Chimera Events for Performance Studies of the MicroBooNE Deep Learning-based Low Energy Excess Search

MicroBooNE Collaboration

June 12, 2020

Abstract

MicroBooNE is a short baseline neutrino oscillation experiment based at Fermilab that employs Liquid Argon Time Projection Chamber (LArTPC) technology. One of its target measurements is to investigate the nature of the excess of low energy electron-like events observed by MiniBooNE. This measurement will require an excellent understanding of systematic uncertainties, obtained through testing the performance of reconstruction algorithms on samples with known properties. However, using exclusively Monte Carlo events for this task is limited by how well the discrepancies between simulation and data are understood. An alternative is to test against samples of "chimera" events, which are made up of separate single-particle components from data that are combined to create neutrino-like events. These chimera events can be used to help quantify systematic uncertainties. This note covers the performance and status of creating and using chimera events that match a target neutrino topology in MicroBooNE.

1 Introduction

1.1 The Deep Learning Low Energy Excess Analysis

A primary aim of MicroBooNE is to investigate the charged current quasi-elastic (CCQE)-like excess of ν_e events observed by MiniBooNE at low energies. This low energy excess signal is expected to predominantly appear in the 200–600 MeV range [1], where CCQE interactions dominate. Our approach is to identify neutrino interactions with a final state of 1 lepton, 1 proton, and 0 mesons, which allows the analysis to use the constraint of two-body scattering kinematics to reduce both cosmogenic and non-CCQE backgrounds. The MicroBooNE detector [2], located 470 m from the Booster Neutrino Beam (BNB) target, utilizes a time projection chamber (TPC), which drifts electrons obtained from the ionization of argon along the trajectories of charged particles using an applied electric field of 273 V/cm through to three wire planes (U, V, Y) that provide the charge read-out. The wire spacing of 3 mm and shaping time of 2 μs results in highly detailed event information which we exploit to create "images" for this analysis. The MicroBooNE detector also employs a set of 32 photomultiplier tubes (PMTs) to collect scintillation light.

The Deep Learning Low Energy Excess (DL LEE) analysis uses "traditional" pattern-recognition algorithms for cosmic ray rejection, vertex-finding, and 3-D reconstruction [3] and uses convolutional neural networks (CNNs) for particle identification [4, 5, 6], for which our images are well-suited. Thus, the key to the DL LEE analysis is to identify events with one electron and one proton meeting at a vertex ($1e1p$) consistent with CCQE events at the LEE signal energies [7]. To constrain systematic uncertainties, we simultaneously fit to the ν_μ interaction counterpart, $1\mu 1p$, since these events occur at similar neutrino energies and share flux and cross-section systematics.

1.2 DL LEE Analysis Chain Steps

The DL LEE analysis consists of the following main steps [8]:

1. Cuts based on observed PMT light to remove non-signal low energy events that are dominated by cosmic rays and are not relevant to the analysis.
2. Conversion of wire response data to an image format in 2D. This is done by filling a 2D array where each entry–or "pixel"–is the amount of charge deposited at a given time and wire [5].
3. Identification of areas of deposited charge in the image as either cosmic ray tracks or a contained region of potential interest for this analysis.
4. Identification of each pixel as track-like (corresponding to a muon, proton, or charged meson), shower-like (corresponding to either an electron or photon), or background using a convolutional neural network [9].
5. Three-dimensional neutrino interaction vertex finding [3].
6. Three-dimensional reconstruction of tracks and showers starting from the vertex [3].
7. Selection of events with two reconstructed particles, taken as $1\ell 1p$ event candidates.
8. Sorting of an event into either a $1e1p$ or $1\mu 1p$ pool based on max shower pixel fraction to proceed to the separate $1e1p$ and $1\mu 1p$ selections. Both selections use the output of a deep-learning-based multi-particle event identification algorithm to determine the particle content at the interaction vertex [6], but each selection has has its own Boosted Decision Tree (BDT) to identify candidates.
9. Constrain the systematics on the $1e1p$ candidates using the $1\mu 1p$ candidates.
10. Determine the consistency of the $1e1p$ candidates with the Standard Model prediction.

