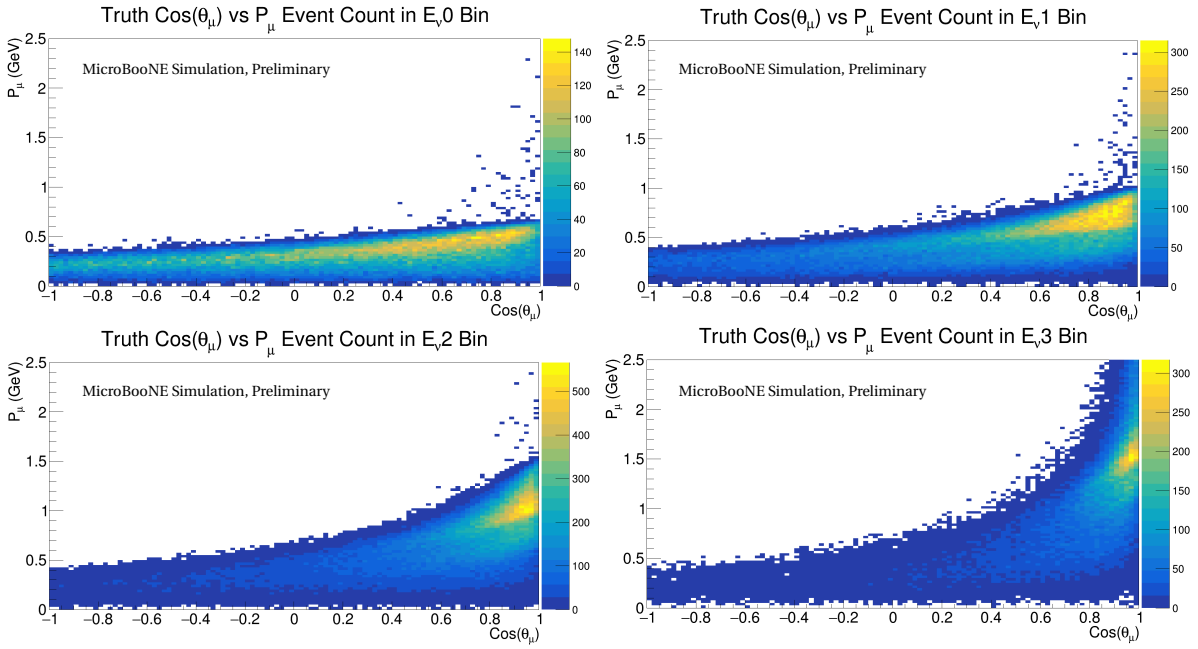

Figure 2: Distribution of simulated events over reconstructed muon momentum and cosine of muon forward angle within each reco E_ν slice. E_ν energy bins are $\{[0.2, 0.705], [0.705, 1.05], [1.05, 1.57], [1.57, 4.0]\}$ GeV.

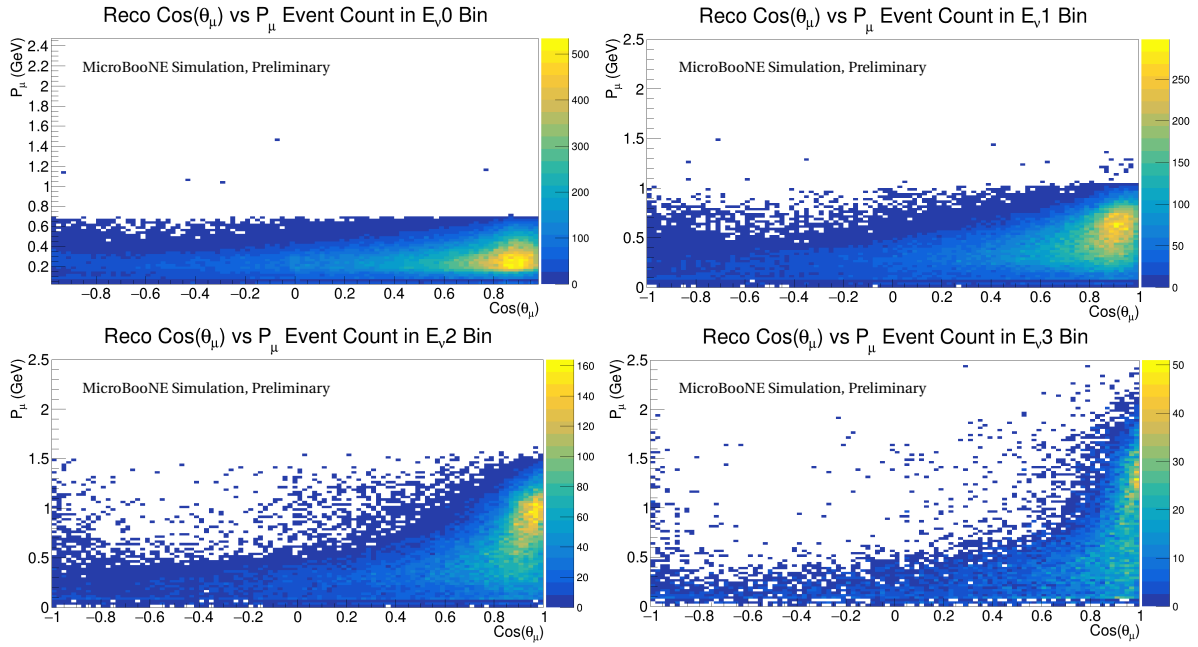

Figure 3: Distribution of simulated events over true muon momentum and cosine of muon forward angle within each truth E_ν slice. E_ν energy bins are $\{[0.2, 0.705], [0.705, 1.05], [1.05, 1.57], [1.57, 4.0]\}$ GeV.

