Progress on a High Sensitivity, High Purity, One-Electron-One-Proton Search for the MiniBooNE Low Energy Excess in the MicroBooNE Detector

MicroBooNE Collaboration
MICROBOONE-NOTE-1039-PUB

MicroBooNE is a large liquid argon time projection chamber (LArTPC) located at Fermilab. One of the experiments primary aims is to investigate the CCQE-like excess of ν_e events observed by MiniBooNE at low energy. A novel approach to analyzing LArTPC data using an amalgam of Deep Learning and traditional algorithms has been developed and is capable of end to end reconstruction and selection. We report here an overview and progress to event selection which yields greater than 99.999% rejection of backgrounds. Statistical, flux, and some cross section systematics have also been evaluated. Detector systematics and additional cross section systematics are still under study.

PACS numbers: 14.60.Pq,14.60.St

I. INTRODUCTION

This public note reports progress on a search for the MiniBooNE Low Energy Excess (LEE) in MicroBooNE data. This unexpected excess of events that reconstruct as ν_e charged-current quasielastic (CCQE) interactions ($\nu_e + n \to e^- + p$) in the MiniBooNE Cherenkov detector has been attributed to short baseline $\nu_\mu \to \nu_e$ oscillations, which necessitates the introduction of a sterile neutrino into the theory [3]. However, the 4.5σ electron-like signal of 381.2 ± 85.2 events is observed in a selection of 1959 electron-like events [6]. Hence, this is a high-statistics but systematics-limited and high-background analysis, as shown in Fig. 1 (top). Before attributing this signal to new physics, the possibility that the excess is due to background must be thoroughly addressed.

The MicroBooNE experiment, now running upstream of MiniBooNE on the Fermilab Booster Neutrino Beamline (BNB), was conceived to investigate a leading Standard Model hypothesis for the LEE anomaly: that the signal is due to an unidentified source of photons ($\gamma \to e^+e^-$). This signature differs from the ν_e CCQE signal in that there is no proton at the vertex, and a e^+e^- pair is produced rather than a single e^-. Unlike MiniBooNE, which is a 450 t Cherenkov detector, the MicroBooNE detector is a 90 t fiducial mass LArTPC [2]. This has two advantages over a Cherenkov detector in isolating ν_e CCQE events from γ backgrounds: (1) in an LArTPC, protons above $\sim$ 25 MeV are reconstructed, while in a Cherenkov detector, all protons below 350 MeV are invisible; and (2) in a LArTPC, we expect to be able to distinguish γ conversion to an e^+e^- pair from a single e^- in about 85% of the events [1], while in a Cherenkov detector, e^+e^- and e^- cannot be distinguished. The strategy of the search is to place strict cuts to isolate a high purity, although statistics-limited, LEE signal. The goal is to reduce photon backgrounds to the MicroBooNE LEE analysis to a negligible level, leaving only the intrinsic ν_e in the beam as an irreducible background. Fig. 1, bottom, shows the expectation at truth-level. As we discuss below, the intrinsic ν_e energy distribution will peak at higher energies than the LEE signal, allowing some separation. Thus it can be reasonably expected that the search can achieve more than 5σ significance.

The MicroBooNE LArTPC [2] is installed 470 m from Fermilab's BNB target and 90 m upstream of the MiniBooNE detector. The beam is primarily ν_μ, with peak energy $\sim$700 MeV. The detector has a total mass of 170 tonnes of liquid argon, with an active region of $2.6 \times 2.3 \times 10.4$ m^3. The system consists of two subdetectors: a time projection chamber (TPC) for tracking, and a light collection system. Light is collected in "time ticks" of 15.625 ns. The TPC drifts ionized electrons using an applied electric field of 273 V/cm to three wire planes (U, V, Y) that provide the charge read-out. The wire spacing is 0.3 mm, and the shaping time is 2 μs, resulting in highly detailed event information that we exploit to form "images" for this analysis.

This analysis concentrates on the low energy events observed in interactions with the liquid argon in the detector. The LEE signal is expected to predominantly appear in the < 600 MeV range, where CCQE interactions dominate. In this region, meson exchange current (MEC) interactions also represent about 20% of the signal. These events often have an additional proton that eliminates them from the 1 lepton-1 proton ($1\ell 1p$) signature. Neutral current production of neutral pions (NCπ^0) and charged current production of neutral and charged pions (CCπ^+ and CCπ^0) have higher production thresholds than CCQE and MEC. Production of more than one pion is collectively called DIS in this analysis.

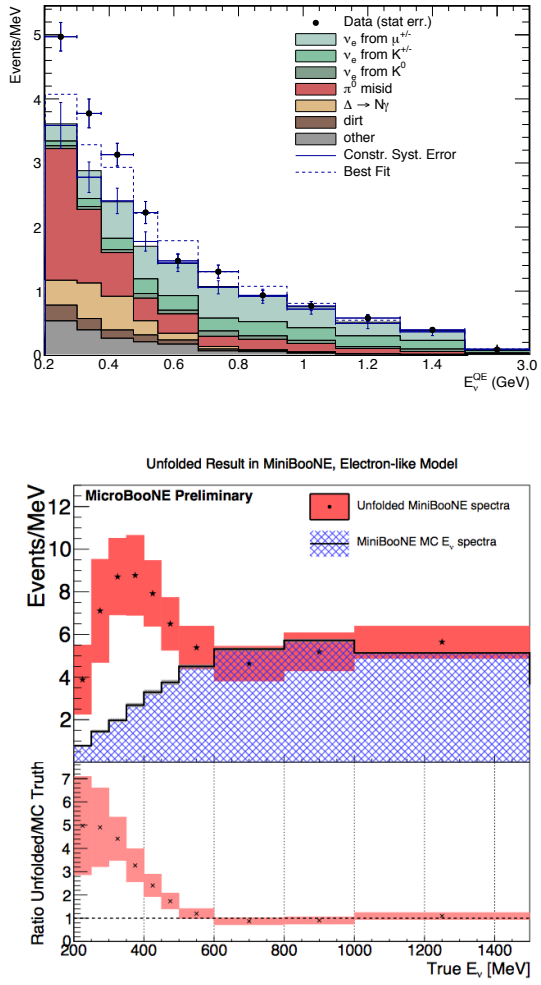


FIG. 1: *Top*: The MiniBooNE electron-like events measured in neutrino mode compared to prediction [6]. The low-energy excess (LEE) is the difference between the points and the stacked histogram of backgrounds.
Bottom: Expectations of the LEE signal seen in MicroBooNE. The excess (here from the 2013 data-release [5]) is unfolded for MiniBooNE detection effects and simulated through the MicroBooNE detector simulation. More detail can be found in [13].

The key to the MicroBooNE LEE analysis described here is to identify events with one electron and one proton meeting at vertex ($1e1p$) which is consistent with CCQE, but not with the other interaction signatures described above. Requiring this distinct topology restricts efficiency, but also offers distinctive features which provide handles with which to greatly reduces backgrounds from cosmic rays, the most significant background on the experiment. To constrain systematic errors, we simultaneously fit to the ν_μ interaction counterpart, $1\mu1p$.

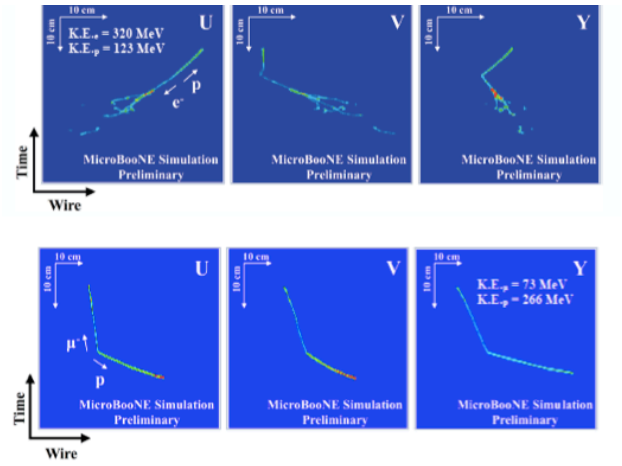


FIG. 2: *Top*: A simulated $1e1p$ event image shown in the U, V, and Y plane;
Bottom: A simulated $1mu1p$ event image in the U, V and Y planes.

Examples of $1e1p$ and $1\mu1p$ events, presented in the form of images used in this analysis, are shown in Fig. 2. The analysis is optimized for the LEE signal energies, which range from about 200 to 600 MeV [13]. In the analysis, we use a combination of C++ algorithms for cosmic-ray rejection, vertex-finding, and 3-D reconstruction and use Deep Learning (DL) algorithms for particle identification [4, 8], for which the "images" are well-suited.

The analysis described here is called the MicroBooNE Deep Learning Low Energy Excess (DL LEE) Analysis. The analysis consists of the following main steps:

1. Cuts based on observed light to remove low-energy events that are not relevant to the analysis.
2. Identify charge observed in the TPC as either cosmic ray tracks or contained regions of charge that are of potential interest in this analysis.
3. Conversion of the event to images, like those shown in Fig. 2.
4. Identification of each pixel as track-like, shower-like or background using semantic segmentation, a deep-learning algorithm.
5. Three-dimensional vertex finding.
6. Three dimensional event reconstruction.
7. Selection of events with two reconstructed particles, hence $1\ell1p$ event candidates, to proceed to the $1e1p$ and $1\mu1p$ selections:

(a) $1e1p$ selection:

i. Use a deep-learning-based multi-particle event identification algorithm to determine the probability that different particles are in the event

ii. Use machine learning algorithms and cuts to identify the $1e1p$ candidates.

(b) $1\mu1p$ selection: Use likelihoods and cuts to identify the $1\mu1p$ candidates.

8. Constrain the systematics on the $1e1p$ candidates using the $1\mu1p$ candidates.

9. Determine the consistency of the $1e1p$ candidates with the Standard Model prediction.

Steps 1 through 6 are already well-summarized in public notes and papers, see Refs. [7–9]. This public note discusses the status of steps 7-9, as of summer, 2019.

Final sensitivity will require the flux and cross section uncertainties, as well as a detailed study of the detector systematic uncertainties and the use of the $1\mu1p$ sample as a constraint. Some, but not all, detector systematics have been evaluated, and so these are not included at this time. Additional plots that support this analysis that are not presented in this main text appear in appendices at the end of this note.

II. OVERVIEW OF THE RECONSTRUCTION

The reconstruction has been described elsewhere [7–9], however a brief explanation here will provide context for the discussion of the selection, constraint and sensitivity. The event reconstruction was tuned on MicroBooNE's open 5E19-protons-on-target sample. Standard MicroBooNE beam simulation is used for comparison, using the GENIE neutrino generator [12]. This analysis uses cosmic data overlaid on the neutrino simulation as described in Ref. [7].

The goal is to isolate $1e1p$ and $1\mu1p$ signals with reconstructed lepton kinetic energy above 35 MeV and reconstructed proton kinetic energy above 60 MeV. The signature must have a vertex with no gaps between the particles and must be located at least 10 cm from the edge of the active volume of the TPC. The outgoing particles are required to be contained in the detector, which greatly reduces backgrounds from cosmic rays, which are significant in MicroBooNE due to the millisecond-scale drift time of a LArTPC.

