Spring 2020

Measurement of Angular Correlation of Two Protons in Quasielastic Neutrino-Nucleus Cross-Section

Matthieu Chalifour , '20

Measurement of Angular Correlation of Two Protons in Quasielastic Neutrino-Nucleus Cross-Section

By: Matthieu Chalifour

Advisor:
Kevin McFarland, Ph.D.
Department of Physics and Astronomy, University of Rochester

and

Catherine Crouch, Ph.D.
Department of Physics & Astronomy, Swarthmore College

Date: March, 2020

Swarthmore College
Department of Physics & Astronomy
2020

Measurement of Angular Correlation of Two Protons in Quasielastic Neutrino-Nucleus Cross-Section

Matthieu Chalifour

Abstract: We search in the MINERvA scintillator detector for multiple proton emission in quasielastic like charged current neutrino scattering. The number of such observable events in MINERvA is predicted to be far greater than currently available samples. We measure the total number of such events, and study the distribution of laboratory frame angles between the multiple protons and the muon, which is sensitive to the production mechanism for such events. We find significant disagreement of signal events between data and Monte-Carlo simulation.

1 Introduction

1.1 History of Neutrino Physics

Neutrinos were first proposed 1930 by Wolfgang Pauli in order to conserve energy during β decays. In β decays a neutron decays into a proton, an electron, and an electron neutrino. Of these particles, the main detected product of a β decay was the emitted electron due to its high energy. Unlike in α or γ decay, in β decay the energy spectrum of the electron was continuous. To Pauli, this implied that energy was being lost in the process to some unknown source. Furthermore, without another particle, angular momentum would not be conserved, since there needed to be another 1/2 integer spin. Pauli believed the source were new neutral particles, giving them the name "neutrons" (now known today as neutrinos). Despite the theoretical backing for the existence of the neutrino, Pauli famously apologized for theorizing the neutrino, saying "I have done a terrible thing, I have postulated a particle that cannot be detected". Neutrinos can only interact via the weak force and have an incredibly small cross section, between 100 MeV and 5 PeV $\frac{sigma}{E} = 10^{-38}\frac{\mathrm{cm}^2}{\mathrm{GeV}}$[8]. Therefore it requires a very large source of neutrinos along with a large detector for detection to be possible.

Neutrinos were first detected by Clyde Cowan and Frederick Reines in 1956 using a nearby nuclear reactor as the source of the neutrinos and a large tank of water and cadmium salt as the target [11]. They used scintillators surrounding the tank to detect the neutrinos. Neutrinos (technically anti-electron neutrinos) from the reactor interacted with protons in the water, producing a positron and a neutron. The positron then annihilates with a electron producing 2 γ rays, which would then be followed by a γ ray from the neutron being absorbed by a secondary nucleus. The characteristic signal of 2 γ rays followed by a third was the first detection of neutrinos, for which Reines was awarded a Nobel Prize.

In 1962 the difference between a muon neutrino v_μ and an electron neutrino v_e was experimentally discovered at Brookhaven [12]. Evidence of a third type of charged lepton and corresponding neutrino, the τ particle and v_τ, occurred in 1975 [17]. Both of these discoveries also earned Nobel Prizes, for Lederman, Schwartz, Steinberger in 1962 and Perl in 1995. In the current Standard Model there are three charged leptons, and three corresponding neutrinos. The Standard Model predicts conservation of both overall lepton number and lepton flavor. Conservation of lepton flavor means that the number of electron neutrinos + number of electrons, for example, must remain the same in any system. While neutrino oscillation has disproven this, it remains true that a neutrino of a certain flavor will only produce a charged lepton of the same flavor. For example if an electron is produced, it could not have been produced by a muon or tau neutrino.

The next mystery in the physics of neutrinos were the results of measuring neutrinos fluxes from the Sun. Models of the nuclear processes within the sun suggested it should produce a large flux of electron neutrinos. In order to test these models, experiments were created which were sensitive to electron neutrinos. The most famous being the Homestake experiment in a gold mine in Lead, South Dakota [10]. The quantity of electron neutrinos measured by detectors of solar neutrinos were a third of theoretical predictions from models of the Sun. These detectors were not sensitive to either muon neutrinos or tau neutrinos. The explanations for this discrepancy were either an issue with the solar model or meant that neutrinos could change flavor. However, the second solution would imply that the Standard Model is incorrect.

The mystery was solved in 1998, when the Super-K experiment announced that neutrinos do have mass and presented one of the first evidence of particle physics outside of the Standard Model [13]. There are currently three different neutrino mass eigenstates: m_1, m_2, and m_3. Since neutrino flavor is a linear combination of these masses, over time neutrinos will oscillate between the different flavors. These masses are much too small to measure directly, but can be measured indirectly by oscillation experiments. Since the rate at which neutrinos oscillate between the three flavors is determined by the differences between the masses, the magnitude of the mass difference can be measured by these experiments. Currently there are measurements of Δm_{12}^2 and Δm_{13}^2. These measurements do not give either the absolute magnitude of the neutrino masses or the order of the three masses.

1.2 What MINERvA is and what questions it answers

Since neutrino oscillations are a recent discovery, there are many open questions within the field of neutrino physics. One important question is the mass ordering; ie the sign of the mass differences. One of the masses is significantly larger than the others, and a current question is whether or not the two masses are lower (normal hierarchy) or higher (inverted hierarchy) than the third. A second question is whether or not neutrinos exhibit CP violation, which is needed to explain the lack of antimatter in the universe. Both of these research questions require precise measurements of the oscillation of neutrinos, which occur over vast distances. There are two types of experiments which measure neutrino oscillation. The first searches for disappearances by comparing the detection rate of one flavor of neutrino close to the beam to the detection rate of the same flavor further away. The second type of experiment are known as appearances searches; using a beam of single flavor of neutrino, they look for appearances of neutrinos of a different flavor further along the beam.

MINERvA, along with most other neutrino detectors, use scintillators to measure neutrino fluxes. Neutrinos are not directly measured by scintillators since they are not charged. Instead, similar to Cowan-Reines experiment, a signal that can be associated with the neutrino is measured. This means that there are false positives when other reactions create the characteristic signal. For precise measurements, better understanding of which interactions can produce characteristic signals and the rate which they occur is instrumental to the analysis.

A second complication with the current generation of neutrino experiments are caused by the energy regime they operate in. While current experiments use neutrinos with energies in the GeV range, early neutrino experiments operated with much lower energy. For example, the largest fluxes of solar neutrinos are in the single MeV range [7]. One disadvantage of using neutrinos in the GeV energy range is that energy reconstruction is much more difficult, since the energy from the neutrino can be transferred either to the products of the reaction or to exciting the nucleus. The two most confounding energies are the ~ 200 MeV to produce a delta baryon (which then decays into a pion) and ~ 50 MeV bonding energy of a neutron in the nucleus. In early experiments the collision is nearly elastic, since the energy remains in the momentum of the particles. Currently the energy of the neutrino can also be transformed into exciting the nucleus, which means that properly modeling nuclear processes is also crucial for energy reconstruction.

Another analysis required before the long range neutrino oscillation measurements can be taken are studies of what occurs within the nucleus after the initial interaction. One difficulty with large nuclei is that the reaction products from the neutrino interact with the rest of the nucleus prior to exiting the atom. This can cause the detected particles to have altered kinematics or be entirely different particles than those in the principal interaction. These collisions are called final state interactions (see Sec 2.2).