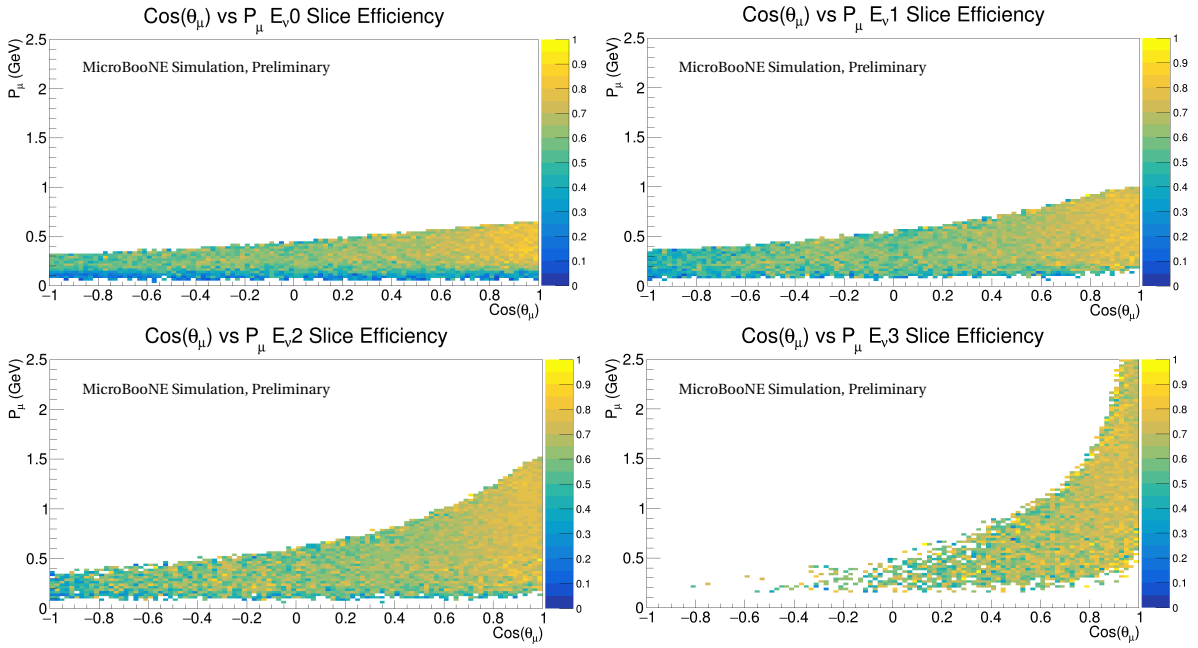

Figure 4: Reconstruction efficiency for simulated events over the 2D truth phase space of muon momentum and cosine of muon forward angle within each truth E_ν slice. E_ν energy bins are $\{[0.2, 0.705], [0.705, 1.05], [1.05, 1.57], [1.57, 4.0]\}$ GeV. An event is considered to be reconstructed if there is a muon track connected to the neutrino vertex.

section model uncertainties [6], the uncertainties on the GEANT4 models [7] used to simulate secondary interactions of protons and charged pions outside the target nucleus [8], the uncertainty on modeling the "dirt" events that originate outside the cryostat [9], the statistical uncertainty of the Monte-Carlo sample [10], the uncertainties on the POT (2% based on in-situ proton flux measurements [11]) and the number of target nuclei (~1%).

The detector response uncertainty follows the work in Ref. [12]. Four types of detector systematic uncertainties are considered: i) variations in the TPC waveform, ii) variations in the light yield (LY) and propagation simulation, iii) variations in the space charge effect simulation, and iv) variations in the recombination modeling of drift electrons. For each variation of detector modeling parameters, a '1σ' sample is simulated and compared with the central value (CV) prediction to assess one standard deviation of each systematic uncertainty. The mean CV-'1σ' difference in event distribution is referred to as the vector V^0, and its uncertainty is described through the covariance matrix M_R, which captures the correlations between bins. The same set of neutrino interactions is used in the comparison to avoid statistical fluctuation, however, some randomness due to the event reconstruction algorithm and limited statistics still exist. As shown in Fig. 5, we utilize a bootstrapping method ("resampling with replacement") [13] to estimate the statistical and correlated systematic uncertainties.

