Search for Electron Neutrinos in Multiple Topologies with the MicroBooNE Experiment

MICROBOONE-NOTE-1085-PUB

The MicroBooNE Collaboration

June 23, 2020

Abstract

This note presents the status of the measurement of electron neutrinos from the Fermilab Booster Neutrino Beamline (BNB) with the MicroBooNE experiment. The analysis is aimed at investigating the nature of the low energy excess of electromagnetic activity observed by the MiniBooNE experiment. The ν_e event selection relies on topological and calorimetric information to characterise particles produced in these interactions, leveraging the Pandora multi-algorithm reconstruction framework as well as custom particle identification and pattern recognition tools. Results presented in this note use 5.88×10^{20} protons on target of data collected between 2015 and 2018. These include the performance of particle identification tools for μ/p and e/γ separation, along with electromagnetic shower calibration with the largest sample of π^0 events measured on argon. Progress towards the completion of this analysis is shown through the measurement of high-energy charged-current ν_e interactions and their kinematic distributions using both inclusive ($1e$X) and exclusive ($1e0p0\pi$ and $1e$N$p0\pi$) channels. Measurements of charged-current ν_μ interactions aimed at constraining flux and cross-section systematic uncertainties are also shown. We present the analysis' preliminary sensitivity of an electron-like signal hypothesis to the MiniBooNE excess which includes flux, cross-section, and detector systematic uncertainties.

Contents

1 Introduction

This document presents the first measurement of electron neutrinos from the Fermilab Booster Neutrino Beamline with the MicroBooNE experiment, and presents the status of the analysis aimed at investigating the low energy excess of electromagnetic activity observed by the MiniBooNE collaboration [1]. MicroBooNE [2] is a liquid argon time projection chamber (LArTPC) collecting data in the same neutrino beamline as MiniBooNE. LArTPC detectors provide detailed topological images of neutrino interactions and calorimetric information about the particles produced in these interactions. This analysis uses the unique capabilities of LArTPCs to perform a high purity measurement of electron neutrinos across multiple signal topologies and a broad range of energies. Background and modeling uncertainties are constrained using data-driven measurements of muon neutrino charged and neutral current interactions. These allow this analysis to test the electron neutrino hypothesis of the MiniBooNE low energy excess. This document presents a brief overview of the analysis, and results from data-sets and sidebands accessed so far in the analysis. The various appendices go into more detail on several key components to this analysis. Each appendix is associated to a specific poster being presented at the Neutrino 2020 conference.

Throughout this document we use $\sqrt{N}$ for the statistical error bar of the data points. plots which show POT-normalized data/MC comparisons show shaded error bands which include ν flux and cross-section uncertainties as well as statistical uncertainties on the MC sample statistics used in the analysis. Detector systematics, while not included in the data/MC ratio plots, are included in the sensitivity estimation reported. Finally, we note that the assessment of systematic uncertainties in this work is preliminary and not finalized.

2 Analysis Approach

Resolving whether the anomalous excess of electromagnetic showers observed by the MiniBooNE collaboration is caused by electrons or photons is the first necessary step in understanding its origin. The primary goal of this analysis is to measure ν_e interactions with the MicroBooNE detector in the same beamline, and thus contribute to testing the hypothesis of an excess of electron-like events as the origin of the MiniBooNE anomaly.

This analysis aims to make a high purity measurement of electron neutrinos across the full spectrum of neutrino energy at the BNB. This analysis does not rely on the kinematics of the ν_e interaction, and the modelling of final-state particle kinematics. Instead calorimetric and topological information from the final state particles are used. This means that if an anomalous signal is observed, the neutrino interaction kinematics may be used to further understand it. Several event topologies, both with and without visible hadronic activity, are selected. The predicted event distribution of events broken down by expected number of protons and pions is shown in Figure 1.
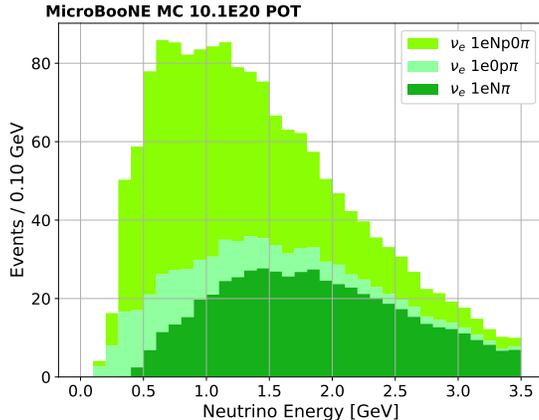
Figure 1: Expected rate of charged current electron neutrino interactions in the BNB broken into topologies with different numbers of protons and pions. Here $N > 0$. The Proton On Target (POT) to which this plot is normalized corresponds to the dataset collected by MicroBooNE in the first four years of running.

Different final-state topologies contribute to the total intrinsic ν_e spectrum. This analysis aims to measure the rate of pion-less ν_e events in order to match MiniBooNE's ν_e signal definition. Two exclusive selections are developed: one selecting events with no hadronic activity ($1e0p0\pi$, or zero-proton), and one with > 0 final state protons and no pions ($1eNp0\pi$, or N-proton, $N > 0$). The two mutually exclusive channels benefit from orthogonal selections which target topological features and backgrounds specific to each one in a tailored way. These two exclusive channels are then combined into a single measurement testing the sensitivity to an electron-like MiniBooNE anomaly. At the same time, the analysis also includes a fully inclusive measurement of ν_e interactions, which is able to more efficiently select ν_e events at higher energy and providing a high-statistics measurement of electron kinematics which provides additional validation to the analysis.

While the analysis presented here aims to perform a broad and kinematically unbiased measurement of ν_e events in the BNB, we rely on a specific model choice to test the analysis' sensitivity to an electron-like MiniBooNE excess. This model is described in ref. [3] and is obtained by unfolding the observed MiniBooNE low-energy excess through the MiniBooNE ν_e energy-response matrix from simulation, and applied to MicroBooNE's simulation as an energy-dependent scaling of BNB ν_e events.

3 Event Reconstruction and Analysis Tools

The analysis relies on a broad set of tools that aim to isolate neutrino interactions and subsequently ν_e events in the MicroBooNE data-set. It benefits from a rich reconstruction and calibration carried out with the MicroBooNE experiment. The reconstruction used as input to the analysis is performed with the Pandora pattern-recognition framework [4].

MicroBooNE's location on the surface means there is a large rate of cosmic-ray interactions which continuously occur in the TPC volume. Combined with the slow charge drift time in the

TPC, this leads to a $10^4 : 1$ cosmic-to-neutrino event rate, making cosmic-ray rejection a critical first step in the analysis. Triggering, together with a dedicated set of tools which aim to combine information from the TPC pattern recognition with timing and position information provided by scintillation light collected with MicroBooNE's PMT array allow for a factor of 10^4 cosmic-rejection. The efficiency in selecting neutrino candidates relative to this level of cosmic-rejection and for the various selections presented in this analysis is shown in Figure 2. These efficiencies are at or above 80% above 0.5 GeV, and decrease to 70% in the range 0.2-0.5 GeV.
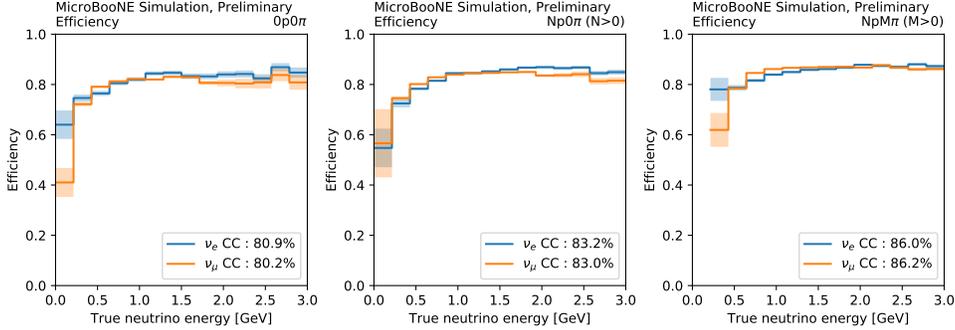

Figure 2: The performance of the cosmic removal tools as a function of the true neutrino energy for the topologies considered in the ν_e CC selection: no vertex activity (left), protons at vertex (middle) and pion production (right) channels. The energy range is $0\,\mathrm{GeV}$ to $3\,\mathrm{GeV}$ and the bin size is $200\,\mathrm{MeV}$. Shaded bands in the plots represent the statistical error on the measured efficiency. The legend shows the integrated efficiency of all BNB simulated ν_e CC and ν_μ CC events.

After having identified neutrino interaction candidates, reconstructing individual final state particles becomes crucial in order to determine the neutrino flavor (ν_e vs. ν_μ), energy, and event kinematics. Identifying electron neutrinos in the BNB beamline is made particularly challenging by the small ν_e beam component, sub-dominant to the ν_μ flux by more than a factor of 10^2 [5]. This analysis in particular, aiming to measure electron neutrino interactions at low energy, needs to overcome the challenges [6] associated to reconstructing $\mathcal{O}(100)$ MeV electromagnetic showers produced by sub-GeV ν_es. This analysis combines the output of Pandora's generic pattern-recognition tools with a robust calibration program [7, 8] and additional 3D particle tracking tools to provide the particle-identification needed to isolate electron neutrino events. High granularity spatial calorimetric information from all three TPC wire-planes provides the input necessary to perform the particle-identification necessary to the analysis. Figure 3a shows the measured energy loss as a function of residual range for a combination of muon and proton candidates in data. This information is employed in a likelihood-based particle ID algorithm focused on muon-proton separation. The performance of this algorithm is shown in Figure 3b, where one can see the clear separation of highly ionizing protons from lower $\mathrm{d}E/\mathrm{d}x$ muons or pions. In order to improve the electron identification, hits in the initial shower trunk are identified and fit with a track hypothesis; this technique allows for a more precise determination of the $\mathrm{d}E/\mathrm{d}x$ at the shower start. Further details on the particle identification (PID) tools developed for this analysis are presented in Appendices A and B.1.
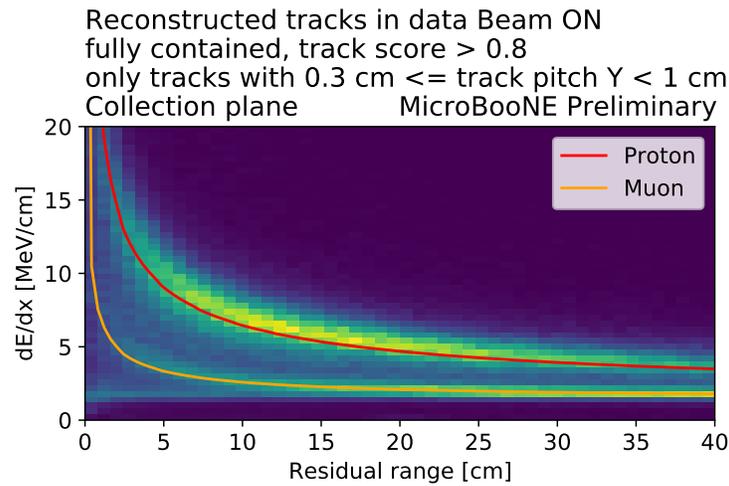
(a) dE/dx vs. residual range for tracks in reconstructed neutrino interactions in data.
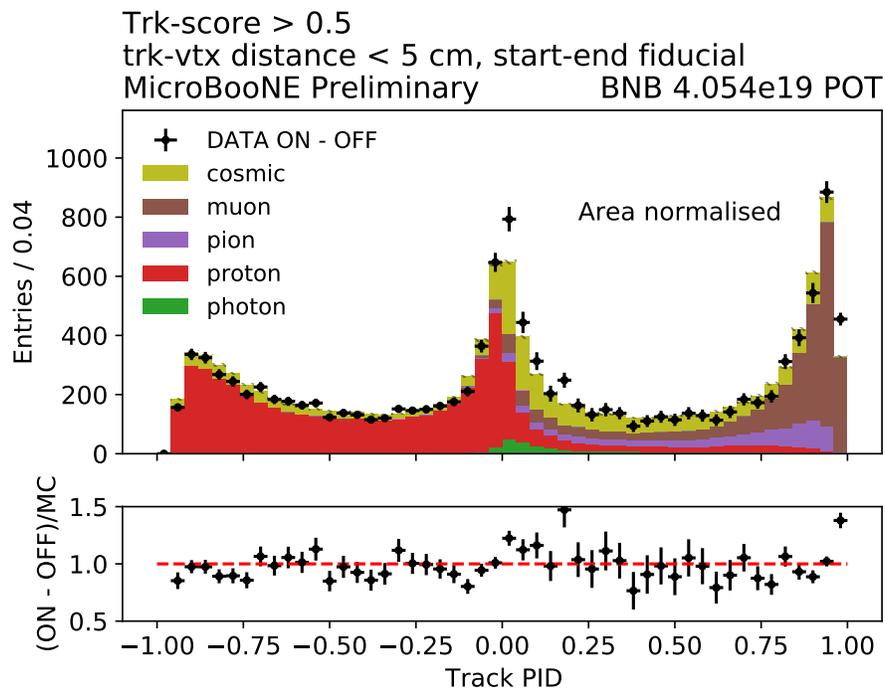
(b) PID variable score on data and simulation.

Figure 3: Top: distribution of dE/dx vs. residual range for contained tracks from neutrino interactions recorded in MicroBooNE's data corresponding to 4.054×10^{19} POT. Bottom: distribution of the log-likelihood-ratio PID variable for neutrino-induced tracks well contained in the fiducial volume. This plot is shown after off-beam subtraction, and area-normalized.

Reconstructing electromagnetic showers is another crucial aspect of the analysis. High statistics π^0 events, though a background to the ν_e selection, provide a valuable source for validating the reconstruction and calibration of electromagnetic showers and their energy scale. Figure 4 shows the reconstructed $M_{\gamma\gamma}$ obtained leveraging the same tools used in the electron neutrino selection. The good agreement between data and simulation, and the alignment between the distribution's peak with the expected 135 MeV π^0 mass help build confidence in the tools used in the analysis.
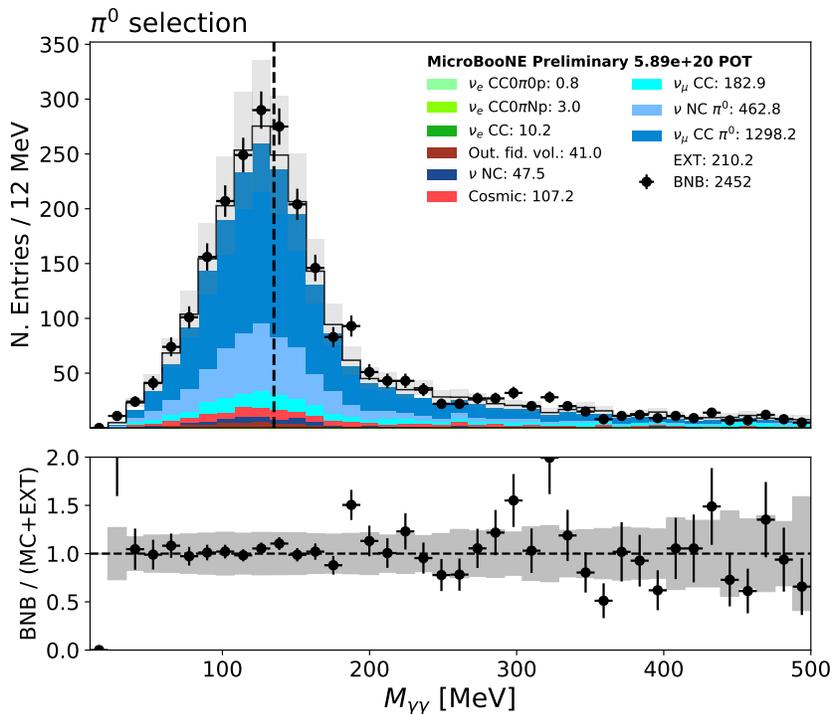

Figure 4: Reconstructed $M_{\gamma\gamma}$ for candidate π^0 events. The shower energy reconstruction and calibration is performed following the methods of ref. [6]. This plot, aimed to validate shower energy reconstruction, is produced by area normalizing the data and MC distributions scaling the MC π^0 contribution.

