Nuclear Effects in Neutrino Interactions at Low Momentum Transfer

A THESIS
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF THE UNIVERSITY OF MINNESOTA
BY

Ethan Miltenberger

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
MASTER OF SCIENCE

Advisor: Dr. Richard Gran

May, 2015

Acknowledgements

Abstract

This is a study to identify predicted effects of the carbon nucleus environment on neutrino - nucleus interactions with low momentum transfer. A large sample of neutrino interaction data collected by the MINERvA experiment is analyzed to show the distribution of charged hadron energy in a region with low momentum transfer. These distributions reveal a major discrepancy between the data and a popular interaction model with only the simplest Fermi gas nuclear effects. Detailed analysis of systematic uncertainties due to energy scale and resolution can account for only a little of the discrepancy. Two additional nuclear model effects, a suppression/screening effect (RPA), and the addition of a meson exchange current process (MEC), are shown to improve the description of the data.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

In the world of neutrino physics there are experiments whose goals require very high accuracy. In order to produce a high accuracy measurement, different systematics will need to be tested and a complete model of the interaction of the neutrino and target nucleus needs to be found. This analysis will provide support for two models which are not yet incorporated into the model used by GENIE but will improve the accuracy.

In MINERvA the large neutrino interaction simulator is called GENIE, which splices together many interaction models in order to produce simulated events that cover the MINERvA energy range. There is a simplification between the model used in GENIE and another model which will be referred to as the Nieves model. The Nieves model allows two different types of reactions that are not accounted for in GENIE.

This analysis is focused on a particular kinematic region of neutrino interactions specific to Nieves' model. To describe this region. the energy and three momentum transfers are going to be used. For this analysis there is a neutrino interacting and producing a negative muon, and transferring energy, momentum, and charge to the nucleon interaction target inside a carbon nucleus. The energy and momentum transfers to the target nucleon or hadron system are defined by

$$q_0 = E_\nu - E_\mu$$
$$q_3 = p_\nu - p_\mu$$
(1.1)

In the equation the energy transfer (q_0) is equal to the energy of the neutrino (E_ν) minus the energy of the muon (E_μ). The second line is the same, but with the magnitude of the three-momentum. These quantities can be thought of as the energy and momentum of a virtual particle being exchanged between the neutrino and a nucleus in the nucleus.
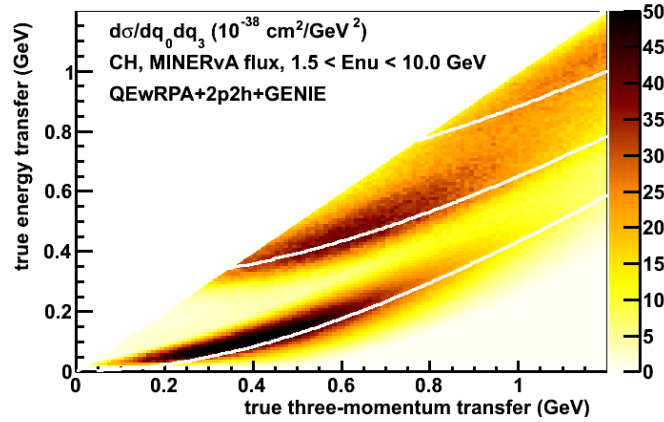
Figure 1.1: Double differential cross section of the events in the sample which is going to be studied. The the white bands are are lines of constant W at the specific values of 938, 1232 and 1520 MeV going from the lowest to the highest band.

The 2D distribution shown in Figure 1.1 is the sample that is going to be examined throughout this paper. The figure shows the double differential cross section for a neutrino interacting with a carbon nucleus. The cross section has units such that when multiplied by incident neutrinos per area per second and the number of target carbon nuclei, then you have calculated interactions per second. This region is limited to the three momentum transfer of less than 1.2 GeV. The analysis will restrict the region even further, even though the cross section itself extends to much higher momentum transfers than this. Energy transfer cannot be higher than the three momentum transfer which is the reason why there are no events above the diagonal.

There is structure within this kinematic region is highlighted by three different white lines cutting through the Figure 1.1. These white lines are lines of constant W, the invariant mass for the hadron system of the interaction. The invariant mass can be

thought of as the rest mass of the hadron system after the interaction (including later breakup or decay). For a single particle, its rest mass is a "relativistic invariant" or a "Lorentz invariant" and has the same magnitude in all reference frames. The relativistic relation between the energy (E), magnitude of the three momentum (p) and mass (M) is

$$W^2 = M^2 = P^2 = E^2 - p^2 \tag{1.2}$$

P^2 is the magnitude of the four momentum. For a multi particle system, with n particles the invariant mass can be calculated using the sum of the four-momentum vectors.

$$W^2 = (\sum_i^n P_i)^2 \tag{1.3}$$

If a single particle decays, W^2 calculated before the decay or after it will give the same result. In Figure 1, the white bands correspond to different constant values of W. The lines in the plot are at the values of 938, 1232, and 1520 MeV from bottom to top. The first line is just the mass of a proton, so it is an elastically scattered event. The second line corresponds to producing a nucleon resonance called the delta.