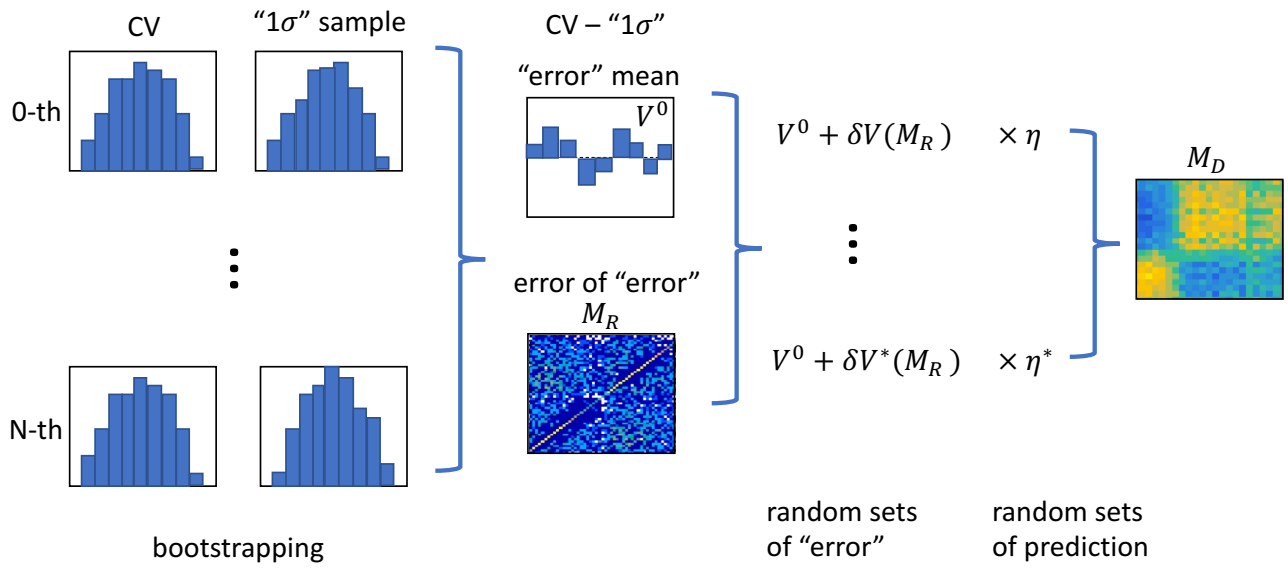

Figure 5: Bootstrapping of detector systematic uncertainties. For each detector modeling variation, we choose the common set of MC events between the CV and the '1σ' samples, and bootstrap MC events to form a distribution of '1σ' deviations. Repeating this procedure many times, we can collect many ensembles and thus establish the mean of the '1σ' deviations, V^0, and its covariance matrix M_R (i.e. the error on the error). In order to determine the final covariance matrix M_D, we adopt a two-step procedure. First, based on M_R, we generate a random set of vectors δV to be applied on the mean difference vector V^0 to obtain a new prediction of the '1σ' error. Second, we generate a random number, η, following the normal distribution $N(0,1)$. The multiplication of the random scalar, η, and the '1σ' error, $V^0 + \delta V$, provides one universe of the detector systematic variation. Repeating the above two-step procedure many times, we can obtain many universes and thus construct the covariance matrix M_D.

This bootstrapping procedure is vulnerable to cases when there are a small number of simulated events relative to the number of bins. Since the simulated events rely on real data to populate the readout with non-beam-window-coincident activity (i.e. cosmic rays), there is ultimately a limited amount of distinct simulated events that can be produced. For double and triple-differential cross sections with a large number of bins, the number of events per bin is low enough that statistical uncertainties dominate the values of V^0 and M_R.

To prevent this over-estimation of the detector variation uncertainty, a smoothing technique based on Gaussian Processes Regression [14] (GPR) is used to improve our estimation of V^0 and M_R. Applying a smoothing algorithm entails computing a smoothed prediction (CV and uncertainty) over a distribution, and using this prediction in place of the raw data that generated it. GPR assumes that measurement bins are jointly Gaussian distributed with a prior distribution based on the bootstrapping covariance matrix M_R, and forms a

prediction given the observed data. In addition to the prior covariance matrix M_R, a Radial Basic Function (RBF) kernel function, K, is computed for each pair of bins and added to the corresponding element in M_R to impose the smoothing intuition that nearby bins are more correlated:

$$K(x_i, x_j) = e^{-|\vec{x}_i - \vec{x}_j|^2 \cdot \vec{s}} \tag{1}$$

$$s_k = \frac{1}{2L_k^2} \tag{2}$$

Here $\vec{x}_i$ ($\vec{x}_j$) denotes the phase space position of the ith (jth) bin along the E_ν, $cos(\theta_\mu)$, and P_μ dimensions. L_k denotes a characteristic length scale for each dimension over which the points are sufficiently correlated. Based on our measured bin resolution, we chose a length scale of 0.1 along the $cos(\theta_\mu)$ axis, 20% of a bin's muon momentum along the P_μ axis, and 20% of a bin's neutrino energy along the E_ν axis.