The MINERvA detector was constructed between 2006-2010 in order to measure these various interactions between neutrinos and nuclear matter. The goal of the experiment is to improve the model of neutrino-nucleus reactions. This is known as a baseline experiment, because a large purpose of the experiment is for calibration for the longer range oscillation experiments. Unlike the near detector in an oscillation experiment, MINERvA is designed to take precise measurements over a large energy range for a variety of nuclear targets.

One of the main classes of neutrino-nuclear reactions are quasielastic events, in which a muon neutrino interacts with a neutron bound in a nucleus to produce a muon and a proton. Due to the few products, it is the easiest reaction to reconstruct the energies of the products, and from there calculate the energy of the incoming neutrino. However an event where only a proton and muon are detected is not necessarily a quasielastic (QE) event. Since MINERvA, and most other neutrino detectors, uses scintillators, they are only ionized by fast-moving charged particles. Therefore neutrons, neutral pions, or low energy particles rarely show up in the detector but carry energy away from the reaction. While studies of neutrino energy focus on QE events, it is important to study any event which can appear to be a QE event.

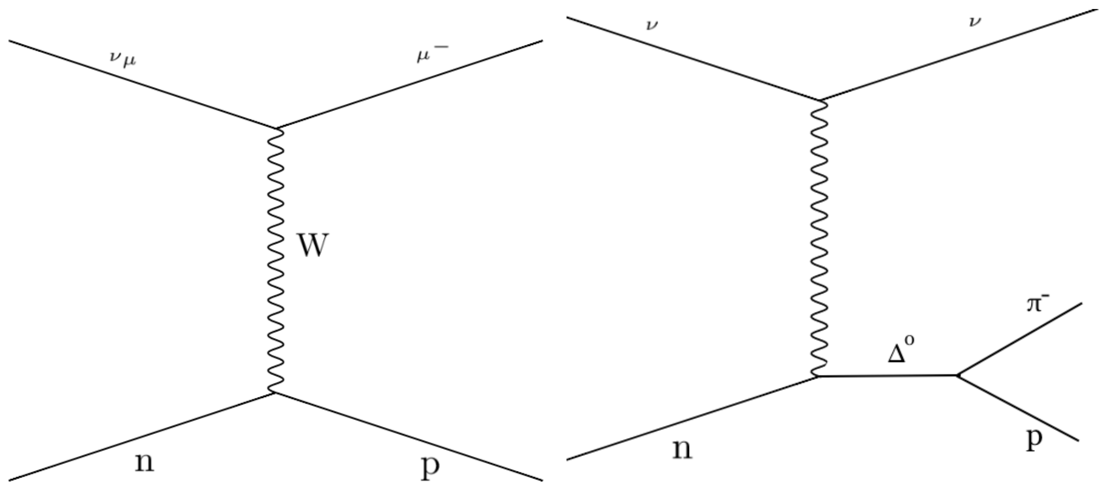
One such nuclear event are 2p2h events, where the neutrino ejects two nucleons instead of one. These events were not included in the models for the initial data release for MINERvA, but were added after observing a significant discrepancy in a energy region where 2p2h events occur. The frequency of these events was then tuned to match the simulation to the measured events. The goal of this project is to examine events with three nucleons (instead of the two nucleons for QE events) to attempt to isolate 2p2h events and determine if this data region either matches simulation or suggests some other tuning or change to the model.

2 Theory

One of the main products of my research is the comparison of the simulations of different nuclear interaction to data from the MINERvA detector. In this section we will discuss the most common categories of neutrino-nucleus interactions and some of the physics that occurs after the initial interaction.

2.1 Different Types of Neutrino Interactions (2p2h pion production, resonance)

For the purposes of this study, we will categorize the various neutrino-nucleus events into three main categories. The first are quasielastic events, in which the neutrino interacts with a neutron to produce a μ^- and a proton (fig 1a).These are labeled as 1p1h events because 1 particle is ejected, leaving 1 hole behind in the nucleus. Because this interaction is mediated by a charged W boson, it is a weak charged current interaction (CC), as opposed to a weak neutral current interaction mediated by the neutral Z boson. Conventionally charged current quasielastic events are known as CCQE events. For events with a visible μ^- and a visible proton, CCQE is the most common source of these events. Some CCQE events will produce other particles in the detector due to interactions after the initial quasielastic collision. One difficulty with reconstructing neutrino-nucleus interactions is that the energy of the neutrinos in the beam are not precisely known. CCQE interactions are therefore useful since the neutrino energy in principle[1] can be calculated by measuring the momentum of the products of the interaction.



(a) Quasi-Elastic

(b) Resonant

Figure 1: Feynman diagram of the different neutrino-nucleus scatterings, images from [16]

The second event type is 2p2h (or npnh) events. In 2p2h, the energy from the W boson is absorbed by the two (or n) nucleons. The 2 correlated nucleons are then ejected from the nucleus, which leaves 2 holes (hence 2p2h). Despite the similarity in name between 1p1h and 2p2h, 2p2h events are not quasielastic. The ejection of the second nucleon can occur by various means, often including a pion which is exchanged between the initially struck and a secondary nucleon [15]. Due to the uncertainty principle, the tight spatial correlation of the nucleons means that they have large relative momentum. The nucleons have a spatial correlation ~ 1 fm, which gives the particles $\approx 1 * 10^{-19}$ kg m/s momenta or a proton with energy ≈ 100MeV. By conservation of momentum the nucleons would then be back to back in the center of mass frame, i.e. the angle between the two. As a result of this argument, the simulation used by MINERvA for 2p2h events has particles fully back to back in the frame of the center of mass. The final common type of event is inelastic pion production. While there are multiple possible pion production processes, the most common is resonance production. In these events the neutrino excites the nucleus into a resonance, a Δ baryon, which then decays into a pion and nucleon, fig 1b. These events are important to model for oscillation experiments because they mimic the features of neutrino appearances or disappearances [18], which are how neutrino oscillations are detected.

2.2 Final State Interactions

Interactions within the initial nucleus are known as final state interactions, or FSI. Interactions between the ejected particles and other nuclei still pose a complexity for analysis, but are considered in a separate category of secondary interactions. The goal of modeling neutrino-nucleus interactions is to describe the results of baseline experiments, but what we measure is what appears in the detector after any further scattering in the nucleus. Due to final state interactions, there is not any guarantee about which initial process produced the particles that produce the tracks we measure.

[1] This is only true if the target, for CCQE interactions a neutron, is at rest and unbound; however, over a large enough dataset these effects can be corrected for on average

The rate of these final state interactions are also relatively high, since within the target molecules there is a high density of nucleons and the cross section between p and π is fairly high at $\sigma \sim 100$ fm^2.

Ideally our signal would be either 1p1h or 2p2h events. In that case our analysis would attempt to isolate these events from the inelastic events. However, since pions that are reabsorbed in the nucleus do not show up in the detector, in order to properly describe the accuracy of the detection and reconstruction process, events without pions in the final state are defined as "quasielastic-like" events. Our signal in practice is defined to be CCQE-like events, regardless of which process they were produced by within the nucleus. Consequently our background (events we want to remove from our sample) are events where there is a pion in the final state.

2.3 Simulation

To simulate the neutrino-nucleus interaction, MINERvA uses a modified version of GENIE 2.12.6 [6]. GENIE is a commonly used Monte Carlo event generator for neutrino-nucleus interactions. Out of the box GENIE does not simulate 2p2h events, but these events have been modeled according to the Nieves 2p2h model [14], with the frequency of 2p2h events tuned to MINERvA data. Parts of the inelastic pion production predictions of GENIE have also been tuned to more closely match experimental results [9]. In order to simulate secondary interactions in the detector (interactions between the products of the initial reaction and other nuclei) MINERvA uses the GEANT4 model [3].