4   Electron Neutrino Event Selection

The two exclusive channels used in this analysis share a common pre-selection which requires the presence of a contained electromagnetic shower with greater than 70 MeV of energy. This pre-selection suppresses a large fraction of ν_μ charged current and off-beam backgrounds, with the energy-cut designed specifically to remove the frequent Michel electrons from muon decay produced in both cosmic and neutrino interactions. At this stage the $1e0p$ and $1e\mathrm{N}p$ selections diverge based on whether a reconstructed track associated with the neutrino interaction is found in the event. This requirement makes the two selections fully orthogonal.

| (a) number of showers. | (b) number of tracks. |
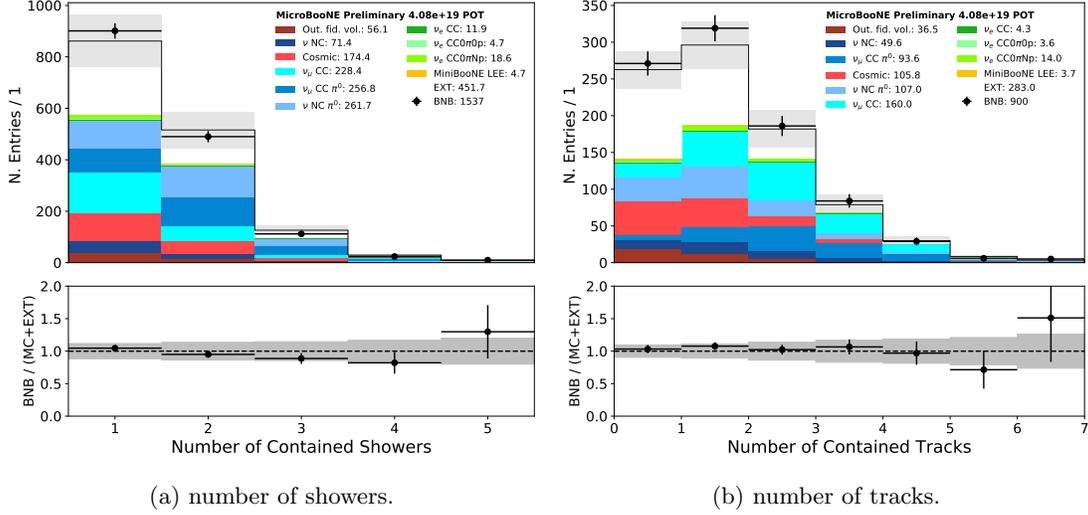
Figure 5: Number of reconstructed showers (left) and tracks (right) after pre-selection. The plot showing the number of tracks is produced further requiring a single reconstructed shower. The number of contained tracks is used to split between the zero and N-proton selections maintaining orthogonality.

After the common pre-selection, backgrounds still dominate over ν_e events by almost two orders of magnitude, with π^0s contributing to about half of the total background events. A series of additional variables and tools are defined to isolate and identify ν_e interactions. Variables are shared across the zero and N-proton selections, though variables related to proton track candidates are not used in the $1e0p$ selection. The choice of variables focuses on isolating ν_e events based on their topological and calorimetric features while limiting reliance on kinematic variables, and possible neutrino interaction modelling dependence.

- **Cosmic Ray rejection** In order to suppress the cosmic background and to select well reconstructed neutrino interactions, all events in this analysis are required to be contained in the TPC. In addition, the proximity of electromagnetic showers to nearby cosmic-ray muons is leveraged as a way to reduce backgrounds induced by cosmic-ray muon Bremmstrahlung showers which are frequently produced in the detector. For events with one or more track, a further requirement that the longest track has a proton-like PID is very effective for the suppression of cosmic-ray muon interactions.

- ν_μ **rejection** At low energies, track/shower separation is challenging due to the track-like appearance of $\mathcal{O}(100)$ MeV electromagnetic showers. Conversely, low-energy muons and pions which decay producing a Michel electron may be incorrectly reconstructed as a single electromagnetic shower. To reject low-energy track backgrounds to the ν_e selections, additional variables are constructed with the goal of characterising electromagnetic showers and adding discriminating power complementary to the track/shower classification available from Pandora. These variables measure the number of subclusters (Figure 23c) in the shower, the shower Moliere angle (an analysis-defined term aimed to measure the transverse dimension of

8

the shower), and the fraction of hits in the track fit of the shower relative to the total number of hits in the shower (Figure 23a).

- π^0 **rejection** The number of reconstructed electromagnetic showers and shower trunk $\mathrm{d}E/\mathrm{d}x$ are powerful in suppressing π^0 backgrounds. For the $1eNp$ selection, the conversion distance between the proton candidate start point and the electron start point provides an additional powerful discriminant. To mitigate reconstruction deficiencies associated with either unidentified low energy sub-leading γ showers or highly co-linear γ pairs, additional variables are defined which aim to identify missing "second showers" and identify merged or overlapping γ showers.

Comparisons of data to MC for two of these variables are shown in Figure 6.

(a) Number of subclusters in the shower.

(b) Fraction of hits in shower-trunk track fit.
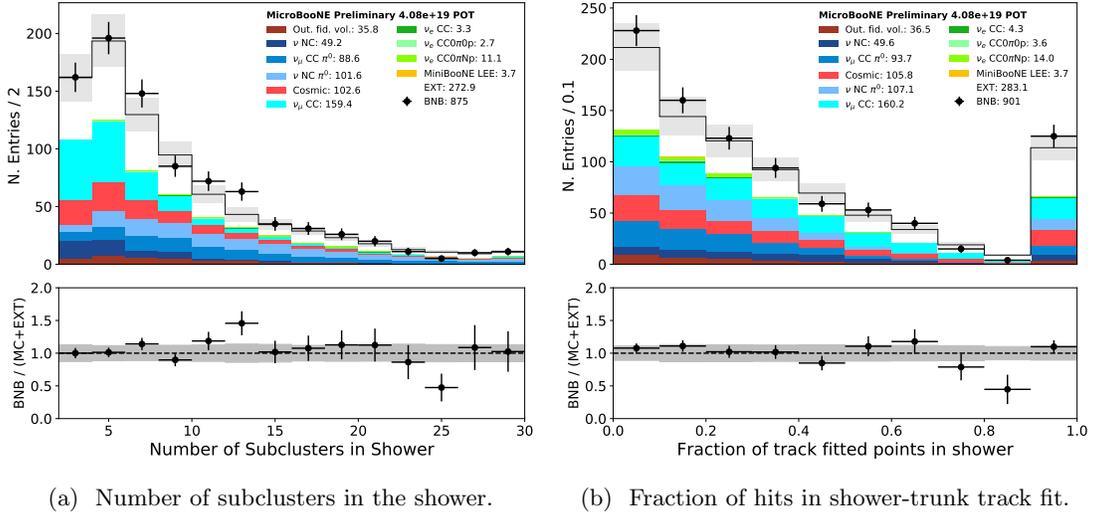
Figure 6: Comparison of data and simulation for selected electron neutrino selection variables at pre-selection for events with one shower, and any number of tracks.

These tools are used to define a series of cuts and are also employed in boosted decision trees (BDTs) used for the final selections. The Np selection makes use of two separate BDTs, one trained against backgrounds associated with π^0 production and the other against all other backgrounds. The response of the former is shown in Figure 7.

(a) Low BDT response (0-0.5). (b) High BDT response (0.5-1).
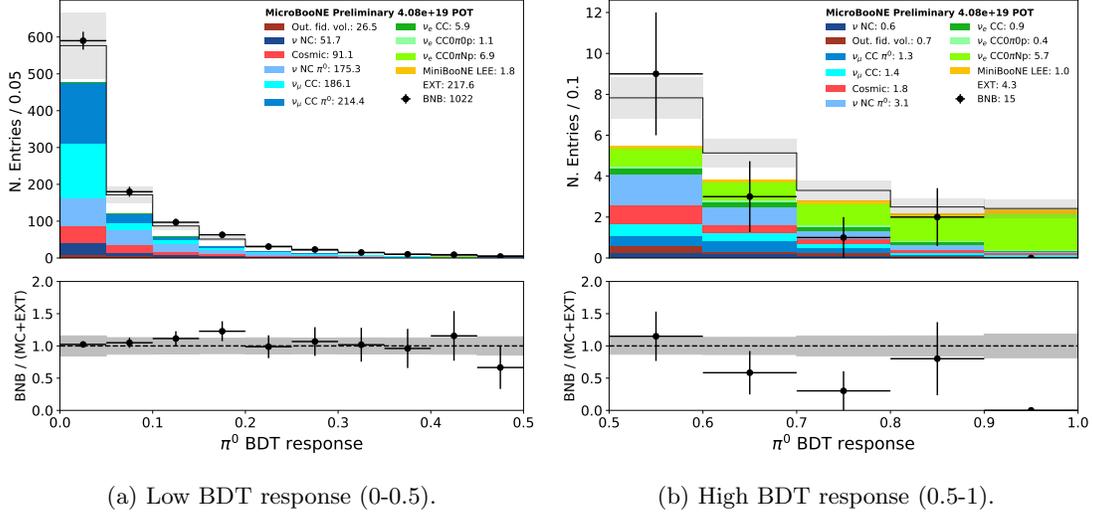
Figure 7: BDT response for the Np selection.

The 0p selection employs a single BDT trained against all backgrounds, and its response is shown in Figure 8


(a) Low BDT response (0-0.5). (b) High BDT response (0.5-1).
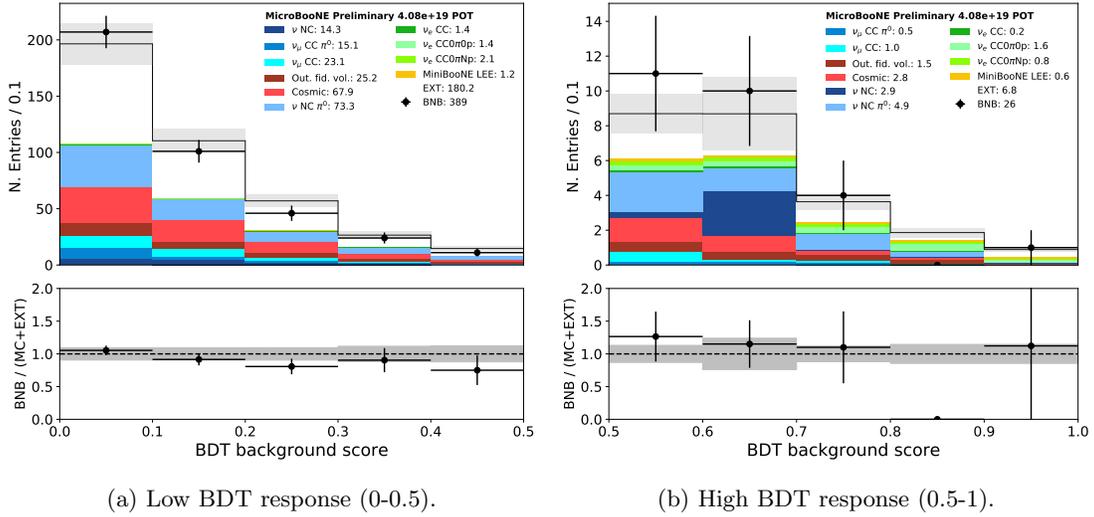
Figure 8: BDT response for 0p selection.

Figure 9 shows the expected reconstructed energy spectrum for the $1e$Np and $1e0p$ selections scaled to the currently analyzed POT. In both plots, data events are shown for the unblinded regions of the spectrum, above 1.05 and 0.9 GeV of reconstructed energy for the two channels respectively.

10

In both plots green stacked distributions denote ν_e interactions expected from the intrinsic BNB ν_e flux, while the orange contribution is the MiniBooNE LEE-signal model, under the ν_e hypothesis, used to evaluate the analysis' sensitivity and described in ref. [3]. For the $1eNp$ selection, in the energy-range $1.05 - 2.05$ GeV, 11 data events are observed with an expectation from the simulation of 18.2, 16.6 of which ν_e. For the $1e0p$ selection in the energy range $0.9 - 2.00$ GeV 3 events are observed with an expectation of 5.5 (4.5 of which ν_e). Systematics on the expected events in these high-energy bins is in the order 20-30%, depending on the bin.


(a) $1eNp$ selection.
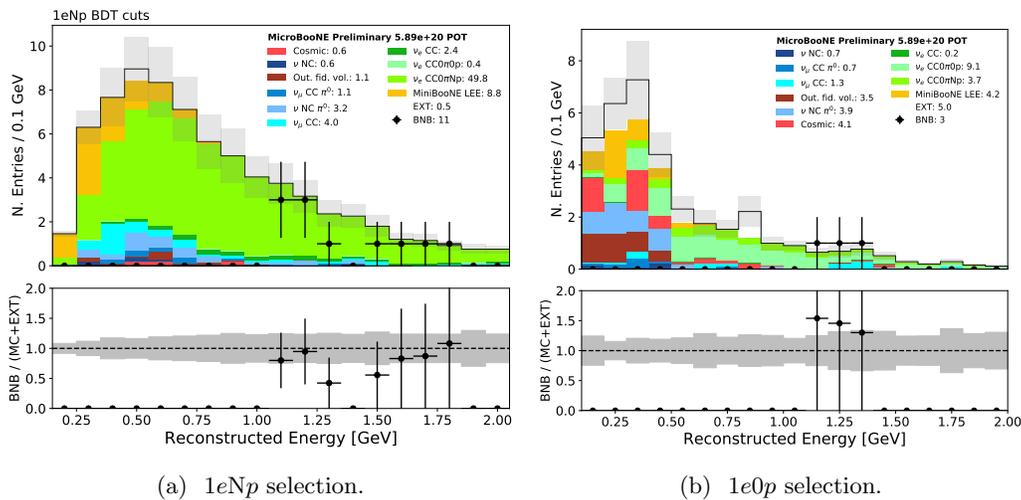
(b) $1e0p$ selection.

Figure 9: Electron neutrino selections for the $1eNp$ and $1e0p$ selections. Data is shown for values of reconstructed energy above 1.05 and 0.9 GeV for the two selections respectively. Data below these thresholds remains blind.