Metric	value
Muon energy resolution	3.3%
Proton energy resolution	2.6%
ν_μ energy resolution	1.8%

TABLE I: Metrics for energy reconstruction in ν_μ events. See [7]

The first step in reconstruction is to reject low energy backgrounds using the light collection system. The requirements are no signal with more than 20 photelectron (PE) within a 6 time-tick window in the 2μs window prior to the beam spill; a signal of more than 20 PE within a single 6 time-tick window or in two adjacent windows during the beam spill; and less than a 60% of the total light in any given photomultiplier tube. These cuts retain more than 97% of the neutrino events while rejecting more than 75% of empty background events.

Next, a set of Cosmic Tagger algorithms is employed to identify the charge in the TPC associated with tracks that cross the boundaries of the active volume of the detector—the signature of an incoming cosmic ray. Following this, sets of remaining, contained charge are identified as contained regions of interest (cROIs) if they have an associated flash in the light collection system.

At this point, the MicroBooNE data expressed as images for input into the DL algorithms. The images are two-dimensional plots, with wire number along the x axis and drift time along the y axis. The intensity of each "pixel" is given by the sum of the the amplitude of the noise-filtered, deconvolved signal from one or six TPC time-ticks (3 μs). The drift velocity of the electrons is around 0.11 cm/μs, therefore, the drift distance spanned over six time ticks is about 0.33 cm, which is roughly equivalent to the detector's 0.3 cm wire pitch.

The cROI images are fed to the first DL algorithm, which performs semantic segmentation using the SSNet track/shower separation algorithm is described in Ref. [8]. This algorithm labels each pixel in the image as track-like, shower-like or background.

Next, vertex finding and 3-D reconstruction algorithms are employed, as described in Ref. [7]. No gap is allowed between particles at the vertex– we will return to this point when we discuss elimination of NCπ0 backgrounds in the $1e1p$ signal. The most important outcome of the 3-D reconstruction is the neutrino energy, which can be defined in three ways:

$$E_\nu^{range} = \mathrm{KE}_p + \mathrm{KE}_\ell + M_\ell + M_p - (M_n - B); \quad (1)$$

$$E_\nu^{QE}[\mathrm{p}] = 0.5 \cdot \frac{2 \cdot (M_n - B) \cdot E_p - ((Mn - B)^2 + M_p^2 - M_\ell^2)}{(M_n - B) - E_p + \sqrt{(E_p^2 - M_p^2)} \cdot \cos\theta_p}; \quad (2)$$

$$E_\nu^{QE}[\mathrm{lepton}] = 0.5 \cdot \frac{2 \cdot (M_n - B) \cdot E_\ell - ((Mn - B)^2 + M_\ell^2 - M_p^2)}{(M_n - B) - E_\ell + \sqrt{(E_\ell^2 - M_\ell^2)} \cdot \cos\theta_\ell}; \quad (3)$$

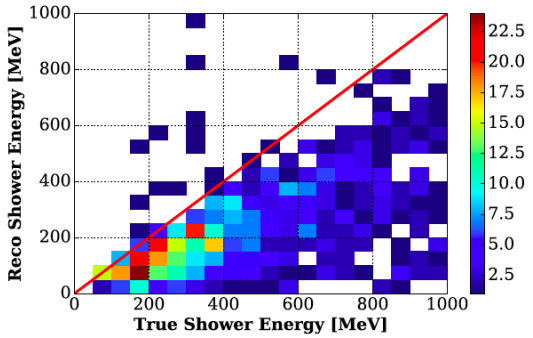
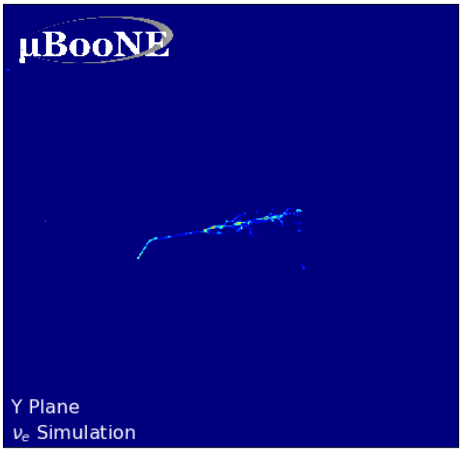


FIG. 3: The resolution of the reconstructed shower energy for ν_e events using range. Reconstructed shower energy in general under estimates the true shower deposited energy, and this will be corrected in the future.



	Proton	Electron	Gamma	Muon
MPID Score	0.74	0.92	0.23	0.35

FIG. 4: MPID example of 1e1p topology with a table output of particle probabilities. Scores indicate higher probabilities of having proton and electon in the image. The image applied to PID is 512 × 512 pixels. Example event is generated by MicroBooNE beam simulation.

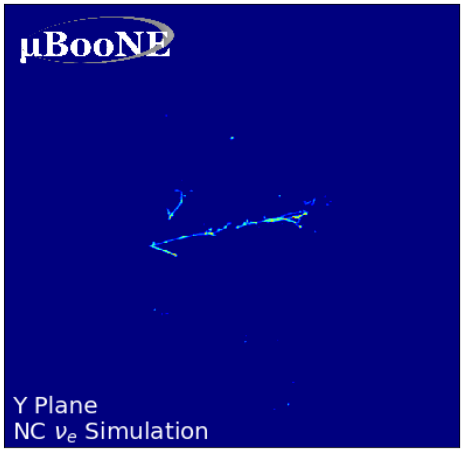
Ref. [7] described reconstruction of ν_μ events. Useful metrics for the reconstruction are reported in Table I for two-track events. The energy resolution is determied by comparing simulated true neutrino energy to E_ν^{range} in this case. The vertexing and tracking algorithms are optimized for identifying two-particle events that are each longer than 3cm. However, the code reports results for up to five particles crossing the 3 cm radius. The events with three particles may be useful in the future for constraint of the cross section systematic errors.

The ν_e event reconstruction is not described in Ref. [7], and so we expand on this here. The particularity of this reconstruction is that it uses the pixels labelled as shower by the SSNet in addition to the track-like pixels. After the vertex location is reconstructed, the shower reconstruction finds the 3D-cone that best clusters the shower-like pixels attached to the same pixel cluster as the vertex point. This procedure returns the direction of the shower and the collected charge is estimated from the intensity of the shower-like pixels in the collection plane. Fig. 3 (Top) shows the shower energy resolution for two-particle events using range. This reconstruction, which is used throughout MicroBooNE, is known to reconstruct the shower energy at lower values than the true energy, as can be seen on this plot. As discussed in Sec. VI, we envision several improvements to the energy estimation in the future, but at present no correction is employed.

At this point, the analysis splits into two related, but independent analysis chains. The first is the 1e1p event selection, which isolates the potential signal and the irreducible intrinsic ν_e background with very high purity, and the second is the 1μ1p event selection, which identifies the sample used for constraining the systematic errors.

III. 1e1p EVENT SELECTION

The 1e1p event selection makes use of an additional DL algorithm that performs particle identification on the images. The algorithm is an extension of work reported in Ref. [4], where we described



	Proton	Electron	Gamma	Muon
MPID Score	0.51	0.64	0.71	0.12

FIG. 5: MPID example of NC event with a table output of particle probabilities. Scores indicate higher probabilities of having proton and gamma in the image. The images applied to MPID are 512 × 512 pixels. Example event is generated by MicroBooNE beam simulation.

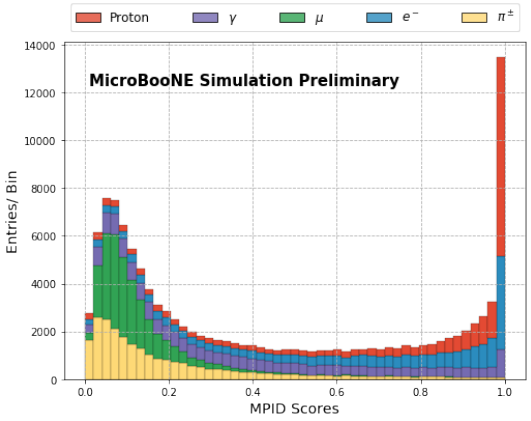


FIG. 6: The MPID scores for simulated $1e1p$ events. A probability is assigned separately for protons, electrons, photons ("gammas"), muons and charged pions to be in the image. The image applied to PID is 512 × 512.

particle ID on images containing only a single particle with the convolutional neural network (CNN). For this analysis, we must perform particle ID on images containing multiple particles. This Multi-particle ID (MPID) algorithm applies the sigmoid function in the output layer, allowing a particle-wise prediction. MPID classifies the probability of finding particles of muon, charged pion, proton, electron or gamma in the image. Each is considered individually, and assigned a probability between 0 and 1. Charged pions are included in MPID training sample to increase traing sample varieties. The output for charged pions is not used in the seleletion.

Training sample for MPID is generated with a "particle gun" simulation inside images are 512×512 pixels. In 80% of the training sample, random number between 1 to 4 of the five particle types are concatenated at the particle starting points. Their energy and momenta are distributed uniformly between 100 MeV and 1 GeV except the protons are generated uniformly between 100 MeV and 400 MeV. The other 20% of training sample is generated with a lower momentum distribution between 30 MeV and 100 MeV except for proton between 40 MeV and 100 MeV.

Example of the MPID network on simulated $1e1p$ event is shown in Fig. 5. The performance of the MPID network on simulated $1e1p$ events is shown in Fig. 6. One can see that the MPID score is highly peaked at ~ 1 for the proton and the electron probabities. There is some small confusion of electrons with photons. We provide comparisons of the MPID result with data in the Appendix to this public note.

The selection comprises a series of cuts followed by the application of a final cut on a score produced by a Boosted Decision Tree which has been trained using the Monte Carlo simulated samples.

The initial cuts isolate a well reconstructed topology that is generally consistent with $1e1p$. The fiducial volume requirements are: 10 cm$< x <$ 246.25 cm; -116.5 cm$< y <$ 116.5 cm; and 10 cm$< z <$ 700 cm or 740 cm$< z <$1026.8 cm. The reconstruction must identify two particles, each with length >5 cm, and both of which must remain $>$ 15 cm from the edge of the active volume of the detector at all points. One particle must correspond to a 3-D shower object.

A pair of cuts remove neutrino events with $\pi 0$ events from the sample. These events, arise mainly from excitation of a Δ. In the Neutral Current (NC) case, especially, the subsequent decays to a π^0 and p can fake a $1e1p$ topology if one of the photons converts to an e^+e^- pair within 3 mm of the interaction. This produces a converted photon that appears to be connected to the proton with no gap between the two particles, because the wire spacing is 3 mm. Many of these events have a second, nearby electromagnetic shower from the second photon from the π^0. The DL SSNet has provided information on if pixels which are nearby, but detached from, the vertex are shower like. If a shower like detached cluster is found, we then check that this cluster is sufficiently large (more than 15 pixels) and that it is not too close to the electron induced shower (opening angle between the two more than 30^o. These requirements prevent a high false positive rate. If such a cluster is

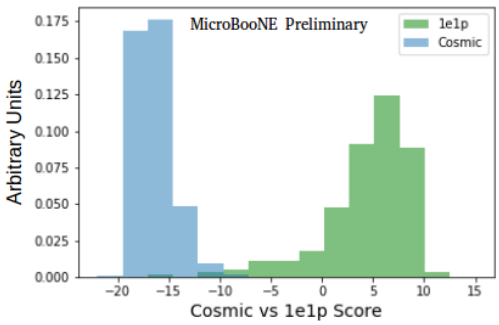


FIG. 7: BDT output scores for ExtBnb cosmic sample vs MC 1e1p signal sample.