3 Methods

This section describes the procedure that the MINERvA collaboration uses to turn a beam of neutrinos into a dataset that can be used for analysis. It is important to note that section 3.1 and most of 3.2 are analyses carried about by other researchers at MINERvA. A portion of the cuts described in section 3.2.1 were developed specifically for a three track sample in the course of the analysis. Finally section 3.3 describes original work that was used to produce the analysis in the Results section (section 4).

3.1 How these events are measured

MINERvA is the name of both the collaboration and the detector at Fermilab. The source of the neutrinos for the detector is the NuMI (Neutrinos at Main Injector) beam. In NuMI 120 GeV protons interact with graphite to produce pions and kaons, which decay into muon and muon neutrinos. The pions and kaons are directed along the direction of the beam by magnetic horns. As the particles travel, they decay, producing a beam of neutrinos. The beam then goes into rock to filter out non-neutrino particles. NuMI produces neutrinos between 1 and 16 GeV, although in two energy ranges, low and medium energy [2]. Our analysis focuses on the medium energy configuration, which peaks around 6 GeV. While the beam is primary v_u, there is $\sim 6\% \, \overline{v_u}$ and $\sim 1\% \, (v_e + \overline{v_e})$ [5].

Fig 2 shows the schematic of the entire detector. The primary component is the tracker region, which is made up of planes of scintillator. Each scintillator is made up of strips which produces photons in proportion to the energy deposited in the tracker. The scintillator can not record where along the strip an event occurred. Therefore in order to reconstruct the exact location of the particle in three dimensions, each plane is rotated 60° angle from the previous. This gives accurate kinematic information for particles moving through the planes of the detector. However, particles that move along a scintillator strip instead of across it are less likely to be reconstructed. As a

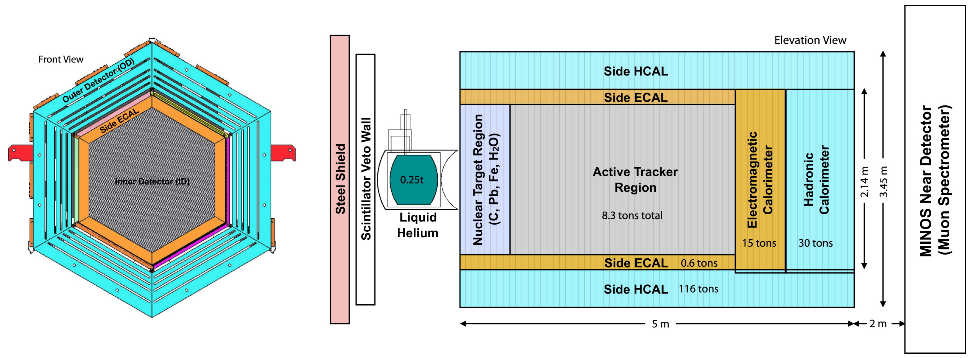


Figure 2: A side profile of the MINERvA detector. On the left is a slice of tracker region perpendicular to the beam. On the right is a lengthwise slice, where the NuMI beam moves from right to left. Figure from [4].

result, the efficiency of the tracker is low for $70 \leq \theta \leq 110°$ [4] where θ is the angle the track is from the beam.

Surrounding the tracker region are calorimeters for both hadrons and charged leptons. These are used to find the remaining energies of particles leaving the tracker region. They also detect particles produced in the rock around the detector.

Along the NuMI beam line after the detector is the MINOS near detector. For the purposes of this experiment, the detector is mainly used to reconstruct muons along the direction of the beam. This is specifically important since muons aligned with the beam are the indication of a muon neutrino interaction. Events where the muon does not reach or misses the MINOS detector are thrown out. Because of the geometry of the experiment, muons $20°$ from the beam angle will not reach MINOS. There also are scintillators before the tracker region to detect muons produced by the neutrino beam prior to the MINERvA detector. These were used for calibration of the detector, but also necessary to track to prevent them appearing as false events.

3.2 Reconstructing Events

The raw dataset from the MINERvA detector is mainly individual energy deposits in the tracking region. The purpose of the reconstruction algorithm is to turn these energy deposits into straight tracks representing a particle's trajectory and energy deposition. The reconstruction algorithm does this by fitting lines to contiguous energy deposits and then attempts to find a vertex that the tracks converge to, which represents where the neutrino-nucleus collision. Tracks need energy depositions to occur in adjacent planes of scintillator and occur within five scintillators in order to be reconstructed as a track.

Fig 3 is an example of a three track candidate. The opening angle θ is defined as the angle in the lab frame between the two candidate proton tracks (which means they do not show up in MINOS as muons). In events with more than three total tracks, the muon track and the two highest energy candidate proton tracks are used. The only measurement from the detector are the triangles representing the energy deposits at that point. Both the lines and the circles are best guesses from the track reconstruction algorithm. Since the reconstruction algorithm finds energy deposits that follow a linear path, tracks which either nearly overlap or are back to back are more difficult to reconstruct. Energies are estimated by fitting the track length to which particle the reconstruction algorithm thinks the track corresponds to.

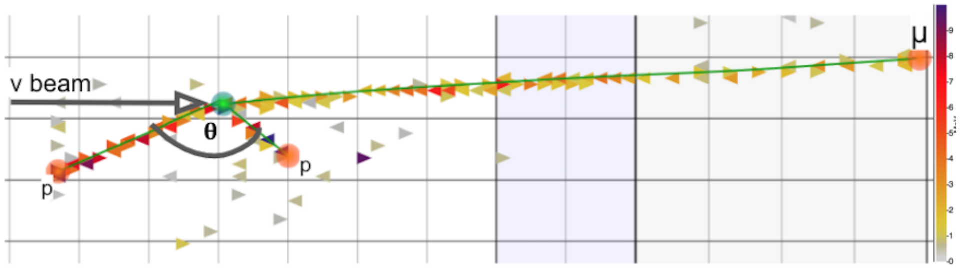


Figure 3: A 3 Track Event in the detector. The triangles represent energy deposits, the orange circles represent end vertices of the tracks, and the blue circle is the reconstruction algorithm's fit for the interaction vertex. The other labels have been added in post for clarity. The ps represent candidate proton events, while μ is a reconstructed muon track and v beam shows the direction of the NuMI beam. Although this figure is a projection into the x-y plane, θ is defined as the angle between the non muon tracks in 3 space.

3.2.1 Base Cuts

Cuts are applied after the track reconstruction phase to remove events with poor reconstruction or improve the purity of signal events. We cut any event that does not occur in the active tracker region and any event that does not have a matching muon in MINOS. Events in the active tracking region have the highest efficiency. MINOS detector of muons are necessary for the eventual energy reconstruction of the neutrino as well as is evidence for a muon neutrino interaction. Events whose primary vertex is too far from the tracks or events with tracks which do not lead to the main vertex are also cut since they are indicative of energy deposits not related to the neutrino collision. Similarly events that have reconstructed tracks either immediately before or after the interaction are removed to lower the rate of scatterings unrelated to the neutrino interaction being reconstructed. Since we are studying multiple proton events, we remove any event with less than three tracks, one for the muon and the two candidate protons.