The prior prediction, including the smoothing correlations, and the observed data are used to form a smoothed posterior prediction (central value and covariance). The posterior prediction central value, V^0, is less driven by statistical fluctuations, and the covariance matrix (M_R representing the error, δV) is consequently smaller. This is seen in Fig. 6 which gives each component as the fraction of the total uncertainty. The detector systematic uncertainty in orange dominates at high energy before and after GPR smoothing, but there is a noticeable 30% reduction in total uncertainty due to the smoothing process.

4 Validation of Model Prediction

MicroBooNE simulates interactions inside the detector using GENIE v3.0.6 tuned [6] to CC0π cross section data obtained by the T2K experiment. This model is used in the unfolding procedure to estimate reconstruction biases such as missing energy due to non-ionizing neutrons. The accuracy of this model in describing the distribution of data is critical to ensuring that the unfolding procedure does not introduce significant bias.

The level of disagreement between data and model prediction is measured using a Pearson χ^2 goodness of fit test by constructing the test statistic:

$$\chi^2 = (M-P)^T \times Cov_{full}^{-1}(M,P) \times (M-P) \tag{3}$$

where M represents the measured data and P describes the model prediction, including contributions from signal and background. First we validate the model prediction describing

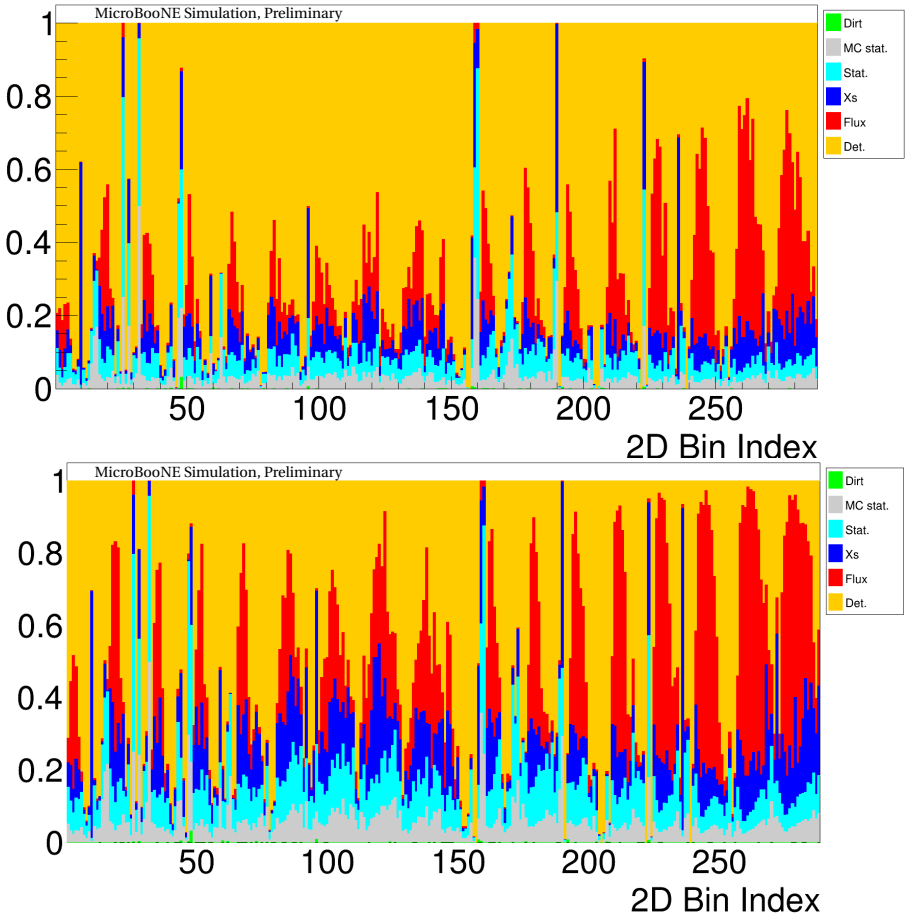

Figure 6: Top (bottom): A breakdown of the full covariance matrix diagonal terms as a fraction of the total uncertainty for each reco-space bin, with each component given in a different color, without (with) smoothing run on the bootstrapping matrix (not directly shown). The 2D binning $\{cos(\theta_\mu), P_\mu\}$ is used, and each plot is split in two with FC $\in [1,144]$ and PC $\in [145,288]$. Within each half there are 9 θ_μ slices ordered from backward facing to forward facing (roughly seen as periodic humps in the plot), and within each θ_μ slice there are 16 P_μ bins. It can be seen that, broadly speaking, detector and flux uncertainties dominate, with lesser contributions from cross-section estimation and statistical uncertainties.

the 2D data distribution over $\{P_\mu, cos(\theta_\mu)\}$ by comparing data to MC. Then we validate the model prediction over the $\{E^{vis}_{had}, cos(\theta_\mu)\}$ distribution, again by comparing data to MC.