The inclusive selection shares the pre-selection with the exclusive $1e0p$ and $1eNp$ selections, and differentiates afterwards so as to be topology independent. Therefore, the ν_e CC selection enables the inclusion of events with no vertex activity, arbitrary proton multiplicity and/or pion production. The selection is not tailored to a specific energy region and is therefore ideal to look at the kinematics of electron showers in charged-current electron neutrino interactions. Results showing the high-statistics measurement of ν_e interaction rate and the reconstructed lepton kinematics with this selection from available data-sets are presented in Figure 10, and described in more detail in Appendix B. The events selected are not orthogonal to the exclusive selections and therefore are not included in a combined low-energy-analysis.
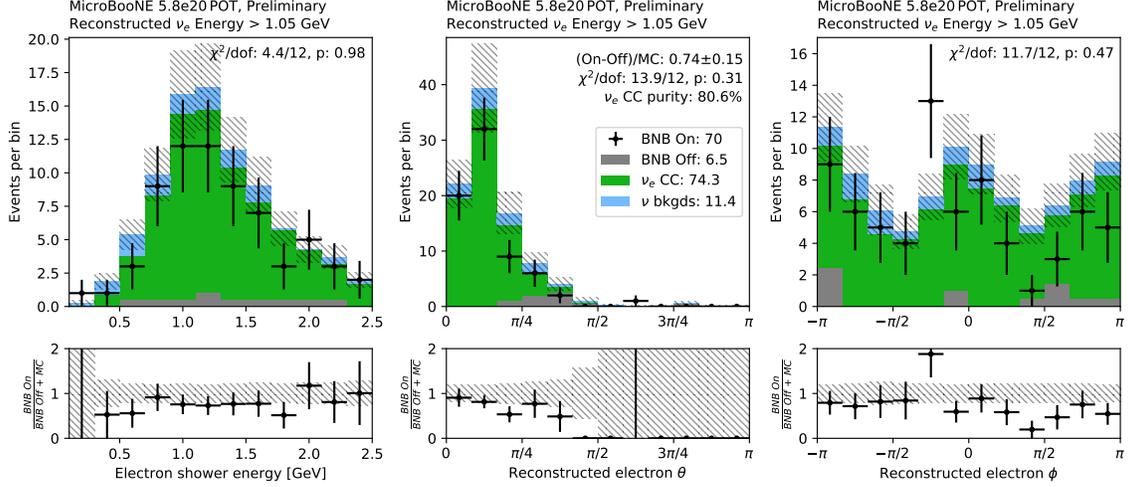
Figure 10: Electron shower energy (left) and directional angles θ (middle) and ϕ (right) of the selected ν_e CC electron candidate. In blue, the neutrino related backgrounds are grouped into one category. The error bars on the data are Poissonian. The errors on the prediction are the flux [5] and cross-section systematic uncertainties [9], combined with the uncertainty originating from the finite simulated sample size. The used data-set correspond to the high-energy sideband 5.8×10^{20} POT and contains neutrino candidates with a reconstructed energy above 1.05 GeV.

Event displays for several selected electron neutrino events are shown in Figure 11, with additional events from the $1e$Np selection available in Appendix E. Additional details on the $1e0p$ selection are presented in Appendix D.

(a) 1eNp candidate event.

(b) 1eNp candidate event.

(c) 1e0p candidate event.

(d) 1e0p candidate event.
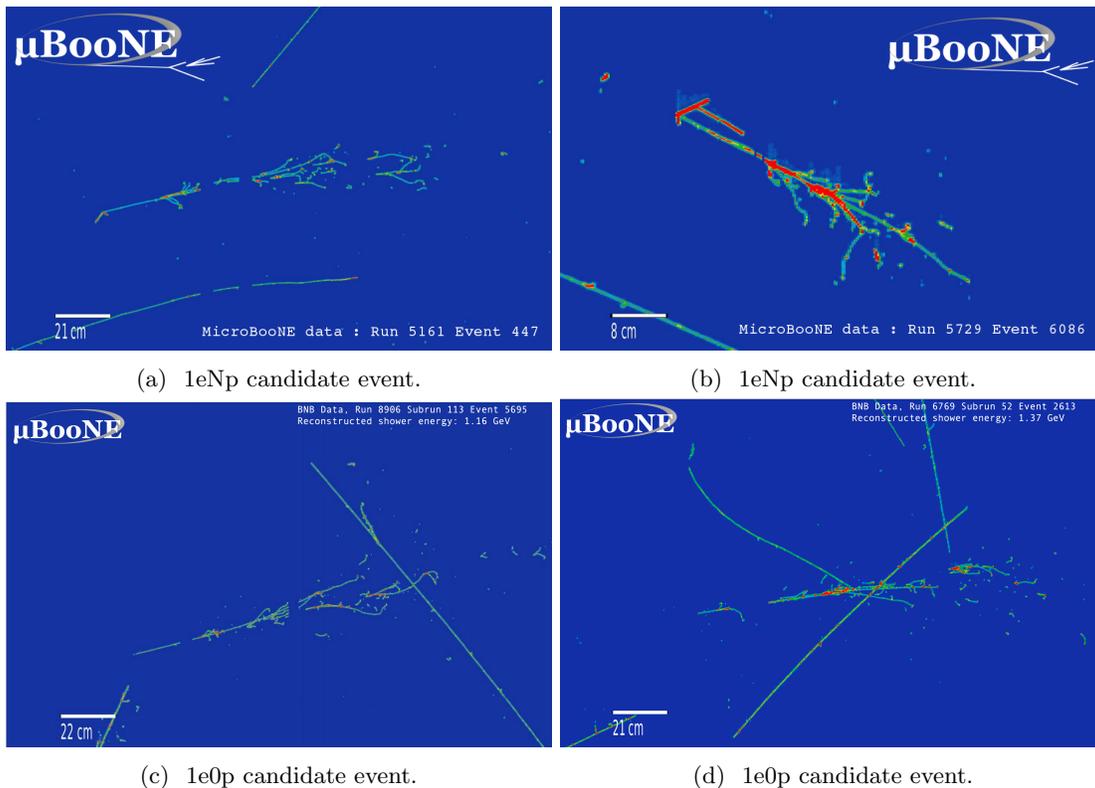
Figure 11: Event displays of selected events.

5 Muon Neutrino Constraint, Systematics, and Analysis Sensitivity

Flux and neutrino interaction cross-section uncertainties limit the ability to predict the expected rate of electron neutrino interactions, and test the hypothesis of the MiniBooNE excess. Muon neutrinos from the BNB can be used to constrain these systematics. Both muon and electron neutrinos are produced by the decay of the same hadrons in the beam, making it possible to constrain ν_e flux uncertainties by measuring high-statistics ν_μ events. Similarly, muon and electron neutrinos undergo charged-current weak interactions on the same target nucleus, making it possible to constrain neutrino interaction cross section uncertainties.

Most of the neutrinos from the BNB are muon neutrinos, so it is possible to make a high purity measurement of muon neutrinos that have well reconstructed neutrino energy with significant statistics [5]. Contained charged current muon neutrinos are selected by identifying events with a contained muon candidate and any number of contained hadron tracks, neither of which exit the detector. The selection as a function of reconstructed neutrino energy is shown in Figure 12.
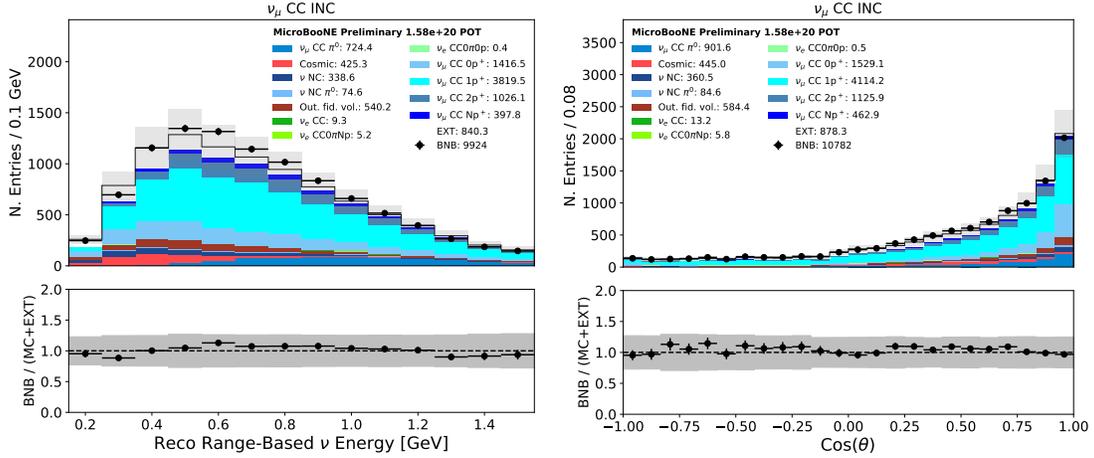
(a) ν_μ energy.

(b) muon angle w.r.t. beam.

Figure 12: Muon neutrino selection kinematics distributions.

Detector systematic uncertainties are sub-dominant compared to flux and cross-section systematics. They are evaluated separately with dedicated variation samples [10]. More detail is included in Appendix C.

The impact of the muon neutrino constraint on the Np selection can be seen by the reduction in errors in Figure 13.

The overall reduction in systematic uncertainties from the muon neutrino constraint is 30-50% in the bins ≤ 0.85 GeV. While the analysis' goal is to perform a comprehensive measurement of electron neutrino events over a broad energy range, sensitivity to a specific model hypothesis [3] for an electron-like excess of events compatible with the MiniBooNE observation is performed. The analysis, folding in detector, cross-section, and flux systematics, and including the constraint from the ν_μ events, obtains a median sensitivity to exclude a standard model in favor of a MiniBooNE-like ν_e excess of 2.3σ with the 6.9×10^{20} POT of BNB neutrino-mode data analyzed to date. With the full expected data set of 12.5×10^{20} POT, the expected median sensitivity with the same analysis is 3.5σ when considering only the statistical errors and 3.0σ after including the systematics. More details on the analysis approach to systematics, the way in which the muon constraint is performed, and the calculation of the sensitivity are provided in Appendix C.
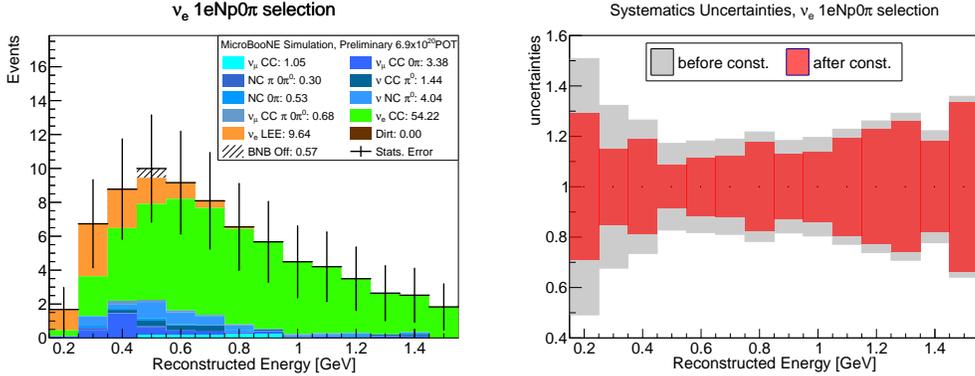
Figure 13: Impact of muon neutrino constraint on Np selection, scaled to 6.9×10^{20}. (Left) Stacked distribution separated by the MC dedicated samples as a function of reconstructed energy. The errors on the plot represent only the statistical uncertainties of the distribution. (Right) The reduction in the systematic uncertainties is visualized as a shrinking in the uncertainty band in the ratio plot. Systematics (including flux, cross-section, and detector uncertainties) before the constraint are shown in grey and after the constraint in red. The dashed line at 1 in the error band is used to guide the eye.

6 Selected Electron Neutrinos at High Energy

This analysis is moving towards a full unblinding in steps in energy and BDT response. To start with, two far side-bands were defined: one at low BDT response, and one at high energy.

The low BDT response side-band is used to validate the background model on high data statistics. The data and simulation are found to agree well in this region for the $0p$ and Np selections.

In the high energy side-band, all of the analyzed data was opened above a reconstructed neutrino energy of 1.05 and 0.9 GeV for the N-proton and 0-proton selections, respectively. This high energy side-band provides the opportunity to study ν_e events with MicroBooNE data for the first time, and provides some of the largest statistics available to date of ν_e interactions in a LArTPC.

The tools and selection cuts developed for the exclusive channels in this analysis have the particular aim of being able to measure ν_e interactions with high purity at low energy, in order to be sensitive to possible new physics in this regime. With the final selection cuts, 11 events pass the full Np selection above 1.05 GeV, and 3 events pass the full $0p$ selection above 0.9 GeV, as shown in Figure 9. Due to both a smaller rate of backgrounds, and a more efficient reconstruction at high energies, the analysis is able to cleanly isolate ν_e events at higher energies with loose cuts on a subset of the variables that are used to train the BDT. The inclusive selection (Appendix B) selects 70 events above 1.05 GeV with a ν_e CC purity of 80.6%. Both the loose-cut exclusive selections and the inclusive selection make it possible to validate the analysis performance with higher statistics of ν_e CC events, and produce plots such as Figures 14 and 30. These showcase the performance of the calorimetric separation of electrons from photons via their characteristic dE/dx signature.
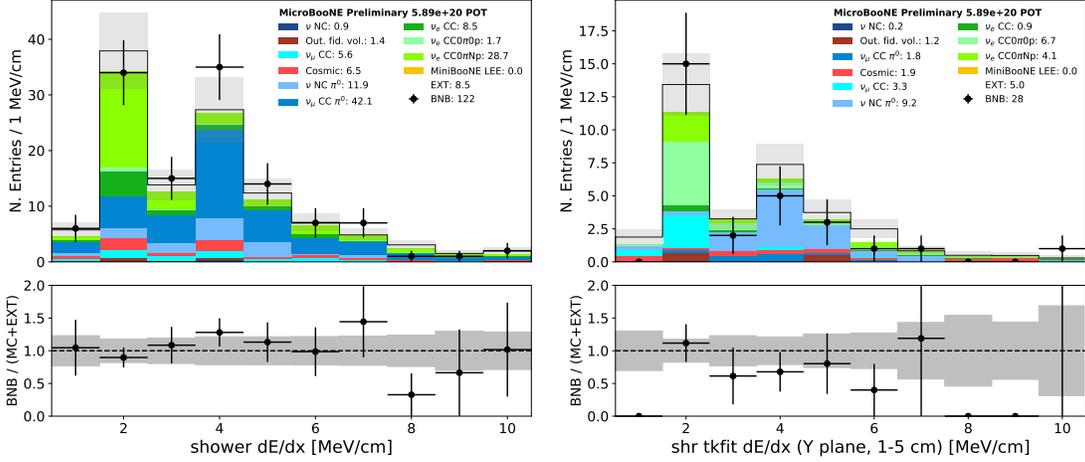
(a) 1eNp dE/dx calculated on the plane with the most hits 0-4 cm from the vertex.

(b) 1e0p dE/dx, calculated on the Y plane 1-5 cm from the vertex.

Figure 14: dE/dx on high energy electron neutrinos after intermediate selection cuts. The left plot comes from the 1eNp selection high-energy sideband ($E_{\mathrm{reco}} > 1.05$ GeV). The plot on the right is from the 1e0p selection high-energy sideband ($E_{\mathrm{reco}} > 0.9$ GeV)

7 Conclusions

This document has presented the status of an analysis which aims to measure ν_e interactions in the BNB beamline with the goal of testing whether the nature of the MiniBooNE excess of EM activity'[1] is associated with electron-like events. Several of the key measurements of the analysis needed to validate EM shower reconstruction (π^0s, e/γ and μ/p separation) and neutrino flux and interaction modeling (ν_μ CC) are presented, together with a full assessment of the analysis' sensitivity to a specific interpretation of the MiniBooNE excess under the electron hypothesis [3]. Results shown in this note include $5.9E20$ POT of data available thus far. The sensitivity estimation is presented for the total $12.5E20$ POT of data collected over MicroBooNE's five years of running. Finally, this document presented first ν_e results from the BNB with an applied threshold on reconstructed energy of $\mathcal{O}(1)$ GeV, which represent the first step in MicroBooNE's unblinding procedure towards finalizing the electron neutrino analyses.