Fig 4 shows the number of proton tracks in our sample or the number of tracks aside from the muon track. This plot contains the sample after the base cuts, so Proton ID cuts are not applied. There are two parts of the plot: the filled histograms represent GENIE simulated measurements, and the data points represent true reconstructed measurements. Our sample is predominantly two proton events. Since our analysis is focused on the opening angle between two protons, we will examine the angle between the protons that have the largest momentum transfer.

Ignoring lower energy protons could potentially cause problems in the analysis: either in cases where the highest energy proton was not reconstructed or due to errors in energy reconstruction the wrong two protons were chosen. However, both of these will be small subsets of the three or more proton events, which are already $< 20\%$ of the total sample, and a much smaller fraction after proton identification cuts.

3.3 Removing Pion Events

3.3.1 Proton ID cuts

Since pions and protons have the same charge and have similar track lengths through the detector, statistical methods have to be used to determine whether a track is a pion or proton in our sample. These are known as proton identification, or PID, cuts. Two PID selection criteria are used. The

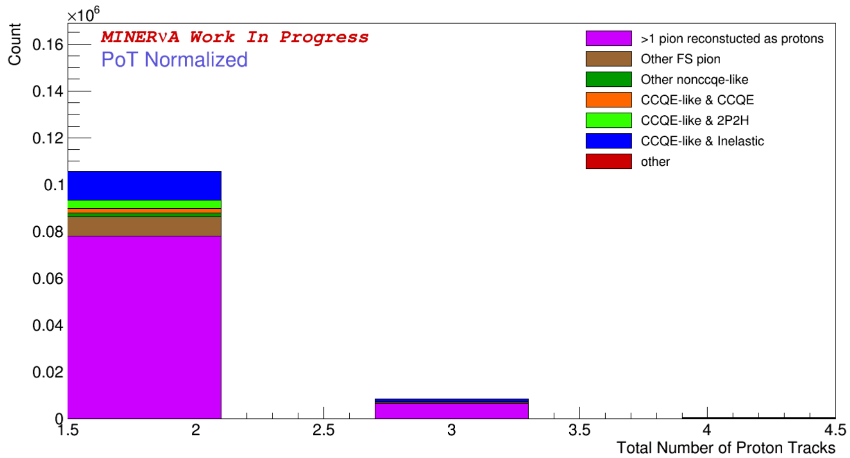


Figure 4: The multiplicity for our sample. The 1.5-2.5 bin represents 2 protons tracks, and the 2.5-3.5 bin represents 3 proton tracks. Histogram represents MC prediction, while data is shown in black with error bars. The histogram is further split up into different subsections to represent physical process that corresponds to the simulated point. Blue, light green and orange represent our signal regions. The purple background represents events where one of the reconstructed proton tracks are actually pions, and brown background represents events where a pion is present in the final state but is not reconstructed.

first is a cutoff in a log-likelihood ratio (LLR for short) between the track being a proton or pion. The likelihood is calculated through k nearest neighbors algorithm on the energy deposition rate $\frac{dE}{dx}$ to probabilistically determine whether the track is more likely a proton or pion. The energy deposition rate from ionization is a function of the mass and momentum. The second PID cut apply limits to energy range that are reconstructed. This cut removes tracks with energies that are not reconstructed well (i.e. most of the events at these energies are not reconstructed or reconstructed with poor accuracy), which improves the accuracy of the overall sample and eliminates events that are likely to trick the other PID cut. The lower energy range removes events who have a low reconstruction efficiency due to low energy. The upper limit removes anomalously large energy interactions. The energy calculation assumes that the particle loses its energy by ionization, and the removed large energy interactions are often the result of a secondary inelastic collision which would otherwise fool the algorithm.

For our analysis, a background event is an event with a pion in the final state regardless of whether or not the pion track is reconstructed. Since the PID cuts are designed to decrease the likelihood that a reconstructed track is a pion track, they are much less effective at removing final state pions that are not reconstructed. These pions are often either lower energy or scatter at an angle with poor reconstruction efficiency. We therefore choose to split our background into events with a final state pion(s) where a pion is reconstructed, and events where there are final state pions but none of them are reconstructed.

3.3.2 Background Sidebands and Subtraction

PID cuts improve the signal purity of the sample, but there still is a significant amount of both types of background events after they are applied. In order to study the accuracy of just signal

MC predictions, we want to fully remove all background data events. While there is not a way to easily remove all pion events from the sample with cuts, if the background simulation exactly matched background data events, then by subtracting the predicted background events from each bin we arrive at the signal MC to data comparison.

In order to determine the accuracy of the background simulation, we will look at background sidebands. Here cuts are used to accentuate the background region instead of suppressing it. One way to produce a sample with high purity of background events is to require that a Michel electron is detected in the sample. Michel electrons are electrons produced by decaying muons. Pions produced by the reaction will decay into muons which then decay into a Michel electron, a muon neutrino and an anti-electron neutrino. Michel electrons are characteristic of pion events, so a sample of events with Michel electrons will have a very high purity of pion events. However only $\sim 1/3$ of $\pi+$ produce a reconstructed Michel electron, so removing events with Michel electrons is not sufficient to have a signal sample with high purity.

If background MC does not match data in the sideband, then a tuning can be applied to the background prior to subtraction to improve the usefulness of the comparison between signal MC and background subtracted data. A tuning applies a scaling factor correction to background MC events so that they exhibit better agreement within the sideband. Assuming that the background MC prediction's differences to data in the sideband are similar to the differences to data in the main plot (with PID cuts to accentuate signal), after tuning and background subtraction the remaining data events should represent the true amount of signal events in our sample, allowing the comparison to signal MC.

4 Results

4.1 Sample Discussion

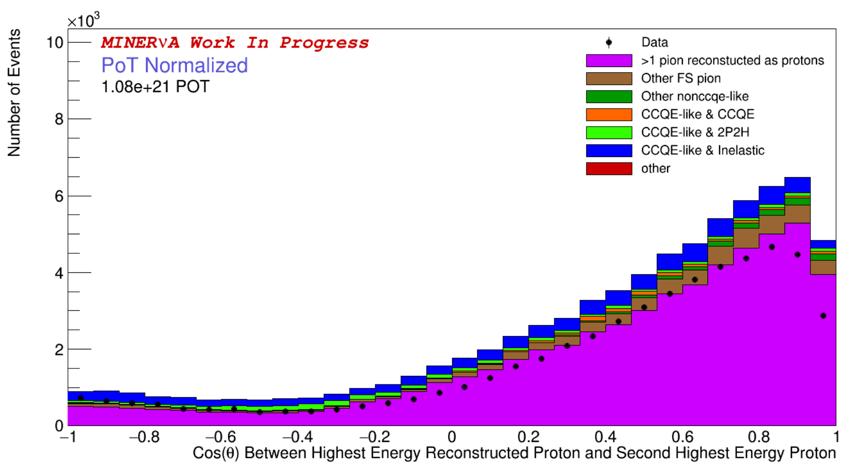


Figure 5: Distribution of cosine of the opening angle of the highest energy candidate protons for three or more track events. This plot has the only the base cuts applied, so the efficiency cuts and michel electron cut are applied but the PID cuts are not. Histogram division is the same as in figure 4 where purple and brown regions are background simulation representing final state pions that are reconstructed and final state pions that are not reconstructed respectively.

h!]

