Search for a 3+1 Sterile Neutrino with the MicroBooNE Experiment using Deep-Learning-Based Reconstruction

MICROBOONE-NOTE-1105-PUB

The MicroBooNE Collaboration
MicroBooNE_Info@fnal.gov

May 24, 2022

Abstract

In this note we present the methods and sensitivity for a search for sterile neutrinos based on the 3+1 model in the MicroBooNE experiment. The recently released results by MicroBooNE show no sign of the MiniBooNE/LSND low-energy-excess anomaly. The 3+1 model examined here expands on the standard model of neutrinos by adding a fourth neutrino flavor and is not necessarily ruled out by the lack of a low energy excess. The search presented relies on Deep-Learning-based reconstruction tools and looks for charged current quasi-elastic-like (CCQE-like) events kinematically consistent with a 2-body interaction. Using two orthogonal samples of CCQE electron neutrino events and CCQE muon neutrino events, we test this model allowing for electron neutrino appearance, electron neutrino disappearance, and muon neutrino disappearance. We present the sensitivity to the oscillation parameters using the Wilks' theorem exclusion confidence levels. We demonstrate the use of a minimizer to find the best fit and discuss the effect of using a Feldman-Cousins' procedure instead of Wilks' theorem to determine sensitivity.

1 The MicroBooNE Detector

MicroBooNE, located in Batavia, Illinois at the Fermilab National Accelerator Laboratory (FNAL), is a short-baseline accelerator neutrino experiment [1]. One of the major goals is to investigate the MiniBooNE low energy anomaly [2]. MicroBooNE receives neutrinos from the Booster Neutrino Beam (BNB). It has a 470 m baseline, and utilizes the same beam as MiniBooNE. The MicroBooNE detector consists of a liquid argon time projection chamber (LArTPC), a light collection system, and an external cosmic ray tagging device. The TPC employs three wire readout planes which enables 3D position reconstruction. As charged particles traverse a volume of liquid argon, they leave trails of ionization electrons. A uniform electric field in the detector volume causes these electrons to drift to one side of the chamber. The anode is parallel to the cathode plane and parallel to the beam direction. The applied

electric field is 273 V/cm, leading to $\approx 0.11\,\mathrm{cm}/\mu\mathrm{s}$ electron drift velocity. The plane-to-plane spacing is 3 mm and the wires in each plane are separated by 3 mm. The collection plane is oriented vertically and the induction planes are at $\pm 60^\circ$ relative to the vertical [1]. The light collection system is a PMT optical system. It collects scintillation light produced in the interaction and is used for beam triggering. The cosmic ray tagger is not utilized in this note as it was not active during much of the data taking for the data used here [1].

2 3+1 Sterile Neutrino Model

One proposed solution to the MiniBooNE anomaly is a fourth light sterile neutrino in addition to the three standard model neutrino flavors [3]. These sterile neutrinos are assumed to have no weak interactions and a mass on the order of 0.1 eV - 10 eV. The mass state is also assumed to have small weak flavor content. These assumptions allow for the currently observed anomalies at LSND and MiniBooNE which feature small oscillation-like signals at a relatively small L/E ($\sim$ 1 m/MeV) [2, 4]. The assumption is made that the three standard model mass splittings degenerate at 0 ev^2 relative to the sterile neutrino mass splitting, as they are known to be $\sim 10^{-3}$ and $\sim 10^{-5} eV^2$. In order for a neutrino to be able to explain the short-baseline anomaly the sterile neutrino mass splitting should be $\sim 1 eV^2$.

$$\Delta m^2_{21} \approx \Delta m^2_{31} \approx \Delta m^2_{32} \approx 0(eV)^2 \tag{1}$$

This allows the oscillation probability equations to be written in a simplified form. This analysis is sensitive to 3 oscillation channels: $\nu_\mu \to \nu_e$ (ν_e appearance), $\nu_e \to \nu_e$ (ν_e disappearance), and $\nu_\mu \to \nu_\mu$ (ν_μ disappearance). The ν_e flux from the BNB is not large enough to observe ν_μ appearance with any statistical significance. The oscillation probabilities are:

$$\begin{aligned} P(\nu_\mu \to \nu_e) &= 4|U_{\mu 4}|^2 |U_{e4}|^2 \sin^2 x_{41} \\ &= \sin^2(2\theta_{\mu e}) \sin^2 x_{41} \end{aligned} \tag{2}$$

$$\begin{aligned} P(\nu_e \to \nu_e) &= 1 - 4|U_{e4}|^2 |(1 - |U_{e4}|^2) \sin^2 x_{41} \\ &= 1 - \sin^2(2\theta_{ee}) \sin^2 x_{41} \end{aligned} \tag{3}$$

$$\begin{aligned} P(\nu_\mu \to \nu_\mu) &= 1 - 4|U_{\mu 4}|^2 |(1 - |U_{\mu 4}|^2) \sin^2 x_{41} \\ &= 1 - \sin^2(2\theta_{\mu\mu}) \sin^2 x_{41} \end{aligned} \tag{4}$$

where $x_{41} = \frac{L\Delta m^2_{41}}{E}$, L is the neutrino travel distance and E is the energy of the neutrino. There are three independent parameters that will be fit over: $|U_{e4}|$, $|U_{\mu 4}|$, and Δm^2_{41} where $|U_{e4}|$ and $|U_{\mu 4}|$ are the explicit PMNS matrix terms. The ranges tested in this analysis are: Δm^2_{41}, 25 logarithmic spaced values [0.01,100] MeV^2; U_{e4}, 25 logarithmic spaced values [0.01,0.5]; and $U_{\mu 4}$, 25 logarithmic spaced values [0.01,0.5]. L for each neutrino is calculated from the 'true' value generated in the MicroBooNE simulations. The E used when scaling the simulation is the reconstructed neutrino energy.

