

1. Introduction

In the early 20th century, physicists were puzzled because energy, momentum, and angular momentum did not seem to be conserved in beta decay (the decay of a neutron). The energy, momentum, and angular momentum of the observed proton and electron did not match those of the original neutron. This observation led some physicists to suggest abandoning conservation laws. In 1930, in order to preserve the conservation laws, Pauli proposed the existence of another particle involved in beta decay that had not been detected yet. This particle would have to have no electric charge, spin 1/2, and was originally thought to be massless.

Pauli's particle would eventually be named the neutrino by Enrico Fermi, which means "little neutral one" in Italian. It was not until 1956 that the neutrino was detected. This is not surprising since neutrinos are very difficult to detect. In fact, the mean free path of neutrinos through Pb is about a lightyear. In 1956 Cowan and Reines performed an experiment in which neutrinos produced from nuclear reactor beta decays interacted with protons, thus undergoing inverse beta decay¹ to produce neutrons and positrons [1]. Until 1962, only one type of neutrino was known to exist, the one that produced electrons and positrons during its reactions. In 1962 Lederman, Schwartz, and Steinberger showed another type existed by creating neutrinos with muons and showing that these neutrinos produced muons when they interacted with Ne nuclei [2].

This experiment also led to the theory of lepton flavor conservation, the idea that only ν_e could produce electrons, while only ν_μ could produce muons. In the Standard Model, the fundamental particles can be divided into 3 categories: force carriers, quarks, and leptons (see

¹ Though unknown at the time, this experiment was actually done with antineutrinos $p + \bar{\nu}_e \rightarrow n + e^+$

Fig 1.1). Each of the four fundamental forces are governed by the exchange of the corresponding force carrier between two particles. Quarks bond together to form hadrons, like the proton and neutron. The last six are known as leptons. As you can see in Fig 1.1, quarks and leptons fall into three families or generations. When quarks decay they can cross generations, however, the theory of lepton flavor conservation prevents this from happening with the leptons. After Lederman et al's experiment, it became clear that there were two different types of neutrinos with flavors corresponding to the leptons produced during their reactions, the electron neutrino (ν_e) and the muon neutrino (ν_μ). After the Stanford Linear Accelerator's discovery of the tau lepton in 1975, a tau neutrino (ν_τ) was expected to also exist and was finally directly detected in 2000 at Fermilab [3].

| | | | | |
|---------|---------|-----------|------------|----------------|
| Quarks | u | c | t | γ |
| | d | s | b | g |
| Leptons | ν_e | ν_μ | ν_τ | Z |
| | e | μ | τ | W |
| | | | | Force Carriers |

Fig 1.1- The fundamental particles of the Standard Model

In the 1960s, Ray Davis set out to measure the most energetic ν_e produced in the sun during the PPIII boron decay, however, these neutrinos only had an energy of a few MeV [4]. Thirty years of attempts to measure the flux of solar ν_e by several experiments all found a significant deficit, which is often referred to as the solar neutrino problem. One possible explanation put forward to explain this deficit was that neutrinos oscillate between flavors. Fewer ν_e were detected than expected because some had oscillated to other flavors and the

detectors were only sensitive to ν_e ². More evidence for neutrino oscillation came from SuperKamiokande, which studied atmospheric ν_μ . SuperKamiokande found that fewer neutrinos were being detected when they traveled through the earth to the detector than those that came from directly above the detector. Though no ν_τ were directly detected, the deficit could be well explained by $\nu_\mu \rightarrow \nu_\tau$ oscillations [5]. In 2001 the Sudbury Neutrino Observatory directly detected ν_e and the other two flavors of neutrinos³ coming from the sun [6]. Only 35% of the detected neutrinos were ν_e and the rest had oscillated from ν_e to one of the other two flavors, confirming both the validity of the standard solar model and neutrino oscillations.

Until the discovery of neutrino oscillations, neutrinos were assumed to be massless. However, if they oscillate to different flavors, then at least two of the three neutrinos must have different, non-zero masses. Neutrinos oscillate because each flavor eigenstate is a linear combination of the three different mass eigenstates (see section 2). This mixing is governed by the unitary Maki-Nakagawa-Sakata matrix. Since it is unitary, the MNS matrix can be specified by three mixing angles and one complex phase. A non-zero complex phase leads to Charge-Parity (CP) violation. CP symmetry states that physics should be the same if matter were changed into antimatter and the mirror image was studied. This, however, was not found to be true for certain weak decays. CP violation is particularly important to understand because it may lead to an explanation of the matter/anti-matter asymmetry in the universe. The MNS matrix is a possible new source of CP violation.

² Davis detected electron neutrinos via the beta decay off chlorine atoms $\nu_e + \text{Cl}^{37} \rightarrow e^- + \text{Ar}^{37}$. The PPIII ν_e just barely had enough energy for this reaction to occur, therefore a similar reaction with ν_μ or ν_τ would not happen since the neutrinos would not have enough energy to create the corresponding muon and tau.

³ SNO could not distinguish between ν_μ and ν_τ

The rate of neutrino oscillations are governed by the three mixing angles and the differences of the squares of the masses ($\Delta m^2_{ij} = m^2_i - m^2_j$). Studies of solar neutrino oscillations ($\nu_e \rightarrow \nu_\mu$ and $\nu_e \rightarrow \nu_\tau$) and atmospheric neutrino oscillations ($\nu_\mu \rightarrow \nu_\tau$) have led to good measurements of two of the mixing angles θ_{12} and θ_{23} and the magnitudes of the mass differences Δm^2_{21} and Δm^2_{32} . However, a good measurement of $\nu_\mu \rightarrow \nu_e$ oscillations has not been made. Measuring this oscillation rate is necessary for determining the third mixing angle θ_{13} . Also, if these neutrinos traveled through the Earth before being detected, the matter effects would determine the sign of the mass difference Δm^2_{32} . θ_{13} is coupled with the CP violating phase in the neutrino mixing matrix and therefore must be well measured before CP violation in neutrino oscillations can be studied. Determining the sign of Δm^2_{32} would determine the mass hierarchy of the neutrino flavors, which is currently unknown. The current mass differences are $\Delta m^2_{21} = (7.9 \pm 0.6) \cdot 10^{-5} \text{ eV}^2$ and $m_{32} = (1.9-3.0) \cdot 10^{-3} \text{ eV}^2$ [7].

Efforts are currently being made to detect these $\nu_\mu \rightarrow \nu_e$ oscillations. One such detector is a Liquid Argon (LAr) Time Projection Chamber (TPC). LAr TPCs are a promising technology that could lead to high-resolution neutrino detection with smaller and cheaper detectors than currently exist. Research and development is being done on detectors that will be exposed to the NuMI beam at Fermi National Accelerator Laboratory. A $\frac{1}{4}$ ton prototype, ArgoNeuT, has been built to perform a near detector study at Fermilab in the coming months. Plans are also in progress to build a 5 kton detector in the next few years in order to do a long baseline ($L \sim 1000 \text{ km}$) study of these oscillations from the NuMI beam.

2. Neutrino Oscillations

The fact that neutrinos oscillate is both surprising and fascinating. What starts out as one neutrino flavor can propagate through a vacuum and have a probability of being detected as a different flavor. This is because the flavor eigenstates of the neutrinos are linear combinations of the different mass eigenstates.

$$|\nu_\alpha\rangle = \sum_{i=1}^3 U_{\alpha i}^* |\nu_i\rangle \quad \alpha = e, \mu, \tau \quad (1)$$

$U_{\alpha i}^*$ are matrix elements of the unitary Maki-Nakagawa-Sakata (MNS) mixing matrix.

$$U = \begin{pmatrix} c_{13}c_{12} & c_{13}s_{12} & s_{13}e^{-i\delta} \\ -c_{23}s_{12} - s_{12}s_{23}c_{12}e^{i\delta} & c_{23}c_{12} - s_{13}s_{23}s_{12}e^{i\delta} & c_{13}s_{23} \\ s_{23}s_{12} - s_{13}c_{23}c_{12}e^{i\delta} & -s_{23}c_{12} - s_{13}c_{23}s_{12}e^{i\delta} & c_{13}c_{23} \end{pmatrix} \quad (2)$$

where $c_{jk} = \cos(\theta_{jk})$ and $s_{jk} = \sin(\theta_{jk})$

As the neutrino evolves through time, the flavor eigenstates becomes

$$|\nu_\alpha(t)\rangle = \sum_{i=1}^3 e^{i(\vec{p}\cdot\vec{x} - E_i t)} U_{\alpha i}^* |\nu_i\rangle \quad (3)$$

Since neutrino's masses are very small, they are ultra-relativistic ($m \ll p$). In this limit (setting $c = 1$), $t \approx x \equiv L$, where L is the distance the neutrino has traveled before being detected. Also, the energy becomes

$$E = \sqrt{p^2 + m^2} \approx p + \frac{m^2}{2p} \approx E + \frac{m^2}{2E} \quad (4)$$