A Measuring Electron Neutrinos with Particle Identification in the MicroBooNE LArTPC for sensitivity to new Physics

In this section we describe the calorimetry-based particle identification tools used for tracks and showers that lead to the 1eNp ν_e selection. First we describe the signature we are looking at using, as an example, one event display coming from the 1eNp high energy sideband (Fig. 15). Afterwards, we describe the way particle identification is performed for tracks and showers in three points. The first one consists in explaining the idea, using a couple of schematics (Fig. 16). The next step is showing plots of the dE/dx versus residual range for muons and protons (Fig. 3a), and of the dE/dx versus distance from the start for electrons and photons (Fig. 17) to show the information available about these particles. The last step describes how the previous information is condensed in one variable, showing plots of the LLR PID for tracks (Fig. 3b), and the shower median dE/dx in the first 4 cm (Fig. 14a). This last couple of plots is shown with selections that emphasise the fact that these variables indeed work to do particle identification. Then, we show the distributions of the particle ID variables used in the analysis in the 1eNp selection sidebands currently explored in two different selection stages in the far sideband (Figs. 18-19). Stage 1 shows the distributions where only a cut on reconstructed energy of 1.05 GeV is applied: it shows how these variables distinguish the signal from the backgrounds. Stage 2 enhances the backgrounds, by selecting only events with low BDT response score (< 0.1). This stage shows how different backgrounds can be classified with these variables, validating the power of particle ID in MicroBooNE. In conclusion, the particle identification tools described in this section are crucial components of the selection chain employed in this analysis; in particular, they are instrumental to achieve levels of purity shown in figure 9a and the expected sensitivity to the MiniBooNE unfolded signal reported in appendix C.
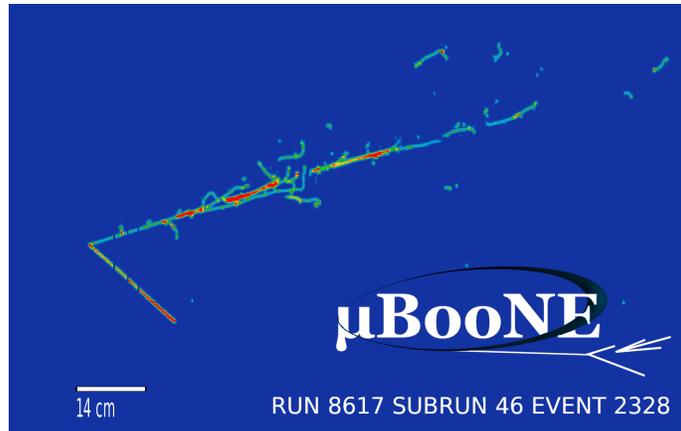

(a) Event display showing a very clear signature of a 1e1p event, with one shwoer and one track in the final state.

Figure 15: Event display used to explain the main signature target of the 1eNp analysis.

17
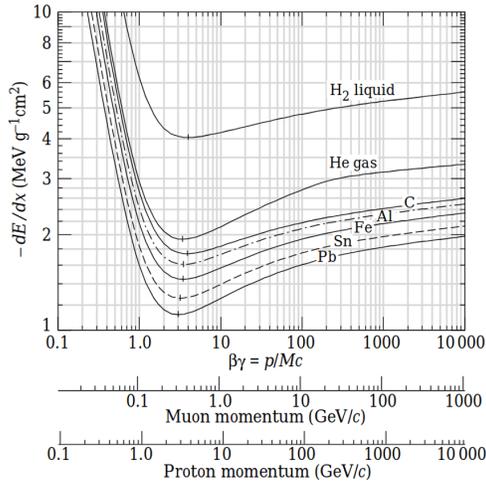
(a) Track PID starts from the Bethe-Bloch function, which describes the energy deposition as a universal function of $\beta\gamma$, and thus the momentum dependence is different for different particle types. The plot is taken from the PDG [11].
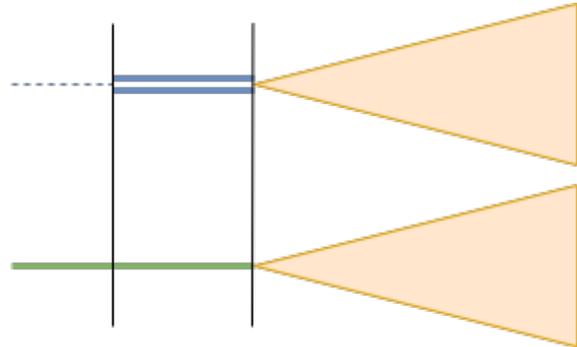
(b) Shower PID considers the beginning of the shower, where a photon shower (in blue) starts with a pair production of an e+/e- pair, implying a double amount of ionisation with respect to an electron shower (in green) which starts with only one electron.

Figure 16: Schematics to explain the ideas behind particle identification for tracks and showers, on the left and right, respectively.
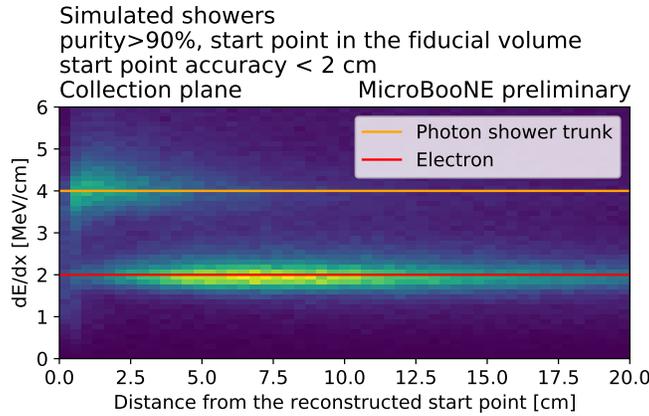
(a) dE/dx versus on the collection plane versus distance from the start point for simulated showers, reconstructed with a start point in the fiducial volume and with good accuracy. The contribution from electron and photon showers is clearly visible and distinct, as well as the dependence on the distance from the shower start point.

Figure 17: Calorimetry plot for showers.

18

(a) Track PID
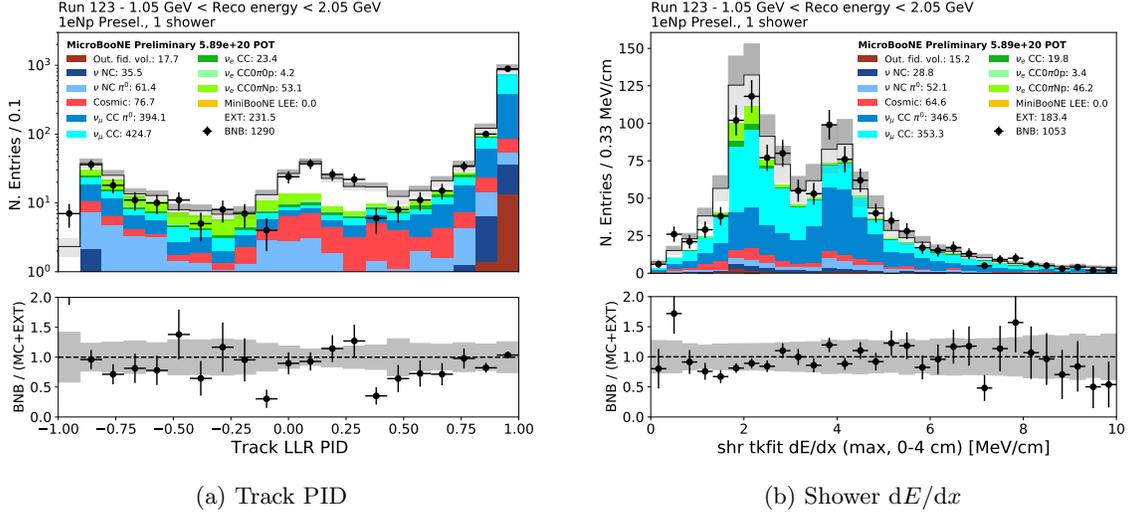
(b) Shower $\mathrm{d}E/\mathrm{d}x$

Figure 18: Distribution of the track PID and shower $\mathrm{d}E/\mathrm{d}x$ in the 1eNp far sideband, after applying the Np preselection cuts, requiring the presence of exactly one shower contained, and applying a cut on the reconstructed energy. It is possible to notice that these variables are already able to separate the $\nu_e CC0\pi0p$ contribution from the background, moving the $\nu_\mu CC$ background at values of the track PID around 1, and the π^0 backgrounds in the peak at 4 MeV/cm.

(a) Track PID
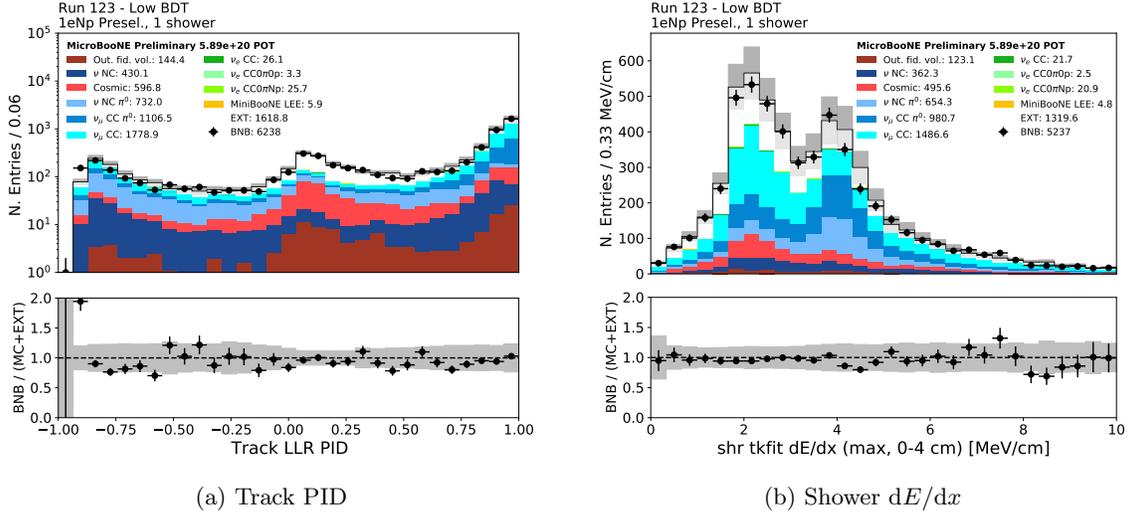
(b) Shower $\mathrm{d}E/\mathrm{d}x$

Figure 19: Distribution of the track PID and shower $\mathrm{d}E/\mathrm{d}x$ in the 1eNp far sideband, after applying the Np preselection cuts, requiring the presence of exactly one shower contained, and requiring low PID - i.e. very background like. Aside for not being able to see any signal, these plots emphasise high statistics backgrounds, again showing the capability of these variables of distinguishing different backgrounds. Interestingly, the track PID is able to separate NC π^0 events, with a proton track, from CC π^0 events, characterised by the presence of a muon track.

19

B Charged-Current Electron Neutrino Measurement with the MicroBooNE Detector

The electron neutrino charged-current (ν_e CC) signal definition used in this section requires a final state lepton with a kinetic energy of at least $20\,\mathrm{MeV}$ and a neutrino vertex inside a fiducial volume (Figure 20). The ν_e CC selection described here selects these events with an efficiency of $(18.1 \pm 0.1)\,\%$ and a purity of $(53.3 \pm 1.0)\,\%$.

The ν_e CC events can be categorised by interaction mode or final state topology. The four interaction modes are quasi-elastic scattering, meson-exchange current, resonant production and deep inelastic scattering. Their relative contributions are shown in the left panel of Figure 21.

The ν_e CC signal can be split into three final state topologies:

- $0\pi0p$: No protons or pions in the final state with kinetic energy (E_{kin}) above $40\,\mathrm{MeV}$; a single electron shower.

- $0\pi\mathrm{N}p$: At least one proton – N> 0 – with $E_{kin}(p) > 40\,\mathrm{MeV}$ in the final state and no pions above threshold.

- $\mathrm{M}\pi\mathrm{N}p$: At least one pion – M> 0 – with $E_{kin}(\pi) > 40\,\mathrm{MeV}$, no restrictions on the number of protons.

Their contribution as a function of the simulated neutrino energy is given in the right panel of Figure 21. In the energy range of the Booster Neutrino Beam (BNB), the dominant interaction mode is quasi-elastic scattering, and the dominant final state is ν_e CC $0\pi\mathrm{N}p$.

The selection strategy is shown in Figure 22. Cosmic rejection is common in all neutrino analyses described in this note and is obtained by combining the accelerator trigger signals with the prompt light signal and the location of the deposited charge in the TPC. Next, the reconstructed neutrino vertex is required to be located in the same fiducial volume used to define the signal events. Attached to the neutrino vertex, there are reconstructed objects, either electromagnetic showers or track-like charge deposits. The event is required to have an electromagnetic shower, reconstructed on the three wire planes with a reconstructed energy of at least $100\,\mathrm{MeV}$. In the next step, the electron candidate shower is identified as electron-like or photon-like, as is further discussed in Appendix B.1. The other particles attached to the neutrino vertex are used to further reject backgrounds relying on the particle identification tools described in Appendix A. Finally, the combination of the particle identification stage is combined into the event selection, as described in Appendix B.2.
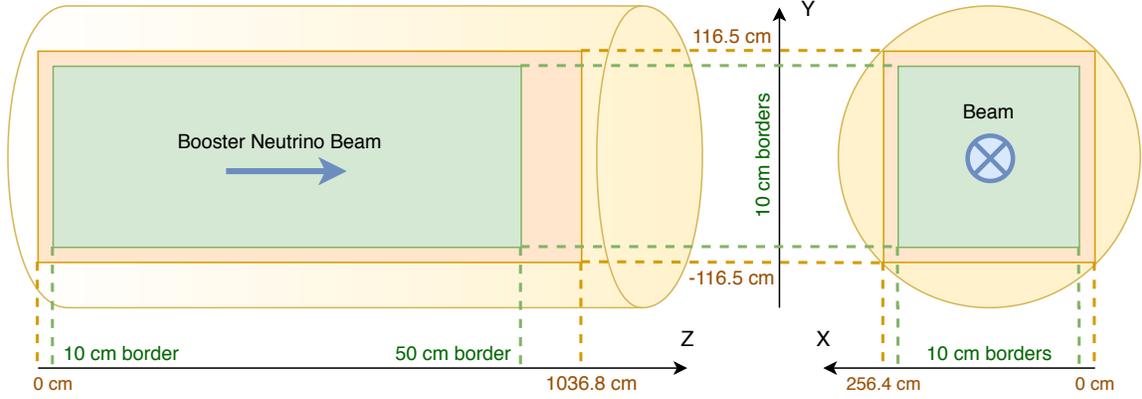
Figure 20: Fiducial volume used in this section for both the signal definition and the ν_e CC selection. The yellow cylinder represents the cryostat, the orange structure is the TPC and the green area is the fiducial volume. Both the YZ (left) and XY (right) projections are shown.
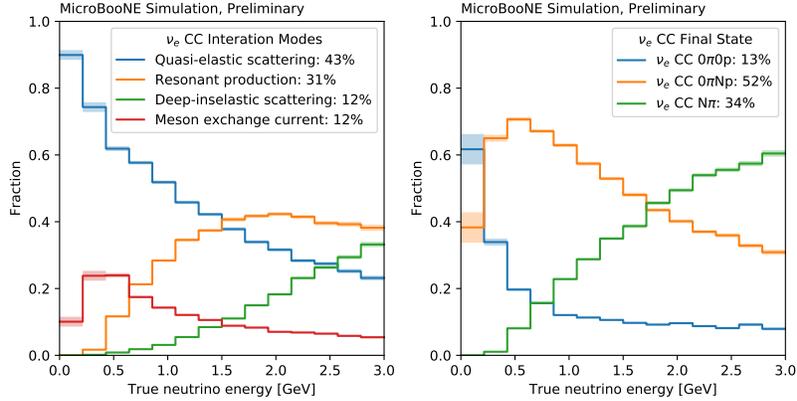