3 CCQE-like $1l1p$ Event Reconstruction and Selection

This study utilizes an event selection which looks for charged current quasi-elastic (CCQE-like) events kinematically consistent with a 2-body interaction with one lepton and one proton ($1l1p$) in the final state. The agreement to the 3+1 sterile neutrino model will be tested using a joint selection of two orthogonal samples of CCQE electron neutrino events ($1e1p$) and CCQE muon neutrino events (1μ1p). This reconstruction chain and the event selections have already been used to search explicitly for the low-energy excess anomaly seen in MiniBooNE and rejected the MiniBooNE signal at 2.4σ [5, 6]. The purity for the correct flavor of neutrinos regardless of the interaction type for the two selections are 90.7% and 97.8% for the $1e1p$ and $1\mu1p$ selections respectively. This purity metric is important as all interactions of a certain neutrino flavor will be oscillated in the same way. The efficiencies are 6.6% and 3.27% for the $1e1p$ and $1\mu1p$ selections respectively.

The data that is used in this study was taken from 2016 to 2018, and recorded over three run periods with 6.67×10^{20} total protons-on-target (POT) in the three runs. The simulation samples are scaled to match this POT.

The kinematic variable this oscillation analysis is performed in is neutrino energy, which is calculated as:

$$E_\nu^{\text{range}} \;=\; \mathrm{K}_p + \mathrm{K}_\ell + M_\ell + M_p - (M_n - B) \tag{5}$$

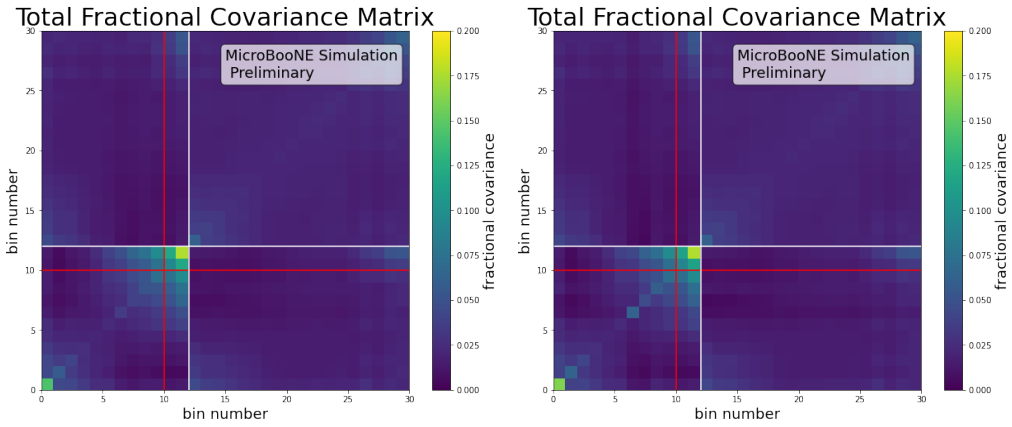
where K is kinetic energy determined from the reconstructed charged lepton (ℓ) or proton (p); M is mass; p is the proton in the nucleus; n is the neutron in the nucleus; and B is the average binding energy, assumed to be 40 MeV [7].

In order to search for the MiniBooNE low-energy excess in Ref. [5], the neutrino energy binning for the 1e1p selection was set to be 200-1200 MeV in 100 MeV bins. The $1e1p$ binning has been expanded for this analysis to capture any high energy oscillation behavior. It has been expanded to: [200-1200] in 100 MeV bins + [1200-1600] in 1 400 MeV bin + [1600-2400] in 1 800 MeV bin.

The uncertainties are incorporated into this analysis using the covariance matrix formalism. The systematic uncertainties are the same as that presented in Ref [5]. However, there are a few key differences from the previous analysis that affect the systematic uncertainties. These are: a change in 1e1p range to allow for sensitivity to higher energy oscillations; changing the $1\mu1p$ selection to include high-statistics π^0 samples (change is within simulation statistical error); and removal of the background fitting performed on the ν_μ background to the $1e1p$ selection (the largest change). The total covariance is regenerated for each oscillation parameter set.