As the neutrino evolves in time, the coefficients in front of the different mass eigenstates oscillate, causing the probability of the neutrino being detected as flavor α to also oscillate. If the neutrino is in the $|\nu_\alpha\rangle$ flavor eigenstate at $t=0$, then the probability of a neutrino oscillating from $\nu_\alpha \rightarrow \nu_\beta$ in time t (after traveling a distance L) is given by

$$P_{vac}(\nu_\alpha \rightarrow \nu_\beta) = \left| \langle \nu_\beta | \nu_\alpha(t) \rangle \right|^2 \approx \left| \sum_i U_{\alpha i}^* U_{\beta i} e^{-im_i^2 L/2E} \right|^2 \quad (5)$$

To simplify Eq5, only mixing between initial and final flavors is considered significant. This causes $P(\nu_\alpha \rightarrow \nu_\beta)$ to depend primarily on $\theta_{\alpha\beta}$ and a specific mass difference. The Sudbury Neutrino Observatory (SNO) [6], SuperKamiokande [5], and KamLAND [8] have studied solar $\nu_e \rightarrow \nu_\mu$ oscillations and give $\Delta m^2_{21} = (7.9 \pm 0.6) * 10^{-5} \text{ eV}^2$ and $\sin^2(2\theta_{12}) = 0.82 \pm 0.07$ at the 90% confidence level. SuperKamiokande has also measured atmospheric $\nu_\mu \rightarrow \nu_\tau$ oscillations along with the K2K experiment [9]. These measurements give $\Delta m^2_{32} = (1.9 - 3.0) * 10^{-3} \text{ eV}^2$ and $\sin^2(2\theta_{23}) > 0.92$ at the 90% confidence levels. These values of θ_{12} and θ_{23} are very high, signifying almost maximum mixing for the corresponding oscillations. The CHOOZ experiment has placed a limit on θ_{13} to be $\sin^2(2\theta_{13}) < 0.19$ at the 90 % confidence level. Unlike θ_{12} and θ_{23} , θ_{13} is very small, corresponding to minimal mixing.

At first one would think to measure θ_{13} by studying $\nu_e \rightarrow \nu_\tau$ oscillations, however, the resulting reaction with the ν_τ would produce a τ , which has a mass of 1.776 GeV. An experiment studying these oscillations would need very energetic ν_e 's, which is not possible for a reactor based experiment. The MNS matrix shows that while $\nu_\alpha \rightarrow \nu_\beta$ oscillations depend primarily on $\theta_{\alpha\beta}$, other mixing angles are also involved. Consequently, studying $\nu_\mu \rightarrow \nu_e$ oscillations would allow for a better measurement of θ_{13} .

Simplifying Eq5 by only considering ν_μ and ν_e mixing gives the $\nu_\mu \rightarrow \nu_e$ oscillation probability to be

$$P_{vac}(\nu_\mu \rightarrow \nu_e) = \sin^2 \theta_{23} \sin^2 2\theta_{13} \sin^2 \left(1.27 \frac{\Delta m^2_{32} L}{E} \right) \quad (6)$$

where Δm_{32}^2 is in eV^2 , L is in km, and E is in GeV. Note that only the magnitude of the Δm^2 term can be measured from pure oscillations.

In order to maximize the probability in Eq6,

$$1.27 \frac{\Delta m_{32}^2 L}{E} = \frac{\pi}{2} + n\pi \quad (7)$$

The NuMI beam energy spectrum peaks around 2GeV. For this energy and $\Delta m_{32}^2 = 2.5 \cdot 10^{-3}$, L should be 990 km. Therefore, a long baseline study of $\nu_\mu \rightarrow \nu_e$ oscillations using the NuMI beam should place its far detector ~ 1000 km from Fermilab. At these energy and distance settings, assuming $\sin^2(\theta_{23}) = \frac{1}{2}$ (maximal mixing with θ_{23}) and optimistically setting $\sin^2(2\theta_{13}) = (0.14 \text{ to } 0.18)$ gives $P(\nu_\mu \rightarrow \nu_e)$ ranging from 7%-9%.

This probability is calculated for the neutrino propagating through a vacuum. There are significant matter effects that contribute to the oscillation probability if the neutrinos were to propagate through matter. For our long baseline experiment, NuMI beam neutrinos will have to travel through the Earth to reach the detector. Taking into account matter effects, the probability becomes

$$P_{\text{matter}}(\nu_\mu \rightarrow \nu_e) \cong \left(1 \pm 2 \frac{E}{E_R}\right) P_{\text{vac}}(\nu_\mu \rightarrow \nu_e) \quad (8)$$

where E_R is the matter resonance energy

$$E_R = \frac{\Delta m_{32}^2}{2\sqrt{2}G_F N_e} \quad (9)$$

G_F depends on the density and the average Z/A of the material the neutrino is propagating through. The important thing about Eq8 is that the probability is now sensitive to the sign of

Δm^2_{32} . Therefore, by measuring the oscillated neutrinos after they have propagated through the Earth, one can determine the mass hierarchy of the neutrinos, that is whether $m^2_3 > m^2_2$ or $m^2_3 < m^2_2$.

3. The Detector

Liquid Argon Time Projection Chambers are a promising technology that we plan to use to study $\nu_\mu \rightarrow \nu_e$ oscillations. The chamber is a box filled with LAr with a constant electric field applied to it. Neutrinos react with the argon nuclei via charged current (CC) or neutral current (NC) reactions. The charged particles from those reactions then travel through the argon ionizing the atoms as they go. The applied electric field causes these ionized electrons to drift with a constant velocity to two different wire planes, the collection and induction planes. Which wires were hit give two spatial points, and the drift time and speed are used to infer the third. Therefore these detectors have the ability to produce a 3D display of an event in the entire volume, as oppose to sandwich detectors like MINOS, which only display the event in parts of the detector volume.

Currently a $\frac{1}{4}$ ton prototype named ArgoNeuT has been built and is ready to be exposed to the NuMI beam. The detector will be placed in front of the NuMI beam at Fermilab for a near detector low energy study. The TPC is 50 cm x 49.9 cm x 82 cm with a 4mm wire spacing, will have a 500 V/cm electric field applied, and an electron drift velocity of 0.16 cm/ μ s. Since these electrons have to drift up to 50 cm through the LAr (see Fig 6.2), high purity is very important. Impurities such as oxygen and water can intermix with the LAr and capture ionized electrons before they reach the detection planes. Without a high purity, we will lose a lot of event

information. For ArgoNeuT the purity must be at least 10 ppb. The argon cycles through two filters and its purity is monitored by using a xenon flash lamp to excite photoelectrons off a photocathode and then measuring how many photoelectrons survive [10].

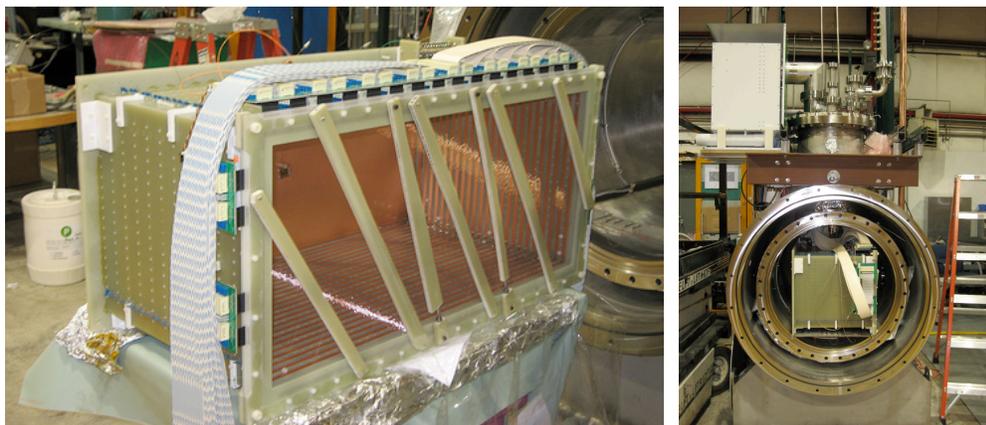


Fig 3.1- ArgoNeuT TPC (right) and TPC in the cryostat (left) Source: <http://t962.fnal.gov/index.html>

The next step after ArgoNeuT is to build a multi-kton detector that will be placed ~ 1000 km away for a long baseline study. The location of the detector and whether or not a new neutrino beam will be built for it are still being determined. If we choose to use the current NuMI beam, the detector will be placed off axis since that causes the neutrinos to be in a narrow energy band (see Fig 3.2). If the energy of the incoming neutrino is so tightly constrained, then the detector can be placed at the optimal L in order to maximize the oscillation probability.