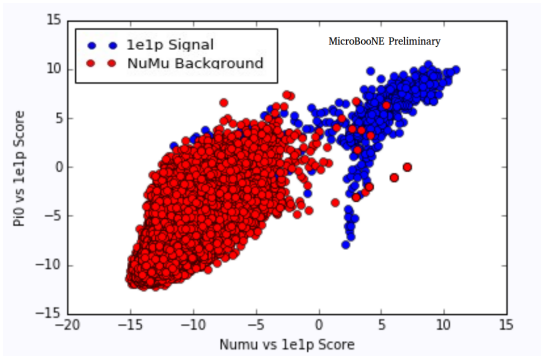


FIG. 8: Simulated events in the plane of scores from the ν_μ and π^0 BDTs. Red: ν_μ events (including NCπ0).

identified, the event is rejected. The second cut uses the PID information from the MPID. The MPID neural net provides particle scores both for attached particles and for the entire region around the vertex. The score from the full surrounding region is compared to the score associated with attached particles only. If the score for γ-like particles Increases by more than 50% when the surrounding region is included, the event is likely to contain a detached photon and is rejected.

At this point, three significant backgrounds remain. The first is events reconstructed on cosmic rays. The second is ν_μ charged-current interactions, where the muon decays to a michel electron, developing an electromagnetic attachment, or where the low energy muon scattered at a level that it appears similar to a low-showering electron. The third is from neutrino events producing π^0's that were not removed by the cuts above.

To address this, we introduce three Boosted Decision Trees (BDTs) that are trained to separate the $1e1p$ signal from each background. In the case of the cosmic events, we use data taken out of beam (EXTBNB) for the training. In the latter two cases, we use simulated data. For the ν_μ we use all ν_μ interactions including those that include π^0's. However, this sample is dominated by $1\mu1p$ where the muon fakes a $1e1p$ signal. In the π^0 sample, the events used for training must contain a π^0 at truth level. The samples used for training did not overlap with the ν_μ interaction training sample.

All three BDTs use the same variables as input, and only differ in the training samples used. The variables used to discriminate are: p_T^{event}/p^{event}; Bjorken x; the difference in average dQ/dx of each track normalized to the sum of average dQ/dx of each track, defined as:

$$\eta = \frac{|(dQ/dx)_1 - ((dQ/dx)_2|}{((dQ/dx)_1 + ((dQ/dx)_2)} \quad (4)$$

(where normalizing to the sum removes detector dependences); the track angles with respect to the z-axis, the track angles around the z-axis; the particle scores from the MPID; and a topology score looking for a branching along the path of the particle. We cut directly on the score the differentiates cosmics from $1e1p$ at score=0, as shown in Fig. 7. This produces better than 99.999% rejection of cosmic events. Then, we plot the remaining events in the plane of the ν_μ and NCπ^0 BDT scores, as shown in Fig. 8. For this analysis, we cut at score=0 in both variables, retaining the events in the upper right of the plot. The cut will be optimized in the future. The agreement between data and the simulation used to train the BDTs is illustrated in the appendix for each variable used in the training.

IV. DISCUSSION OF THE $1e1p$ RESULT AND THE METHOD OF $1\mu1p$ CONSTRAINT

Fig. 9 shows the $1e1p$ events as a function of E_ν^{range} (eq 1). The green entries are the intrinsic ν_e which represent an irreducible background to the LEE signal, but also illustrate the ability of MicroBooNE and this analysis to identify significant quantities of ν_e like events. At this point, cosmic and other backgrounds are expected to be sub-dominant to the intrinsic electron neutrino contribution. Further studies to fully assess these backgrounds and associated uncertainties are underway. The error bars are statistical only.

The gray hatched region in Fig. 9 indicate the absolute systematic error, before and after $1\mu1p$ constraints, from the flux and GENIE cross section parameters. We use the same systematic variations as other analyses and the experiment and the details of the models appear in Ref. [10, 11]. The cross section uncertainties shown here are derived from reweight-

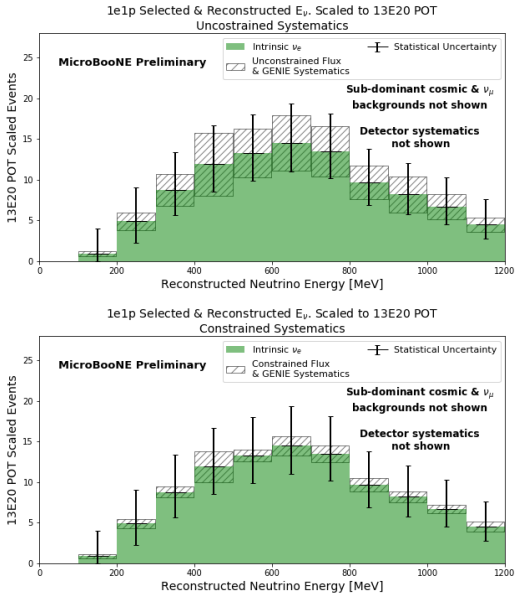


FIG. 9: *Top*: Selected and reconstructed ν_e spectrum with statistical and *unconstrained* flux GENIE cross section systematics illustrated. Backgrounds are not illustrated, but are subdominant to the ν_e contribution. *Bottom*: Selected and reconstructed ν_e spectrum with statistical and ν_μ *constrained* systematics.

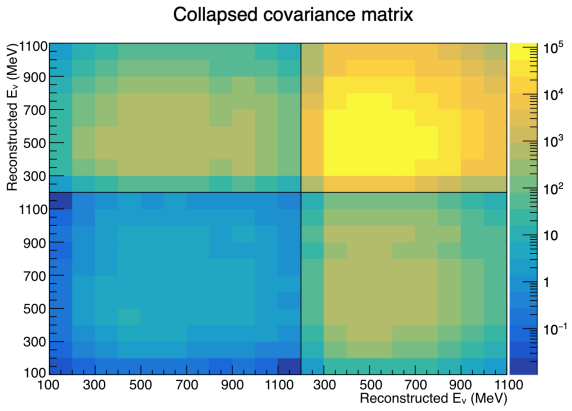


FIG. 10: Covariance matrix for the flux and GENIE cross section systematic uncertainties on the $1e1p$ and $1\mu1p$ events. The first 11 bins correspond to the $1e1p$ events from 100 MeV to 1200 MeV in 100 MeV increments; the remaining 9 bins correspond to the $1\mu1p$ events from 200 MeV to 1100 MeV in 100 MeV increments. Note the logarithmic scale on the z-axis.

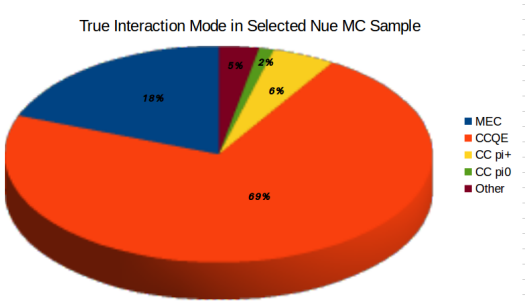


FIG. 11: Breakdown of interaction types for the simulated LEE sample.

ing of GENIE parameters and therefore do not include systematic effects associated with the random phase approximation (RPA) effect and with MEC interactions. These two systematic effects will be quantified in a future version of the analysis. The resulting covariance matrix is shown in Figure 10. The lower left block shows the ν_e events as in Fig. 9 and the upper right block shows the ν_μ events as in Fig. 13. The off-diagonal blocks show the correlations between the ν_e and ν_μ. The relatively large values in the off-diagonal blocks indicate that there are substantial correlations, which we leverage when we use the $1\mu1p$ events to constrain the systematic uncertainties on our $1e1p$ prediction.

The breakdown of interaction modes that are selected by the $1e1p$ requirements described above are shown in Fig. 11. The events are majority CCQE, followed by MEC. Together these comprise over 90% of the total. Although the remainder, which are from CCπ^+, CCπ^0, and other minor modes, appears to be a small background, they will be important to the analysis.

The gray, hatched region shows the absolute systematic uncertainty. Because the $1\mu1p$ sample is affected by many of the same systematic effects as the $1e1p$ sample, the relatively high-statistics measurement of the $1\mu1p$ constrains these uncertainties. We are re-using a technique was first applied first in the MiniBooNE analysis [14]. The $1\mu1p$ sample is able to constrain the flux kinematics because the ν_μ and the majority of the intrinsic ν_e come from the same same originating pion:

$$\pi^+ \rightarrow \mu^+ + \nu_\mu \quad (5)$$
$$\hookrightarrow e^+ + \bar{\nu}_\mu + \nu_e \quad (6)$$

Because the initial decay is two-body, the muon neutrino energy is directly proportional to the pion energy: $E_{\nu_\mu} = 0.43E_\pi$. Thus the pion energy spectrum is well-constrained by the ν_μ energy spectrum. This, in turn, constrains the ν_e from muon decay. This is described in detail in Ref. [15]. The cross section uncertainties are also constrained by the high-statistics $1\mu1p$ measurement.

The constraint is introduced in a combined fit to the $1e1p$ and $1\mu1p$ data, which incorporates the cor-

relations between them encoded by the systematic covariance matrix. The constrained systematic errors shown in Figure 9 are obtained in the following way. We perform a simple fit to the number of events in each of the $1e1p$ and $1\mu1p$ bins. The fit minimizes a χ^2 that has two terms: one that enforces agreement of the $1e1p$ and $1\mu1p$ bins with the simulated prediction within systematic errors, and one that enforces agreement of the $1\mu1p$ bins with the $1\mu1p$ data within statistical errors. Given that the available 5E19 data sample is fairly small, we substitute in the simulated $1\mu1p$ prediction scaled to expected final data sample size of 13×10^20 POT for the $1\mu1p$ data. The constrained uncertainties are extracted from the effective covariance matrix associated with the χ^2 used in the fit.

V. $1\mu1p$ EVENT SELECTION AND RESULTS

The $1\mu1p$ selection has been presented in Public Note 1051 [9]. Therefore, we only briefly describe the selection here.

This analysis begins with a set of cuts that follow the 3-D Reconstruction. These additional cuts require: energy consistency (range-based and CCQE hypothesis must agree within 1 GeV); well reconstructed tracks (ad defined by 3D reconstruction or by an additional energy consistency requirement of agreement with 30%; transverse momentum consistent with a planar event; four-momentum transfer, $Q^2 = 2E_\nu(E_\mu - p_\mu^L) - m^2 > 0$; no detached shower clusters (determined in the same manner of $1e1p$); and SSNet shower fraction $< 50\%$.

These initial cuts were designed to remove the cosmic rays while retaining $> 95\%$ efficiency. Nevertheless, by far the largest background to the contained $1\mu1p$ interactions still comes from cosmic rays after these cuts. Cosmic rays mimic neutrino interactions because of scattering, dead wire regions, distortion due to space charge, and other detector effects. To address these, we develop a cosmic discrimination likelihood which is designed to separate signal from background. This likelihood was developed by comparing $1\mu1p$ CCQE simulated events to cosmic events taken off-beam, called EXTBNB.