This analysis chain succeeds in isolating a high-purity sample of ν_e CCQE events with sufficient energy resolution to separate the low energy signal from the intrinsic ν_e background [8].

1.3 Chimera Events

As part of studies on systematic uncertainties in MicroBooNE, the performance of algorithms must be tested on event samples with known properties. Examples of existing samples are Monte Carlo events, data events found through hand scanning, and cosmic muon events identified by the Muon Counter System [10]. Chimera events consist of single-particle components from cosmic data that are selected and combined to create neutrino-like events. They provide a sample of events made using data that cover all the physics parameter space of interest for the signal region in the low energy excess analysis. For example, a ν_μ-like chimera would be a cosmic stopping muon, with the entering portion of the track removed so that it appears to be contained inside the fiducial volume, and either a proton track from a neutrino interaction or a proton produced by a neutron placed together so that they appear to be outgoing from a common vertex. We plan to create both ν_μ-like and ν_e-like chimera events.

Chimera events are created out of component particles that have already been reconstructed [3] and are placed at a known location within the event, so that the track vs. shower content of the hits and location of the vertex are known. We can evaluate the performance of the algorithms by comparing the results of reconstruction to the known "truth" information of the chimera. For example, the vertexer in step 5 of the DL LEE analysis chain has a dependence on opening angle; chimeras can be used to check for data and Monte Carlo disagreement here. Furthermore, chimeras allow us to produce a data sample where we can evaluate the selection's efficiency and the reconstruction's visible energy resolution for events similar to our target final state.

However, we do expect the chimera sample to have some biases. Known biases include the angle of the component particles and imperfect modeling of activity at the interaction vertex. We also expect these biases to be different from event samples obtained by other methods. Hence the chimera events provide a complementary event sample that will be useful as an additional data-driven test for understanding our systematics.

We will use chimera events to test several components of the analysis chain. In particular, we are interested in testing the track vs. shower labeling by the semantic segmentation network and the vertex reconstruction algorithm.

2 Creating Chimera Events

2.1 Overview

This note demonstrates a proof-of-principle method for the creation of ν_μ chimera events. Our algorithms have so far only been implemented for track objects, not shower objects, limiting the scope to ν_μ events. However, a future expansion to ν_e events is planned.

The prototype for the creation of a chimera event works as follows. We first choose an input event sample including muon and proton tracks. This sample can in theory contain a neutrino event or be entirely made up of cosmic events; the muon and proton in question do not have to be attached at a vertex. We then choose a desired 3D vertex position, length of muon and proton tracks, and θ and ϕ angles of the tracks, where θ refers to the angle of the track from the beam direction ($\theta = 0$ corresponds to the track being along the beam direction) and ϕ refers to rotation ($\phi = 0$ is along the drift direction). This choice can be made by the user manually, or can be taken from an existing event. The "closest matching candidate" for both a muon and proton is then chosen from the input pool of events given the desired parameter values. In the case of a two-prong event, we determine whether a track is a candidate muon or proton by taking the track with the higher average ionization to be the proton, and the other to be a muon. The selection of the "best candidate track" of the pool for each respective particle is then made by maximizing a log likelihood product of these variables, with individual weights that can be parameterized.

The likelihood is constructed in the following manner. For each of the parameters, $\psi_j \in \{x, y, z, \phi, \theta, l\}$, we define the target values, $\hat{\psi}_j$. We use a Gaussian with means $\hat{\psi}_j$ and variances $\sigma^2_{\psi_j}$ to assign a score to each i^{th} track, which is a function of the value of parameters for that track, $\psi_{i,j}$. The score, f_i, is then

$$f_i = \frac{1}{\sqrt{2\pi\sigma^2_{\psi_j}}} e^{-\frac{(\psi_{i,j}-\hat{\psi}_j)^2}{2\sigma^2_{\psi_j}}}. \tag{1}$$

We choose the track by finding the instance in our pool that maximizes

$$ln(f_i) = \sum_j \frac{(\psi_{i,j} - \hat{\psi}_j)^2}{2\sigma^2_{\psi_j}}. \tag{2}$$

The $\sigma^2_{\psi_j}$ values allow us to adjust the relative importance of each variable when identifying a matching track.