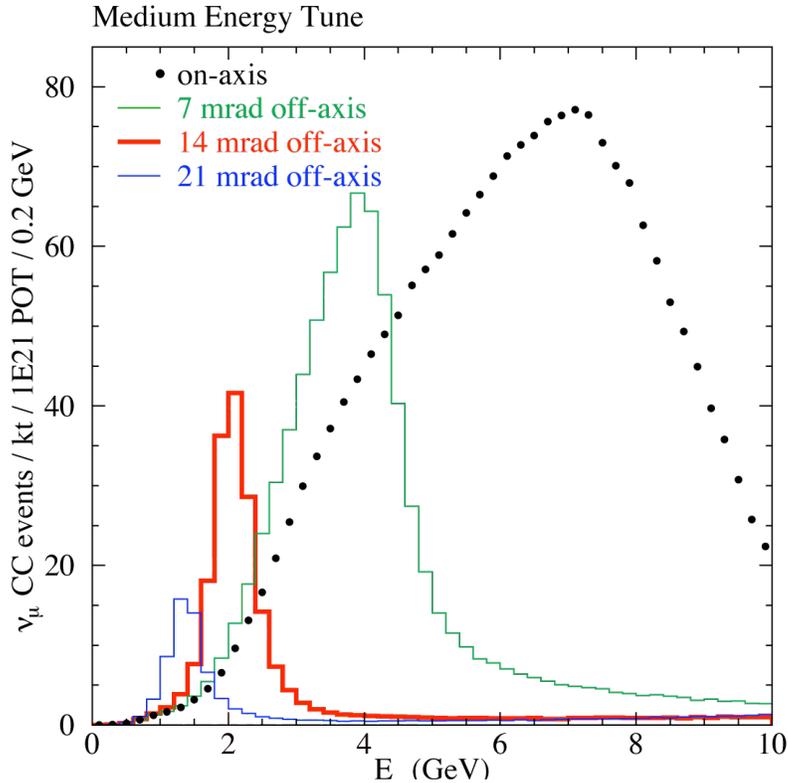


Fig 3.2- CCmu event rates expected if no oscillation at a distance 800km for the NuMI medium-energy beam configuration. Source: NoVA Proposal [11]

4. Types of Events

Neutrinos will interact weakly with the argon nuclei by exchanging either a W^- or Z^0 to yield charged current (CC) or neutral current (NC) reactions respectively. The basic CC reaction is $\nu_{\alpha} + n \rightarrow p + l_{\alpha}$ where l_{α} is the lepton corresponding to the neutrino flavor α . Since we are looking for $\nu_{\mu} \rightarrow \nu_e$ oscillations, $\nu_e + n \rightarrow p + e^-$ is our signal reaction, this is called a charged-current electron (CCe) event. It is characterized by having an electron produced at the primary vertex along with a proton (see Fig 4.1). The quasi-elastic scattering is the most basic, however, resonance scatterings (RES) and deep inelastic scatterings (DIS) can have additional protons and pions coming from the primary vertex. The key thing to look for to distinguish CCe events is the primary vertex electron. Electrons are characterized by the shower they produce as they travel

through the argon and interact with the argon atoms, however the length and size of these showers can vary based on the energy of the electron.

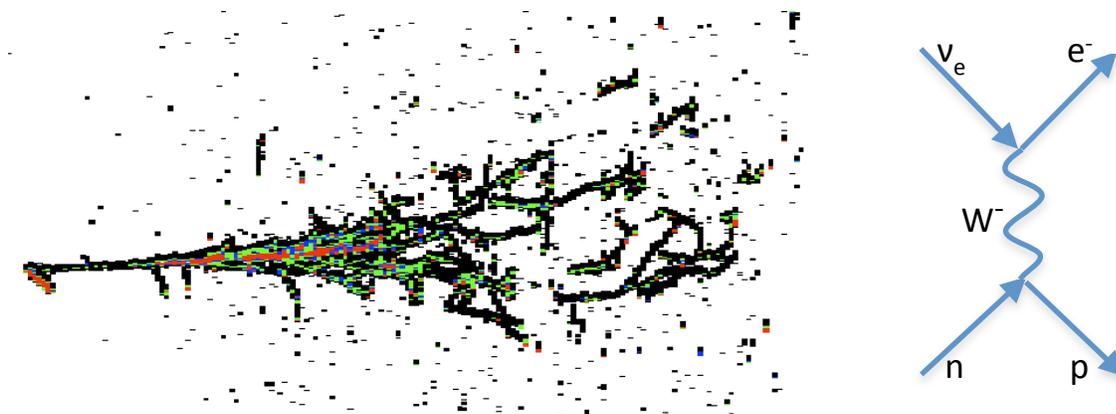


Fig 4.1- Monte Carlo simulated CCE event (right) and corresponding Feynman diagram (left)

The other two events that will serve as background are charged-current muon events (CCmu) and neutral current events (NC). The CCmu events are similar to the CCE events except a muon is produced instead of an electron. $\nu_\mu + n \rightarrow p + \mu^-$. Since the muon is much heavier than the electron, it will just pass through the argon atoms leaving a long, straight, minimum-ionizing track (see Fig 4.2). Although, the muon can sometimes decay into an electron, which could then shower a little, and pions and protons can be produced at the primary vertex during more complicated scatterings.

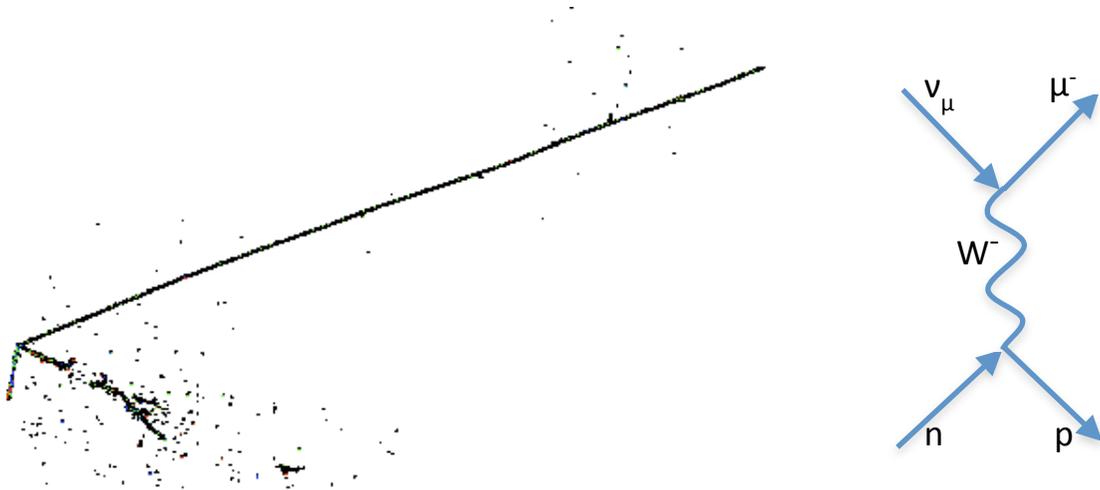


Fig 4.2- Monte Carlo simulated CCmu event (right) and corresponding Feynman diagram (left)

NC events can occur from an Ar reaction with any type of neutrino, but leave no signature as to which flavor caused the reaction and are therefore a background. The basic reaction is $\nu_\alpha + N \rightarrow \nu_\alpha + N$. Typically, no charged particles are produced in the basic reaction so NC events are characterized by not having any signature tracks coming from the primary vertex (since the detector is only sensitive to charged particles). However RES and DIS scatterings could cause protons and pions to come from the primary vertex. A significant source of background is from π^0 production since they can decay into two photons, which could then convert to a e^-e^+ pair (see Fig 4.3). Since no magnetic field is applied, the photon conversion pair could travel on top of each other and shower like a single electron. If this photon conversion occurs close to the primary vertex, the NC event could easily be mistaken for a CCE.

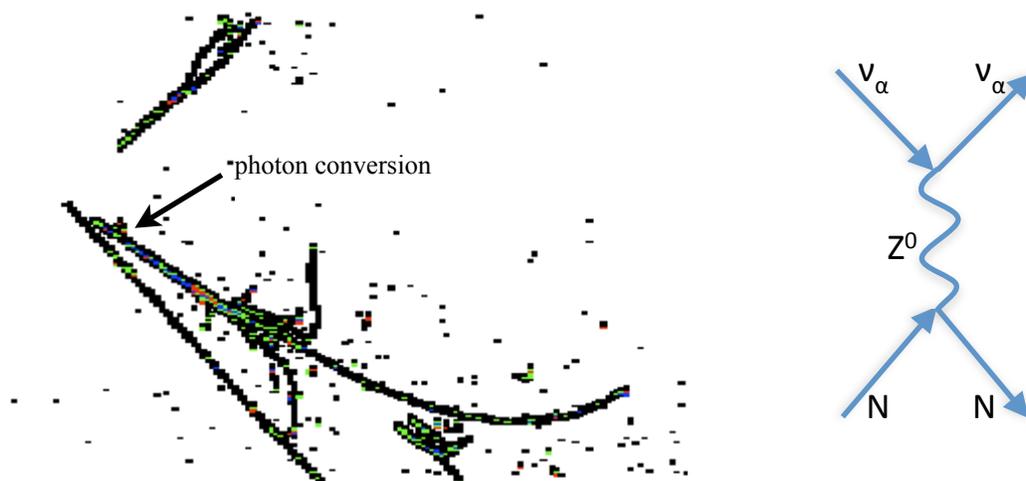


Fig 4.3- Monte Carlo simulated NC event (right) and corresponding Feynman diagram (left)

5. Signal Detection Difficulties

As previously discussed, the probability of $\nu_\mu \rightarrow \nu_e$ oscillations is very low (7%-9% at best). Since neutrinos are so difficult to detect, the CCE signal will be small in an already small sample of events. Also, while the NuMI beam is designed to produce ν_μ , a small ν_e background exists (0.5%-1%) from muon and kaon decays. These intrinsic beam ν_e 's will be characterized by having a broader energy spectrum (see Fig 5.1), but they still present a significant background. Also, a near detector study, like ArgoNeuT, can help give a measurement of this beam background. Long baseline studies that will study $\nu_\mu \rightarrow \nu_e$ oscillations from the NuMI beam will need to be able to pull the small CCE signal from the CCmu, NC, and beam ν_e background. NOvA is one such long baseline (810 km) experiment that uses an above ground liquid scintillator detector [11]. Fig 5.1 shows their predicted background and signal strengths, using a low $\sin^2(2\theta_{13})$. As you can see, the background events will dominate and so it is very important to have good background reduction in order to pull the small CCE signal above the

intrinsic beam CCE events. At 2 GeV, the intrinsic beam background is about 20% of the total CCE signal. The intrinsic beam background also has a clearly different energy spectrum, which can be used to distinguish it from the signal CCE events.