Figure 21: Breakdown of ν_e CC events in interaction modes (left) and final state topologies (right) as a function of neutrino energy. Note that these are relative fractions and have to be multiplied by the BNB flux to retrieve the interaction rates. The relative contributions to the predicted ν_e CC interaction rates from the BNB is given in the legend. The shaded bans represent the statistical error arising from limited simulation statistics.
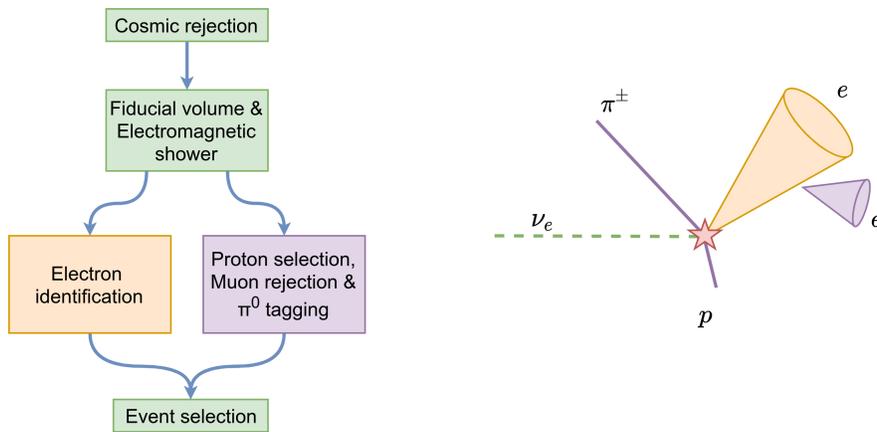
Figure 22: (Left) Flowchart indicating the different steps in the ν_e CC selection. (Right) A hypothetical ν_e CC event with multiple reconstructed showers and tracks. At the pre-selection stage, an electron candidate shower is identified (orange) and electron particle identification is performed on this shower. The other reconstructed objects (purple) are classified to improve background rejection before merging the outputs and perform the final event selection.

B.1   Electron identification

After the electromagnetic shower requirement described in the introduction of this appendix, electrons make up less than $10\,\%$ of the selected showers. The majority of showers are caused by neutrino interactions containing $\pi^0 \to \gamma\gamma$ decays. A gradient boosted decision tree [12, 13] is used to discriminate between electron and photon induced showers.


(a) Track fit to shower trunk   (b) $\mathrm{d}E/\mathrm{d}x$ at the start of the shower   (c) Number of subclusters
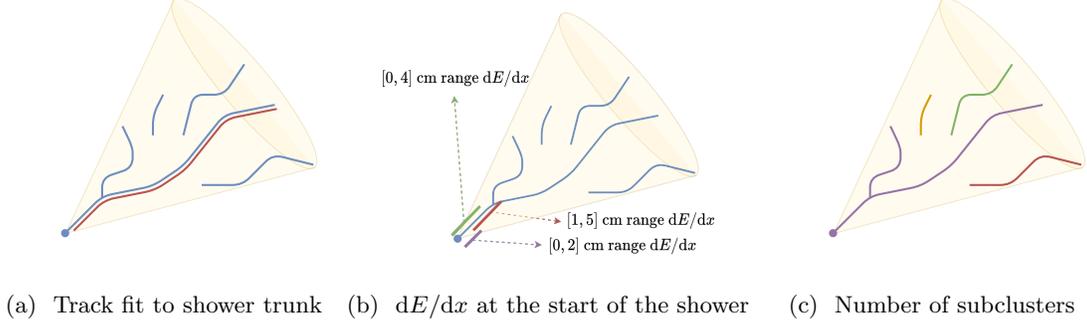
Figure 23: Subset of the variables used for electron identification.

The nine variables used to discriminate electron showers, and their correlation, are shown in top-left block of Figure 24. The trunk of the shower is fit as a track. The fraction of calorimetric hits in this shower trunk compared to the whole shower is used as one of the variables (see Figure 23a). The start of the trunk is used to measure the ionisation. For reasons introduced in Appendix A, different but highly correlated variables are constructed (Figure 23b):

- The $\mathrm{d}E/\mathrm{d}x$ on the collection plane in the first $4\,\mathrm{cm}$.

- The $\mathrm{d}E/\mathrm{d}x$ on the collection plane in the first $2\,\mathrm{cm}$.

- The $\mathrm{d}E/\mathrm{d}x$ on the collection plane from $1\,\mathrm{cm}$ to $4\,\mathrm{cm}$.

- The weighted mean of the $\mathrm{d}E/\mathrm{d}x$ in the first $4\,\mathrm{cm}$ using both the two induction planes and the collection plane.

As shown in Figure 17, it is possible that an the electron-positron pair at the start of a photon shower only overlaps for a fraction of the first $4\,\mathrm{cm}$. Therefore, the variable using the first $2\,\mathrm{cm}$ is important albeit the additional statistical fluctuations. The third variation introduces a $1\,\mathrm{cm}$ gap to avoid overlapping charge deposits close to the vertex. The last variation relies on a combination of the three planes and is important when the collection plane has few hits at the beginning of the shower.

The number of sub-clusters in the shower (Figure 23c), the Moliere angle of the shower and the distance of the shower start point from the reconstructed neutrino vertex are also used. Finally, the variable *Second shower hits* takes into account the presence of another shower in the event that might be a signature of a neutral pion decay.

The correlations between the input variables and the *Electron PID score*, seen in Figure 24, indicate a preference for showers with a smaller vertex distance, smaller Moliere angle, and larger
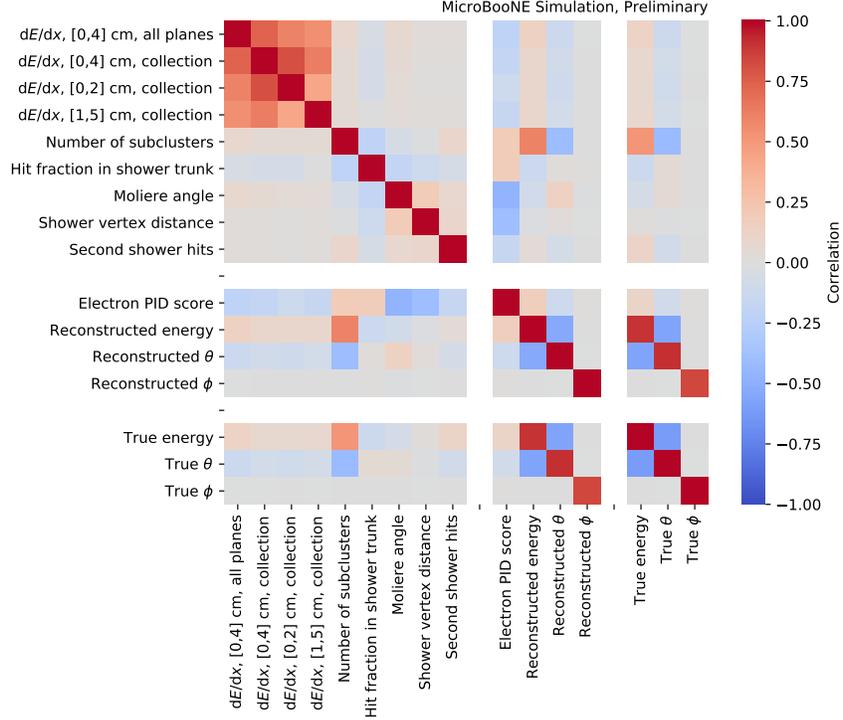
Figure 24: Correlation matrix of the variables of interest in the electron identification step. The first group (top-left) contains the variables used in the electron PID. The second group (middle) are the BDT score and the reconstructed lepton kinematics that can be compared with their simulated counterparts in the third group (bottom-left).

number of sub-clusters. The correlation with the second shower tagging variable demonstrates that second showers with a substantial amount of hits is unlikely to occur in the case of an electron. The negative correlation between the electron score and the dE/dx variables is explained by realising that electrons have a specific dE/dx expectation of $\sim 2.1\,\mathrm{MeV\,cm^{-1}}$, instead of the double values expected for the start of photon showers.

Finally, it is encouraging to see that the correlation between the electron score and the truth-based kinematics is low. This serves as a demonstration that the electron identification is kinematically unbiased, covering a large portion of the lepton kinematics phase space.

In Figure 25, simulated neutrino events passing the *pre-selection* are used to demonstrate the obtained electron-gamma separation. It is found that one can reject 90 % of photon showers, while keeping 75 % of electron showers, corresponding to a 8 : 1 increase of signal-to-background ratio. This solely serves as a demonstration and no one-dimensional cut was placed on this value in the final selection, as described in Appendix B.2.
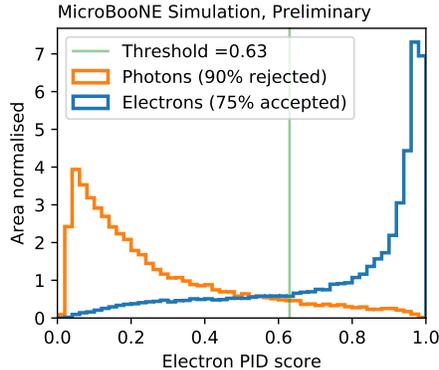
Figure 25: $e - \gamma$ separation performance obtained by the electron classifier on simulation. The electron BDT response is shown for simulated electrons (blue histogram) and for background photons (orange histogram). The green line indicates the score at which 95% of the background photons are rejected.

B.2 Charged-Current Electron Neutrino Event Selection

The final event classification is performed using a gradient boosted decision tree, building on the particle identification performed in Appendices A and B.1. Figure 26 shows the efficiency in function of the simulated neutrino energy for the three different signal topologies introduced in the beginning of this section. The efficiency is given at three different stages in the selection, corresponding to different levels in Figure 22. *NeutrinoID* refers to the set of tools introduced in Section 3 and corresponds to the first step of cosmic removal in Figure 22. The *Pre-selection* includes the additional requirements defined by the second block in Figure 22. At this stage, the ν_e CC purity is $(5.74 \pm 0.04)\,\%$.

The final selection combined the output of the electron shower identification with the additional background rejection obtained from the classification of the other reconstructed particles in the event. The response of the event classification is shown for all events passing the *pre-selection* in the left panel of Figure 27. Here, the unblinded data is used, corresponding to approximately 4% of the total POT collected by MicroBooNE . The event selection, the final block in Figure 22, corresponds to the requirement of a ν_e CC event classification score of 0.87 or higher. The purity obtained by the ν_e CC event selection is $(53.3 \pm 1.0)\,\%$.

In Figures 27 and 28 the errors shown on the *Beam On* data are statistical and modelled by a Poisson distribution. The errors on the prediction, represented by the shaded areas, include the uncertainty arising from the limited simulation and *Beam Off* data statistics, added in quadrature with the systematic uncertainties arising from the flux and cross-section modelling. The systematic uncertainties are taken into account through a covariance matrix approach. On the figures, the *(beam on - beam off)/ simulation* ratio is given with corresponding uncertainty. Additionally, the p-value obtained from a two-sample Kolmogorov–Smirnov test is given. The former is indicative for the normalisation agreement while the latter quantifies the shape agreement of the distributions. Additionally, the χ^2-test with the corresponding degrees of freedom and p-value is quoted, as defined in [14].

The right panel of Figure 27 shows the reconstructed shower energy of the electron candidate
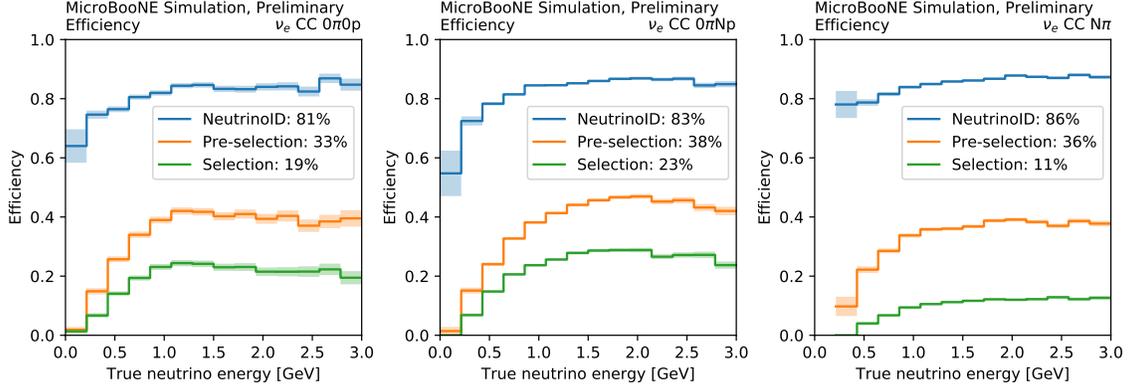
Figure 26: Efficiency of the different stages in the ν_e CC selection as a function of the simulated neutrino energy. The three panels correspond to different final state topologies. The shaded band is the statistical uncertainty. The statistical uncertainty on the integrated efficiency in the legend is below 1 % for all cases.