1.1 Quasi Elastic Interaction

The main reaction for this study is the Quasi-Elastic interaction (QE) which is the lowest W line in Figure 1.1. This type of interaction is named this because it can be modeled using elastic collision techniques, which is why this type of interaction is often referred to as the billiard ball interaction. Looking back at Figure 1.1, the QE populates the events along the bottom white line which has a value of 938 MeV/c^2. The "quasi" refers to the fact that the kinetic energy before and after are not exactly the same, as in a perfectly elastic collision. Some energy is lost to removal of a nucleon from the nucleus and also to become the mass of the muon.

What happens is a neutrino collides with a neutron in the nucleus of the atom and produces a proton and a muon. This reaction can be written in a reaction style equation

$$\nu + n \rightarrow p^{+} + \mu^{-} \qquad (1.4)$$

and illustrated with a Feynman Diagram, shown below in Figure 1.2.
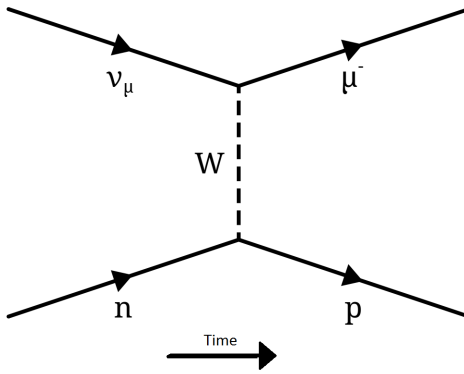
Figure 1.2: Feynman diagram for the QE style interaction with forward time being the horizontal axis going to the right.

To explain the diagram in terms of the experimental setup starts with the incoming muon neutrino. The muon neutrino is coming down the beam line and it collides with a neutron in the nucleus of one of the carbon atoms in a target made of scintillating plastic material. At that point the two particles exchange a virtual boson (carrying the energy q_0 and momentum q_3 to be transferred and one unit of charge) called the W for "weak". The result is that the muon neutrino becomes a muon coming out of the nucleus. The neutron is transformed into a proton. This reaction conserves energy, momentum, charge, lepton number and the baryon number.

Looking at this reaction as a pure elastic collision, a lot of the information about it is measured using the detector; two which are not measured are the incoming neutrino energy and momentum. The energy of protons coming out of the nucleus is measured along with the muon's angle and energy. Once all the other measurements are known, through conservation of energy and momentum the neutrino's initial energy or momentum can be found. Then through the relativistic expression for energy, momentum and mass the other can be calculated.

There is a complication, the nucleons in the nucleus are not stationary, they are moving at random, slightly relativistic velocities. This comes from a fundamental argument about the wave nature of quantum systems. Heisenberg's uncertainty relation for the position and momentum says that since we know the position is localized to the nucleus, there must be some range of momenta. The Heisenberg uncertainty relation for position and momentum is

$$\Delta p_r \Delta r \geq \frac{\hbar}{2} \tag{1.5}$$

Which means that since the nucleon is confined to the carbon atom, the range of momenta can be calculated.

$$\Delta p_r c \geq \frac{\hbar}{2\Delta r} = \frac{6.582119x10^{-22} MeV \cdot s}{2 \cdot (2 \cdot 2.7x10^{-15}s)} c = 18.28 MeV \tag{1.6}$$

The radius of a carbon nucleus is $2.7x10^{-15}$m and the particle is confined to the whole nucleus which has a diameter of twice the radius. This is where the additional factor of two comes from. Taking the classical approximation that momentum is equal to mass times velocity results in

$$v = \frac{pc}{mc} = \frac{18.28 MeV}{938.27 \frac{MeV}{c}} = 0.0195c = 5.84x10^6 \frac{m}{s} \tag{1.7}$$

The mass of the nucleon was taken to be the mass of a proton. This means that the nucleons are moving almost two percent of the speed of light in the nucleus.

The movement of the nucleons in the nucleus is tackled two different ways between GENIE and Nieves. Both of the paths follow the procedure of pulling a random magnitude and direction from a distinct distribution for the neutron's momentum and then generate for each individual event the billiard ball collision. The difference is in the actual distributions that they are pulling from.