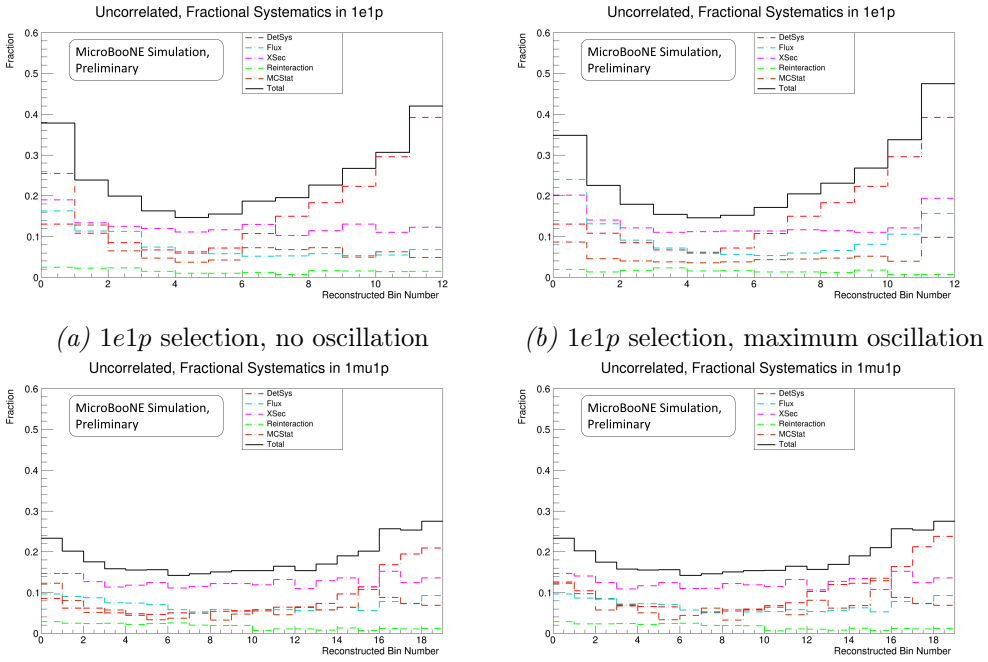
The total fractional covariance matrix is shown in Figure 1 for both no oscillation and maximum oscillation. The contributions of each type of uncertainty to the total spectrum is shown in Figure 2. This has been broken up into the $1e1p$ and $1\mu1p$ spectra.

Fig 3 shows the results of applying the selection to both simulation and data as used in Ref [5]. The simulation has no sterile neutrino signature in this figure. Note, the spectrum shown in Fig 3 (a) does not have the two extra high energy bins as this is from a previous data/simulation comparison and data is not yet included in this study.



(a) null model *(b)* maximum oscillation

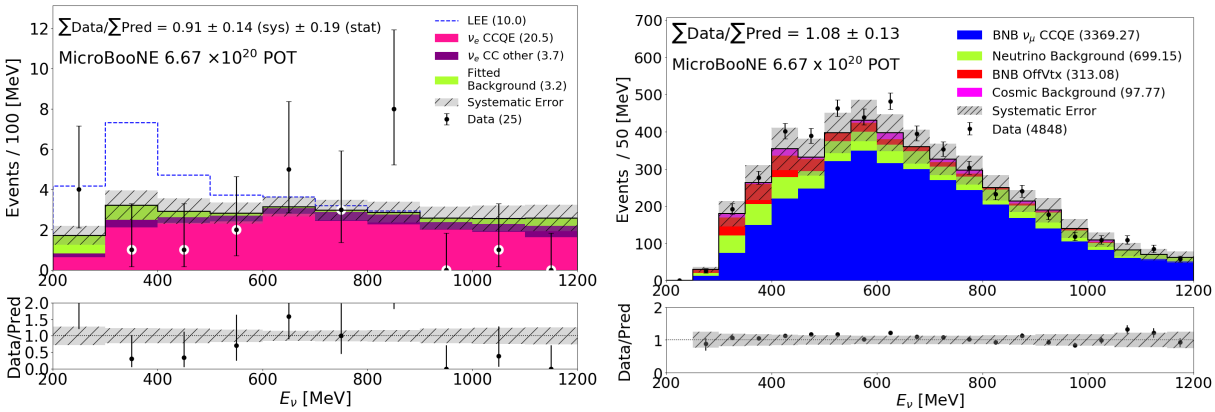
Figure 1: Total fractional covariance matrices for (a) the null model and (b) maximum oscillation. Bins 0-9 are the original bins of the 1e1p selection (indicated by red line). Bins 10-11 are the additional high energy 1e1p bins (indicated by white line). Bins 12-29 are the $1\mu1$p bins.



(a) 1e1p selection, no oscillation *(b)* 1e1p selection, maximum oscillation

(c) $1\mu1p$ selection, no oscillation *(d)* $1\mu1p$ selection, maximum oscillation

Figure 2: Systematic uncertainties on the 1e1p selection (a,b) and the $1\mu1p$ selection (c,d) broken up by contribution type. The y axis has been set to the same value (0.6) in all plots.



(a) 1e1p *(b)* 1e1p

Figure 3: Predicted E_ν^{reco} spectrum for the $1e1p$ (a) and $1\mu1p$ (b) selections, with the data selection represented by black points. Adapted from Ref [5].