The second most important remaining source of background is neutrino events that are misreconstructed as contained $1\mu1p$. To address this, we have developed a neutrino discrimination likelihood. This likelihood was developed by comparing $1\mu1p$ CCQE simulated events to all other simulated ν_μ events.

Both likelihoods rely on the same variables. These are: the difference between the average dQ/dx for each track divided by the sum, η, discussed in Sec. III; the 3-D opening angle between the tracks; the sum of the θ angles; the difference between the ϕ angles; and p_T, α_T and Φ_T, as defined in Sec. III. In the Appendix to this note, we show the distributions for the signal and background for each of these inputs.

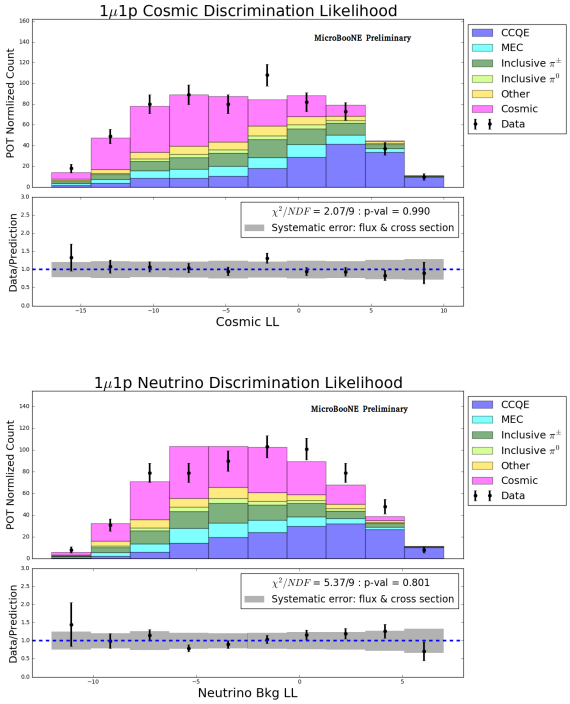


FIG. 12: Top: Cosmic Discrimination Likelihood data to simulation comparison. Contributing neutrino interaction sources and the cosmic ray background are indicated. The gray hatched region indicates the flux and GENIE cross-section systematic error. The red line is the cut used in this analysis. Bottom: Neutrino Discrimination Likelihood data to Simulation Comparison.

Fig. 12 compares the two likelihood distributions between the 5E19 data set and the prediction. The neutrino interaction contributions to the prediction are noted in the stacked plot, as well as the cosmic ray contribution. The cuts applied to the likelihoods are shown by the red line. In this public note, we extend beyond Public Note 1051 [9] by presenting absolutely normalized distributions with systematic uncertainties from flux and GENIE cross section parameters, shown in gray (hatched) for the first time. Detector systematic uncertainties and additional cross section systematic uncertainties are not yet included.

A template fit is further performed on the cosmic likelihood distribution. The normalization of the extbnb relative to the neutrino contribution is

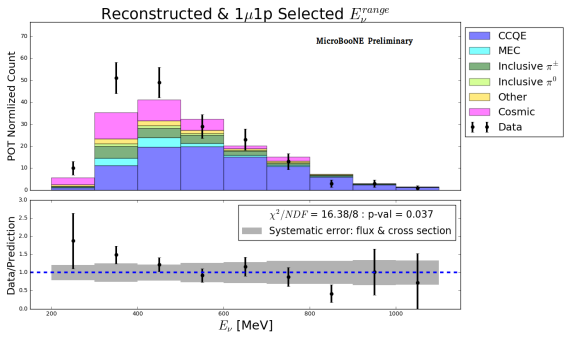


FIG. 13: Distribution of events versus neutrino energy determined by the range of the muon and proton tracks, E_ν^{range}.

permitted to float and a minimum chi2 sought. This optimum is located where the relative contribution of extbnb vs neutrino content is 8% larger. This adjustment to the normalization is fixed and propagated to all subsequent plots generated after the likelihood cuts.

A collection of comparisons of data to Monte Carlo simulation are shown in the Appendix to this paper. Good agreement is seen for all distributions. The most important distribution is the one which will be used to constrain the errors of the $1e1p$ result: E_ν^{range}, which is shown in Fig. 22(top). Although the ratio of data to prediction shows a slope, this is within the flux and cross section uncertainties.

VI. DISCUSSION OF CONSTRAINED $1e1p$ RESULTS AND PLANS

After applying the $1\mu1p$ constraint, as discussed above, we obtain significantly reduced systematic errors as illustrated by Fig. 9. However, one notes that the strength of this method is limited for interactions which are not CCQE and MEC events because of lower statistics. For the analysis at this time, these backgrounds are brought in with the uncertainty on the absolute predicted rate. This can be addressed in the future, as we discuss below.

We plan a number of improvements to the analysis in the future. These goals include:

- Reducing the systematic uncertainty on the ν_μ background events through a measurement of the rate in the MicroBooNE detector.
- Improving the energy resolution for the ν_e events.
- Determining systematic errors from the detector.
- Converting to use of the cutting-edge simulation and signal processing, which is about to become available, called MCC9

[1] H. Chen et al., MicroBooNE Collaboration, *Proposal for a New Experiment Using the Booster and NuMI Neutrino Beamlines: MicroBooNE*, FERMILAB-PROPOSAL-0974

[2] C. Adams et al., MicroBooNE Collaboration, *Design and construction of the MicroBooNE Cosmic Ray Tagger system*, JINST 14, P04004 (2019)

[3] A. Ereditato, J. Conrad and M.Shaevitz, *The State of the Art of Neutrino Physics*, Chapter 10 : Sterile Neutrinos: An Introduction to Experiments, `https://doi.org/10.1142/10600`

[4] R. Acciarri et al., MicroBooNE Collaboration, *Convolutional neural networks applied to neutrino events in a liquid argon time projection chamber JINST*, `P03011(2017)`.

[5] A. A. Aguilar-Arevalo et al., MiniBooNE Collaboration, *Improved Search for* $\bar{\nu}_\mu \longrightarrow \bar{\nu}_e$ Oscillations in the MiniBooNE Experiment,Phys. Rev. Lett. 110, 161801

[6] A. A. Aguilar-Arevalo et al., MiniBooNE Collaboration, *Significant Excess of Electronlike Events in the MiniBooNE Short-Baseline Neutrino Experiment*, Phys. Rev. Lett.,**121**,221801,2018

[7] MicroBooNE Collaboration, *First Deep Learning based Event Reconstruction for Low-Energy Excess Searches with MicroBooNE*, MICROBOONE-NOTE-1042-PUB, `https://microboone.fnal.gov/wp-content/uploads/MICROBOONE-NOTE-1042-PUB.pdf`

[8] C. Adams et al., *A Deep Neural Network for Pixel-Level Electromagnetic Particle Identification in the MicroBooNE Liquid Argon Time Projection Chamber*, arXiv:1808.07269

[9] MicroBooNE Collaboration, *Selection of* ν_μ *Events for the MicroBooNE Deep Learning Low Energy Excess Analysis*, MICROBOONE-NOTE-1051-PUB, `https://microboone.fnal.gov/wp-content/uploads/MICROBOONE-NOTE-1051-PUB.pdf`

[10] A.A. Aguilar-Arevalo et al., MiniBooNE Collaboration, *The Neutrino Flux prediction at MiniBooNE*, Phys.Rev. D79 (2009) 072002

[11] MicroBooNE Collaboration, *First Muon-Neutrino Charged-Current Inclusive Differential Cross Section Measurement for MicroBooNE Run 1 Data*, MICROBOONE-NOTE-1045-PUB, `https://microboone.fnal.gov/wp-content/uploads/MICROBOONE-NOTE-1045-PUB.pdf`

[12] C. Andreopoulos et al.,*The GENIE Neutrino Monte Carlo Generator*, Nucl. Instrum. Meth., vol. A614, pp. 87104, 2010.

[13] MicroBooNE Collaboration, *MicroBooNE low-energy excess signal prediction from unfolding MiniBooNE Monte-Carlo and data*, MICROBOONE-NOTE-1043-PUB, `https://microboone.fnal.gov/wp-content/uploads/MICROBOONE-NOTE-1043-PUB.pdf`

[14] A. A. Aguilar-Arevalo et al., MiniBooNE Collaboration, *Unexplained Excess of Electronlike Events from a 1-GeV Neutrino Beam*, Phys. Rev. Lett. 102, 101802

[15] J. R. Monroe. *A Combined* ν_μ *and* ν_e *Oscillation Search at MiniBooNE*, Phys. Rev. Lett. 102, 101802

Appendix A: Additional MPID plots

The following additional plots are used to characterize the capability of the MPID algorithm in talks.

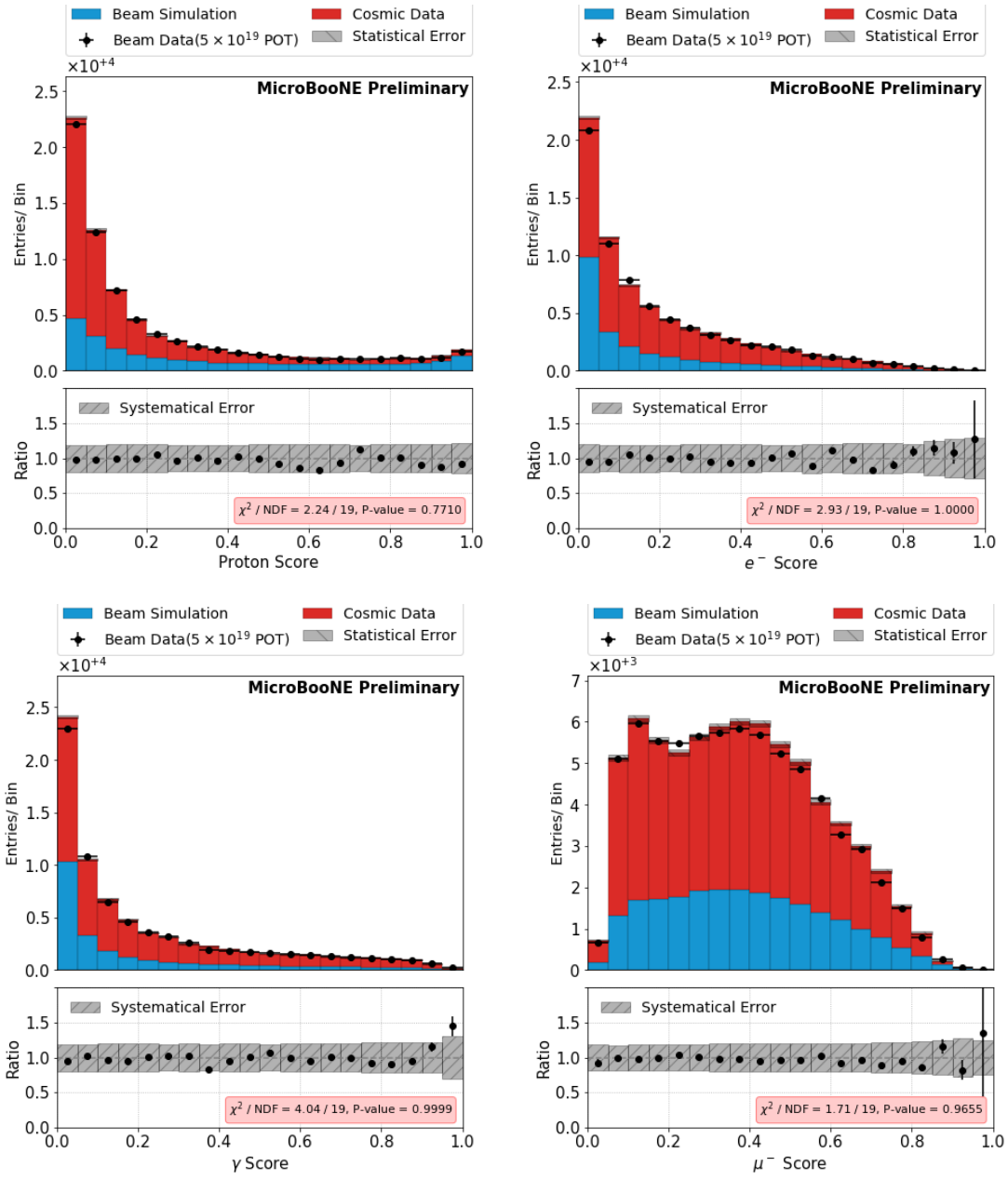


FIG. 14: Plots showing the data-to-Monte Carlo simulation agreement for the MPID algorithm. Top to bottom, left to right: proton score, electron score, gamma score, muon score. Data points are from the 5E19 open data set, for contained 2-particle events. The stacked plot is: beam simulation in blue, misreconstricted cosmic background in red. of These are the inputs to the $1e1p$ selection. Error bars in bin plots are statistical uncertainties. Error bars in ratio plots are systematic uncertainties.