GENIE models the nucleus as a global Fermi gas. The global Fermi gas says that in a 3-D momentum space there is a sphere with a radius of 250 MeV with a uniform momentum distribution inside. After integrating over the two non-radial dimensions to a plot of the magnitude of the momentum there are fewer events with zero momentum and more with higher momentum, which produces the momentum distribution with a quadratic shape plot with the cut off at 250 MeV. The global Fermi gas model produces the simplest distribution for the momenta.

The Nieves model treats the nucleus as a local Fermi gas. This starts the same way as the global Fermi gas but different nucleons see a different local nuclear density depending on how far from the center they are. Higher density environments have a higher Fermi limit. So Nieves pulls from two distributions and the convolution is the final momentum distribution. The two distributions used for the different models are shown below in Figure 1.3.
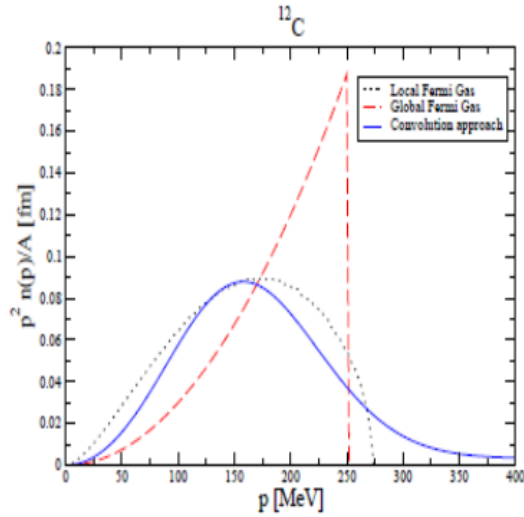
Figure 1.3: Momentum distribution for the global and local Fermi gas models[1].

Both of these methods for tackling this problem produce an effect which causes the QE events to not lie on a perfect thin line as W=938 MeV. This is an important feature seen in Figure 1.1. These variations themselves are only slightly different and are not what we are testing. Both models approximately correct for the movement, but not having this effect in the model at all would produce a discrepancy with the data, but both models have it. The additional screening effect described later significantly overlaps with this effect.

1.2   Delta Production

The middle line shown in Figure 1.1, which is equally important for the region of interest as the QE, is the resonance delta production often referred to as "the delta".

This is when the neutrino interacts with either a neutron or proton in the nucleus and produces a muon and either a Δ^+ or a Δ^{++}. The Δ^+ is simply the lowest excited state of the three quark system we normally call a proton; so it is a UUD excited state. The Δ^{++} is a UUU system with practically the same energy. These are often referred to as resonances because the Delta never quite becomes a separate particle, just a temporary excited state excited state.

For the delta production interaction, the decay time of a delta particle is $0.6x10^{-23}$s. That is less than the travel time across the nucleus at the speed of light. What is actually seen in the detector is the decay products of the delta, however the invariant mass is constant through the reaction. When the invariant mass of the system is estimated, its value should be the rest mass of the delta, which is 1232 MeV/c^2.

Like the QE, the Δ is kinematically broadened away from the W=1232 MeV line in Figure 1.1. In addition to the Fermi motion it picks up an intrinsic width due to its very short lifetime. The broadening is related to the short lifetime can be directly related to the other version of Heisenberg's "Uncertainty principle". The energy-time version of the uncertainty principle is

$$\Delta E \Delta t \geq \frac{\hbar}{2} \tag{1.8}$$

The Δ decays quite quickly, and with a very small range of lifetimes, which means there is a large range of the decay energies. This causes the the events to be spread out around the invariant mass line. Some of this broadening would happen if the reaction happened on a free neutron or proton with no Fermi motion.

One potential decay path for a Delta is shown below in Figure 1.4.
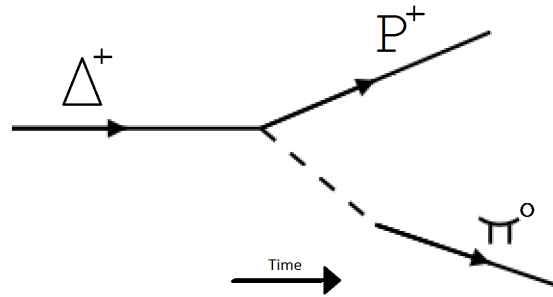
Figure 1.4: Feynman diagram for the decay of a delta with a positive charge into a proton and a neutral pion with the forward base time being in the horizontal direction to the right.