4 Sensitivity to the 3+1 Model

This analysis combines the $1e1p$ selection with the $1\mu1p$ selection in a joint fit and tests three oscillation channels concurrently. ν_e appearance, ν_e disappearance, and ν_μ disappearance are all allowed to occur. To find the best fit oscillation parameters in this analysis for a given x, we seek the maximum likelihood, or equivalently, the minimum $-\ln(L)$, where L is the likelihood. We also multiply by a factor of 2 to match up with Wilks' theorem since it will be used to determine sensitivity. The test metric is then:

$$-2\ln(L) = \sum_{i,j}^{i=N}[(x_i - \mu(\theta)_i)M^{-1}_{tot,ij}(x_j - \mu(\theta)_j)] + \ln(|M|) \tag{6}$$

where x is the observation, $\mu(\theta)$ is the expectation for a given set of oscillation parameters, $|M|$ is the determinant of the covariance matrix, and N is the number of bins. There are 31 bins coming from the 12 $1e1p$ energy bins and the 19 non-zero $1\mu1p$ bins. This is the metric used to test agreement to various fit parameters in this analysis.

The method used to determine sensitivity in this study utilizes Wilks' theorem [8]. Wilks' theorem states that, as long as some general conditions are satisfied, $-2ln(L)$ approaches a χ^2 distribution in the limit where the data sample is large [9]. In other words,the maximum $ln(L)$ ratio (R) follows a χ^2 distribution with n degrees of freedom (DOF). R is defined here as:

$$R = -2\ln\frac{L(\theta)}{L(\theta_{bf})} \tag{7}$$

where $L(\theta)$ is L for a given set of oscillation parameters, and $L(\theta_{bf})$ is the L for the best fit oscillation parameters. R written out fully for this analysis is therefore:

$$\begin{aligned} R = &\sum_{i,j}^{i=N}[(x_i - \mu(\theta)_i)M(\theta)^{-1}_{tot,ij}(x_j - \mu(\theta)_j)] + \ln(|M(\theta)|) \\ -&\sum_{i,j}^{i=N}[(x_i - \mu(\theta_{bf})_i)M(\theta_{bf})^{-1}_{tot,ij}(x_j - \mu(\theta_{bf})_j)] - \ln(|M(\theta_{bf})|) \end{aligned} \tag{8}$$

where $M(\theta)$ is the total covariance matrix for a given set of oscillation parameters.

To build the Wilks' theorem exclusion sensitivity, R is calculated at each point in parameter space with $\mu(\theta)$ and $M(\theta)$ as the expected number of events and covariance matrix for the parameters, x is the null oscillation spectrum, and $\mu(\theta_{bf})$ and $M(\theta_{bf})$ are the oscillation parameters giving the minimum $-2\ln(L)$ when compared to the null oscillation. This value is then compared at each point in parameter space to the Wilks' theorem 90% confidence level (CL) for 3 degrees of freedom (6.25).

As 3D confidence levels are not intuitive to visualize, the Wilks' theorem result is instead displayed in 2D slices at each of the 25 parameter values tested for each of the 3 parameters. The result is shown in Figures 4, 5, and 6 for Δm^2_{41}, $|U_{e4}|$, and $|U_{\mu4}|$ respectively. For a given value of a parameter the 2D exclusion contour in terms of the other 2 parameters is shown. The color scale represents the R value at each parameter grid point and has been capped