Appendix B: Additional $1\mu1p$ plots

1. Likelihood Inputs

The following additional plots compare signal to background for the inputs to the $1\mu1p$ log likelihoods. Left are distributions of inputs to the cosmic log likelihood, with simulated signal (CCQE $1\mu1p$ events at truth level) in blue and cosmic data (EXTBNB) in green. Right are distributions of inputs to the neutrino background log likelihood, with simulated signal in blue and simulated ν_μ events that are reconstructed as two-particle, but are not CCQE $1\mu1p$, in red.

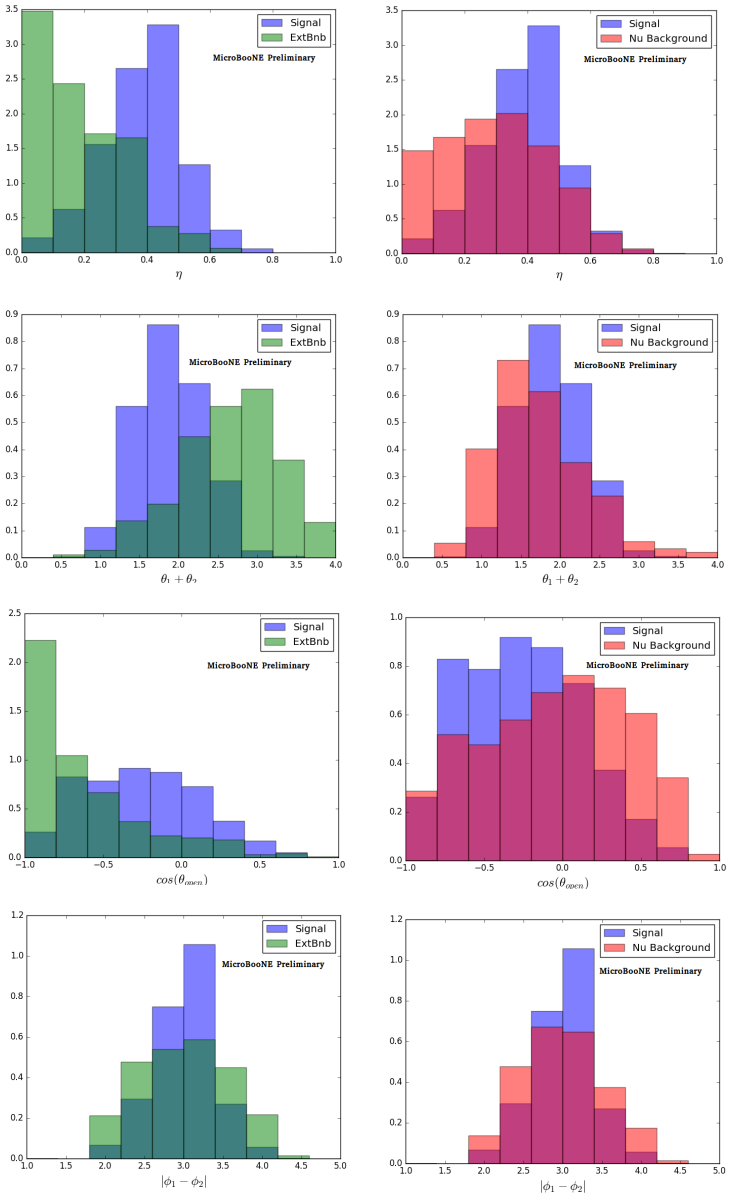


FIG. 15: (Top) η–dQ/dx difference normalized to sum; (Upper Middle) cosine of the 3D opening angle; (Lower Middle) Sum of two track θs; (Bottom) absolute difference between two track ϕs.

2. Data-to-Monte Carlo simulation comparisons of the selected $1\mu1p$ events

Unit normalized data-to-Monte Carlo comparisons of the selected $1\mu1p$ events has already been presented in Public Note 1051 [9], and the reader is referred to that document for more details. The plots shown here update this information to provide absolutely normalized comparisons. Systematic errors for flux and GENIE cross section parameters are also added to these plots in gray (hatched). The stacked plots show the contributing interactions.

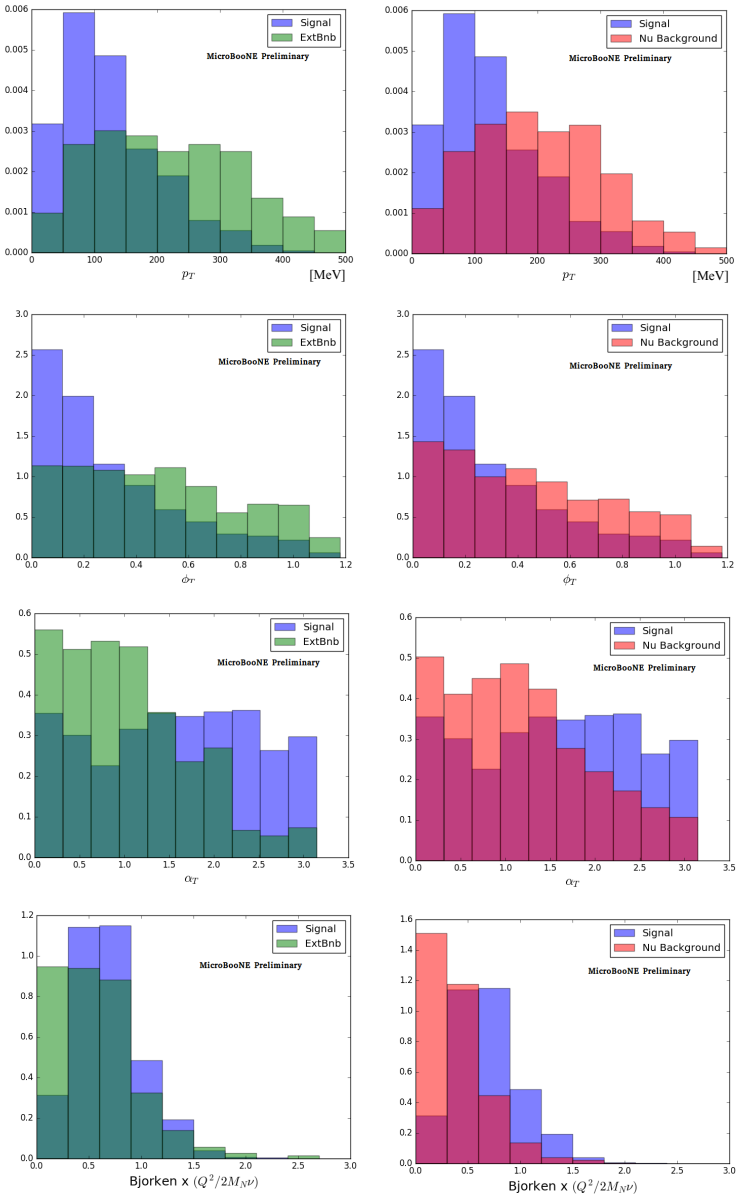


FIG. 16: (Top) p_T – 3 momentum imbalance in transverse plane; (Upper Middle) ϕ_T– Angular discrepancy from back to back expectation; α_T– Transverse momentum imbalance relative to the muon axis; x-Bjorken.

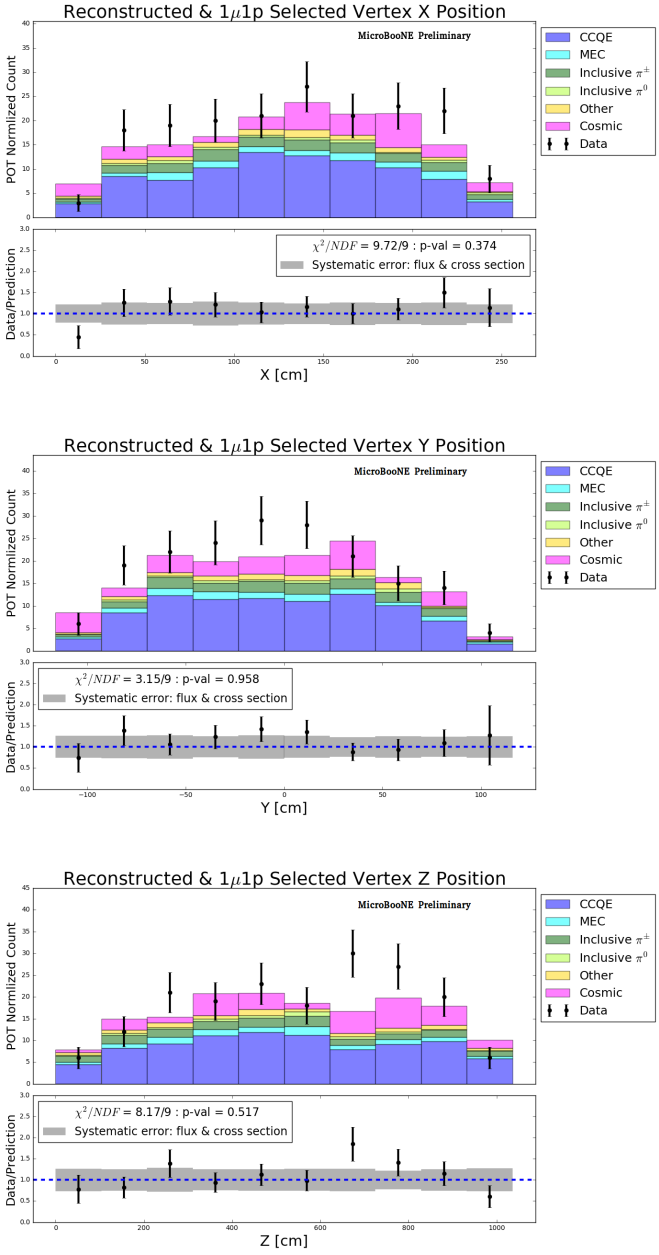
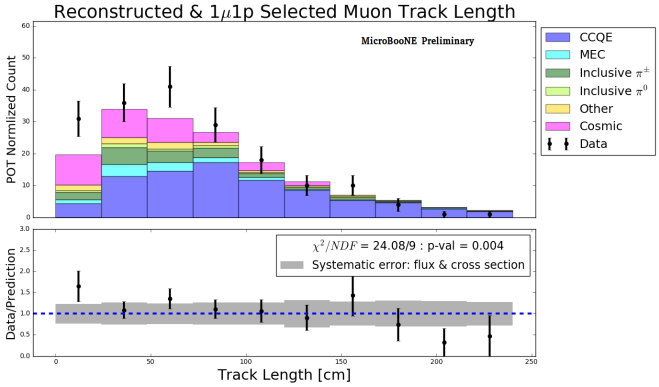
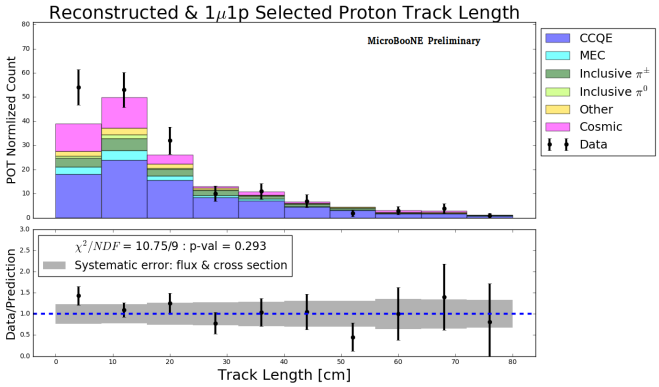
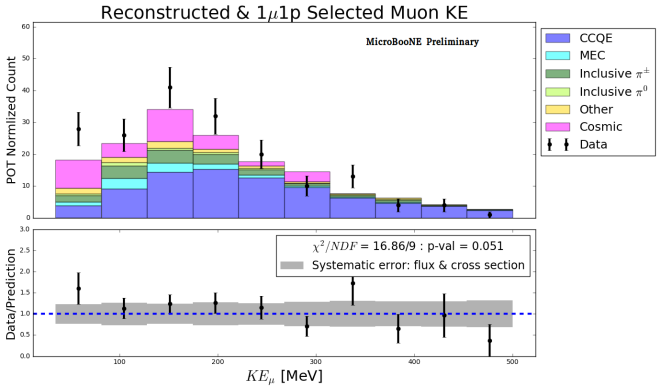