in the selected events, defined as events with a ν_e CC identification score of 0.87 or higher. In Figure 28, the track multiplicity at vertex is shown before and after the event classification boosted decision tree. It can be appreciated that even in the small unblinded data-set, events with a wide variety in track multiplicity are selected. A further indication of the scope of the selection is given by the collection plane view of three selected events in Figure 29.
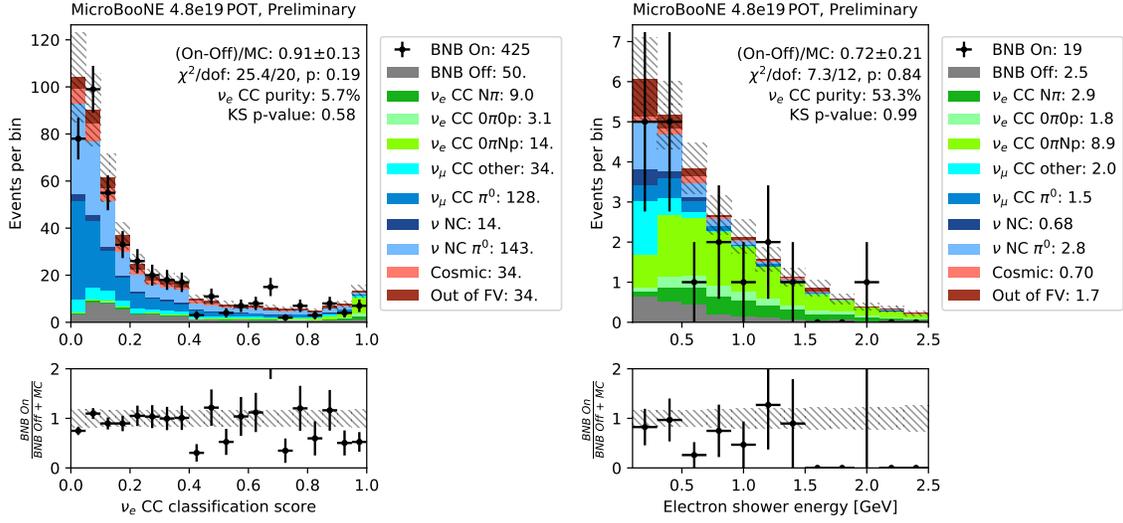
Figure 27: (left) BDT response of the ν_e CC inclusive event classifier. (right) Reconstructed electron shower energy distribution after the selection. The data shown corresponds to the unblinded subsample.
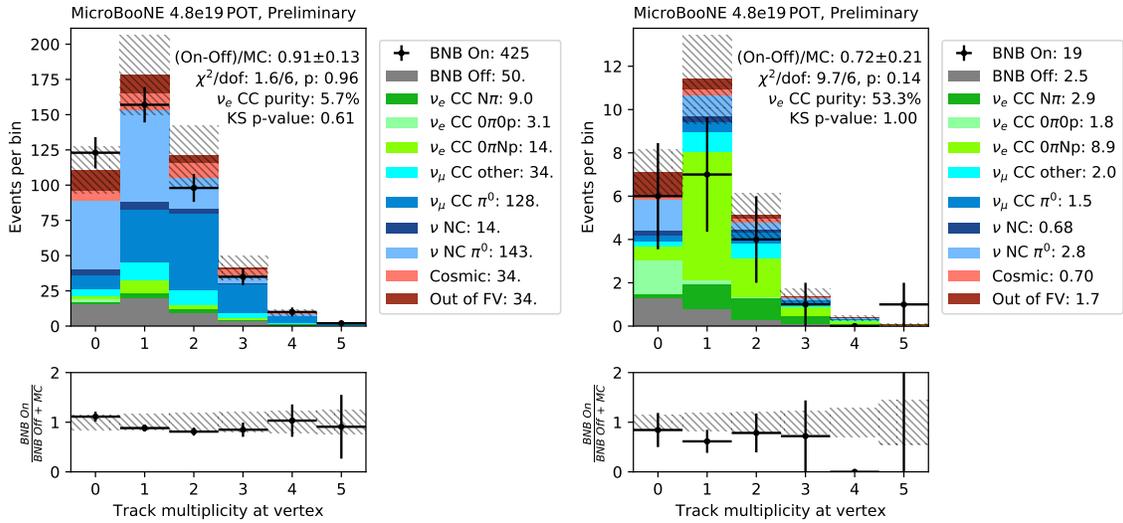
Figure 28: Track multiplicity at vertex at the pre-selection stage (left) and after the final selection stage (right).
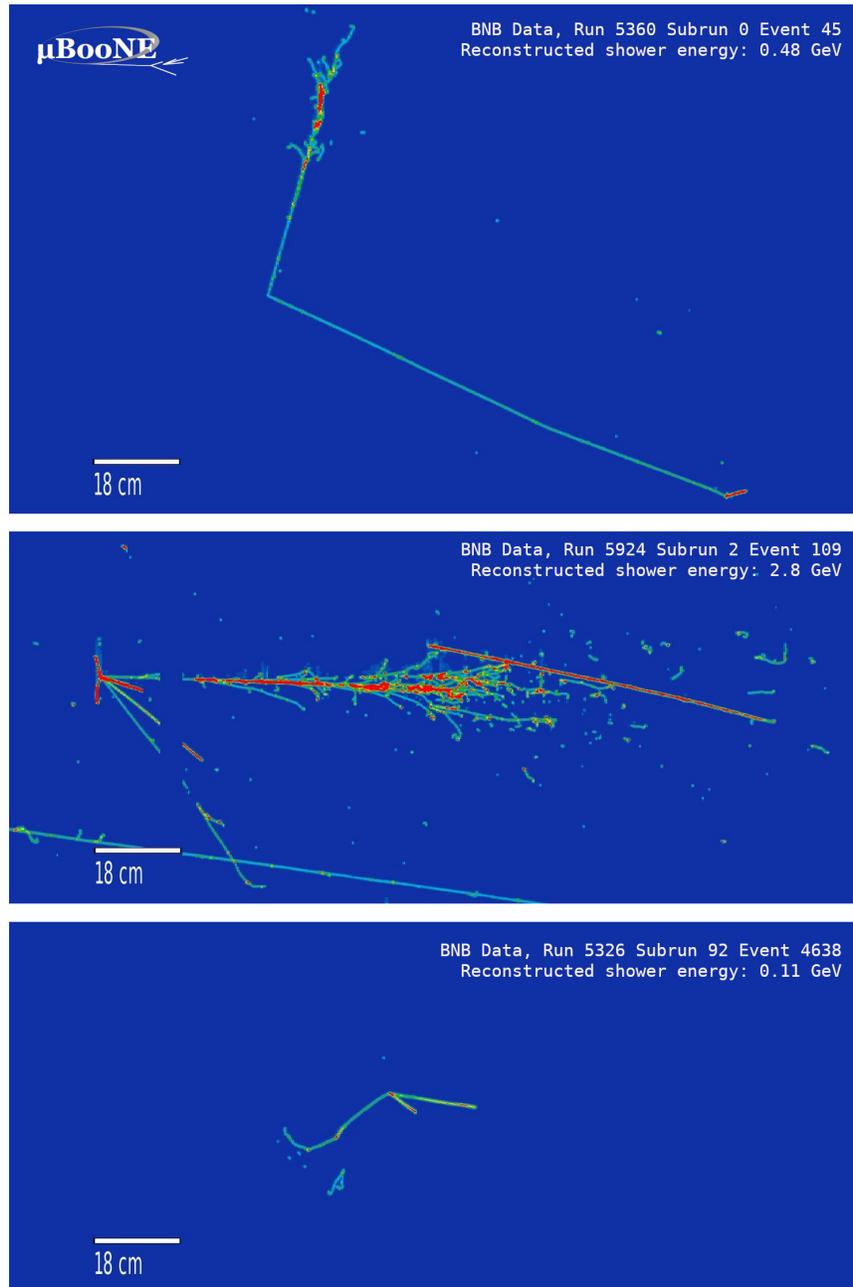
Figure 29: Collection-plane displays of selected events by the ν_e CC selection. Note the differences in the reconstructed angle and energy of the electromagnetic shower in the three events, demonstrating the wide range of kinematics covered by the selection.

B.3 High-Energy Sideband

In this section, a sideband containing neutrinos with a reconstructed energy above $1.05\,\mathrm{GeV}$ is studied. The exposure corresponds to $5.8 \times 10^{20}\,\mathrm{POT}$, roughly $50\,\%$ of the collected data. Figure 30 shows the two variables with the strongest $e - \gamma$ separation power after the *pre-selection*.

As stated in Appendix B.2, the selection does not rely directly on the kinematics of the event, and is therefore suited to explore the kinematics of ν_e candidate events. The resolution, obtained from simulation and the data-simulation comparisons for the three kinematic variables of the electron after selection are given in 31 and 10 respectively.

After the unblinding of a significant data-set, the inclusive ν_e CC selection described in this section will serve as a validation of the electron neutrino modelling in MicroBooNE and be used to measure the intrinsic electron flavour component in the Booster Neutrino Beam.
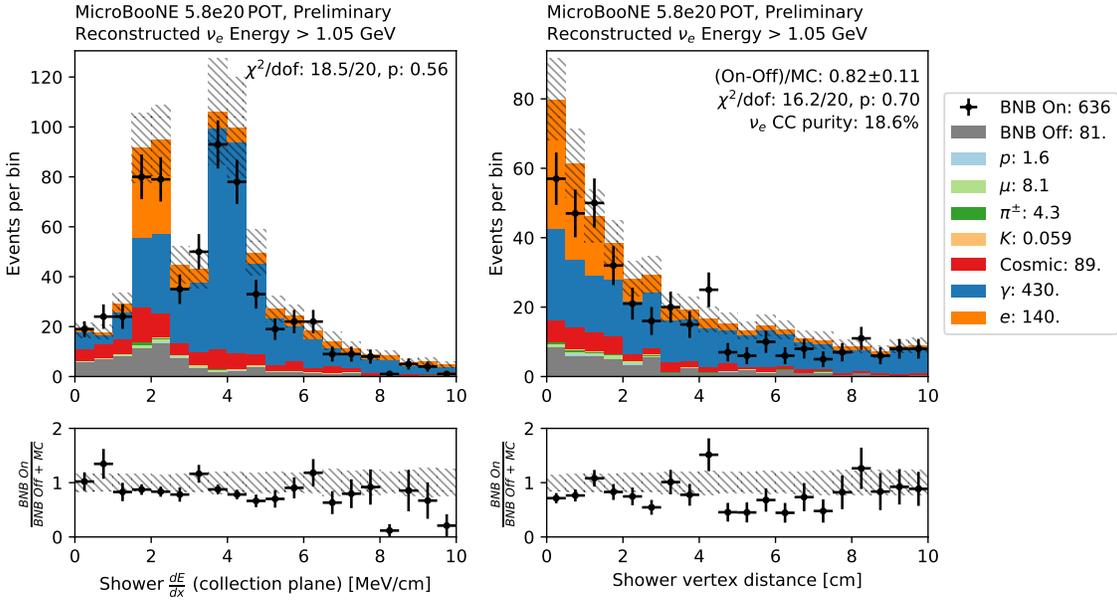

Figure 30: Data/MC comparisons for the two variables of the electron shower classifier with the strongest separation power: The $\mathrm{d}E/\mathrm{d}x$ at the start of the shower (left) and the shower vertex distance (right). The used data-set corresponds to the high-energy sideband $5.8 \times 10^{20}\,\mathrm{POT}$ and contains neutrino candidates with a reconstructed energy above $1.05\,\mathrm{GeV}$.
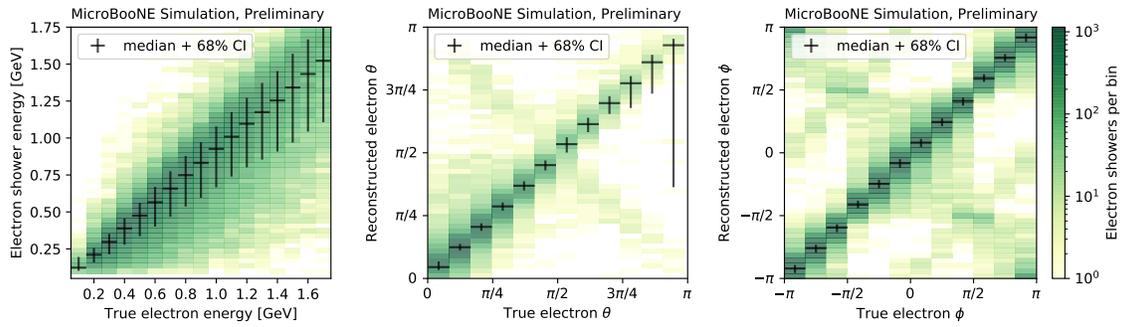
(a) Resolution for the reconstructed electron variables after the ν_e CC selection. The colour scale is logarithmic. In black, the median and 68 % confidence interval are given, binned in the simulated electron kinematics.

Figure 31: Resolution of ν_e CC electron kinematic variables.

C Constraining Systematic Uncertainties for the Electron Neutrino Search at MicroBooNE

MicroBooNE is designed to investigate the anomalous excess in the electron neutrino event spectrum observed by the MiniBooNE experiment. MicroBooNE runs in the same Booster Neutrino Beamline (BNB) as MiniBooNE and is located at a baseline of 470 m. To achieve the level of sensitivity to probe this low energy excess, it is necessary to ensure a significant reduction of uncertainties related to the expected neutrino energy spectrum. This section outlines the strategy to constrain the systematic uncertainties in electron neutrino modeling using a high-statistics muon neutrino sample. It also discusses the different sources of systematic uncertainties, the constraint approach that accounts for known correlations between different electron final-states and muon neutrino fluxes and cross-sections, and the implementation of the constraint given this correlation to improve the sensitivity.

C.1 Event Selection

The strategy to constrain the systematic uncertainties leverages a variety of channels in the ν_e selection. The ν_e selection is divided into two final state topologies:

- $1e\mathrm{N}p0\pi$: One single electron shower and at least one proton in the final state and no pions above 40 MeV kinetic energy. This selection is the most sensitive to the unfolded LEE signal (see Fig.9a)

- $1e0p0\pi$: One single electron shower and no protons or pions in the final state above 40 MeV kinetic energy. This selection is utilised to constrain the less understood protons kinematics and reconstructions (see Fig. 9b)

A high-statistics ν_μ sample and the correlations to the ν_e sample is leveraged to constrain flux and neutrino cross-section uncertainties. These correlations are due to the ν_μ and ν_e sharing the same neutrino parents from the BNB and interacting via the same mechanism in the liquid argon medium. The ν_μ's selection is designed to maximize its phase-space overlap with the ν_e's selection and to prioritize low-energy performance where the constraining power is most needed. Requiring contained tracks, for example, filters out higher energy muons and ensures the use of both range-based and multiple-coulomb scattering-based calorimetry techniques on muon tracks for higher precision energy reconstruction (see Figure 12a).

C.2 Systematics and Constraint

Systematic uncertainties in this analysis are associated to the level of precision of the model used in the event generation (argon cross section/GENIE ref. [15]), to the knowledge of the neutrino beam (flux), and to the limitation of our detector simulation (detector response).

Two approaches are used to estimate the uncertainties: unisim and multisim.

- Unisim: a single variation of given analysis input parameter. The difference between number of events in the variation and the central value is taken as the uncertainty in that bin. This approach is implemented to estimate the detector response systematics by using dedicated variation samples. These systematics are treated as uncorrelated.

- Multisim: several variations of given analysis input parameter, called universes. The correlations in the various bins of the distributions of the selected events are preserved. This approach is implemented when calculating the flux and cross-section uncertainties.

Systematics which impact TPC's responses can impact calorimetric energy reconstruction important for shower, and thus ν_e event energy reconstruction. The impacts of these systematics are generally at the few percent as shown in Fig 32.
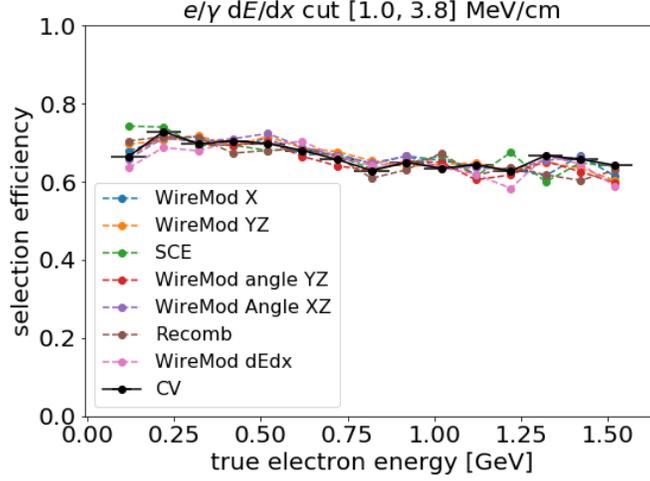

Figure 32: This figure displays the impact of the TPC variation on the selection efficiency as a function of $\mathrm{d}E/\mathrm{d}x$ for true backtracked electrons. TPC variations which impact spatial and calorimetric reconstruction can impact particle ID discriminant variables such as the `LLR trkpid` used for μ/p separation and $\mathrm{d}E/\mathrm{d}x$ used for e/γ separation. The impact of TPC detector variations (after $1e\mathrm{N}p0\pi$ pre-selection) on electron PID cuts in $1e\mathrm{N}p0\pi$ box-cuts is subdominant overall at few percent variation.

For each uncertainty, a covariance matrix correlating the variation "sys" universe $k = 1, \ldots, N$ in the measured number of events (N_i^k) and the central value (N_i^{CV}) between bins is calculated as follows:

$$V_{ij}^{sys} = \frac{1}{N} \sum_k^N (N_i^{CV} - N_i^k)(N_j^{CV} - N_j^k), \tag{1}$$

where only the diagonal elements are considered to calculate the systematics in the unisim approach.