In a reaction style equation, the entire reaction with the neutrino and the nucleus would be

$$\nu_\mu + n \rightarrow \Delta^+ + \mu^- \rightarrow p^+ + \pi^0 + \mu^- \tag{1.9}$$
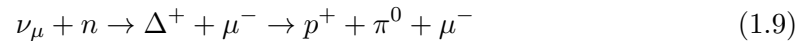
Once a delta decays there are different decay products that come out. Figure 1.4 is an example of a positive charged delta being decayed into a proton and a neutrally charged pion.

1.3   Meson Exchange Current

Nieves model provides another type of interaction called the Meson Exchange Current or the 2p2h interaction[2]. The previously described QE and Delta interactions would be classified as a 1p1h interaction, or when one particle is knocked out of the nucleus leaving one empty unoccupied energy state. With the MEC the neutrino collides with two coupled particles in the nucleus and knocks them out, leaving two holes. The name "Meson Exchange Current" refers to how the initial state nucleons are correlated because they are exchanging a pion. The resulting Feynman diagram for this type of event is shown below in Figure 1.5.
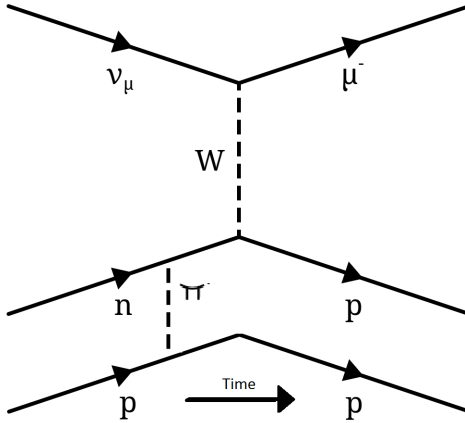
Figure 1.5: Feynman diagram for the MEC interaction with a proton and neutron with the forward base time being in the horizontal direction to the right.

This type is not modeled in GENIE, but is in the Nieves model. This effect will be tested in order to see if it improves the agreement between the data and the simulation. A large component is predicted to lie near a line of W=1100MeV. Another produces a Δ between the π and W and lies near the 1232MeV Δ process, but without a π in the final state.

1.4   Random Phase Approximation

The other big effect that the Nieves model includes that is not included in GENIE is the Random Phase Approximation or the "RPA effect" for short[2][3]. This is a many-body effect due to the 12 nucleons in the carbon nucleus. What this does is it modifies the low-lying energy states of the nucleus by changing the potential. Simultaneously, this produces a screening effect which reduces the probability for an electromagnetic and weak interaction with those states. Nieves calculated this only for the Quasi-Elastic interaction at this time. The result is fewer QE interactions since there is less of a probability to interact via the weak force.

In order to see how these two affect the sample being looked at, a ratio plot can be taken of Nieves's model to a model without those enhancements. The caveat is that there are only two effects in interest, the RPA and MEC. In order to look at just those

two and not all the other differences in the models, a patchwork combination was used. For the cross section using the MEC and RPA what was done is the MEC was added to Nieves's QE with RPA and then added to GENIE's non QE. This was taken as the numerator on a ratio with the Nieves non RPA QE and GENIE's non QE as the denominator. Each cross section, numerator and denominator looks roughly like Figure 1.1, and the ratio is defined as

$$\text{ratio} = \frac{\text{MEC} + \text{NievesQEw/RPA} + \text{GENIEnonQE}}{\text{NievesQEwithoutRPA} + \text{GENIEnonQE}} \tag{1.10}$$

Even with the resolution effects described in a later chapter, predicted discrepancies of 20% would likely be apparent in a comparison of data to a simplistic model. Plotting this ratio of models, Figure 1.6, these discrepancies would be visible.
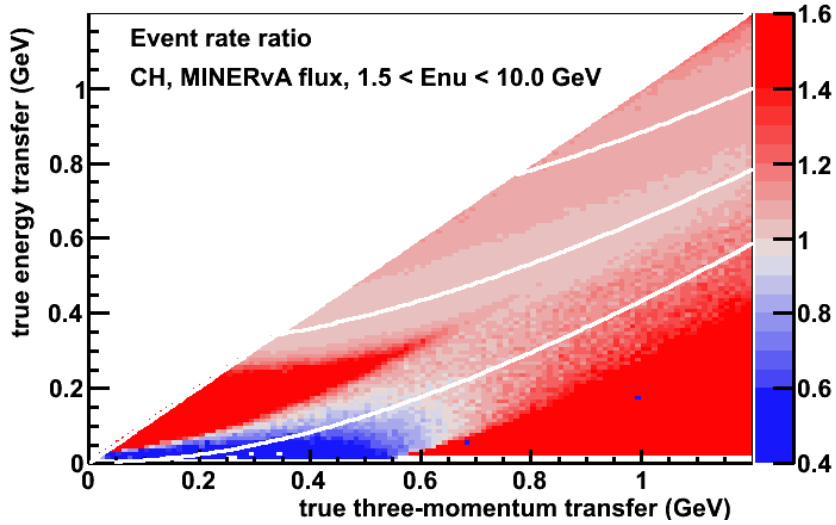
Figure 1.6: Ratio of a model with the MEC and RPA to the model without with constant lines of W.