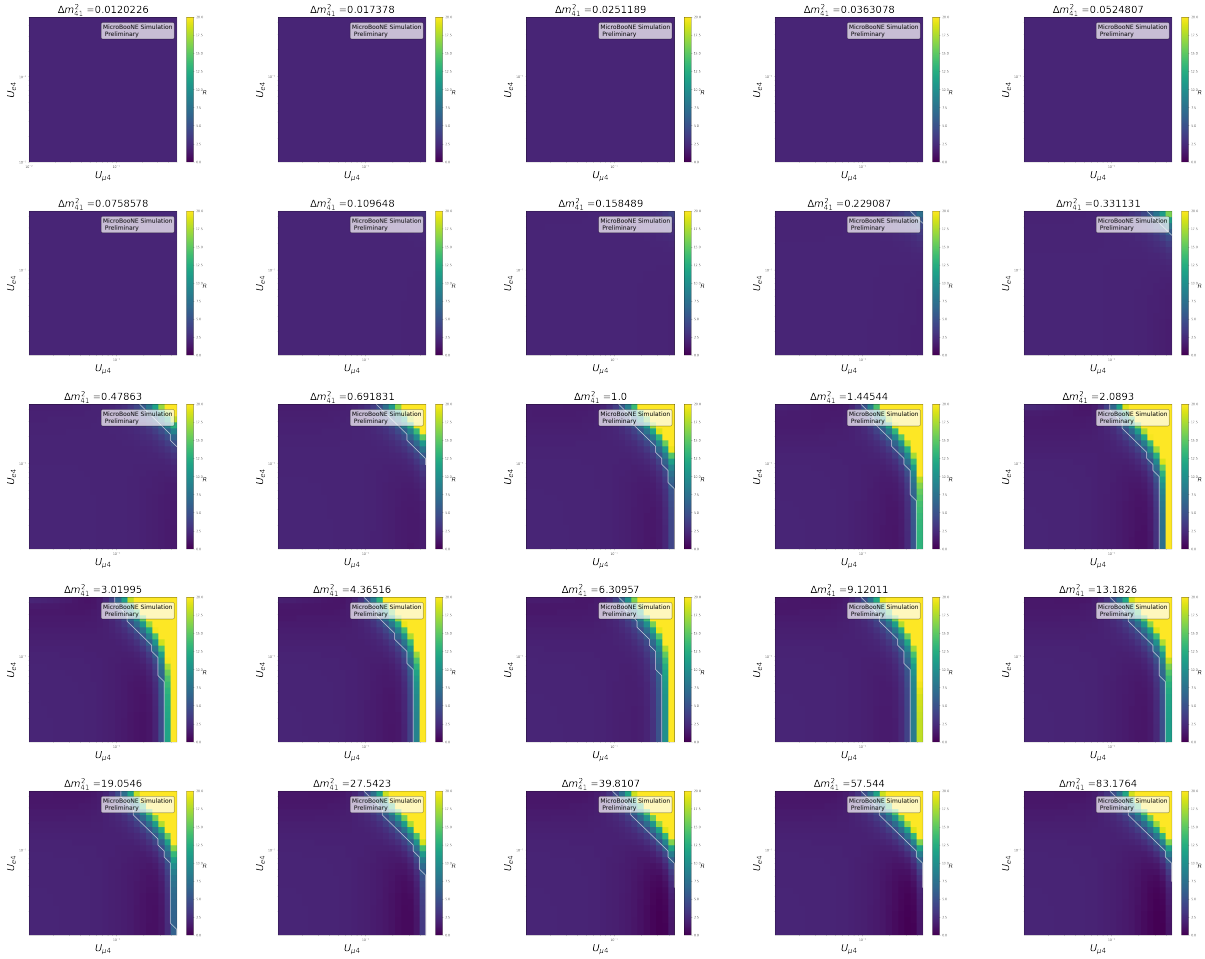


Figure 4: 2D slices for each Δm^2_{41} value tested to construct the Wilks' 90% confidence excluded region. The Δm^2_{41} value in each case is indicated by the plot title.

at 20.0 to allow for better visualisation for R close to 6.25 (Wilks' 90% exclusion level for 3 degrees of freedom). The exclusion sensitivity contour is indicated by the white lines.

A region of particular interest to analyze is the global best fit from previous experiments. The global best fit for these parameters as found in Ref. [10] are: $\Delta m^2_{41} = 1.32(eV)^2$, $|U_{e4}| = 0.116$, and $|U_{\mu 4}| = 0.116$. The 2d slices for these parameters are shown again in Fig 7, with the global best fit point shown by the red star. Note that the 2d slices are not exactly at the global best fit, but at the closest parameter value tested.