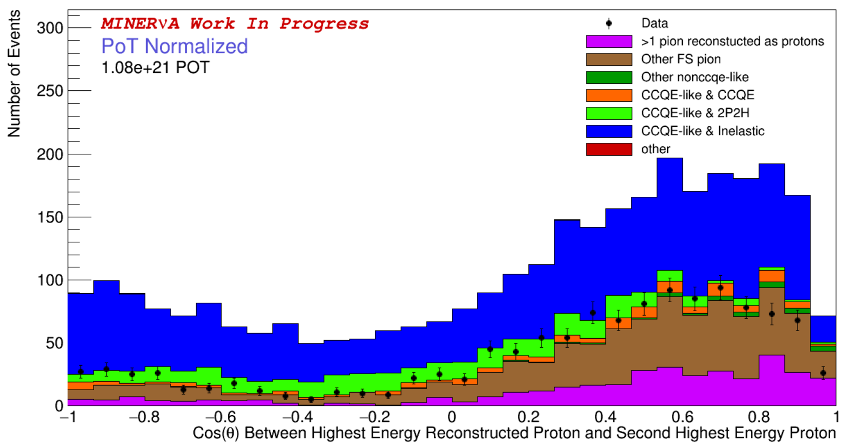


Figure 6: Distribution of cosine of the the opening angle, but with additional Proton ID cuts applied. The labels of the histogram are the same as in figure 4.

Figure 5 shows both the distribution of GENIE simulation and reconstructed data of three track events with respect to the opening angle between the two highest energy protons. The plot has the base cuts applied but without the additional PID cuts. The variable $\cos(\theta)$ is of interest since it indicative of the kinematics of the interaction, as well as is the common variable of interest in literature of two proton events. Figure 5 shows that there is a disagreement between our simulated and measured events. The disagreement is both in overall area normalization and increases at high $\cos(\theta)$. The disagreement at high $\cos(\theta)$ is likely related to the reconstruction process. Since the reconstruction algorithm attempts to group energy deposits into lines in order to reconstruct tracks, tracks with $\theta \sim 0°$ are more difficult to divide into two tracks.

Figure 6 shows the sample presented in fig 5, but with the PID cuts applied. Applying the PID cuts there is a substantial decrease in the background contribution from reconstructed pions, and a slight decrease in the relative amount of remaining unreconstructed pion background. In Fig 6, we have identified a sample with a relatively high purity of signal events with a large amount of predicted signal events. This represents a significant improvement in statistics from previous 2p2h studies ([1]). The MC and data disagreement has also significantly increased following the application of the PID cuts, there are between 2x and 10x the amount of predicted events from MC to measured events.

4.2 Sidebands and Tuning

Fig 7 shows the opening angle in the sideband where Michel electrons were detected. Since the purpose of the plot is to compare the simulation of the background region instead of the signal region it is a sideband. This sideband has high background purity. It also shows overall the best qualitative agreement between our simulation and measurement than figure 6 or figure 5. Below $\cos(\theta) = 0.6$ there is good accuracy between data and MC; however, the agreement is not as good at high $\cos(\theta)$. There also is an oversimulation of MC around $\cos(\theta) \sim -0.4$ and slight undersimulation elsewhere. Fig 7 implies that most of disagreement in the previous plots is from the signal region, with the possible exception at high $\cos\theta$.

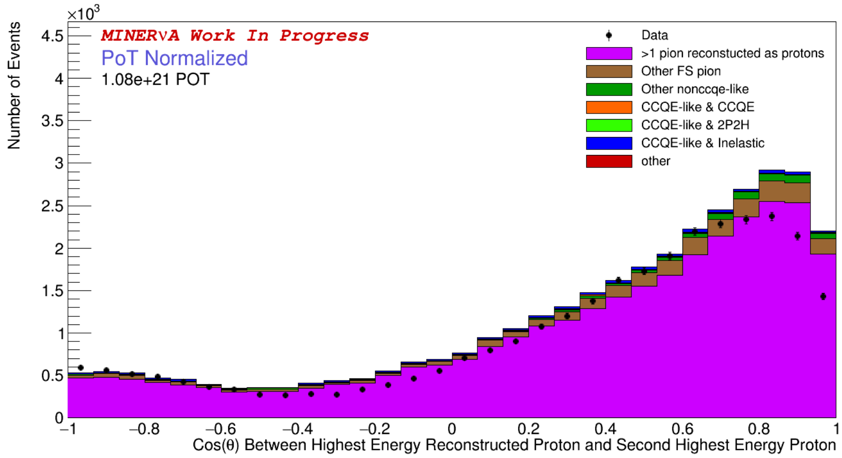


Figure 7: This plot contains the same axis definitions and legend as previous plots. This plot requires a detected michel electron instead of prohibiting them; all other base cuts are applied and PID cuts are not present. This plot represents the background sideband of events with Michel electrons

Since there is evidence of a $\cos(\theta)$ dependency is the prediction accuracy of the background MC, the scaling ratio of the tuning between MC and data will be a function of $\cos(\theta)$. Since figure 6 shows that data is slightly higher than our background prediction but substantially lower than the overall MC prediction, we will assume that, despite the signal MC prediction in the figure **??**, there are no signal data events in the sample. That means that our tuning function will attempt to fit $\frac{\text{\#Data events in bin i}}{\text{Background MC in bin i}}$. We chose a to fit the ratio with a quartic polynomial to capture the larger trends while not overfitting statistical variation.

Fig 8 is a per bin ratio plot between data and background MC, with the fitted tuning overlaid. In this plot signal MC is discarded and the two sources of background are combined. Since this plot is predominantly reconstructed pions, it is necessary to check the tuning on a sideband of unreconstructed pions. At the moment since the source of these events is the same interactions, we will assume that the simulation accuracy will be similar for both so we will applying the tuning to both reconstructed and unreconstructed pions.

Figure 9 shows another sideband, events with Michel electrons that also pass the PID cuts. Michel electrons are indicative of events that produce pions, and PID cuts remove reconstructed pions, so this should a sample primarily made up of unreconstructed pions. After applying the tuning, the reconstructed pion background should represent an accurate prediction of the respective true events . Figure 9 supports that events with unreconstructed pions are being modeled accurately after tuning. While the statistics are not as good as the michel electron sideband without PID cuts, the sample is still primarily unreconstructed pion events and it features much better agreement than our signal region (fig 6). The good agreement combined with low purity and statistics mean that we will not apply a second tuning based on the results of this sideband.

We are using a log likelihood calculation to obtain a qualitative measure of MC accuracy.

$$ln(\lambda) = \sum_{i}^{N} \mu_i - n_i + n_i ln \frac{n_i}{\mu_i} \quad (1)$$

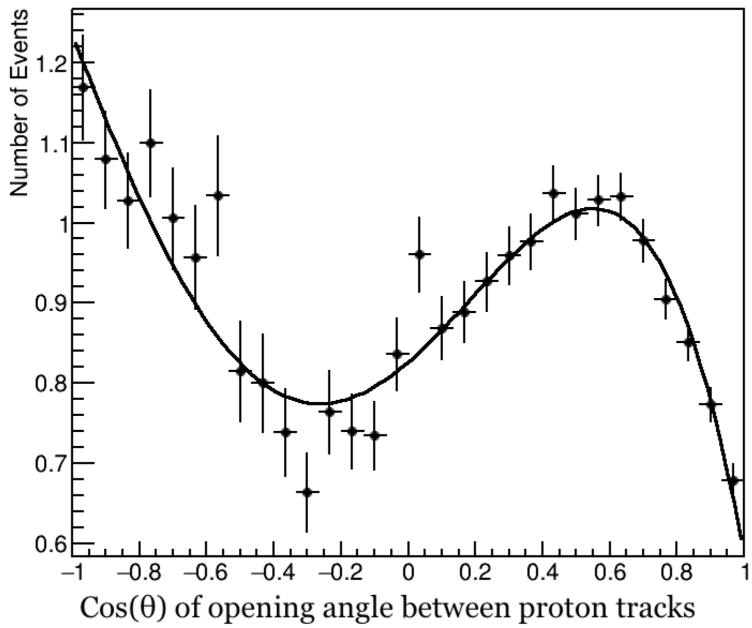


Figure 8: Plot of the $\frac{\text{\# of Data events}}{\text{\#of predicted background events}}$ per bin in the Michel electron sideband with error bars. Both reconstructed and unreconstructed pion background are Overlaid in red is our polynomial tuning factor that will be applied.