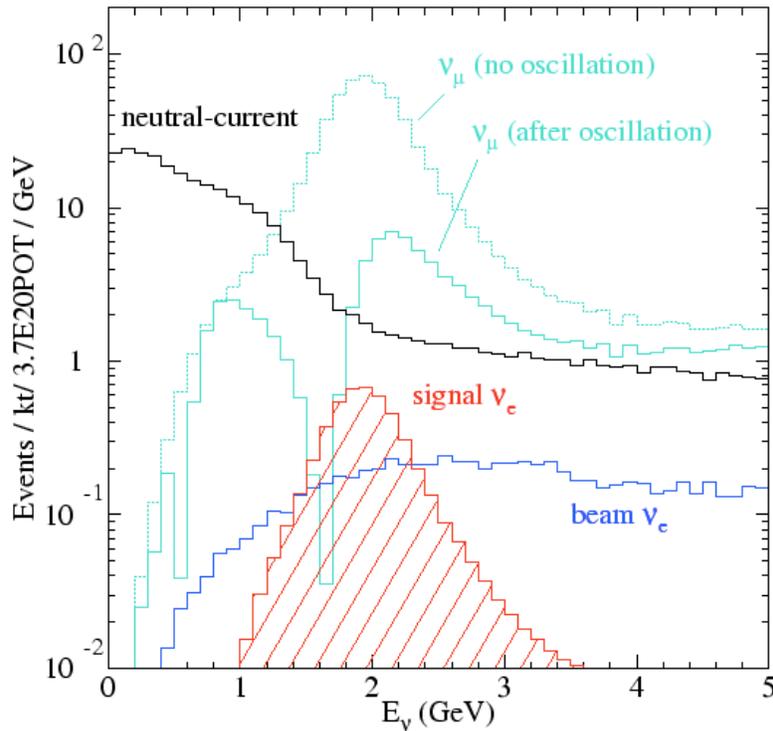


Fig 5.1- Simulated energy distributions for the oscillation signal, intrinsic beam events, neutral-current events and ν charged-current events with and without oscillations. The simulation used $\Delta m_{32}^2 = 2.5 \cdot 10^{-3} \text{ eV}^2$, $\sin^2(2\theta_{23})$, and $\sin^2(2\theta_{13})=0.04$. An off-axis distance of 12 km at 810 km was assumed. Source: NOvA proposal [11]

6. Multivariate Analysis

Since the expected CCE flux is so low it is crucial to have an event identification algorithm that efficiently reduces the CCmu and NC background, while keeping enough CCE signal to be seen above the intrinsic beam ν_e background. A Monte Carlo (MC) program written by Dr. Bruce Baller was used to simulate the events the LAr TPC is expected to detect [12]. The user defined what fraction of the neutrinos were each flavor and what fraction of the reactions were CC or NC. Then, using the NuMI neutrino spectrum (see Fig 6.1) to pick the neutrino's

energy, the program had a random neutrino interact with the Ar nuclei. It then used GEANT to simulate the corresponding tracks through the LAr. These files were then read into another program that analyzed the signal as will be done by the actual the TPC's and outputted several variables into an Ntuple. An Ntuple is a giant array of variables that can be read and analyzed by The Physics Analysis Workstation (PAW).

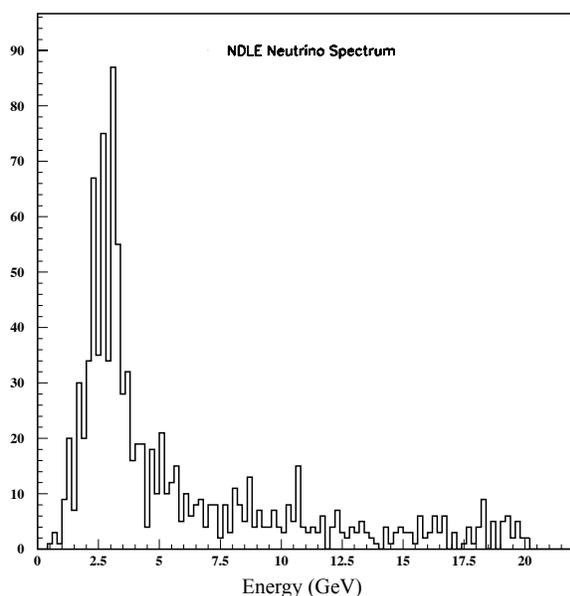


Fig 6.1- NuMI beam energy spectrum used in Monte Carlo simulations

The Monte Carlo events generated were done so using ArgoNeuT's geometry. The TPC was 50 cm x 49.5 cm x 82 cm with the two wire planes (U and V) placed 60° and 120° from the drift direction (X) with a wire spacing of 10mm (see Fig 6.2). The electron lifetime was set to 0.15s and the drift velocity was set at 0.16 cm/ μ s. The electron lifetime is the average time an electron can travel before being absorbed by an impurity (like oxygen). This factor will be important when modeling events in the larger kton detectors, however, for this study, the electron lifetime was set high so no ionized electrons were lost. Also, only events with a true visible

energy < 4 GeV were studied since these events will be in the optimal energy range of the oscillated ν_e .

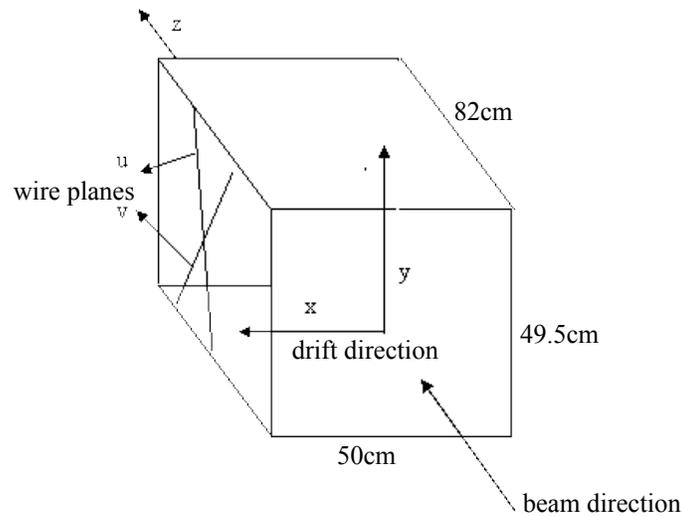


Fig 6.2- Geometry of the TPC used in Monte Carlo simulation

Examples of the analyzed and plotted MC events can be seen in Fig 4.1-Fig 4.3 from the previous section. The wire and drift position of each hit is plotted along with a color depicting its minimum ionizing particle (MIP) value (see Table 6.1). A MIP is a particle that has lost the minimal amount of energy in the LAr. Basically MIP is a measure of how much energy was deposited at that position by the passing particle. Therefore, at each wire coordinate position and drift coordinate position, whether or not a hit was recorded and how much energy was deposited is measured.

| | |
|-------|---------------------------|
| black | $0 < \text{MIPs} < 1.5$ |
| green | $1.5 < \text{MIPs} < 2.5$ |
| blue | $2.5 < \text{MIPs} < 5$ |
| red | $\text{MIPs} > 5$ |

Table 6.1- MIP color ranges for Monte Carlo plots

7. Distinguishing Variables

First, a group of distinguishing variables needed to be found that could distinguish the background CCMu and NC events from the signal CCE events. Then, simple cuts on these variables would be used to decide if an event was signal or background. Also, using these variables in a multivariate analysis tool called a decision tree (see section 9) was explored. 1000 events of each type were created and analyzed in PAW. Recall that CCE events are characterized by having an electron coming out of the primary vertex, which leave a signature shower. It is this shower that served as the inspiration for most of the variables. The six variables used in this study were *ncluster*, *notused*, *pvelec*, *long*, *nhit*, and *tmip*.

- *ncluster*

The particles produced in these showers travel and leave several tracks. Instead of creating a 3D construction of the event, Dr Baller's code plotted two 2D projections of the event with the wire coordinate as the horizontal axis and the drift coordinate as the vertical axis. His analysis code, *tanal*, was able to fit 2D projections of the tracks called clusters for each view. *ncluster* is a variable that is the sum of the number of clusters fitted in both 2D views. As you can see, the background events (CCMu and NC) typically have only a few clusters, while the CCE events have more because of the electron shower.