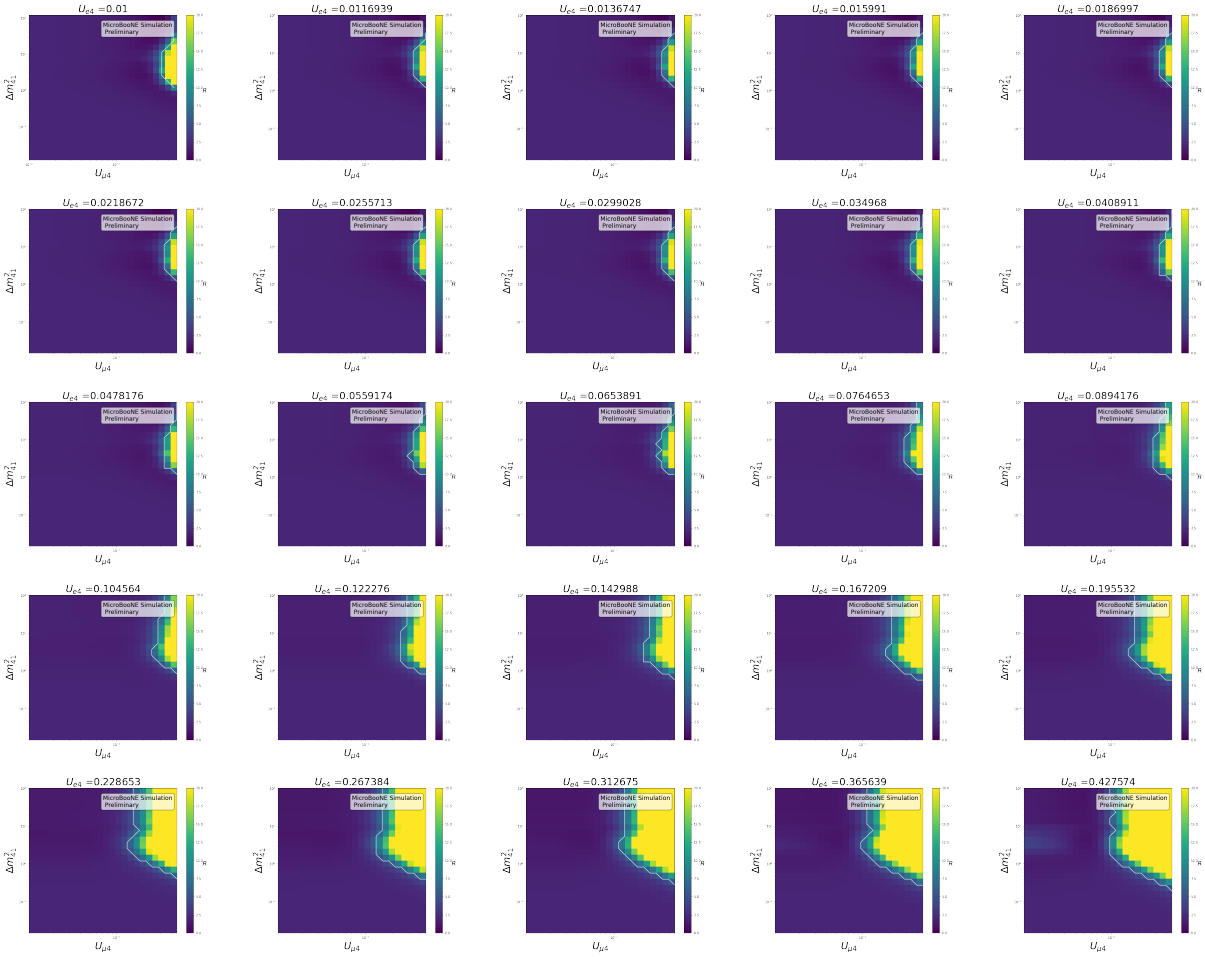


Figure 5: 2D slices for each U_{e4} value tested to construct the Wilks' 90% confidence excluded region. The U_{e4} value in each case is indicated by the plot title.

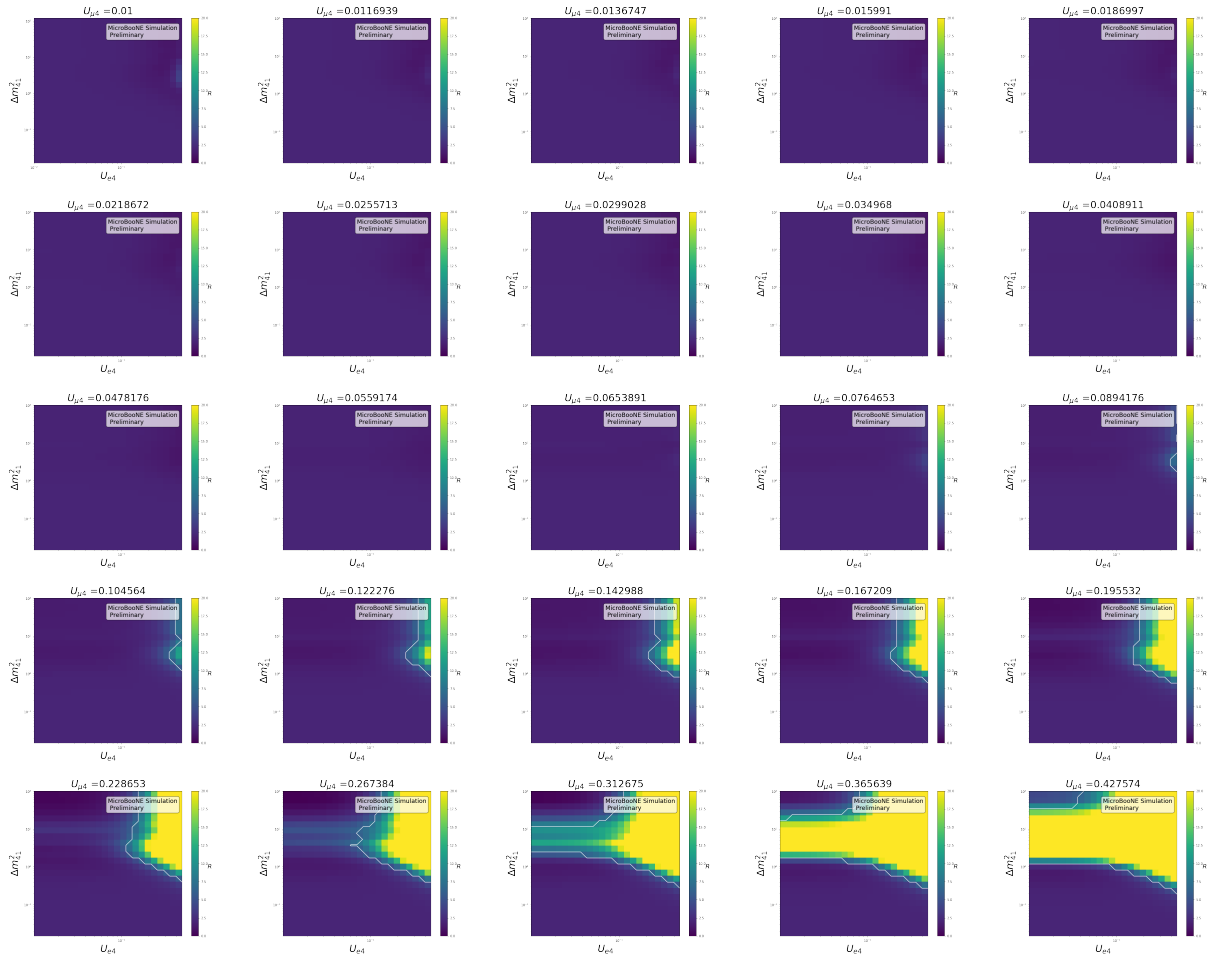
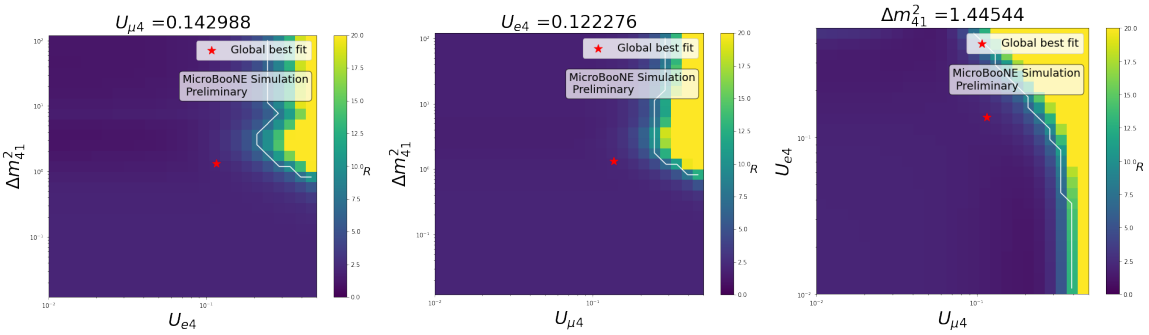