A separate algorithm isolates tracks from a given event and creates a pool of isolated track images to later piece together to create a chimera event. Here, we take as input a charge deposition image and a track object associated with an event, and return a set of charge deposition images each containing only a single track taken from the original event. This "separation" is done by looping through all identified neutrino interaction vertices in the input event, and then all tracks associated with the given vertex, followed by removing all pixels apart from those associated with the track. We will note here that currently, we do not make corrections due to extra charge around the vertex belonging to another track; this work is currently in progress.

After a "best candidate lepton" and "best candidate proton" are found by maximizing the product of the likelihoods of the six parameters of interest for each particle, we then place the tracks together onto one image at their vertex. This algorithm takes as input selected track information containing run, subrun, and event number; kinematic variable values for the muon and proton; and the images of isolated tracks corresponding to each selected particle. The output is a charge deposition image of the newly constructed chimera event. After the individual tracks are placed into one image, one of the tracks is shifted to align with the vertex of the other track. We have chosen to only shift the candidate proton track to align with the candidate muon vertex to minimize the amount of overall shifting of pixels, since proton tracks tend to overall be shorter in length. This pixel shifting is shown in Figure 1. In the final form of the procedure, the tracks will instead be drawn into a data event with cosmics and noise in the background. The result will be an event with a ν_μ-like topology with all tracks, cosmic backgrounds, and noise from detector data. The first chimera event produced using this prototype process is shown in Figure 2.

2.2 Current Limitations and Strengths

There are some limitations to the application of chimera events, which we describe here. For one, optical information from the MicroBooNE PMTs will not match the chimera interaction. We also do not locate the absolute x-position of tracks, though this can be accounted for in future work. Similarly, differences in detector response in the YZ plane are also not currently taken into account. The effects of space charge on the shifted tracks are also not corrected for. This is an inherent limitation; a track shifted in the x-direction

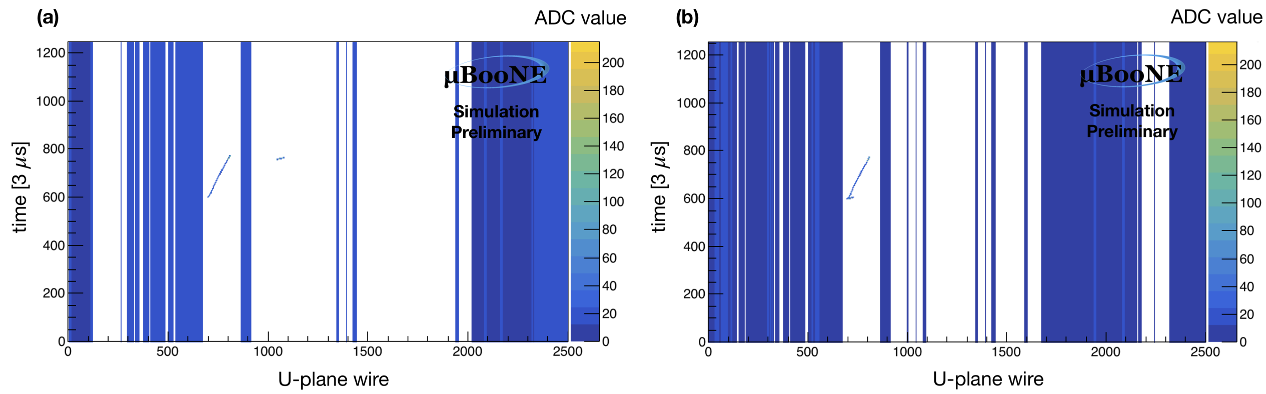

Figure 1: Demonstration of pixel shifting. The vertical axis indicates time tick and horizontal axis indicates U-plane wire number. Note that time = 0 here represents 2400 ticks in the MicroBooNE readout plane. The z-axis corresponds to amount of charge deposition. In image (a), the chosen muon and proton tracks are in their original positions. In image (b), the proton track has been shifted to align with the beginning point of the muon track to create a vertex. Note that temporarily unresponsive and "dead" wire regions, represented by blue vertical stripes, are shifted as well. The union of these regions is taken when creating the final image.

will not have the correct effect from space charge as would be expected in that area of the detector. We currently allow movement in the x-direction during chimera creation only due to the limited number of events available. With chimera events, we can, in the future, determine a measure of efficiency as a function of both opening angle and track length. However, it should be noted that this would be done for an acute opening neutrino angle in the collection plane only, since we can simply combine images by element-wise addition in this plane. We are limited in the induction planes, where this is more complex.