Included MC regions	All MC regions	Only Reconstructed Pions	Only Background	All but Inelastic Signal
λ	86	865	81	74

Table 1: This table corresponds to figure 9, and shows the λ values for different combinations of simulated regions calculated from equation 1. The columns represent which Monte-Carlo regions are turned on or off. The first column is the overall comparison of MC to data. The second compares the simulated amount of reconstructed pions to the entire data sample. The third column compares the simulated background to data; and the last column compares all Monte-Carlo except for the inelastic signal region.

where i refers to the bin, μ_i represents the predicted events in the i^{th} bin and n_i represents the true data events in that bin. Lower λ means better agreement between μ and n. λ is related to χ^2 but without including the degrees of freedom calculation. This means that λ is useful to compare different hypothesis for the same plot, but since it is not χ^2 there is no conversion between a λ value and p values.

The different hypotheses that will be compared will involve "turning off" different regions of the Monte-Carlo simulation, where turning off means that we will essentially be scaling the MC points that are part of that region by 0, points in the blue histogram are part of the CCQE-like but inelastic region, for example. Our hypotheses are: the entire MC prediction; turning off just the inelastic signal region; turning off all signal MC events; and turning off everything but reconstructed pions (so all signal MC and unreconstructed pions are turned off). The source of the discrepancy is unlikely solely within a single category, but these hypothesis give information

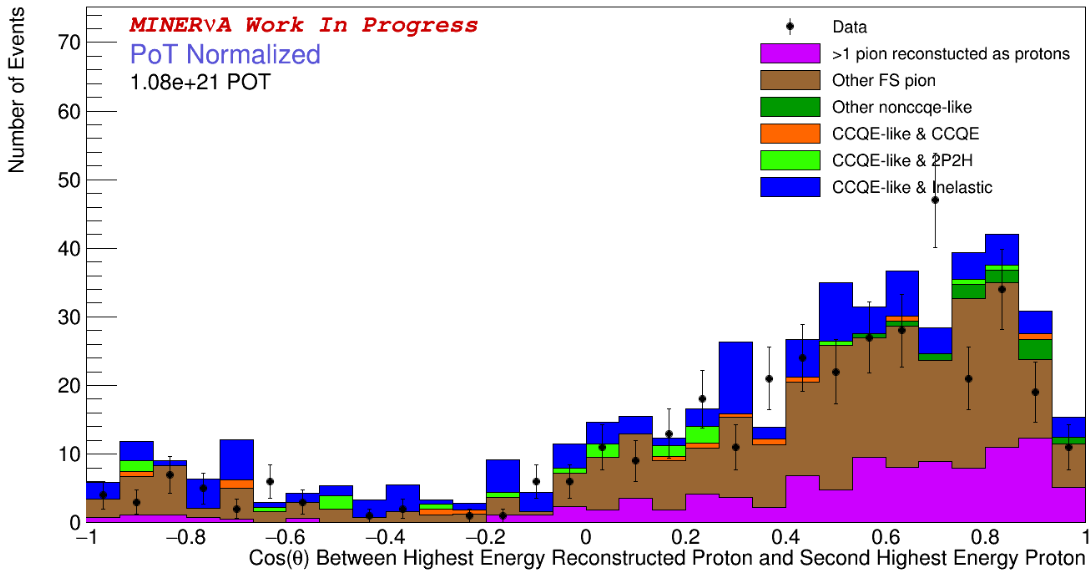


Figure 9: Distribution of $\cos(\theta)$ plot. Cuts for the plot are base cuts and PID cuts, but with the reversed michel cut (requiring a michel electron instead of removing michel electrons). This plot is a sideband specifically for unreconstructed pions in order to verify that after the background tuning simulation matches closely with data.

about what predictions of the Monte-Carlo are the most supported by our data. This calculation is useful only after the background is tuned since all of the hypothesis assume the reconstructed pion background has good agreement.

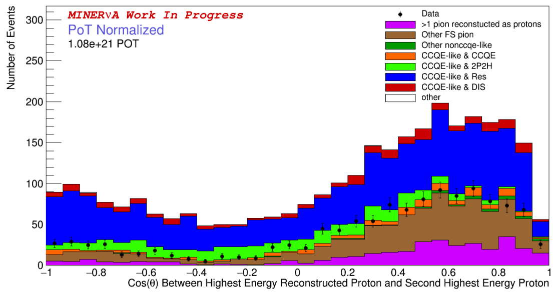
Table 1 shows the λ agreement between different parts of the simulation and the data for the unreconstructed pion sideband. From table 1, the best agreement is the hypothesis that there are no inelastic CCQE-like events, with the table strongly opposing that only reconstructed pions are present in the sample.

Since the Monte-Carlo of figure 9 is predominantly unreconstructed pion events, the λ values are mainly indicative of the accuracy of the unreconstructed pion simulation. From the table 1 there is significant support that unreconstructed pions are being simulated correctly, and thus are not likely the primary source of the significant disagreement within our signal region (figure 6).

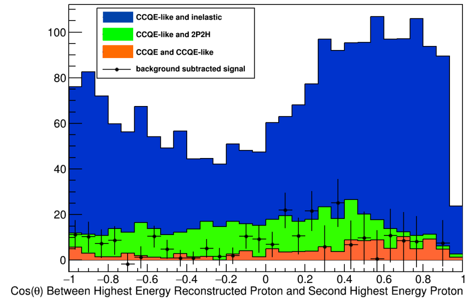
4.3 Main Result

In section 4.2, we presented the argument that after background tuning the background simulated events correspond very closely to true data points. Under this argument, the data points above the tuned simulated background correspond to the true signal events in our dataset.

Fig 10 shows the cosine plot with PID cuts and with tuned backgrounds. On the left is the overall plot, and on the right we show signal MC against data after background subtraction. In fig 10a, in the region $-1 \leq \cos(\theta) \leq 0$, there are more inelastic simulated events than total data events prior to background subtaction, so independent of the background efficiency there is an issue in the simulation of those events. Based on figure 10b there is a significant disagreement between the signal prediction and the events that correspond to signal events. There is evidence that there are signal events in the sample ($\sim$ 200), but substantially less than predicted by Monte-Carlo.



(a) Figure 6 but with background tuning applied.



(b) Figure 10a but with both background MC and data events corresponding to background MC removed

Figure 10: Distribution of $\cos(\theta)$ with PID cuts and background tuning. Left is the entire sample and right shows compares just the signal region to data after background subtraction.

Even comparing just quasi-elastic and 2p2h simulation to the entire signal data suggests that these regions do not occur or are not reconstructed as much as predicted even in the event that no inelastic signal events are measured.