The constraint is exercised through the covariance matrix constructed from the energy spectrums of the ν_e and ν_μ selections described above. Figure 33 displays the covariance matrix between $1e\mathrm{N}p0\pi$ ν_e, $1e0p0\pi$ ν_e, and ν_μ channels for the full systematics accounting flux, cross-section, and detector response systematics. The correlation between the elements in the matrix allow to constrain ν_e flux and cross-section modelling uncertainties through the measurements of correlated ν_μ events as shown in Figure 34.
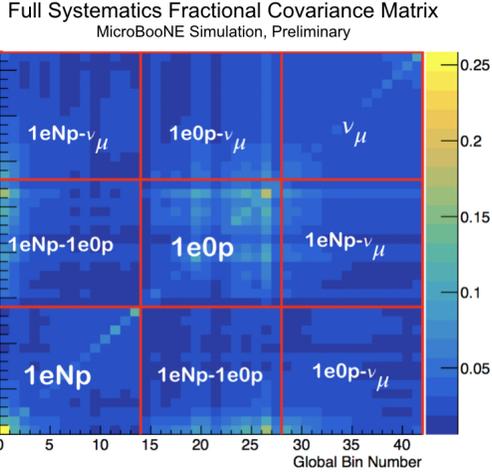
Figure 33: Total uncertainties of the flux modelling, neutrino interaction modelling, and detector modeling systematics. The detector response systematics are treated as uncorrelated and increase the size of the fractional errors of the diagonal elements. This has the effect of degrading the power of the correlation between the ν_μ channels and the $1eNp0\pi$ and $1e0p0\pi$ channels.
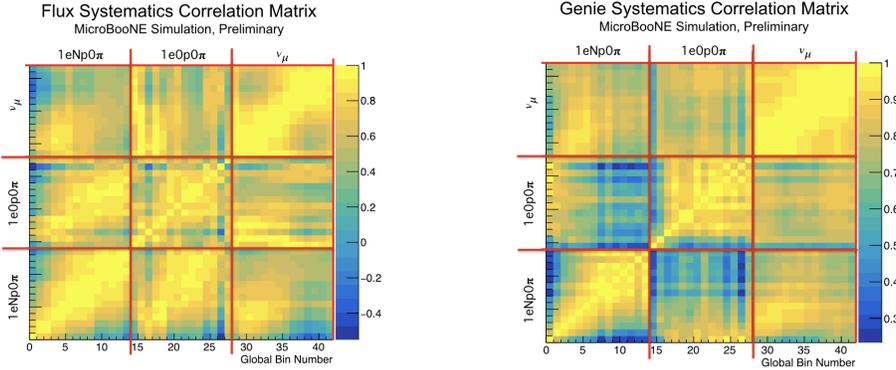

(a) Flux Correlation Matrix

(b) Cross-section (GENIE) Correlation Matrix

Figure 34: (a) Flux modelling systematics has the highest correlation in the low-energy, where the uncertainties from model parameterization for short Berilium target π^+ production cross section dominates. (b) Significant uncertainty in ν_μ/ν_e CC cross section models due to limited experimental data in low neutrino energy. Good correlation at very low energy between $1eNp0\pi$ and $1e0p0\pi$ channels selections provided additional constraint in the combined fit.

The maximum systematics constraint can be estimated using the correlation elements between

the ν_e and ν_μ events as follows

1. Minimize the χ^2 between the data and MC in selections by using the covariance matrix inversion, $(V_{ij})^-1$.

2. Assume that the observed events (N_{Data}) is equal to the predicted events (N_{MC}) within the statistical uncertainties, and fit the N_{MC} to the statistical uncertainties in each bin of the spectrum, as follows

$$(V_{ij}^{-1})^{new} = V_{ij}^{-1} + \frac{1}{N_i^{MC}}. \tag{2}$$

3. Invert back the new inverse matrix, $(V_{ij}^{-1})^{new}$ to give the new constrained matrix, $(V_{ij})^{new}$. The new constrained matrix is plotted as the red shade in Figure 13.

This method demonstrates the best estimation of the constraint power that we can expect from utilizing the ν_μ constraint. Overall, systematics reduction of up to 50% in flux, cross section, and detector response systematics at the low energy region where it is most needed. The impact of the constraint is shown on Figure 13.

C.3 Sensitivity Estimation

The sensitivity of the selection to the MiniBooNE unfolded LEE signal is calculated by generating toy experiments under two hypotheses, which definitions can be used interchangeably:

- H_0 is the Standard Model hypothesis and H_1 is the MiniBooNE unfolded LEE hypothesis.

- H_0 is the MiniBooNE unfolded LEE hypothesis and H_1 is the Standard Model hypothesis.

The median, $\pm 1\sigma$, $\pm 2\sigma$ of the sensitivity are calculated relative to the H_1. Following this convention, a test statistic is employed to rule out the H_0 hypothesis in favor of the H_1 hypothesis.

The calculation uses the χ^2_{CNP} formalism [14] that approximates Poisson statistical errors for the covariance matrix diagonals and it is defined as follows:

$$(\chi^2_{CNP})_{COV} = (N - \mu(H))^T (V_{CNP}^{stat} + V^{syst})^{-1} (N - \mu(H)), \tag{3}$$

where N is the observed number of entries in a bin, $\mu(H))$ is the expected number of entries under the H_0 or H_1 hypothesis, V^{syst} is the systematic covariance matrix on the H_0 hypothesis, and the statistical covariance matrix is defined as

$$V^{stat}_{CNP\,ij} \equiv 3 / \left(\frac{1}{N_i} + \frac{2}{\mu_i(H)} \right) \delta_{ij}. \tag{4}$$

The median sensitivity of the selection is estimated for the current open dataset of 6.9×10^{20} POT and the full dataset of 12.5×10^{20} POT. The systematics uncertainties is only assumed for the Standard Model (SM) hypothesis, and no systematics uncertainties is attached to the MiniBooNE LEE signal model. These estimates are reported considering the following scenarios:

- The median sensitivity projected for 6.9×10^{20} POT to rule out the Standard Model (SM) hypothesis in favor of the LEE hypothesis is 2.3σ, shown in Figure 35

- The median sensitivity projected for 6.9×10^{20} POT to rule out the LEE hypothesis in favor of the SM hypothesis is 1.9σ.

- The median sensitivity projected for 12.5×10^{20} POT to rule out the SM hypothesis in favor of the LEE hypothesis is 3.0σ shown in Figure 36.

- The median sensitivity projected for 12.5×10^{20} POT to rule out the LEE hypothesis in favor of the SM hypothesis is 2.4σ.
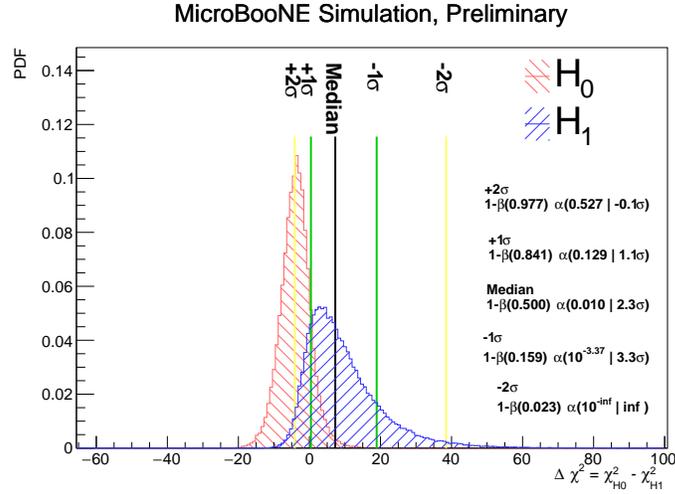

Figure 35: Sensitivity estimate to the MiniBooNE LEE unfolded signal on the combined $1e$N$p0\pi$ and $1e0p0\pi$ selection with the ν_μ constraint for an integrated POT of 6.9×10^{20}. In this plot, H_0 is the Standard Model hypothesis and H_1 is the MiniBooNE unfolded LEE hypothesis

Leveraging multiple selections in the fit brings an overall improvement to the sensitivity. For the current open dataset of 6.9×10^{20} POT, incorporating the $1e0p0\pi$ with the $1e$N$p0\pi$ in the combined fit improves the median sensitivity to rule out the standard model in favor of the LEE from 1.8σ to 2.0σ. And including the ν_μ selection in the combined fit, we observe an improvement from 2.0σ to 2.3σ.
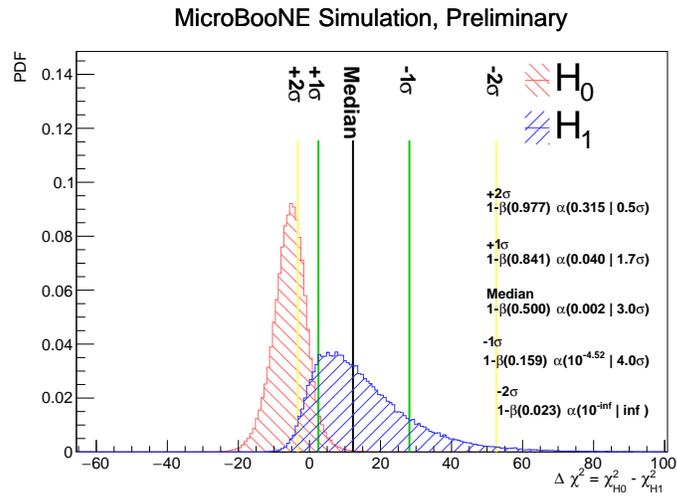
Figure 36: Sensitivity estimate to the MiniBooNE LEE unfolded signal on the combined $1eNp0\pi$ and $1e0p0\pi$ selection with the ν_μ constraint for an integrated POT of 12.5×10^{20}. In this plot, H_0 is the Standard Model hypothesis and H_1 is the MiniBooNE unfolded LEE hypothesis

D   Measurement of Single Electron Events in MicroBooNE

Electron neutrinos can create a wide variety of topologies when interacting in liquid argon, and this analysis measures events both with (1eNp0π) and without (1e0p0π) visible protons. This section presents the measurement of single electron events in the MicroBooNE detector, which includes events at low energies and are important for understanding the excess observed by MiniBooNE. The single electron selection is orthogonal to the 1eNp0π selection which makes it possible to allow migrations between the two channels in a joint fit, and constrain uncertainties associated with low energy protons such as those related to reconstruction, multiplicity and their kinematics. A selection of plots summarizing the development of the $1e0p0\pi$ selection, and not already included in the main body of this document, are shown in Fig. 37.

(a) Reconstructed Visible Energy

(b) BDT Response
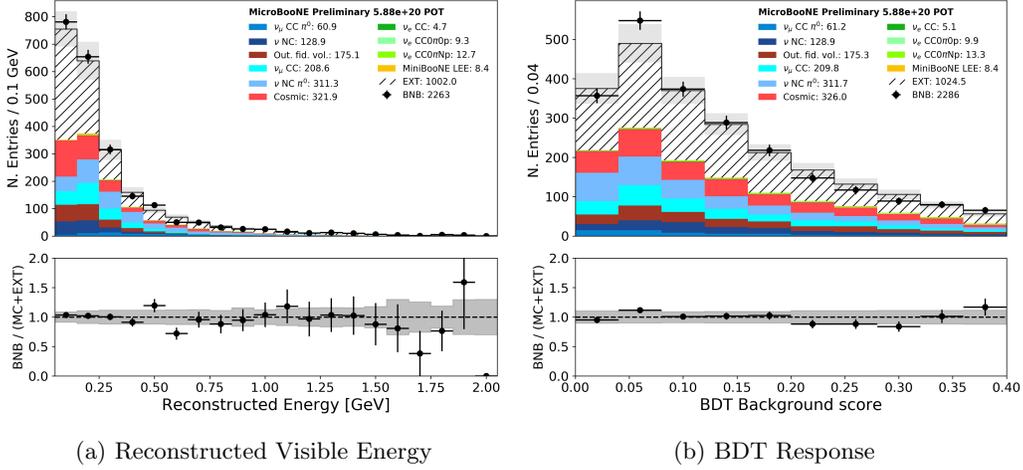
Figure 37:   Reconstructed visible energy and BDT score in the low BDT score region of the far sideband after the $1e0p0\pi$ pre-selection stage. Additional cuts were applied to reduce the EXT contribution. Simulation in agreement with data measurement.

In liquid argon, a minimum ionizing particle (MIP) has an energy loss per unit distance about of 2 MeV/cm [16]. Separation of showers of one MIP (mostly $1e0p0\pi$ events) and showers of two MIPs (mostly NCπ^0 events) is observed in 14b. The selection framework shows great capability at selecting true $1e0p0\pi$ events in the high energy sideband region, as can be seen from the event displays in Fig. 11.

37
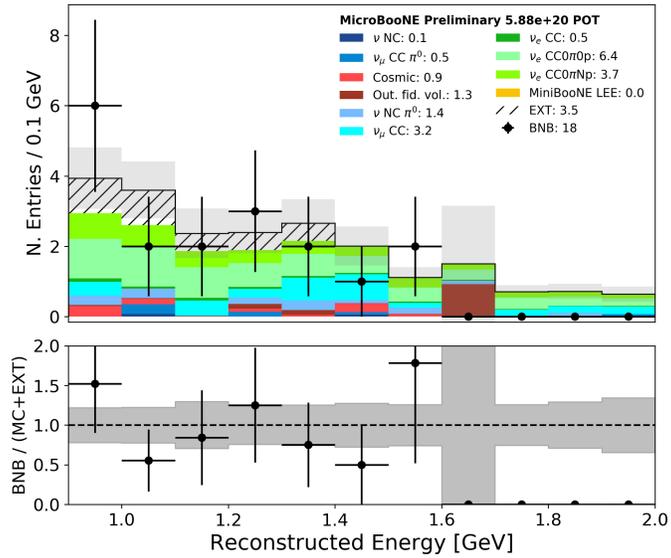
(a) High Energy and BDT score > 0.4

Figure 38: Selection of single shower events in the high energy region (> 0.9 GeV) and a BDT score > 0.4. This is at the $1e0p0\pi$ pre-selection and after applying cuts on the shower/track fitter start and end points in Y.