FIG. 17: Vertex distribution comparisons.







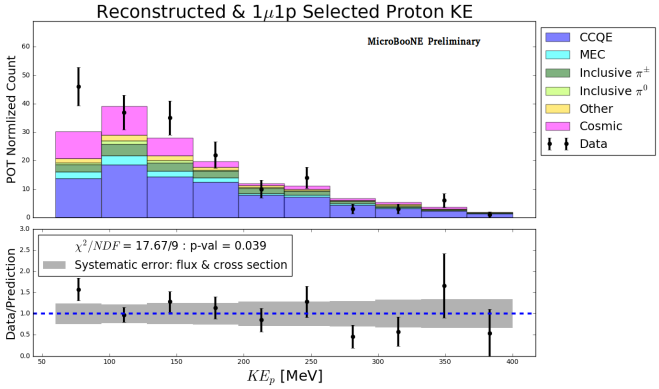


FIG. 18: Energy and length comparisons.

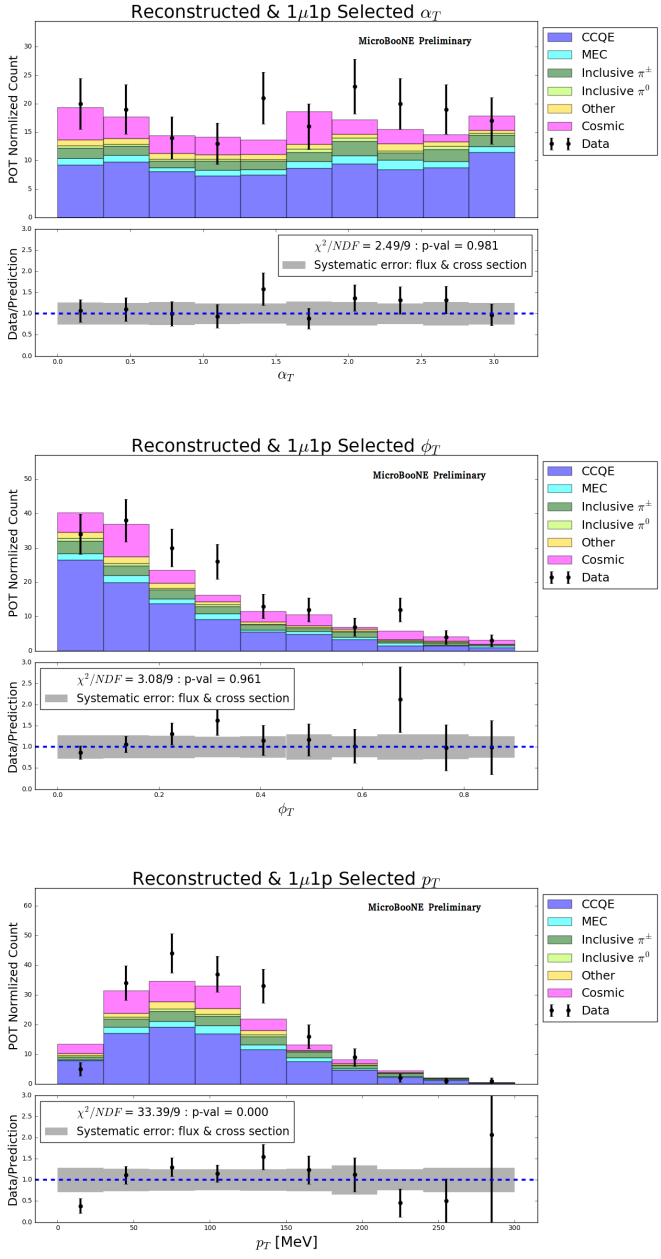


FIG. 19: Transverse variable comparisons.

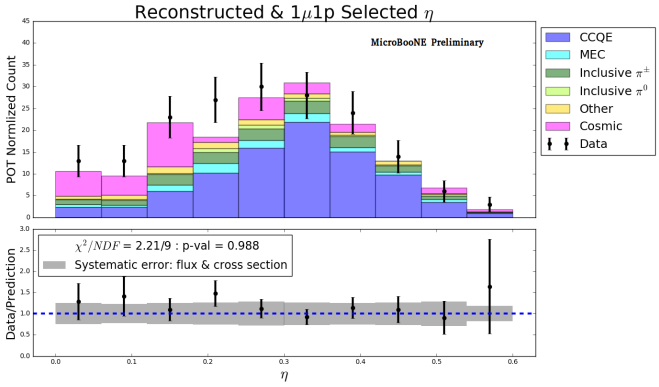


FIG. 20: η variable comparison.

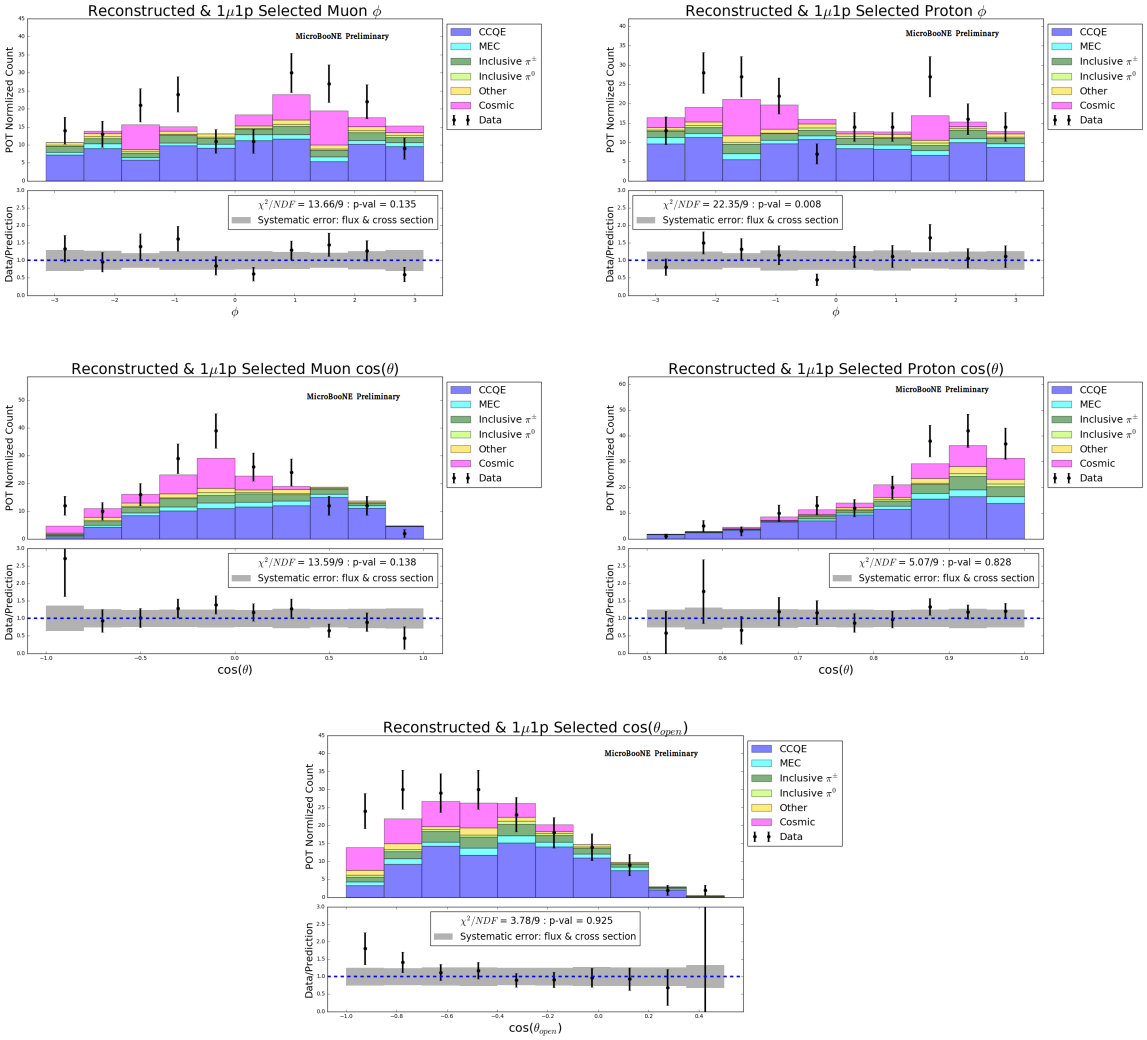


FIG. 21: Reconstructed track angles comparisons.