Compared to other methods of studying uncertainties, like changing wire signals in Monte Carlo, the use of chimera events has some strengths. One major advantage is that we work completely with data. Since chimera tracks are created from data components, no Monte Carlo, or "truth," information enters the creation of these events, with the intention to sidestep any discrepancies we see in simulating the MicroBooNE detector. Another advantage of this method is that the effects on shortening tracks can be explored; for example, we can take an event and delete the middle portion of a muon track to study the effect this has on reconstruction. In a similar manner, we can also zero-out various wires of a chimera image to "simulate" dead wire channels of different sizes and explore the effect this has on efficiency and reconstruction. This specific test will allow us to study how close dead regions can be to a vertex before they introduce complications with our vertexer. A third application is summing a chimera image with one or more cosmic rays, which we can use to test the efficiency and reconstruction of both the vertexer and tracker.

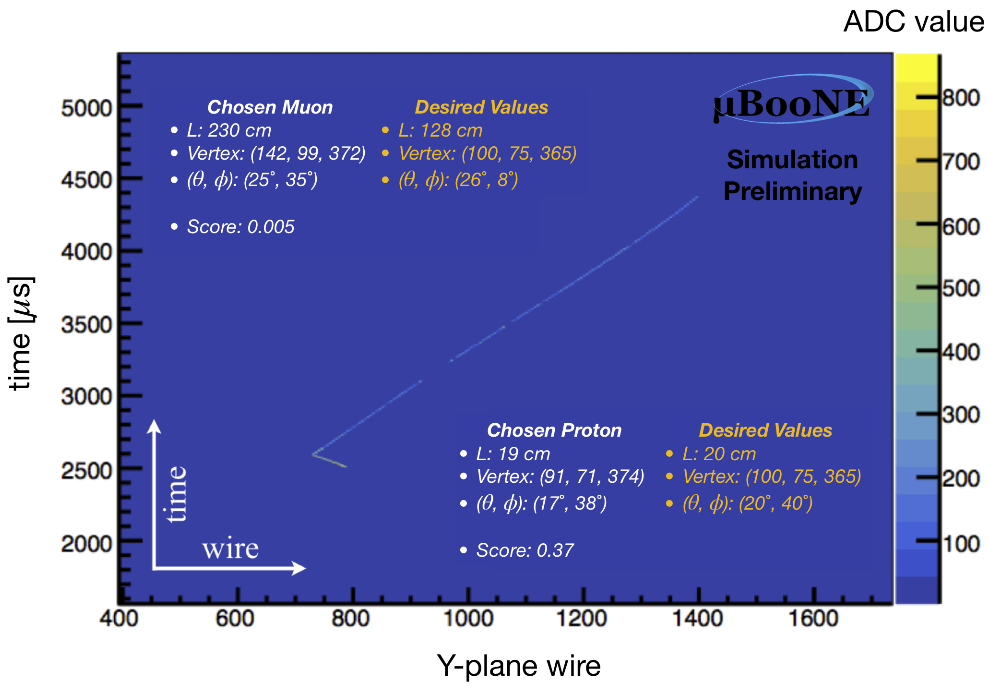

Figure 2: The first chimera event produced using the prototype chimera process. The horizontal and vertical axes indicate Y-wire and time, respectively. The color scale corresponds to charge deposition. The text on the figure shows the actual and requested parameters for the two tracks, in white and yellow respectively. The likelihood score measures how "well" the selected particle matches all requested parameters; for example, the muon and proton in this image have the highest scores compared with other candidate muons and protons in the pool.