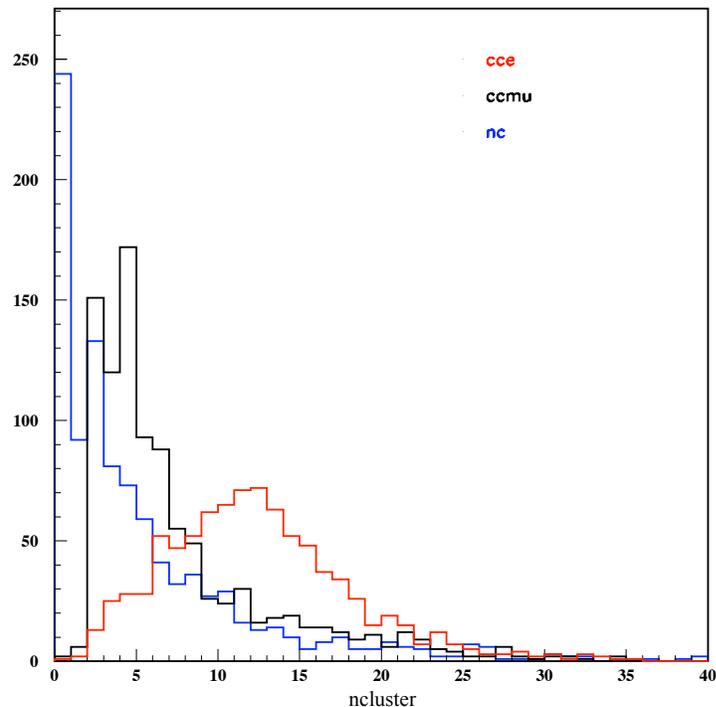


Fig 7.1- Histogram of ncluster for CCE (red), CCmu (black), and NC (blue) events.

- *notused*

The photons produced from the bremsstrahlung radiation of the electrons produced in the electromagnetic shower ionize LAr electrons that do not travel long enough to be associated with any cluster. These electrons produce the speckle found around the shower clusters. *notused* is a variable that is the total number of hits from both 2D views that were not fitted to any cluster. As you can see in Fig 7.2, like *ncluster*, the background events peak with low numbers of unused hits, while the CCE events tend to have more unused hits.

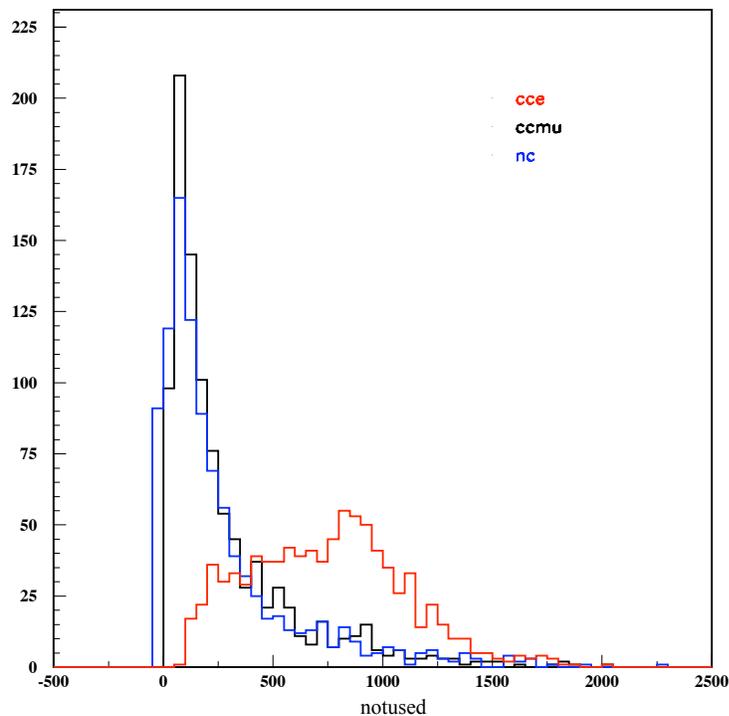


Fig 7.2- Histogram of notused for CCE (red), CCmu (black), and NC (blue) events.

- *pvelec*

pvelec attempts to find electron events coming from the primary vertex. Dr Baller's code also finds vertices and does a good job finding the primary vertex. In both 2D views, a triangle with an opening angle of 10^0 is drawn centered on each cluster coming out of the primary vertex. The number of hits in that triangle are then added up, and the hits from the original cluster are subtracted. Then, the primary vertex cluster with the most hits in that triangle from both 2D views is found. *pvelec* is the number of hits within the triangle centered on that cluster. Since particles like protons, pions, and muons do not shower like electrons, *pvelec* was expected to be greater than 0 if the primary vertex cluster chosen was an electron. However, this is subject to how good the vertexing is at not fitting photon conversions to the primary vertex. Also, since the triangle is drawn all the way to the edge of the plot (since the shower continues after the original

electron has stopped), messy events with lots of photon conversions could produce a false electron ID. However, since only low energy events (true visible energy $< 4\text{GeV}$) are used in this study, these false electron IDs will be scarce. As you can see in Fig 7.3, cutting on $p_{\text{velec}} > 1$ would cut away $\sim 71\%$ of the NC events and $\sim 45\%$ of the CCmu events while keeping $\sim 82\%$ of the CCE events. Therefore, despite the possibility of false electron ID from photon conversions, this variable does a good job at reducing the background events while keeping the signal events.

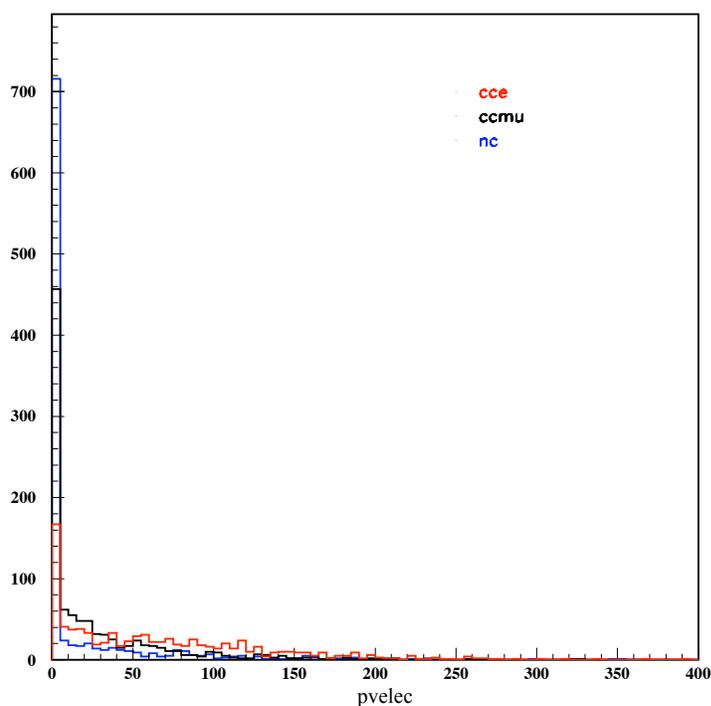


Fig 7.3- Histogram of p_{velec} for CCE (red), CCmu (black), and NC (blue) events.

- *long*

Since p_{velec} did a good job of cutting away the NC events, the *long* variable was considered in order to reduce the CCmu events. Recall that CCmu events are characterized by having a muon produced at the primary vertex, which then typically travels in a long, straight

track. long is simply the number of hits in the longest cluster. It should be noted that while the Monte Carlo program randomized the position of the primary vertex, it did so in an area less than half way down stream. Therefore, the muon (which typically travels downstream) had at least half the length of the TPC to travel through. In real life this will not be the case. However, events where the primary vertex occurs very far downstream will most likely be rejected since (unless there are a lot of backwards moving particles) the particles will not have traveled enough to give any useful information.

As you can see in Fig 7.4, long is not only a great variable for reducing the CCmu, but also for cutting away the NC events that produced few to no clusters.

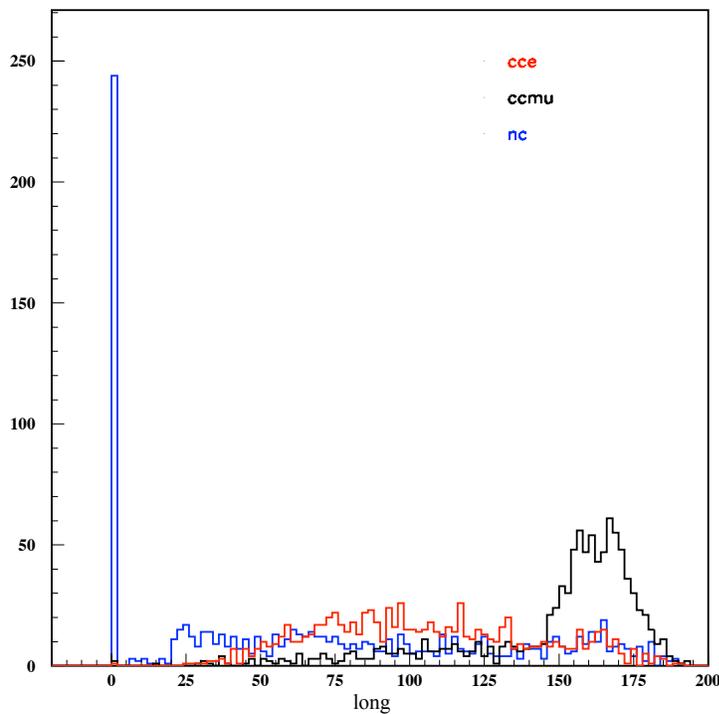


Fig 7.4- Histogram of long for CCE (red), CCmu (black), and NC (blue) events.