E ν_e 1eNp Event Displays

This section shows a collection of event displays for selected 1eNp events. In the images, reconstructed showers are displayed as cones, while tracks are displayed as lines. The recosntructed neutrino interaction vertex is represented with a white dot, and has been slightly offset from the actual point for better readability.
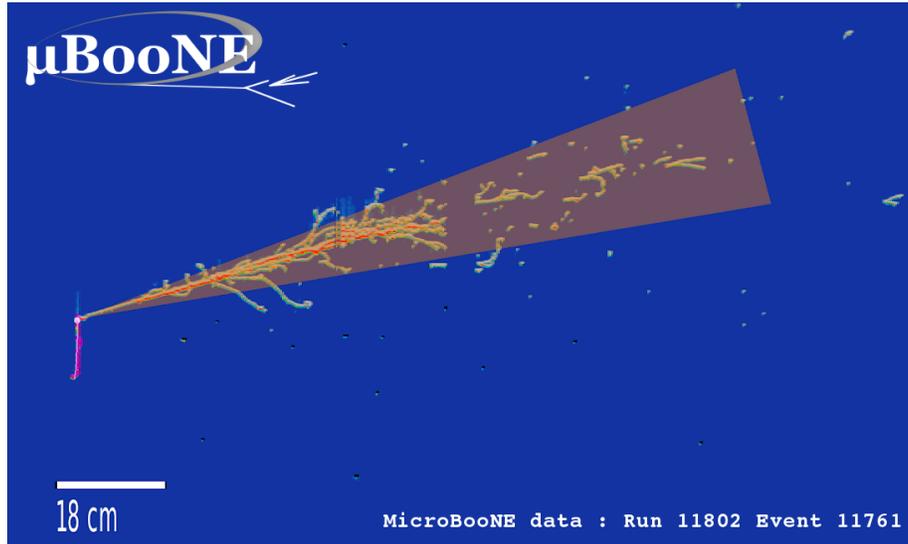
Figure 39: Selected 1eNp ν_e candidate from high-energy (> 1.05 GeV) sideband.
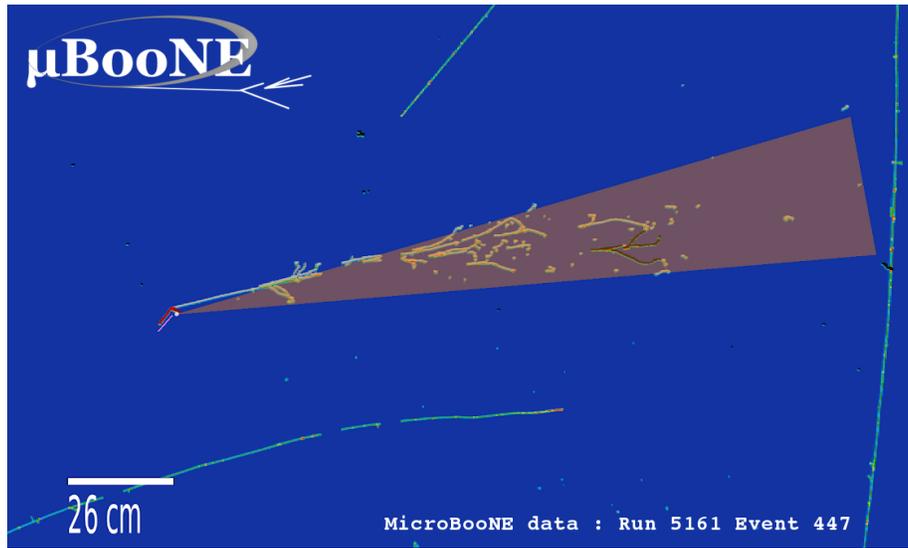
Figure 40: Selected $1eNp$ ν_e candidate from high-energy (> 1.05 GeV) sideband.
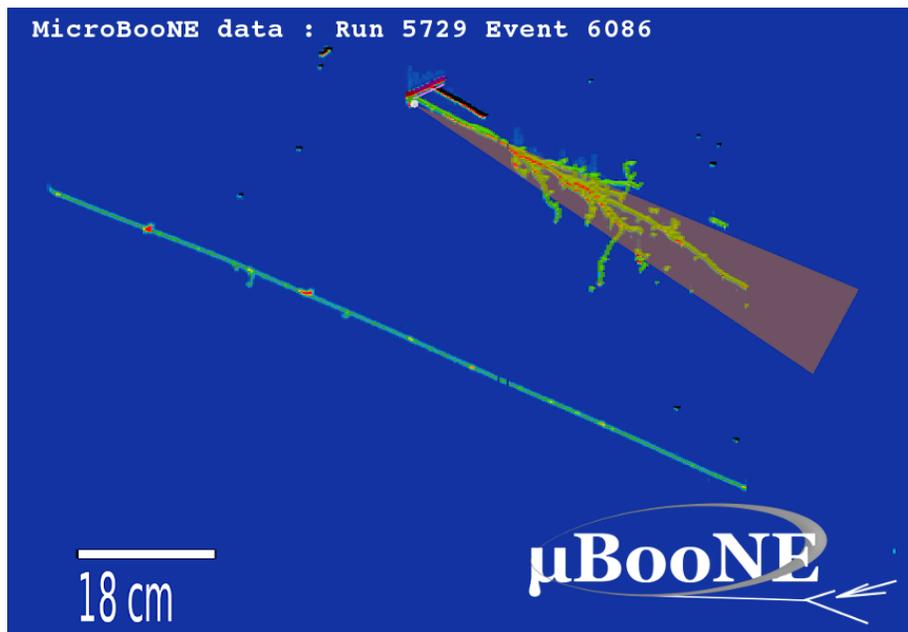

Figure 41: Selected $1eNp$ ν_e candidate from high-energy (> 1.05 GeV) sideband.
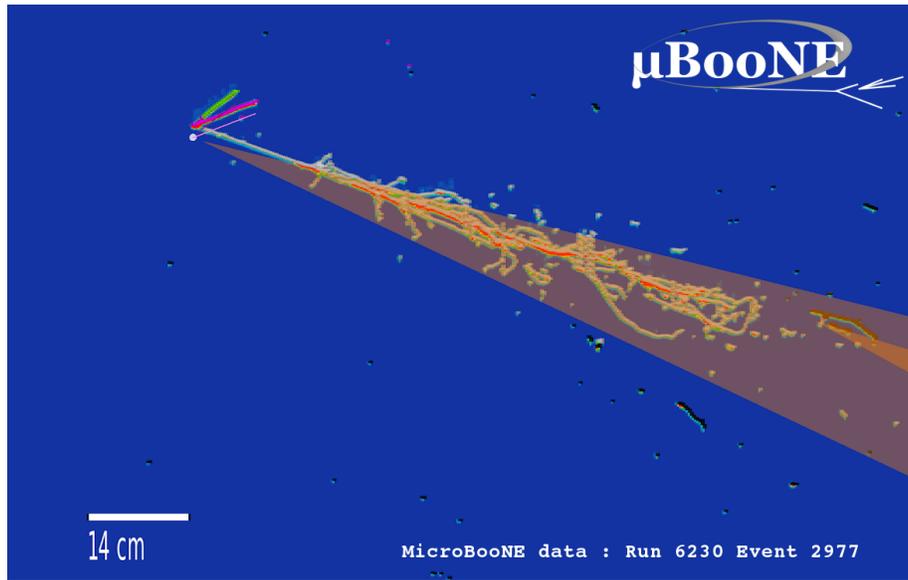
Figure 42: Selected $1e$Np ν_e candidate from high-energy (> 1.05 GeV) sideband.

Figure 43: Selected $1e$Np ν_e candidate from high-energy (> 1.05 GeV) sideband.

41
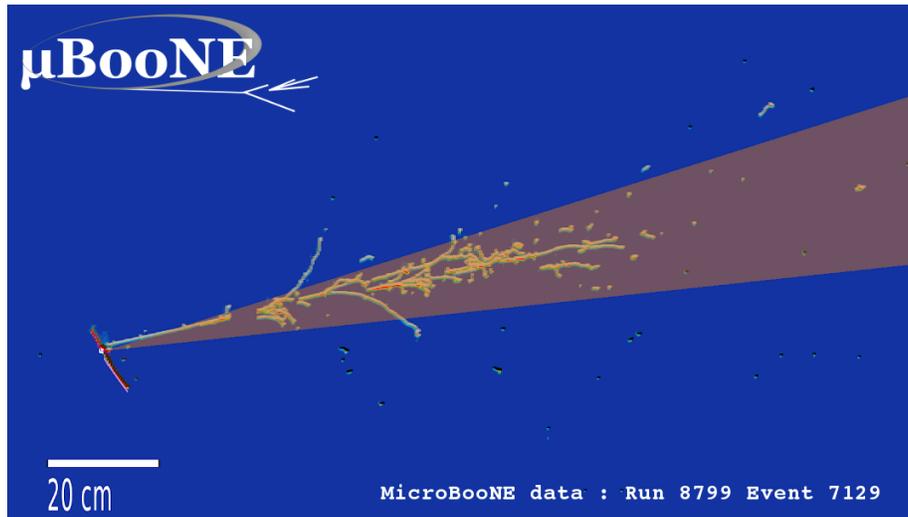
Figure 44: Selected $1eNp$ ν_e candidate from high-energy (> 1.05 GeV) sideband.
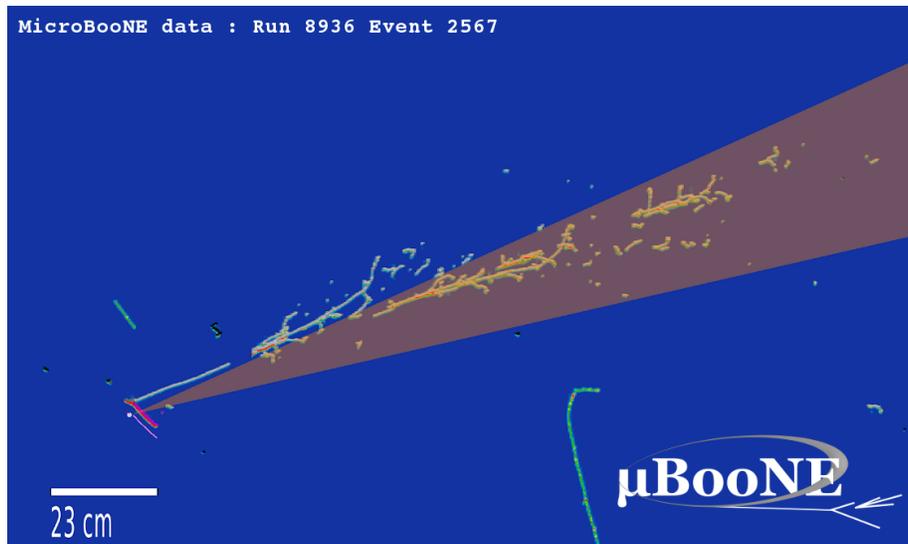
Figure 45: Selected $1eNp$ ν_e candidate from high-energy (> 1.05 GeV) sideband.
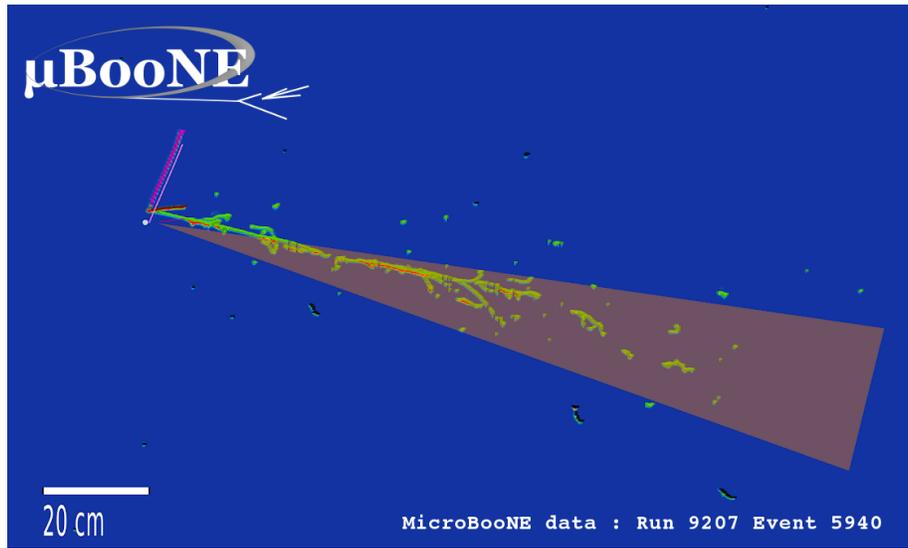
Figure 46: Selected $1eNp$ ν_e candidate from high-energy (> 1.05 GeV) sideband.

References

[1] A. A. Aguilar-Arevalo et al. "Significant Excess of Electronlike Events in the MiniBooNE Short-Baseline Neutrino Experiment". In: *Phys. Rev. Lett.* 121 (22 Nov. 2018), p. 221801. DOI: 10.1103/PhysRevLett.121.221801. URL: https://link.aps.org/doi/10.1103/PhysRevLett.121.221801.

[2] MicroBooNE Collaboration. "Design and Construction of the MicroBooNE Detector". In: *JINST* 12.P02017 (2017). URL: https://arxiv.org/abs/1612.05824.

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

[4] MicroBooNE Collaboration. "The Pandora multi-algorithm approach to automated pattern recognition of cosmic-ray muon and neutrino events in the MicroBooNE detector". In: *Eur. Phys. J.* C.78 (2018), p. 1. URL: https://arxiv.org/abs/1708.03135.

[5] MicroBooNE Collaboration. *Booster Neutrino Flux Prediction at MicroBooNE.* URL: https://microboone.fnal.gov/wp-content/uploads/MICROBOONE-NOTE-1031-PUB.pdf.

[6] MicroBooNE Collaboration. "Reconstruction and Measurement of $\mathcal{O}$O(100) MeV Energy Electromagnetic Activity from $\pi^0 \to \gamma\gamma$ Decays in the MicroBooNE LArTPC". In: *JINST* 15 (2020), P02007. DOI: 10.1088/1748-0221/15/02/P02007.

[7] MicroBooNE Collaboration. "Calibration of the charge and energy loss per unit length of the MicroBooNE liquid argon time projection chamber using muons and protons". In: *JINST 15 (2020) 03, P03022* (). DOI: 10.1088/1748-0221/15/03/P03022. URL: https://arxiv.org/abs/1907.11736.

[8] MicroBooNE Collaboration. "A Method to Determine the Electric Field of Liquid Argon Time Projection Chambers Using a UV Laser System and its Application in MicroBooNE". In: *JINST (accepted)* (). DOI: FERMILAB-PUB-19-496-ND. URL: https://arxiv.org/abs/1910.01430.

[9] MicroBooNE Collaboration. *Neutrino Interaction Model and Uncertainties for MicroBooNE Analyses.* URL: https://microboone.fnal.gov/wp-content/uploads/MICROBOONE-NOTE-1074-PUB.pdf.

[10] MicroBooNE Collaboration. *Novel Approach for Evaluating Detector Systematics in the MicroBooNE LArTPC.* URL: https://microboone.fnal.gov/wp-content/uploads/MICROBOONE-NOTE-1075-PUB.pdf.

[11] M. Tanabashi et al. "Review of Particle Physics". In: *Phys. Rev. D* 98 (3 Aug. 2018), p. 030001. DOI: 10.1103/PhysRevD.98.030001. URL: https://link.aps.org/doi/10.1103/PhysRevD.98.030001.

[12] Tianqi Chen and Carlos Guestrin. "XGBoost: A Scalable Tree Boosting System". In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining.* KDD '16. San Francisco, California, USA: ACM, 2016, pp. 785–794. ISBN: 978-1-4503-4232-2. DOI: 10.1145/2939672.2939785. URL: http://doi.acm.org/10.1145/2939672.2939785.

[13] Jerome H. Friedman. "Greedy Function Approximation: A Gradient Boosting Machine". In: *The Annals of Statistics* 29.5 (2001), pp. 1189–1232. ISSN: 00905364. URL: http://www.jstor.org/stable/2699986.

[14] Xiangpan Ji et al. "Combined Neyman–Pearson chi-square: An improved approximation to the Poisson-likelihood chi-square". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 961 (May 2020), p. 163677. ISSN: 0168-9002. DOI: 10.1016/j.nima.2020.163677. URL: http://dx.doi.org/10.1016/j.nima.2020.163677.

[15] Costas Andreopoulos et al. *The GENIE Neutrino Monte Carlo Generator: Physics and User Manual*. 2015. arXiv: 1510.05494 [hep-ph].

[16] MicroBooNE Collaboration. "Calibration of the charge and energy loss per unit length of the MicroBooNE liquid argon time projection chamber using muons and protons". In: *Journal of Instrumentation* 15.03 (Mar. 2020), P03022–P03022. DOI: 10.1088/1748-0221/15/03/p03022. URL: https://doi.org/10.1088.