Since the measurement of E^{vis}_{had} is more complicated than for P_μ, a more stringent test is created in which we improve the model prediction by using the muon kinematic measurements as a constraint. Because there are systematic uncertainties in common between measurements of P_μ, $cos(\theta_\mu)$, and E^{vis}_{had}, the muon kinematics measurements are capable of restricting the model prediction for E^{vis}_{had}. The simplest example of this would be if we detected an overall rate abundance in the P_μ and $cos(\theta_\mu)$ distributions, which would require

flux model parameters to lie within a narrow range to generate the rate abundance prediction. This would then carry over to the E_{had}^{vis} prediction, where the restricted flux parameters would predict a corresponding rate excess, and the restricted parameter phase space would have lower overall uncertainties.

In practice, this is implemented using a Bayesian process known as a "conditional constraining procedure". More details are documented in the section II-B of the Supplemental Material of Ref. [3]. The default model prediction forms the prior, the muon kinematics measurements are the data that constrain the parameter space and are used to generate the posterior distribution. Due to the significant uncertainties in common which are eliminated upon applying the muon kinematics constraint, this process demonstrates that most of the data/MC disagreement in the $\{E_{had}^{vis}, cos(\theta_\mu)\}$ distribution can be explained by the disagreement in the $\{P_\mu, cos(\theta_\mu)\}$ distribution, which is relatively well reconstructed. Validating the model's prediction of E_{had}^{vis} gives us confidence that the unfolded transfer energy $\nu = E_{had}^{vis} + E_{had}^{missing}$ and neutrino energy $E_\nu = \nu + E_\mu$ are estimated in a fairly unbiased manner over the 3D phase space. Furthermore, it is clear both visually and in the χ^2/ndf calculations that on the whole we overestimate our uncertainties. This will reduce the rejection power of the unfolded result slightly, but gives the model more tolerance to account for potential inaccuracies, such as in modeling $E_{had}^{missing}$, without introducing significant bias in the unfolding.

Figs. 7 and 8 show the comparison between data and prediction over the E_μ^{vis} and E_{had}^{vis} distributions respectively. Fig. 9 shows the comparison between data and prediction over the 2D reco-space binning of $\{P_\mu, cos(\theta_\mu)\}$ for FC and PC events separately. Fig. 10 shows the comparison between data and prediction over a 2D binning of $\{E_{had}, cos(\theta_\mu)\}$ for FC and PC events separately, before and after applying the muon kinematics constraint. All of these comparisons yield GoFs (χ^2/ndf) below unity and corresponding p-value above 0.9, demonstrating that the model and its uncertainties are able to explain the distribution observed in data over $\{P_\mu, E_{had}^{vis}, cos(\theta_\mu)\}$.

5 Conclusion

This public note describes the methodology that will be used to produce a triple-differential numu charged-current inclusive cross-section measurement in neutrino energy, muon forward angle, and muon momentum. The Wire-Cell reconstruction and event selection algorithms allow for sufficient kinematic phase space coverage and reconstruction to accurately unfold results in multiple dimensions. The work in this note is crucial in performing an