Figure 6: 2D slices for each $U_{\mu 4}$ value tested to construct the Wilks' 90% confidence excluded region. The $U_{\mu 4}$ value in each case is indicated by the plot title.



(a) $U_{\mu 4}$ *(b)* U_{e4} *(c)* Δm^2_{41}

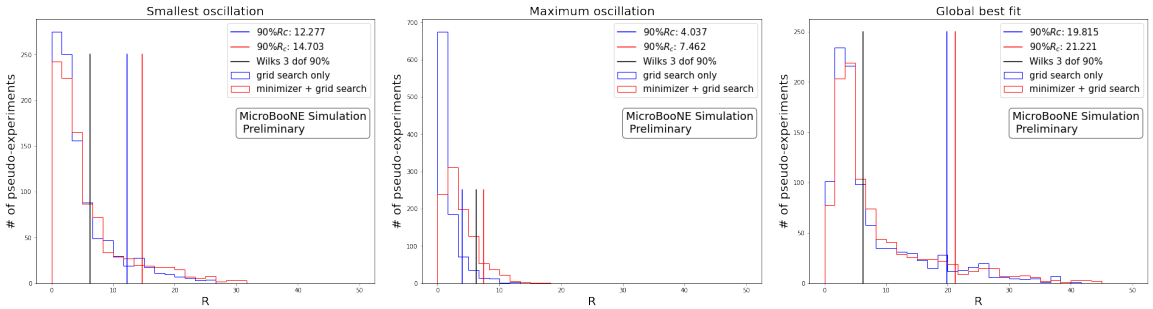
Figure 7: 2D slices of the Wilks' theorem exclusion sensitivity close to the global best fit point. The 90% CL is shown by the white line. The color scale (capped at 20) indicates the R value.

5 Finding the best fit with a Minimizer

In order to determine the best fit parameters with data, this analysis utilizes a minimizer. The 90% confidence level is then found via a grid search. The minimizer used in this study is TMinuit from Root [11]. The minimizer requires specification of a start point. The best performance is found when the starting point is the minimum from a grid search. This process allows for a finer resolution when finding the best fit oscillation.

Fig. 8 shows the R distribution for 1000 pseudo-experiments for a variety of test points. Both the grid search only and the grid search + minimizer are shown. The corresponding vertical lines represent the value for which 90% of R values are below for that test point. This value is referred to as R_{crit}. For each test point, R_{crit} of the combined grid search and minimizer is higher than the grid search on its own. The Wilks' 90% test points are a series of points which lie along the 90% sensitivity contour line found using Wilks' theorem.

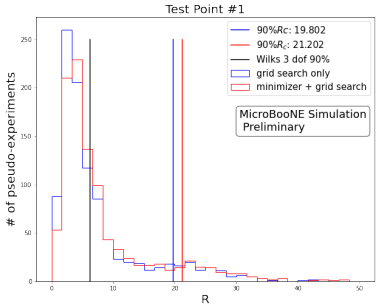
Additionally, the Wilk's theorem 90% confidence level value for 3 degrees of freedom is shown as a black vertical line in Fig. 8. If this analysis instead used a Feldman-Cousins method (Ref. [12]) to determine sensitivity, the R values found in the sensitivity study would instead be compared to the R_{crit} in these plots. R_{crit} is consistently higher than 6.25 when using the grid search combined with the minimizer. This indicates that using Wilks' theorem to draw the exclusion contour lines is not a valid approximation. Wilks' theorem and the Feldman-Cousins method are both frequentist methods of evaluating agreement. Wilks' theorem relies on assumptions that the Feldman-Cousins method does not. Perhaps most importantly, the use of Wilks' theorem assumes a large number of events. This assumption likely is broken in the $1e1p$ selection. The Feldman-Cousins method does not make this assumption. For first result on data then, the best fit will be found as described in this section. The agreement with the null model and the global best fit will instead need to be evaluated using the Feldman-Cousins' method.