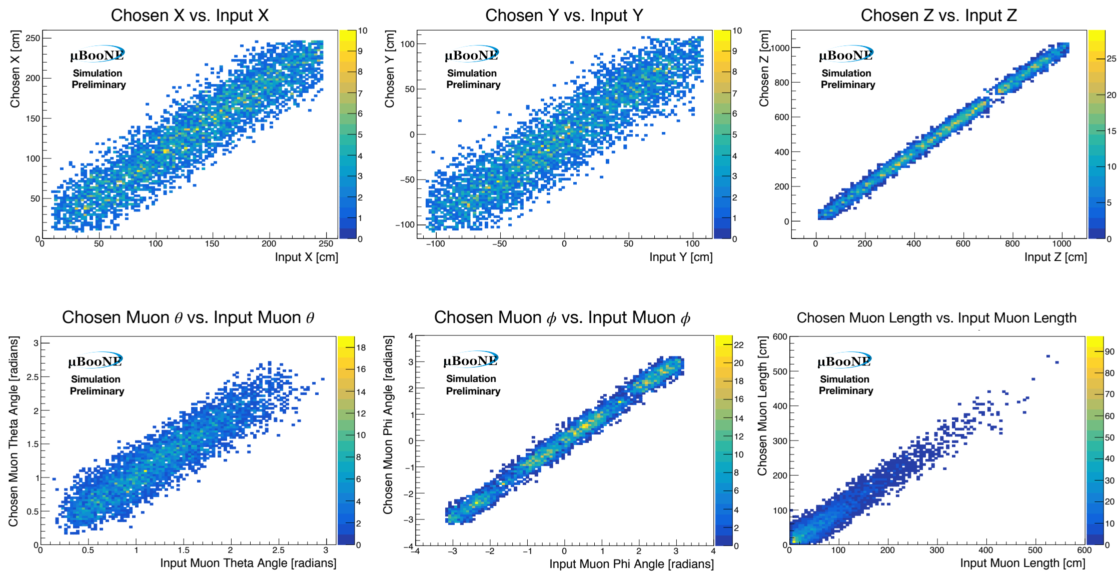

Figure 3: Correlation between the selected values of the position, lengths and angles of the muon and the desired input values.

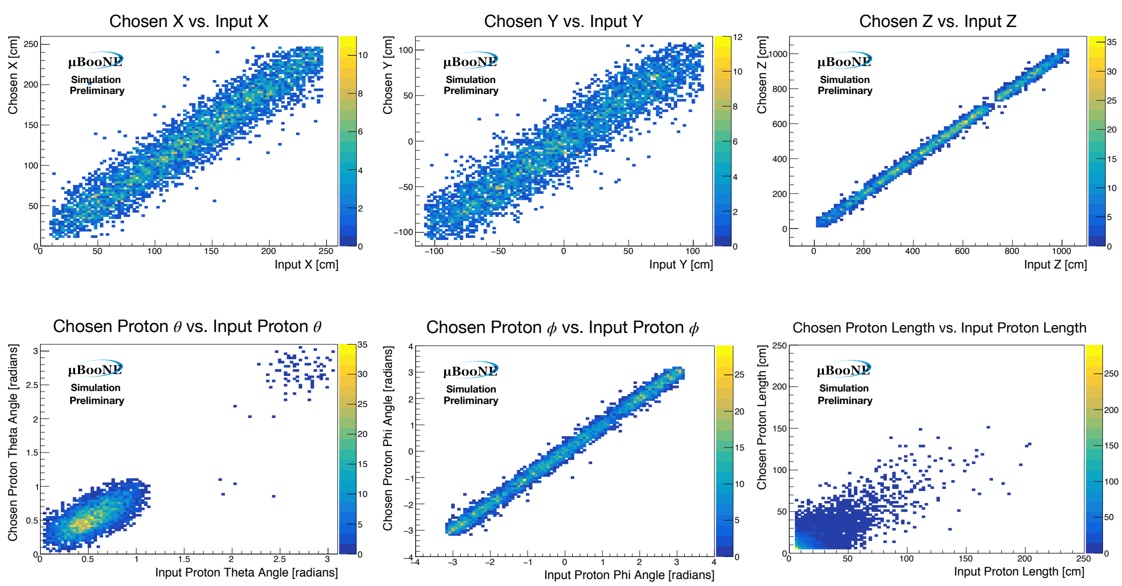

Figure 4: Correlation between the selected values of the position, lengths and angles of the proton and the desired input values.