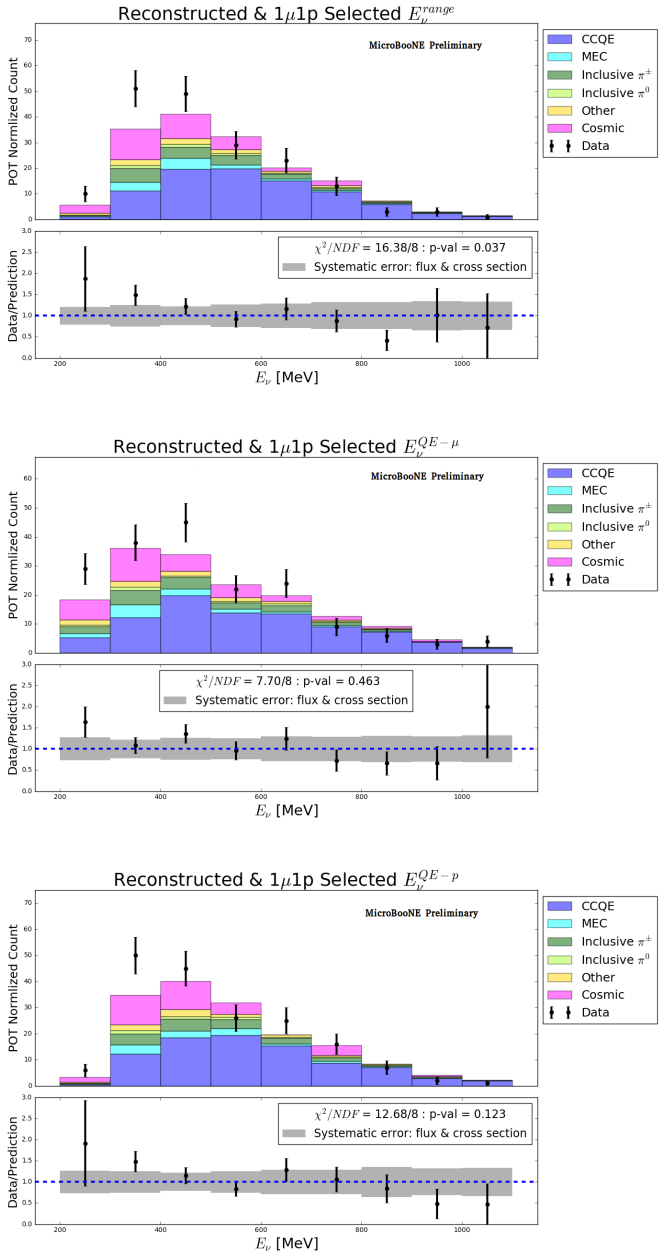


FIG. 22: Reconstructed energies. Top : range-based energy (Eq.1), bottom left : neutrino energy estimated from the proton kinematics only, assuming a CCQE interaction (Eq.2), and bottom right : neutrino energy estimated from the muon kinematics only, assuming a CCQE interaction (Eq.3).

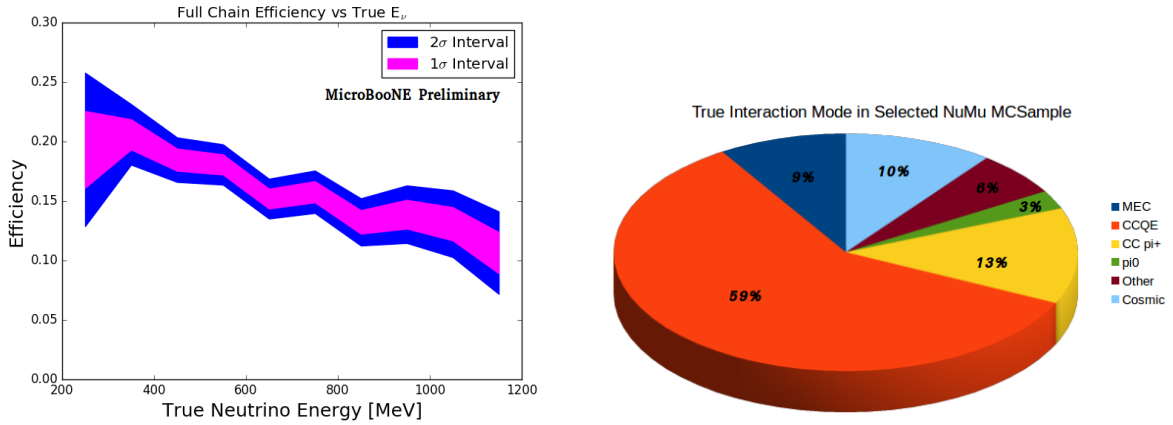


FIG. 23: Left : Full chain efficiency of the ν_μ candidate selection from 1μ1p contained events. Right : distribution of the components of the ν_μ candidates.

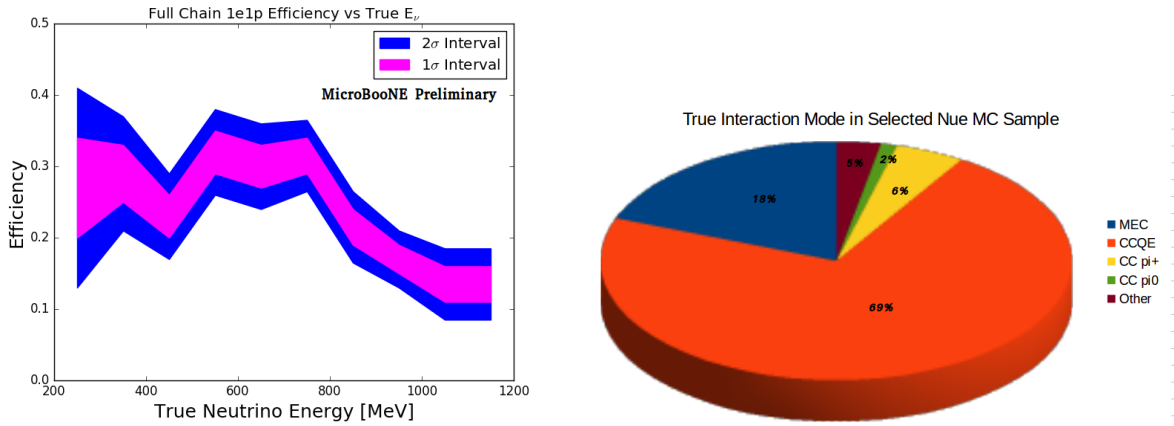
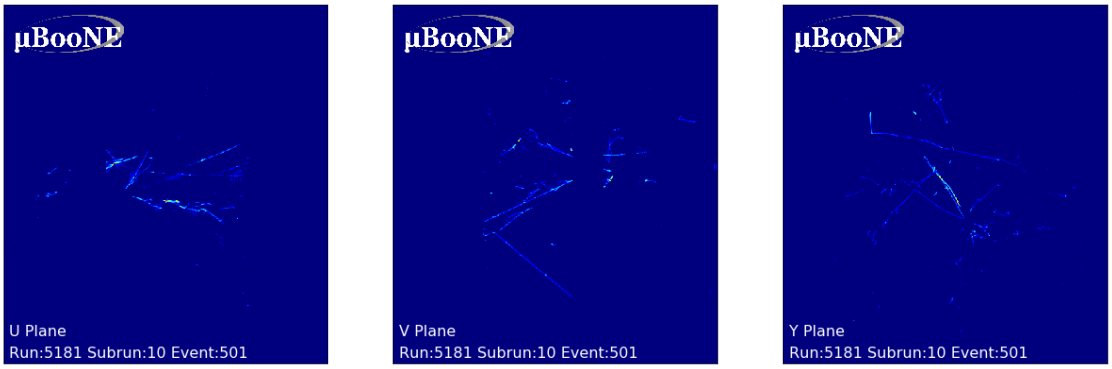


FIG. 24: Left : Full chain efficiency of the ν_e candidate selection from 1e1p contained events. Right : distribution of the components of the ν_e candidates.

3. Event displays of three selected $1e1p$ candidates

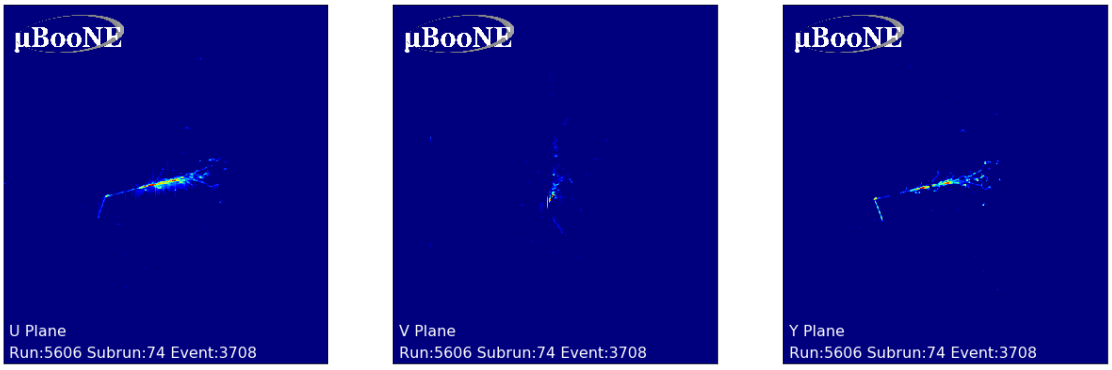


	Proton	Electron	Gamma	Muon
MPID Score	0.25	0.35	0.97	0.58

FIG. 25: Left to right, first selected $1e1p$ candidate on U,V and Y readout wire planes of MicroBooNE. MPID score is calculated based on image of plane Y. (run 5181, subrun 10, event 501)

	Proton	Electron	Gamma	Muon
MPID Score	0.59	0.85	0.45	0.14

FIG. 26: Left to right, second selected $1e1p$ candidate on U,V and Y readout wire planes of MicroBooNE. MPID score is calculated based on image of plane Y. (run 5328, subrun 30, event 1515)



	Proton	Electron	Gamma	Muon
MPID Score	0.29	0.96	0.27	0.18

FIG. 27: Left to right, third selected $1e1p$ candidate on U,V and Y readout wire planes of MicroBooNE. MPID score is calculated based on image of plane Y. (run 5606, subrun 74, event 3708)