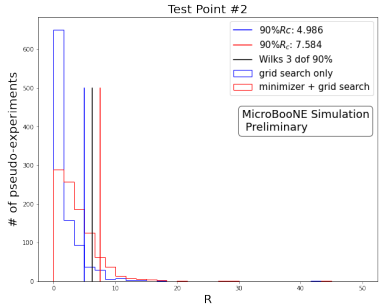
(a) Null Oscillation

(b) Maximum Oscillation

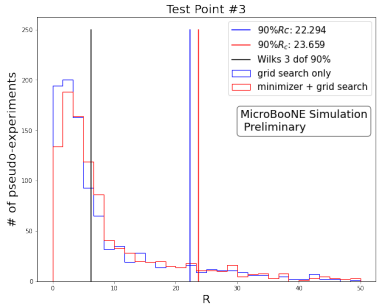
(c) Global best fit



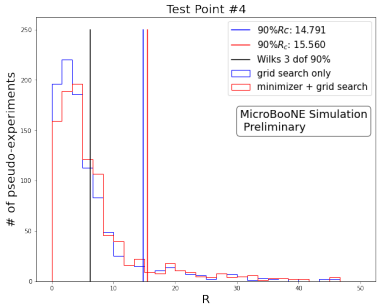
(d) Wilks' 90% point 1



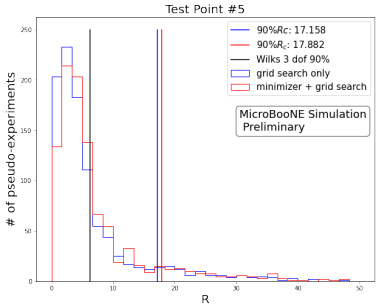
(e) Wilks' 90% point 2



(f) Wilks' 90% point 3



(g) Wilks' 90% point 4



(h) Wilks' 90% point 5

Figure 8: The R distribution for 1000 pseudo-experiments thrown from the specified set of oscillation parameters. The red spectrum is the result using only a grid search. The blue spectrum is the result of using a grid search followed by the minimizer. The Wilks' theorem 90% CL value is shown and compared to the 90% R_{crit} from the given spectrum.

6 Conclusions

In this note, we have provided detailed documentation of the deep-learning based joint 1e1p+1μ1p 3+1 sterile neutrino search. The physics motivations, reconstruction, selections, and systematic uncertainties were reviewed. The exclusion sensitivity utilizing Wilks' theorem when creating confidence levels was found. This showed what the excluded region is expected to be if the data matches the null oscillation spectrum. Next a minimizer to find the best fit oscillation parameters was described. This was used to show the difference between Wilks' theorem and a Feldman-Cousins method for this analysis.

References

[1] R. et al. (MicroBooNE Collaboration) Acciarri. "Design and construction of the MicroBooNE detector". In: *Journal of Instrumentation* 12.02 (Feb. 2017), P02017–P02017. ISSN: 1748-0221. DOI: 10.1088/1748-0221/12/02/p02017.

[2] A. A. et al. (MiniBooNE Collaboration) Aguilar-Arevalo. "Updated MiniBooNE neutrino oscillation results with increased data and new background studies". In: *Physical Review D* 103.5 (Mar. 2021). ISSN: 2470-0029. DOI: 10.1103/physrevd.103.052002.

[3] K. N. Abazajian et al. *Light Sterile Neutrinos: A White Paper*. 2012. arXiv: 1204.5379 [hep-ph].

[4] A. et al. (LSND Collaboration) Aguilar. "Evidence for neutrino oscillations from the observation of $\overline{\nu}_e$ appearance in a $\overline{\nu}_\mu$ beam". In: *Phys. Rev. D* 64 (11 Nov. 2001), p. 112007. DOI: 10.1103/PhysRevD.64.112007.

[5] P. Abratenko et al. (MicroBooNE collaboration). *Search for an anomalous excess of charged-current quasi-elastic ν_e interactions with the MicroBooNE experiment using Deep-Learning-based reconstruction*. 2021. DOI: 10.48550/ARXIV.2110.14080.

[6] P. Abratenko et al. (MicroBooNE collaboration. *Search for an Excess of Electron Neutrino Interactions in MicroBooNE Using Multiple Final State Topologies*. 2021. DOI: 10.48550/ARXIV.2110.14054.

[7] Artur M. Ankowski and Jan T. Sobczyk. "Argon spectral function and neutrino interactions". In: *Phys. Rev. C* 74 (2006), p. 054316. arXiv: nucl-th/0512004.

[8] S. S. Wilks. "The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses". In: *The Annals of Mathematical Statistics* 9.1 (1938), pp. 60–62.

[9] P.A. et al. (Particle Data Group) Zyla. "Review of Particle Physics". In: *Progress of Theoretical and Experimental Physics* 2020.8 (Aug. 2020). 083C01. ISSN: 2050-3911. DOI: 10.1093/ptep/ptaa104.

[10] A. Diaz et al. "Where are we with light sterile neutrinos?" In: *Physics Reports* 884 (Nov. 2020), pp. 1–59. ISSN: 0370-1573. DOI: 10.1016/j.physrep.2020.08.005. URL: http://dx.doi.org/10.1016/j.physrep.2020.08.005.

[11] F. James. *Minuit Function Minimization and Error Analysis Reference Manual*. Report 94.1. CERN, 1994.

[12] Gary J. Feldman and Robert D. Cousins. "Unified approach to the classical statistical analysis of small signals". In: *Physical Review D* 57.7 (Apr. 1998), pp. 3873–3889. ISSN: 1089-4918. DOI: 10.1103/physrevd.57.3873.