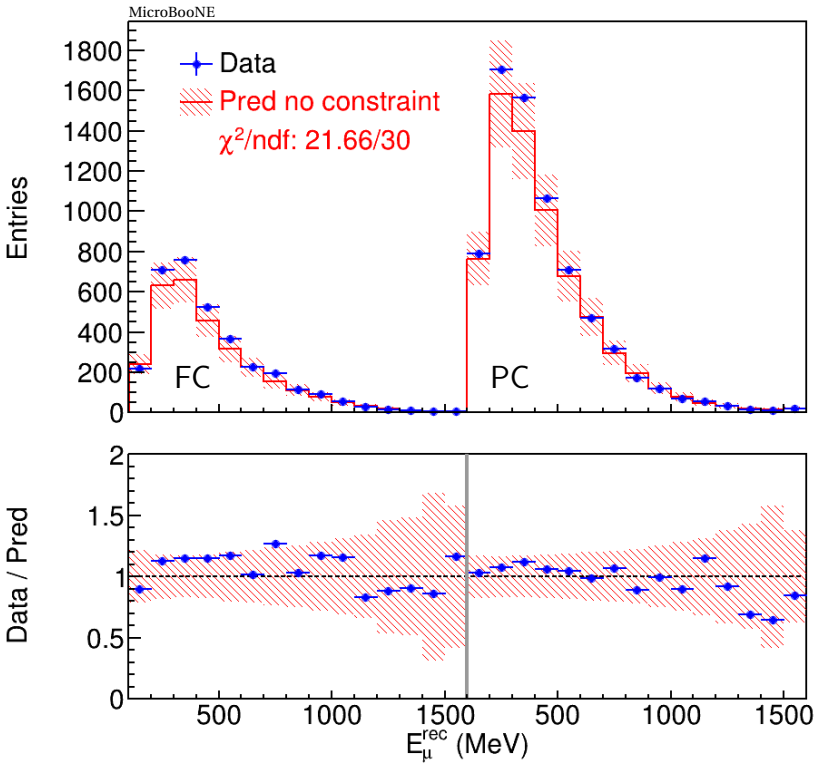

Figure 7: Comparison between data and prediction as a function of E_{μ}^{rec} (also referred to as E_{μ}^{vis}). The statistical uncertainties of the data and Monte Carlo simulation are also included in the bands. For each plot, the first 14 bins with 100 MeV per bin correspond to the fully contained events from 0.1 GeV to 1.5 GeV. The 15th bin is the overflow bin corresponding to fully contained events above 1.5 GeV. The next 14 bins with 100 MeV per bin correspond to the partially contained events from 0.1 GeV to 1.5 GeV. The last bin is the overflow bin corresponding to the partially contained events above 1.5 GeV.

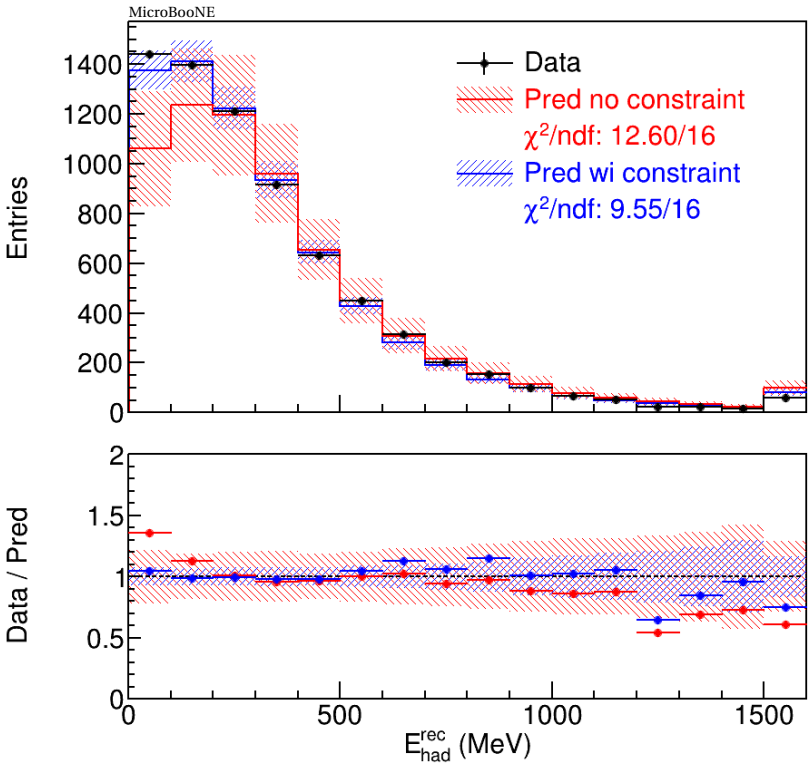

Figure 8: Comparison between data and prediction as a function of E_{had}^{rec} (also referred to as E_{had}^{vis}) (right). The statistical uncertainties of the data and Monte Carlo simulation are also included in the bands. The red (blue) lines and bands show the prediction without (with) the constraints from the fully contained event sample. For each plot, the first 15 bins correspond to 0 GeV to 1.5 GeV. The last bin correspond to overflow bin above 1.5 GeV.

accurate unfolding, and serves as a stepping stone towards triple-differential cross section results.