	All MC regions	Only Reconstructed Pions	Only Background	All but Inelastic Signal
λ	1193	2970	124	84

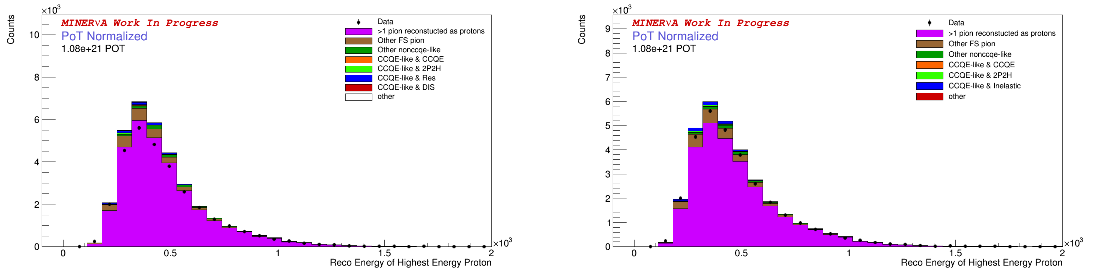
Table 2: λ table corresponding to fig 10a. Each columns corresponds to a hypothesis of different MC regions are not present. The columns represent the same regions as in table 1. These values are calculated prior to background subtraction, so from figure 10a.

From table 2, the data agrees best with there being no inelastic events in our sample, followed closely by no signal events. By comparison the table rejects either that there are only reconstructed pions in our sample or that the overall MC simulation agrees with data. As a result we find very strong disagreement between MINERvA's simulation of three track events without final state pions and data taken from MINERvA.

4.4 Proton Kinetic Energy

While the main conclusions have been drawn from $\cos(\theta)$, comparing to other variables is important to validate our previous results and help find further trends in the signal disagreement. It is also used to verify the background tuning applied in the previous subsection and search for possible further tuning to improve the accuracy of the background subtraction process. The largest proton energy is of particular interest since energy is indicative of the nuclear process that produced the proton.

Fig 11 shows the Michel sideband without (left) and with (right) the background tuning on cosine of the opening angle. All MC regions (as well as data) peak around 300MeV. At higher energy there is relatively more background events than signal. This matches theoretical analysis since QE processes peak at lower energy than inelastic events [18]. From $\cos(\theta)$ data we expect that the michel sideband, after background tuning, should have fairly high data/MC agreement. The agreement supports that the background tuning overall improves MC accuracy for the background region.



(a) Untuned Michel sideband of larger proton energy

(b) Tuned Michel sideband of larger proton energy

Figure 11: Plot of the distribution of the largest proton energy. Both plots are michel sidebands, the right applies the background tuning used in section 4.2.

	All MC regions	Only Reconstructed Pions	Only Background	All but Inelastic singal
Untuned λ	780	267	583	640
Tuned λ	317	527	246	263

Table 3: Compares the different MC region hypotheses for the michel sideband between the tuned background and untuned background of figure 11.

Table 3 compares the λ values for the different hypotheses with and without the $\cos(\theta)$ tuning. From the $\cos\theta$ analysis, we expect the "Only background" and "All but Inelastic Signal" MC regions to have the best agreement. Since both of these improve upon tuning, this supports that tuning improves agreement, but a significant portion of that is likely due to fixing the overall normalization between data and MC. From table 3 there is moderate evidence that an energy tuning on top of the cosine tuning may help agreement. Because there is more correlation between the disagreement and $\cos(\theta)$ than proton 1 KE, the tuning in energy would be on top of the cosine tuning instead of in replacement.

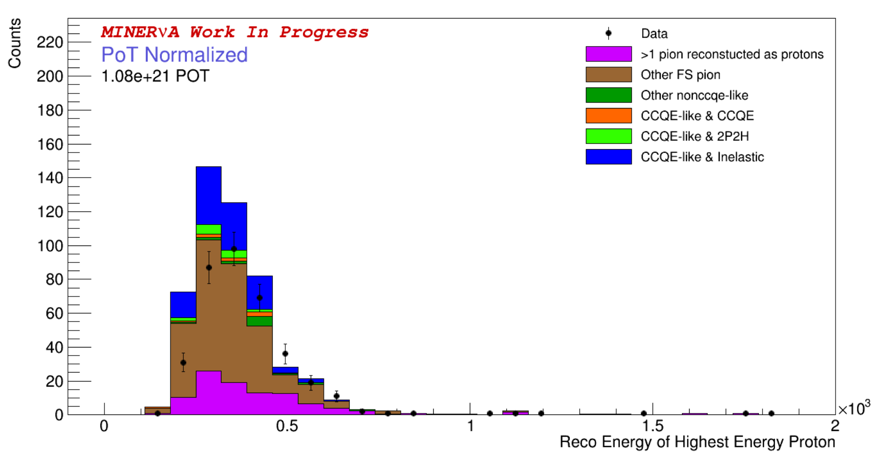
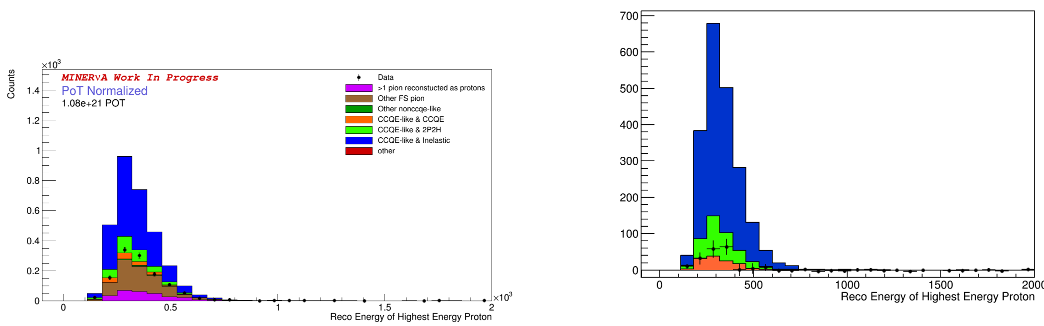


Figure 12: Sideband of largest proton energy with PID cuts and Michel electron requirement

Fig 12 shows our sideband for the unreconstructed pion background (the brown background) in the leading proton energy variable. This plot has both the PID cuts and the michel electron requirement as well as the cosine tuning. This plot supports the argument that there is a sizable population of unreconstructed pion events, but it does challenge the quality of our simulation of these events. Below ~ 200 MeV there is a significant disagreement between simulation and data. Furthermore above ~ 500 MeV there is some undersimulation of MC, although with full systematic error applied the MC would mostly likely be within the error bounds of the data. As a result, this is may be a good candidate for a secondary tuning in a future work, although that would be complicated by the low statistics.



(a) Signal region of largest proton energy with tuned background

(b) Left plot but with pion background subtracted

Figure 13: Leading proton energy with PID cuts and tuning

Fig 13 shows the signal region of the larger of the proton kinetic energies, on the left is with background and on the right the background has been subtracted. From fig 11 and 12, the background subtraction likely overestimates the amount of signal at low energy and underestimates it at higher energy. Since this matches where the data is, it is not possible from this plot to determine where the signal is. Visually there is still the large inelastic prediction which is not supported by the data. The disagreement between the unreconstructed background and data is less present in this plot, so the previous plot may just be an issue with small statistics. This plot also supports that the true signal appears to be at higher energy.

	All MC regions	Only Reconstructed Pions	Only Background	All but inelastic region
λ	1209	3085	54.5	58.8

Table 4: λ values corresponding to figure 13.