2.3 Track-Matching Performance

The track-matching algorithm was tested on an "overlay" input sample, consisting of simulated neutrino interactions overlaid onto data taken when beam is off. The likelihood algorithm takes in two sets of events, with the intention that each set contains a different pool. This particular test was done with the same pool of events, skipping the exact event being matched to. We note that all input into chimera creation relies on reconstructed information rather than truth. Figure 3 shows the correlation between final selected parameter values and the respective desired input values. The individual weights of each variable in the likelihood are yet to be optimized, but the selection seems to work well for the current prototype given that each distribution follows a similar strong linear correlation. We note that any one of the variables can be optimized to be extremely precise by increasing the value of the weight, at the cost of the performance of the matching of other variables. Here, the selection is set to weights of $\sigma_\theta = 0.01\,\mathrm{rad}$, $\sigma_\phi = 0.01\,\mathrm{rad}$, and all other $\sigma = 1.0\,\mathrm{cm}$ for both muons and protons to optimize correlation for all six parameters in parallel. Figure 4 shows the same comparison as Figure 3, but for protons, where a similar correlation can be found.

2.4 Future Plans

The chimera event studies described in this note have so far been limited to ν_μ-like events in overlay. A quick extension is to run the track isolation and chimera-creation algorithms on a sample of purely cosmic data events.

We would also like to pursue ν_e-like chimera events. The main challenge for this is obtaining a large enough sample of single electrons in cosmic data. One possible source of electrons is Michels from stopping cosmic muons. Although these are generally lower in energy than the electrons we would expect to observe from BNB ν_e interactions, there is some overlap with the lower end of the relevant energy range. We expect the low energy electrons to be more difficult to reconstruct, so testing our performance at the lower end of the energy region is especially important and Michel electrons could be used for this. Another method for constructing ν_e-like chimeras would be to use photons from the cosmic data. Although we expect the dE/dx to be approximately doubled, photons have the same electromagnetic shower profile as electrons and may be a suitable stand-in for testing some algorithms. We will also study the possibility of dividing the charge deposition at the beginning of shower by a factor of two, to make it appear more like a single minimum ionizing particle (MIP). A third possible source of electrons is delta rays branching off from cosmic muons. We need to study the energy range and possible selection methods for the delta rays in order to determine whether this is feasible. Other possible sources of electrons include NuMI beam ν_e events, if they can be obtained in sufficient quantities, and Monte Carlo.

3 Conclusions and Outlook

We have demonstrated a proof-of-principle approach for the creation of ν_μ chimera events in this note. A small sample of such events has successfully been produced using overlay BNB events. Extensions to expand this method in order to produce chimera events useful as a systematics sample have been defined, including the addition of shower objects to create ν_e events, as well as working purely with data rather than an overlay sample. A procedure to generalize the creation of chimera events and create a sample of thousands of events in data is currently being processed. Once this sample of ν_e chimeras is made, it will allow for the next steps of systematics testing for the DL LEE analysis and general LArTPC reconstruction in MicroBooNE.

References

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

[2] R. Acciarri et al., MicroBooNE Collaboration, *Design and construction of the MicroBooNE detector*, JINST 12, P02017 (2017)

[3] MicroBooNE collaboration, *Vertex-Finding and Reconstruction of Contained Two-track Neutrino Events in the MicroBooNE Detector*, arXiv:2002.09375

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

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

[6] MicroBooNE collaboration, *A Convolutional Neural Network for Multiple Particle Identification in the MicroBooNE Liquid Argon Projection Chamber*, PUBLIC NOTE CURRENTLY IN REVIEW

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

[8] MicroBooNE collaboration, *Event Selection in the MicroBooNE Deep Learning Based Low Energy Excess Analysis Using Two-Body Scattering Criteria*, PUBLIC NOTE CURRENTLY IN REVIEW

[9] MicroBooNE collaboration, *A Deep Neural Network for Pixel-Level Electromagnetic Particle Identification in the MicroBooNE Liquid Argon Time Projection Chamber*, Phys. Rev. D99, 092001 (2019).

[10] S. R. Soleti, *The Muon Counter System for the MicroBooNE experiment*, arXiv:1604.07858