Appendix C: Additional $1e1p$ plots

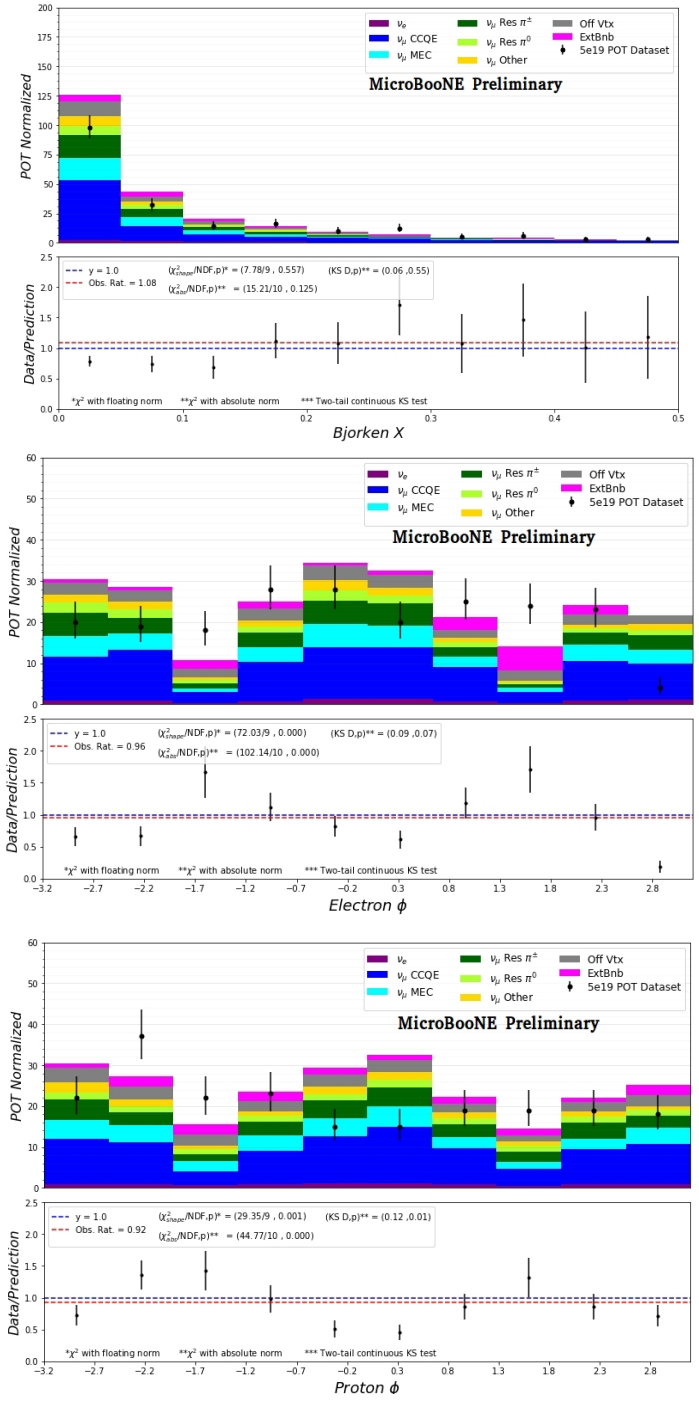


FIG. 28: Candidate proton and electron distributions. Part of the BDT inputs.

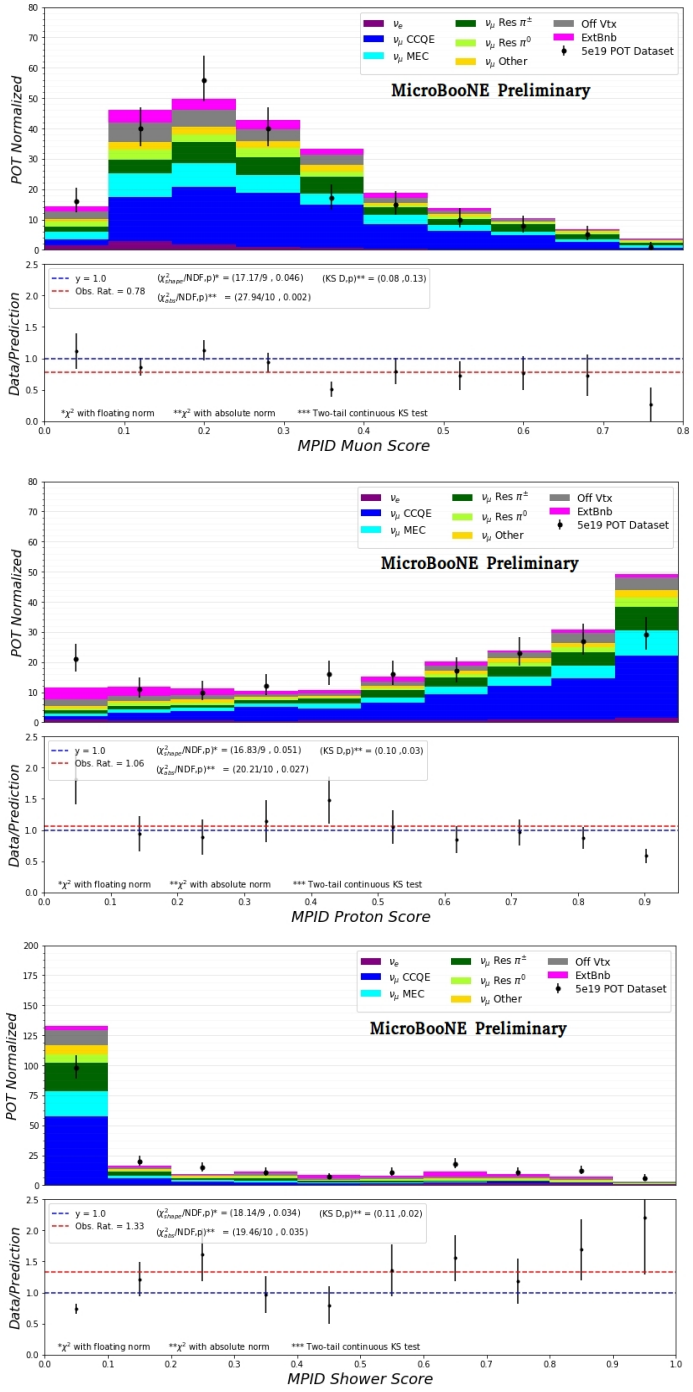


FIG. 29: MPID scores used in the 1e1p BDT

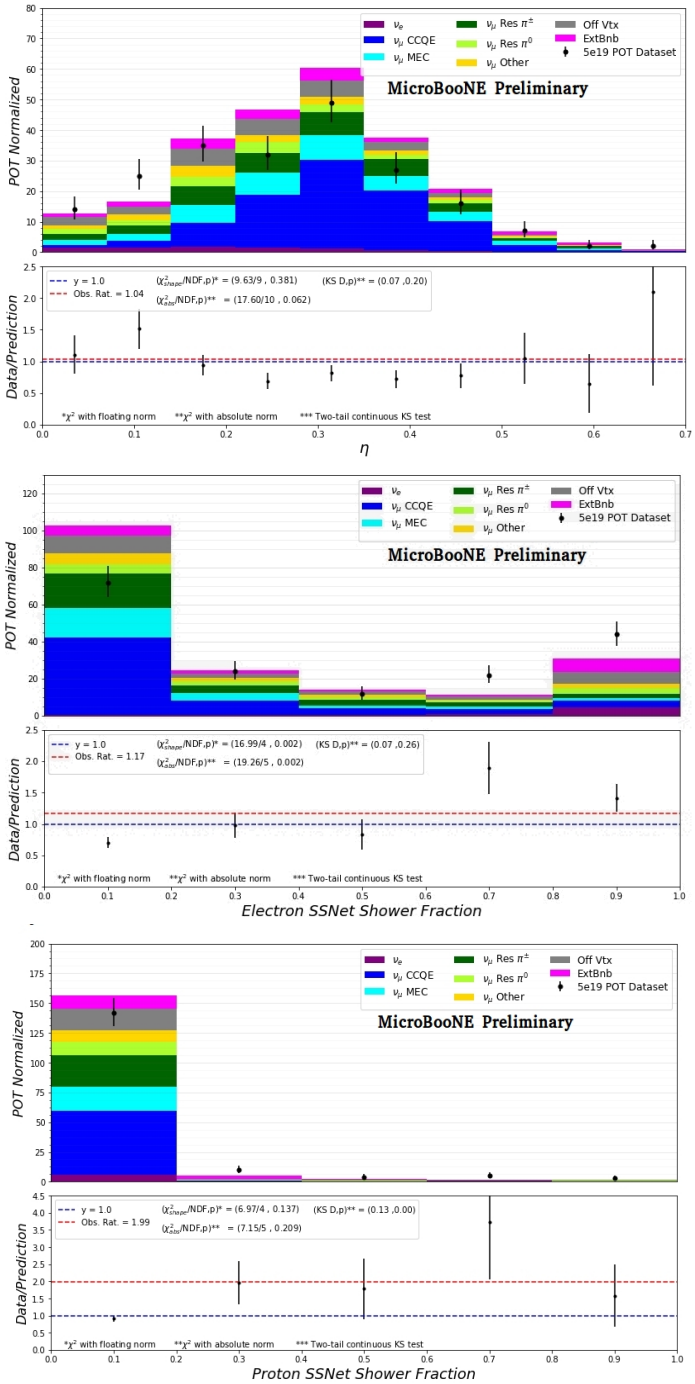


FIG. 30: Some additional 1e1p BDT input variables.

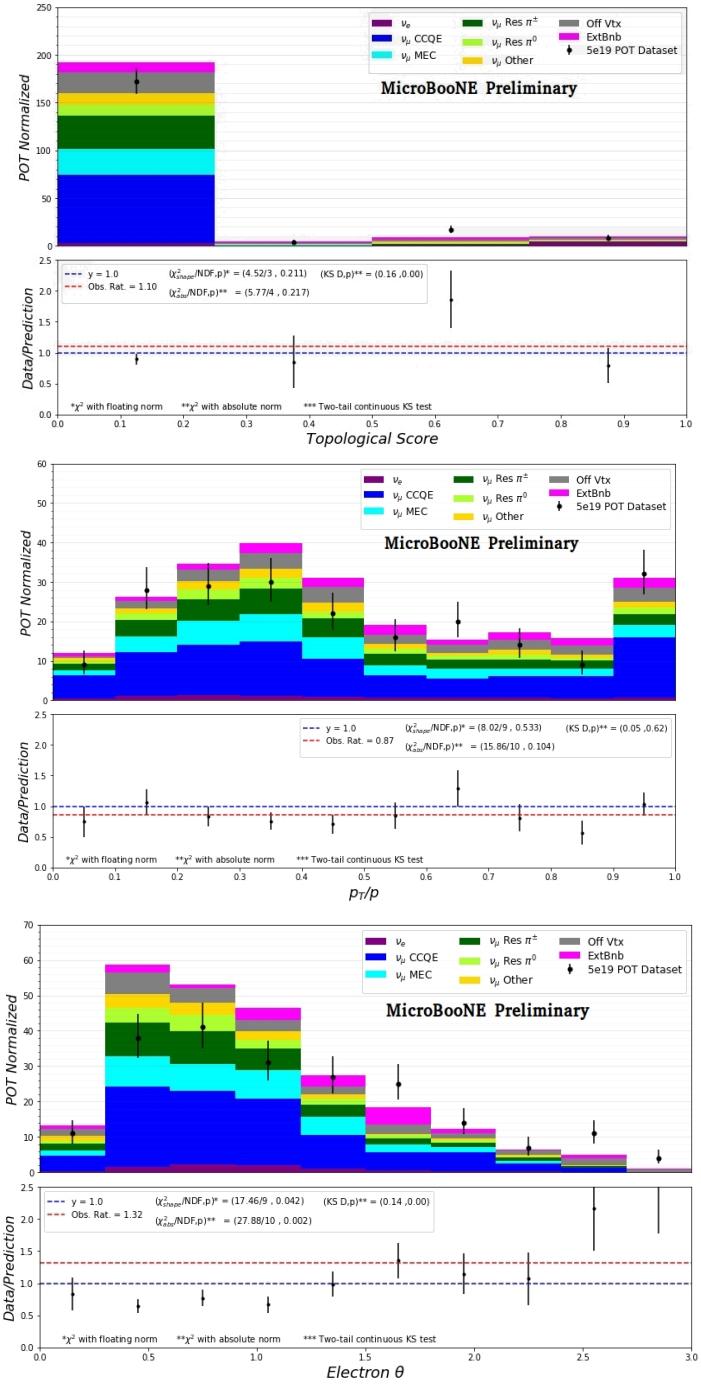


FIG. 31: Some additional 1e1p BDT input variables.