Table 4 supports the results from $\cos(\theta)$: that the backgrounds are reconstructed fairly well; the inelastic signal region oversimulated and not presenting strong evidence on the other signal regions. While there is less evidence supporting the simulation accuracy of the unreconstructed pion region compared to the reconstructed pion region, both table 4 and figure 12 support that the unreconstructed pion region is present at about the rate predicted by simulation with tuning. The significant disagreement at lower energies is worth further investigation but does not change the evidence for the disagreement in the inelastic signal region. Even if there were no unreconstructed pions below 200MeV, there still would be an excess of inelastic events throughout the plot.

5 Conclusion

We have produced a comparatively large sample of two proton and single muon events from MINERvA's mid-energy data runs. Within this sample, Monte-Carlo simulation predicted a signal region with little angular correlation, excluding for a large dip at high $\cos(\theta)$. We found that applying a background tuning on high $\cos(\theta)$ improved the data/MC agreement for both $\cos(\theta)$ and the leading proton energy. When comparing the simulation to the results from the MINERvA detector we find almost none of the predicted signal events in our data. Inelastic interactions were the leading predicted signal region and even ignoring the other signal prediction disagrees significantly with data. The data also suggests that the quasi-elastic regions are being oversimulated; however we do not know to what extent. Background events, regardless of whether or not the pion was reconstructed in the final state, seem to be predicted fairly well, especially in consideration of the prediction of the inelastic signal. Our results do not strongly confirm or reject the simulation of 3 track quasielastic events or 2p2h events. While the project was initially planned to examine the production mechanisms of these events, the surprising lack of signal events means that this analysis is not possible. However the disagreement is itself a significant result and likely will necessitate further research to find its source and hopefully imply improvements to MINERvA's simulation. Systematic uncertainties, mainly originating from the beam, muon reconstruction or pion re-interactions, were not calculated for these distributions so a rigorous account of these uncertainties is necessary.

6 Acknowledgements

There are really so many that belong here. Foremost Professor Kevin McFarland and Jeffrey Keyklamp for supplying me with a great research project fairly last minute and being incredibly supportive through this whole process. Similarly working the MINERvA collaboration has been an incredible experience, and I am very thankful to everyone involved who has given suggestions to this project. At Swarthmore a huge thanks goes to Professors Catherine Crouch and Tristan Smith for working incredibly hard to make this work a reality. Finally thanks to Professor Wing-Ho Ko for a very useful round of last minute edits.

References

[1] R Acciarri, C Adams, J Asaadi, B Baller, T Bolton, C Bromberg, F Cavanna, E Church, D Edmunds, A Ereditato, et al. Detection of back-to-back proton pairs in charged-current neutrino interactions with the argoneut detector in the numi low energy beam line. *Physical Review D*, 90(1):012008, 2014.

[2] P Adamson, K Anderson, M Andrews, R Andrews, I Anghel, D Augustine, A Aurisano, S Avvakumov, DS Ayres, B Baller, et al. The numi neutrino beam. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 806:279–306, 2016.

[3] Sea Agostinelli, John Allison, K al Amako, John Apostolakis, H Araujo, P Arce, M Asai, D Axen, S Banerjee, G 2 Barrand, et al. Geant4?a simulation toolkit. *Nuclear instruments and methods in physics research section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 506(3):250–303, 2003.

[4] L Aliaga, L Bagby, B Baldin, A Baumbaugh, A Bodek, R Bradford, WK Brooks, D Boehnlein, S Boyd, H Budd, et al. Design, calibration, and performance of the minerva detector.

Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, 743:130–159, 2014.

[5] L Aliaga, M Kordosky, T Golan, O Altinok, L Bellantoni, A Bercellie, M Betancourt, A Bravar, H Budd, MF Carneiro, et al. Neutrino flux predictions for the numi beam. *Physical Review D*, 94(9):092005, 2016.

[6] Costas Andreopoulos, Christopher Barry, Steve Dytman, Hugh Gallagher, Tomasz Golan, Robert Hatcher, Gabriel Perdue, and Julia Yarba. The genie neutrino monte carlo generator: Physics and user manual, 2015.

[7] John N. Bahcall, Aldo M. Serenelli, and Sarbani Basu. New Solar Opacities, Abundances, Helioseismology, and Neutrino Fluxes. *apjl*, 621(1):L85–L88, Mar 2005.

[8] P Berge, A Blondel, P Böckmann, H Burkhardt, F Dydak, JGH De Groot, AL Grant, R Hagelberg, EW Hughes, M Krasny, et al. Total neutrino and antineutrino charged current cross section measurements in 100, 160, and 200 gev narrow band beams. *Zeitschrift für Physik C Particles and Fields*, 35(4):443–452, 1987.

[9] M Betancourt, A Ghosh, T Walton, O Altinok, L Bellantoni, A Bercellie, A Bodek, A Bravar, T Cai, DA Martinez Caicedo, et al. Direct measurement of nuclear dependence of charged current quasielasticlike neutrino interactions using minerva. *Physical review letters*, 119(8):082001, 2017.

[10] Bruce T Cleveland, Timothy Daily, Raymond Davis Jr, James R Distel, Kenneth Lande, CK Lee, Paul S Wildenhain, and Jack Ullman. Measurement of the solar electron neutrino flux with the homestake chlorine detector. *The Astrophysical Journal*, 496(1):505, 1998.

[11] Clyde L Cowan, Frederick Reines, FB Harrison, HW Kruse, and AD McGuire. Detection of the free neutrino: a confirmation. *Science*, 124(3212):103–104, 1956.

[12] G. Danby, J-M. Gaillard, K. Goulianos, L. M. Lederman, N. Mistry, M. Schwartz, and J. Steinberger. Observation of high-energy neutrino reactions and the existence of two kinds of neutrinos. *Phys. Rev. Lett.*, 9:36–44, Jul 1962.

[13] Y Fukuda, T Hayakawa, E Ichihara, K Inoue, K Ishihara, Hirokazu Ishino, Y Itow, T Kajita, J Kameda, S Kasuga, et al. Measurement of the flux and zenith-angle distribution of upward throughgoing muons by super-kamiokande. *Physical Review Letters*, 82(13):2644, 1999.

[14] R. Gran, J. Nieves, F. Sanchez, and M. J. Vicente Vacas. Neutrino-nucleus quasi-elastic and 2p2h interactions up to 10 gev. *Phys. Rev. D*, 88:113007, Dec 2013.

[15] Richard Gran, J Nieves, F Sanchez, and MJ Vicente Vacas. Neutrino-nucleus quasi-elastic and 2p2h interactions up to 10 gev. *Physical Review D*, 88(11):113007, 2013.

[16] OctoMiao and The Gitter Badger. Neuphysics/neutrino: Linear stability analysis, November 2016.

[17] M. L. Perl, G. S. Abrams, A. M. Boyarski, M. Breidenbach, D. D. Briggs, F. Bulos, W. Chinowsky, J. T. Dakin, G. J. Feldman, C. E. Friedberg, D. Fryberger, G. Goldhaber, G. Hanson, F. B. Heile, B. Jean-Marie, J. A. Kadyk, R. R. Larsen, A. M. Litke, D. Lüke, B. A. Lulu, V. Lüth, D. Lyon, C. C. Morehouse, J. M. Paterson, F. M. Pierre, T. P. Pun, P. A. Rapidis,

B. Richter, B. Sadoulet, R. F. Schwitters, W. Tanenbaum, G. H. Trilling, F. Vannucci, J. S. Whitaker, F. C. Winkelmann, and J. E. Wiss. Evidence for anomalous lepton production in $e^+ - e^-$ annihilation. *Phys. Rev. Lett.*, 35:1489–1492, Dec 1975.

[18] Sam Zeller. Neutrino cross sections, 2015.