- *nhit*

nhit is simply the number of hits in both 2D views. Similarly to *ncluster* and *notused*, the background events peak for low *nhits* while the CCE events typically have more hits because of the electron shower (see Fig 7.5)

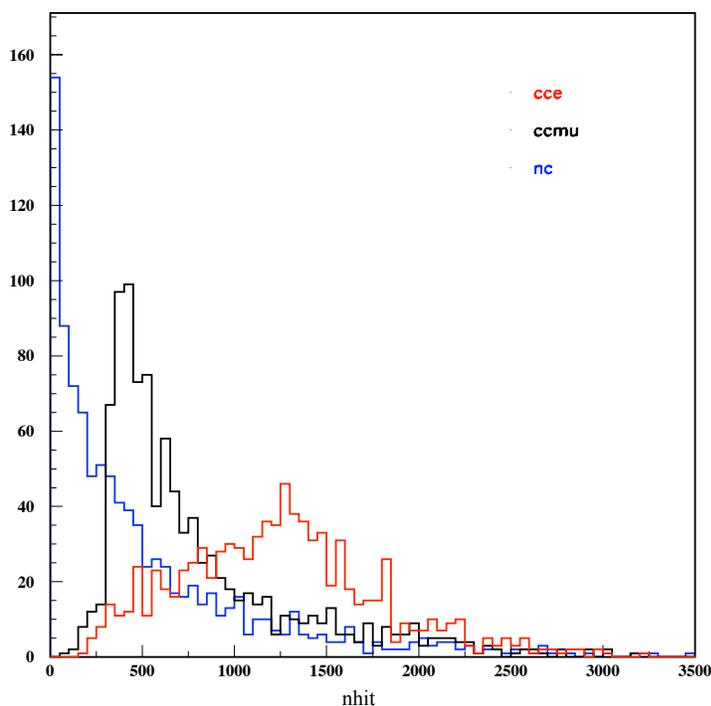


Fig 7.5- Histogram of *nhit* for CCE (red), CCmu (black), and NC (blue) events.

- *tmip*

tmip is the sum of all the MIP values from each hit in both 2D views. Recall that MIP is a measure of how much energy was deposited by the particle at each position. The electron showers associated with CCE events create a lot of hits, and therefore a large total MIP value. As you can see in Fig 7.6, the background events peak with low MIP values, while CCE events tend to have higher *tmip* values.

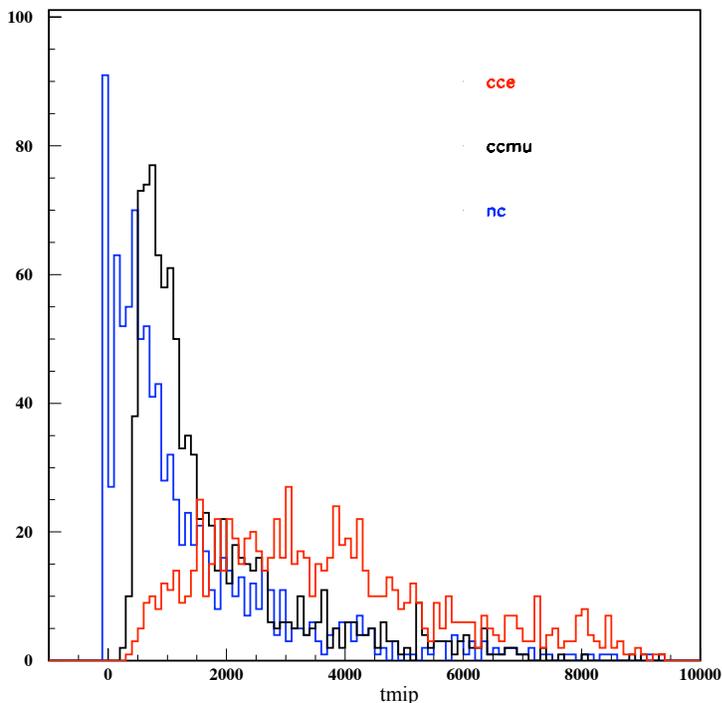


Fig 7.6- Histogram of tmip for CCE (red), CCmu (black), and NC (blue) events.

- *variable correlation*

An important characteristic of a variable is its correlation with another variable. If your variables are correlated, then making a cut on one is essentially making a cut on both, so you should just pick the best one. The variables ncluster, notused, nhit, and tmip are all correlated (see Fig 7.7)

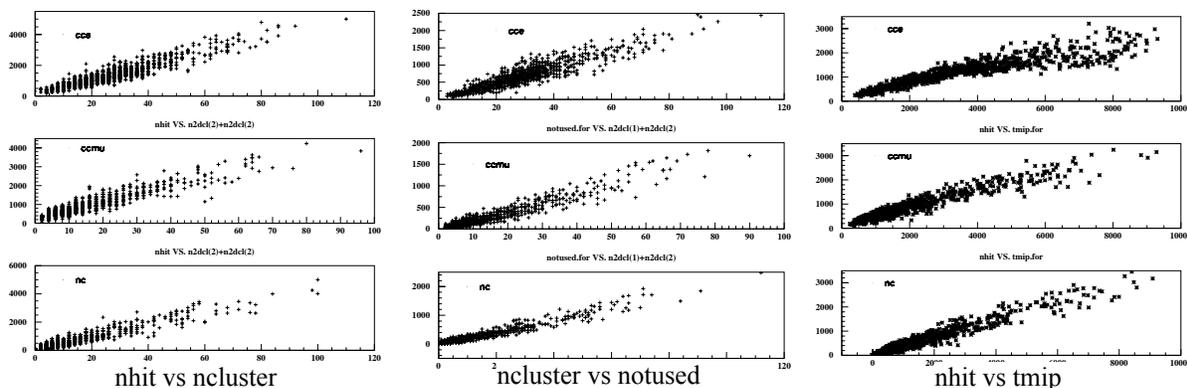


Fig 7.7- Plots of correlated variables for CCE (top panels), CCmu, (middle panels), and NC (bottom panels) events.

The variables `ncluster`, `long`, and `pvelec`, however, are not correlated (see Fig 7.8)

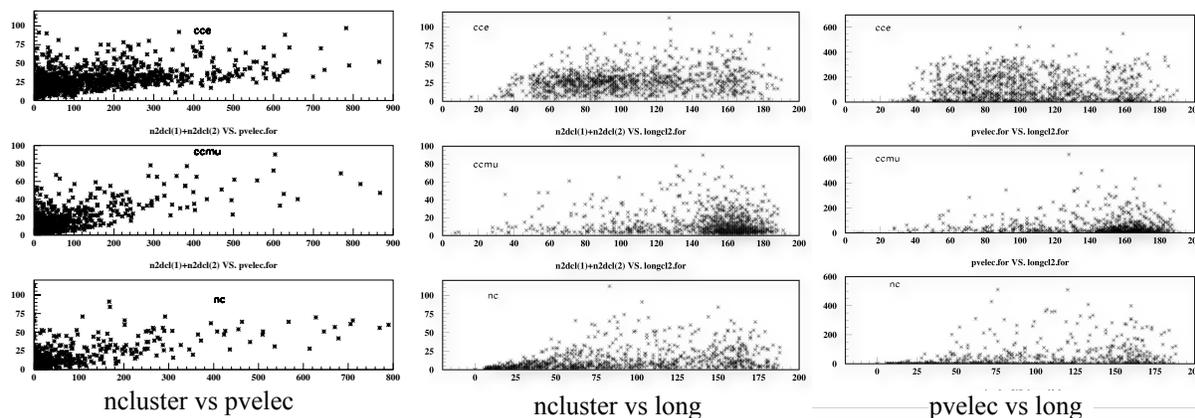


Fig 7.8- Plots of uncorrelated variables for CCE (top panels), CCmu, (middle panels), and NC (bottom panels) events.