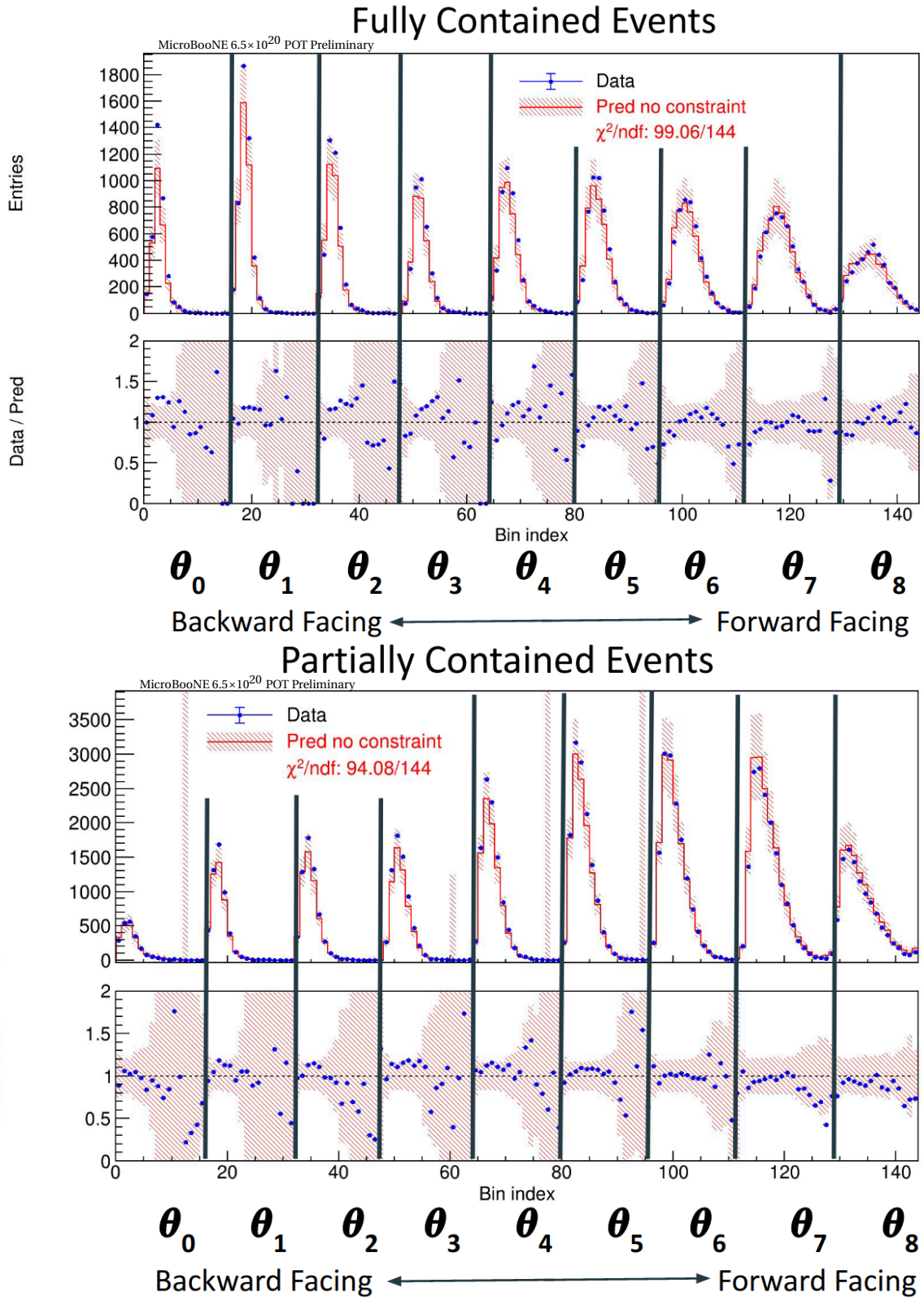

Figure 9: Top (Bottom): Distribution of data and prediction over the 2D reco-space binning of P_μ and $cos(\theta_\mu)$ for FC (PC) events. GPR smoothing is run on the detector systematic bootstrapping matrix. The x-axis bin index is an enumeration of the 2D binning $\{P_\mu, cos(\theta_\mu)\}$ where the 9 θ_μ slices are placed side by side with 16 P_μ bins in each slice.