8. PAW cuts analysis

A multivariate identification analysis was done by making simple cuts with the variables `ncluster`, `notused`, `pvelec`, and `long`. These were chosen because they are mostly uncorrelated. `notused` and `ncluster` were both used, even though they are correlated, because between the two of them, every hit is utilized. In order to decide where to cut, a brute force code was written to step through each possible cut and optimize $\frac{\text{signal}}{\sqrt{\text{signal} + \text{bkgd}}}$ using 1000 events of each reaction type (CCE, CCmu, and NC). The optimal cuts found in the code were `ncluster` > 2, `notused` > 235, `pvelec` > 0, and `long` < 192. Table 8.1 shows the results of using these cuts on 4 data sets containing roughly equal numbers of events for each type. CCE efficiency is the percentage of the original signal events that survived the cut and were classified as signal. Background misidentification is the percentage of the original background events that were misclassified as signal.

| <i>Data Set 1</i> | | | | <i>Data Set 3</i> | | | |
|----------------------------|------------|-------------|----------------|-------------------|------------|-------------|----------------|
| | CCe | Bkgd | ←classified as | | CCe | Bkgd | ←classified as |
| CCe | 654 | 227 | | CCe | 335 | 104 | |
| CCmu | 282 | 703 | | CCmu | 150 | 350 | |
| NC | 179 | 806 | | NC | 85 | 415 | |
| ↑actual | | | | ↑actual | | | |
| <i>Data Set 2</i> | | | | <i>Data Set 4</i> | | | |
| | CCe | Bkgd | ←classified as | | CCe | Bkgd | ←classified as |
| CCe | 680 | 203 | | CCe | 118 | 71 | |
| CCmu | 310 | 672 | | CCmu | 78 | 222 | |
| NC | 169 | 818 | | NC | 53 | 247 | |
| ↑actual | | | | ↑actual | | | |
| avg CCe efficiency | | | | 75.04% | | | |
| avg CCmu misidentification | | | | 29.05% | | | |
| avg NC misidentification | | | | 17.49% | | | |
| avg bkgd misidentification | | | | 23.27% | | | |

Table 8.1- Results from optimized cuts: ncluster > 2, notused > 235, pvelec > 0, long < 192

As you can see, these optimized cuts do a good job keeping CCe signal events, but only cut away about 3/4 of the background events. Most the the background that got through were CCmu events. However, the variable long was not really utilized with these optimal cuts, even though it has the potential to cut away a lot of the CCmu background. Therefore, a study was done making the cut $25 < \text{long} < 145$ in order to be classified as signal. Using the same cuts as before, but changing the long cut gave the results in Table 8.2

| <i>Data Set 1</i> | | | | <i>Data Set 3</i> | | | |
|----------------------------|------------|-------------|----------------|-------------------|------------|-------------|----------------|
| | CCe | Bkgd | ←classified as | | CCe | Bkgd | ←classified as |
| CCe | 533 | 348 | | CCe | 259 | 180 | |
| CCmu | 81 | 904 | | CCmu | 54 | 929 | |
| NC | 94 | 891 | | NC | 94 | 891 | |
| ↑actual | | | | ↑actual | | | |
| <i>Data Set 2</i> | | | | <i>Data Set 4</i> | | | |
| | CCe | Bkgd | ←classified as | | CCe | Bkgd | ←classified as |
| CCe | 553 | 330 | | CCe | 154 | 105 | |
| CCmu | 98 | 884 | | CCmu | 22 | 278 | |
| NC | 88 | 898 | | NC | 29 | 271 | |
| ↑actual | | | | ↑actual | | | |
| avg CCe efficiency | | | 60.17% | | | | |
| avg CCmu misidentification | | | 7.76% | | | | |
| avg NC misidentification | | | 8.37% | | | | |
| avg bkgd misidentification | | | 8.07% | | | | |

Table 8.2- Results from cuts: $n_{cluster} > 2$, $notused > 235$, $p_{elec} > 0$, $25 < long < 145$

As you can see, while the signal efficiency is reduced by 14.87%, significantly more CCmu events are cut.

As mentioned before, at best the signal CCe events will be ~7% of the total signal. Since this is the case, a study was done on four realistic data sets that contained 117 CCe, 1000 CCmu, and 550 NC events each. Since this will require very good background reduction, the optimized cuts were used along with $25 < long < 125$. The results are presented in Table 8.3

| <i>Data Set 1</i> | | | | <i>Data Set 3</i> | | | |
|----------------------------|------------|-------------|----------------|-------------------|------------|-------------|----------------|
| | CCe | Bkgd | ←classified as | | CCe | Bkgd | ←classified as |
| CCe | 70 | 47 | | CCe | 64 | 53 | |
| CCmu | 84 | 916 | | CCmu | 96 | 904 | |
| NC | 58 | 492 | | NC | 63 | 487 | |
| ↑actual | | | | ↑actual | | | |
| <i>Data Set 2</i> | | | | <i>Data Set 4</i> | | | |
| | CCe | Bkgd | ←classified as | | CCe | Bkgd | ←classified as |
| CCe | 68 | 49 | | CCe | 67 | 50 | |
| CCmu | 98 | 902 | | CCmu | 85 | 915 | |
| NC | 53 | 497 | | NC | 54 | 498 | |
| ↑actual | | | | ↑actual | | | |
| avg CCe efficiency | | | 57.48% | | | | |
| avg CCmu misidentification | | | 9.08% | | | | |
| avg NC misidentification | | | 10.35% | | | | |
| avg bkgd misidentification | | | 9.72% | | | | |

Table 8.3- Results from optimized cuts using 7% CCe data sets: ncluster > 2, notused > 235, pvelec > 0, long < 192

As you can see the efficiency and background misidentification are about the same as before. However, from this study you can see that while these cuts do a good job reducing the background, it is still dominant over the small signal. The signal to noise ratio after these cuts is ~1:2, without taking into account the intrinsic beam background. This method is not good enough to detect the small CCe signal we expect.

9. Decision Trees

Decision Trees are another multivariate method that was explored to see if it could improve on the previous results. Decision trees are predictive algorithms that perform similarly to the analysis described above, but are more sophisticated. The user supplies a training set of data that the program creating the tree uses to optimize a series of cuts (see Fig 9.1). A new test set of data is then subjected to those cuts and (hopefully) well classified by the tree. Fig 9.1 is an example of a decision tree created with data from this study. As you can see, a cut is made at

each node that decides which branch the particular event is to continue traveling along. It then encounters another node, which makes a cut leading it down another branch, and so on. The ends of the branches are called leaves. In the training set, whatever type of event was dominant at each leaf is what future test set events will be defined as if they end up on that leaf.

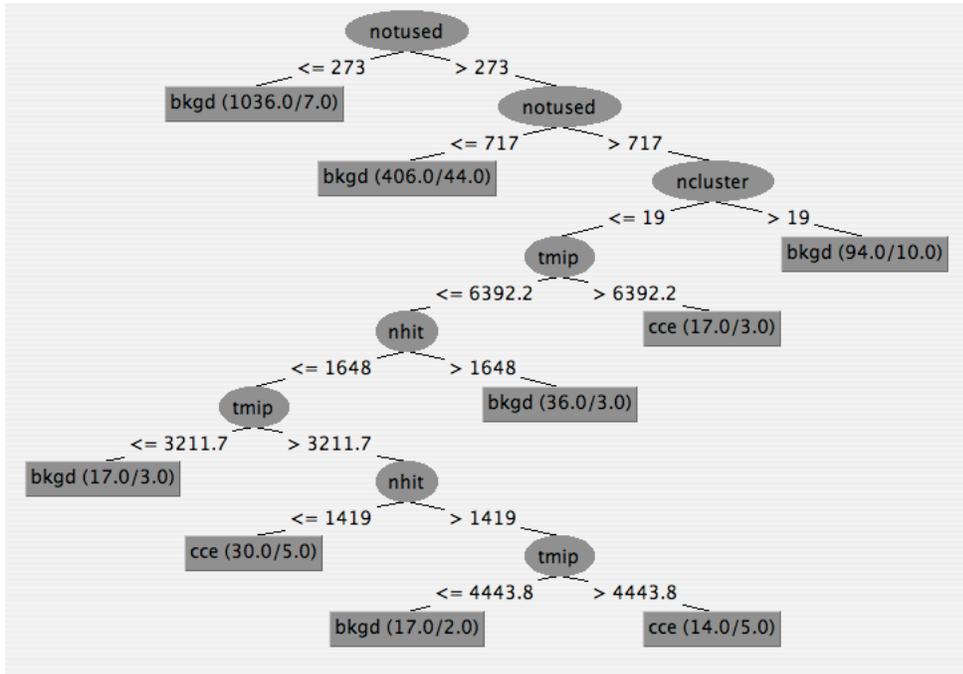


Fig 9.1- Decision tree created by train on a 7% CCE data set

In this study, a Java program developed by The University of Waikato called Weka was used to create and test decision trees [13]. A tree-type called J48 was used, which uses the C4.5 algorithm developed by Ross Quinlan to make the decision tree. C4.5 decision trees work to optimize the amount of information gain at each node [14]. If you have S total events in a set of data and N_k of those belong to a certain class k (are labeled as signal or background in our case), information, which is given in bits, is defined as

$$Info([N_1, N_2, \dots, N_k]) = - \sum_k \frac{N_k}{S} \log_2 \frac{N_k}{S} \quad (10)$$

For example, Fig 9.1 is the decision tree created by training with a 7% CCE events data set (117 CCE and 1550 bkgd). The amount of information before any cut is made is

$$Info([117,1550]) = -\left(\frac{117}{1667}\log_2\frac{117}{1667} + \frac{1550}{1667}\log_2\frac{1550}{1667}\right) = 0.367bits$$

As you can see in Fig 9.1, the first cut on notused separates the data into two sets: one containing 1036 bkgd and 7 signal, and another with 514 bkgd and 110 signal. The information from this cut is

$$Info([7,1036]) + Info([514,110]) = -\left(\frac{7}{1043}\log_2\frac{7}{1043} + \frac{1036}{1043}\log_2\frac{1036}{1043} + \frac{514}{624}\log_2\frac{514}{624} + \frac{110}{624}\log_2\frac{110}{624}\right) = 0.730bits$$

Therefore, by making this cut, the amount of information gained was 0.363 bits. The C4.5 tree chooses which variable to cut on and what cut to make at each node by maximizing the amount of information gained at each node.

First, the ability of these decision trees was explored using the data sets that contained roughly the same number of CCE, CCmu, and NC events. The CCmu and NC events were merged together and labeled as bkgd since distinguishing between the two is unnecessary. Also, the best success came from only using the correlated variables (nhit, notused, tmip, and ncluster) to build the tree. 66% of each set was used to train the tree, which was then tested on the remaining 34%. Table 9.1 shows the results from this study

| | | | | | | | |
|----------------------------|------------|-------------|----------------|-------------------|------------|-------------|----------------|
| <i>Data Set 1</i> | CCe | Bkgd | ←classified as | <i>Data Set 3</i> | CCe | Bkgd | ←classified as |
| CCe | 213 | 92 | | CCe | 122 | 20 | |
| Bkgd | 99 | 534 | | Bkgd | 84 | 246 | |
| ↑acutal | | | | ↑acutal | | | |
| <i>Data Set 2</i> | CCe | Bkgd | ←classified as | <i>Data Set 4</i> | CCe | Bkgd | ←classified as |
| CCe | 212 | 77 | | CCe | 69 | 20 | |
| Bkgd | 80 | 572 | | Bkgd | 30 | 162 | |
| ↑acutal | | | | ↑acutal | | | |
| avg CCE efficiency | | 76.66% | | | | | |
| avg bkgd misidentification | | 17.25% | | | | | |

Table 9.1- Decision tree results from data sets with equal numbers of CCE, CCmu, and NC events

The signal retention is high, but a lot of the background still gets through, which could mask the small expected signal.

As done before, more realistic data sets were used that contain 7% CCE events, 33% NC and 60% CCmu. The best results came when the first of these sets was used to train the tree, which is depicted above in Fig 9.1. The results from using that tree on the remaining three data sets is shown in Table 9.2.

| <i>Data Set 2</i> | | | | <i>Data Set 4</i> | | | |
|------------------------|------------|-------------|----------------|-------------------|------------|-------------|----------------|
| | CCe | Bkgd | ←classified as | | CCe | Bkgd | ←classified as |
| CCe | 32 | 85 | | CCe | 30 | 87 | |
| Bkgd | 16 | 1534 | | Bkgd | 19 | 1531 | |
| ↑acutal | | | | ↑acutal | | | |
| <i>Data Set 3</i> | | | | | | | |
| | CCe | Bkgd | ←classified as | | | | |
| CCe | 33 | 84 | | | | | |
| Bkgd | 21 | 1529 | | | | | |
| ↑acutal | | | | | | | |
| avg CCe efficiency | | | 27.07% | | | | |
| avg bkgd contamination | | | 1.20% | | | | |

Table 9.2- Decision tree results from train on 1 7% CCe data set and training on the other 3. Decision tree used is depicted in Fig 9.1

Though the purity of the signal is low, the background rejection is very high, which produces a signal to noise ratio of $\sim 2:1$. About $2/3$ of the events classified as signal by the tree are actually signal. This method does a much better job at pulling the small CCe signal out while reducing the background.

10. Discussion

A study of $\nu_\mu \rightarrow \nu_e$ oscillations is needed in order to get a better measurement of the mixing angle θ_{13} and to determine the sign of Δm^2_{32} from matter effects. These parameters are necessary for determining if neutrino oscillations can be used to study CP violation and to determine the mass hierarchy of the neutrinos. So far the detectors looking at these oscillations have detected the neutrino events using a liquid scintillator. A promising new generation of detectors currently being researched and developed are Liquid Argon Time Projection Chambers. The high density of LAr allows for smaller and cheaper detectors with similar event rates as the

liquid scintillator detectors. Also, LAr TPCs will be able to reconstruct the entire 3D event within the volume. Currently a $\frac{1}{4}$ ton prototype named ArgoNeuT has been built and will be exposed to the NuMI beam at Fermilab to do a near detector, low energy study. Plans are underway to build a 5 kton detector ~ 1000 km from Fermilab to do a long baseline study in the next few years.

The current limits of θ_{13} are quite low compared to the other two mixing angles. This means that $\nu_{\mu} \rightarrow \nu_e$ oscillations will have a low probability of occurring. At best, $P(\nu_{\mu} \rightarrow \nu_e) = 7\%$. This means that a very efficient event identification algorithm will be necessary to pull the small CCE signal over the CCmu, NC, and intrinsic NuMI beam background. First, 6 variables were found that distinguished the CCmu and NC background from the CCE signal. The efficiency of making simple cuts in PAW with ncluster, notused, pvelec, and long were studied. Using the optimized cuts for ncluster, notused, and pvelec, and $25 < \text{long} < 145$, about 60% of the CCE signal was kept while 91% of the background was reduced. However, even with this high background reduction, the signal to noise ratio using the 7% CCE data sets was only about 1:2. Therefore these simple cuts are not very good at pulling the small CCE signal above the background.

The efficiency of decision trees was also explored. When a 7% CCE signal data set was used to train a decision tree, the remaining 7% CCE data sets were well classified by that tree. Though there was a low signal purity, the background reduction was enough to have the majority of the events classified as signal actually be signal. A signal to noise ratio of about 2:1 was found using this tree, making it an efficient classification algorithm that could be used to pull the small CCE signal above the CCmu and NC background. Even if 20% of the correctly identified

CCe events were removed as intrinsic beam background, the true CCe signal would still rise above the misclassified background with a signal/noise of 1.38.

Though the decision tree used above had much success with the 7% CCe data, that 7% oscillation probability was calculated using an optimistically high value of $\sin^2(2\theta_{13}) = 0.14$. Further tests need to be made assuming even lower oscillation probabilities to see if the decision tree trained on the 7% CCe data continues to do such a good job. Since the CCe signal will be much lower, the signal-to-noise ratio will most likely drop, however hopefully the same number of background events will be misidentified. If it can be well determined what percentage of the total number of analyzed events are going to be misidentified background events, then those can be subtracted from the classified signal events and still (hopefully) leave a detectable signal.

Also, the tree found in this study can of course be improved upon. The distinguishing variables used in this study were mainly developed to look for the electron shower. Other variables that look more closely at what is happening near the primary vertex can be imagined and could possibly lead to better classification.

In conclusion, the multivariate analysis performed in this study with C4.5 decision trees resulted in a signal to noise ratio of 2:1 using data sets that were originally 7% CCe signal. Therefore this could be a scheme used to identify $\nu_\mu \rightarrow \nu_e$ in a multi-kiloton LAr TPC long baseline study. Future analysis is necessary to determine if this multivariate scheme will be as successful if the $\nu_\mu \rightarrow \nu_e$ oscillation probability is lower than 7%. It is possible new distinguishing variables may need to be created and a better scheme implemented. However, this study has shown that using `ncluster`, `notused`, `tmip`, and `nhit` in a C4.5 decision is a good starting point for CCe signal identification in a $\nu_\mu \rightarrow \nu_e$ long baseline oscillation experiment.

References:

- [1] C. Cowan et al. *Science*. 124, (1956) 103-104
- [2] G. Danby et al. *Phys. Rev. Lett.* **9** (1962) 36-44
- [3] DONUT Collaboration, K. Kodama *et al.* *Phys. Lett.* **B504** (2001) 218-224
- [4] B.T. Cleveland *et al.* *ApJ* **496** (1998) 505-526
- [5] SuperKamiokande Collaboration, Y. Fukuda *et al.* *Phys. Rev. Lett.* **81** (1998) 1562-1567
- [6] SNO Collaboration, Q. R. Admad *et al.* *Phys. Rev. Lett.* **87** (2001) 071301
- [7] The Particle Data Group, C. Amsler *et al.* *Phys. Lett.* **B667** (2008)
- [8] KamLAND Collaboration, E. Eguchi *et al.* *Phys. Rev. Lett.* **90** (2003) 021802
- [9] K2K Collaboration, M. H. Ahn *et al.* *Phys. Rev. Lett.* **90** (2003) 041801
- [10] ArgoNeuT Collaboration, B. Fleming *et al.* <<http://t962.fnal.gov/index.html>>
- [11] NovA Collaboration, D. S. Ayres *et al.* “Proposal to Build a 30 Kiloton Off-Axis Detector to Study $\nu_{\mu} \rightarrow \nu_e$ Oscillations in the NuMI Beamline” (2005)
<<http://www-nova.fnal.gov/>>
- [12] B. Baller. “Liquid Argon TPC Simulation & Reconstruction Code”
<http://home.fnal.gov/~baller/lartpc/LAr_TPC_Simulation.htm>
- [13] Weka Collaboration, E. Frank *et al.* <<http://www.cs.waikato.ac.nz/ml/weka/>>
- [14] E. Frank and I. H. Witten. *Data Mining*. Morgan Kaufmann Publishers, 2000