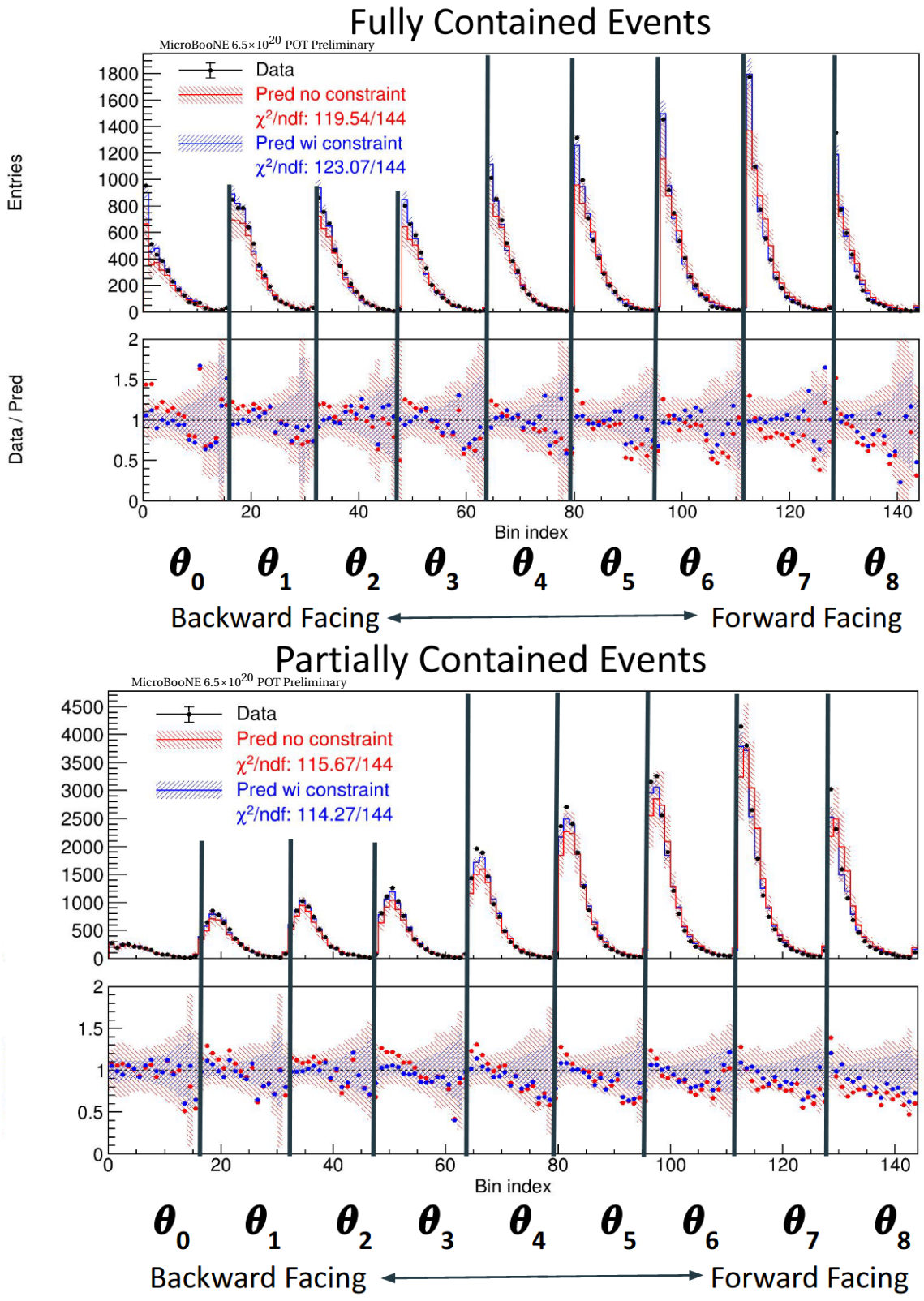

Figure 10: Top (Bottom): Distribution of data and prediction over the 2D reco-space binning of E_{had} and $cos(\theta_\mu)$ for FC (PC) events. GPR smoothing is run on the detector systematic bootstrapping matrix. The x-axis bin index is an enumeration of the 2D binning $\{E_{had}, cos(\theta_\mu)\}$ where the 9 θ_μ slices are placed side by side with 16 E_{had} bins in each slice.

References

[1] R. Acciarri et al. (MicroBooNE Collaboration), Design and construction of the MicroBooNE detector. Journal of Instrumentation, 12(02):P02017–P02017, February 2017.

[2] P. Abratenko et al. (MicroBooNE collaboration), First Measurement of Inclusive Muon Neutrino Charged Current Differential Cross Sections on Argon at E_ν 0.8 GeV with the MicroBooNE Detector. Phys. Rev. Lett., 123(13):131801, 2019.

[3] P. Abratenko et al. (MicroBooNE Collaboration), First Measurement of Energy-Dependent Inclusive Muon Neutrino Charged-Current Cross Sections on Argon with the MicroBooNE Detector. Phys. Rev. Lett., 128(15):151801, 2022.

[4] Xin Qian, Chao Zhang, Brett Viren, and Milind Diwan. Three-dimensional Imaging for Large LArTPCs. JINST, 13(05):P05032, 2018.

[5] P. Abratenko et al. (MicroBooNE Collaboration), Neutrino event selection in the MicroBooNE liquid argon time projection chamber using Wire-Cell 3D imaging, clustering, and charge-light matching. JINST, 16(06):P06043, 2021.

[6] P. Abratenko et al. (MicroBooNE Collaboration), New $CC0\pi$ GENIE model tune for MicroBooNE. Phys. Rev. D, 105(7):072001, 2022.

[7] J. Allison et al. Recent developments in Geant4. Nucl. Instrum. Meth. A, 835:186–225, 2016.

[8] J. Calcutt, C. Thorpe, K. Mahn, and Laura Fields. Geant4Reweight: a framework for evaluating and propagating hadronic interaction uncertainties in Geant4. JINST, 16(08):P08042, 2021.

[9] P. Abratenko et al. (MicroBooNE Collaboration), Search for an anomalous excess of inclusive charged-current ν_e interactions in the MicroBooNE experiment using Wire-Cell reconstruction. arXiv:2110.13978 [hep-ex] [Phys. Rev. D (to be published)].

[10] Carlos A. Argüelles, Austin Schneider, and Tianlu Yuan. A binned likelihood for stochastic models. JHEP, 06:030, 2019.

[11] A.A. Aguilar-Arevalo et al. (MiniBooNE collaboration), The Neutrino Flux prediction at MiniBooNE. Phys. Rev. D, 79:072002, 2009.

[12] P. Abratenko et al. (MicroBooNE Collaboration), Novel Approach for Evaluating Detector-Related Uncertainties in a LArTPC Using MicroBooNE Data. Eur. Phys. J. C, 82:454, 5 2022.

[13] Bootstrapping (statistics). https://en.wikipedia.org/wiki/Bootstrapping_(statistics). Accessed: 2022-04-28.

[14] Jie Wang. An intuitive tutorial to gaussian processes regression. arXiv preprint arXiv:2009.10862, 2020.