One big effect occurs at low energies along the QE line. There is along the lines of a 60 percent suppression in the QE line at low energy transfers, this is the blue region in Figure 1.6. This specific area is the direct cause of implementing the RPA effect.

The other two big differences is the "dip region" in between the QE and delta line at low momentum transfer and under the QE line at higher energies. The default QE

and Delta predicts almost no cross section in these areas; with the MEC even a small addition of events in these areas has a large fractional effect.

1.5 Analysis Strategy

The triangle plot in Figures 1.1 and 1.6 are very useful to see the bigger picture. With the predictions made to this point the main focus will be on just the lowest three-momentum transfer region, below 0.4 GeV. Focusing on this region reduces the impact from the other types of interactions that the neutrino can have like the coherent pion production and Deep Inelastic Scattering, which happen at higher energies. In the low three-momentum region, both the RPA and MEC are predicted to have large impacts.

There are several different options for isolating this region, for example a cut in q_3 or Q^2 could do the job. The square of the four momentum transfer, Q^2, is another common value to frame an analysis or a theory. The Q^2 is a relativistic invariant that is defined to be the difference between the space-like and time-like components of a four vector.

$$Q^2 = q_3^2 - q_0^2 \tag{1.11}$$

A line of constant Q^2 traces out lines in the energy-momentum transfer space, previously seen in Figure 1.1, that look like the white lines shown in Figure 1.7. Cut regions in q_3 are separated by vertical lines.
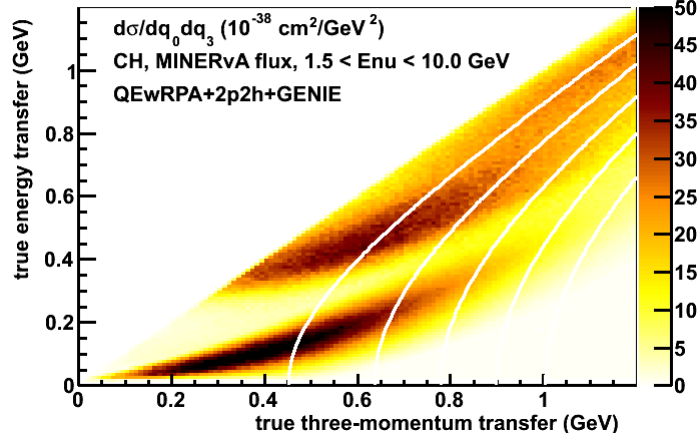
Figure 1.7: Triangle plots of the data with lines of constant Q^2

A nice effect seen here and not in the figure with constant W lines is that the point at which the lines cross the horizontal axis is the value of Q. If q_0 is zero then the four momentum transfer is the three momentum transfer. By simply taking the square of the horizontal value would give the constant Q^2 value of that line.

Through the shape of the cuts in Q^2 several potential problems arise. The first is that there is a huge tail in the sample, which follows the diagonal. This detracts from our interest in isolating the tip of the triangle. A cut that has low energy and three momentum transfer is grouped into a region which has very high energy and three momentum transfer. This long tail also causes a single cut to contain most if not all of the different type of interactions.

A potential solution to both of these problems is to simply not cut on Q^2 but to cut on something else, like the three momentum transfer. There is a part of my analysis which will consider Q^2 cuts, to better understand the advantages of each. By taking cuts on the three momentum transfer, the long tail is eliminated, because the cut lines are vertical. A cut in this value to isolate the low region would pick up the delta and QE region but, none of the higher energy interactions.

For those reasons the cut parameter that is going to be utilized for the latter work will be a cut in the three momentum transfer. The exact cuts that are going to be looked

at are the low, mid, high and very high regions. The values of the cuts that separate these regions are 0.4, 0.8 and 1.2 GeV. Two examples of histograms of the events in the low q_3 region(0 to 0.4 GeV) are shown below in Figure 1.8.
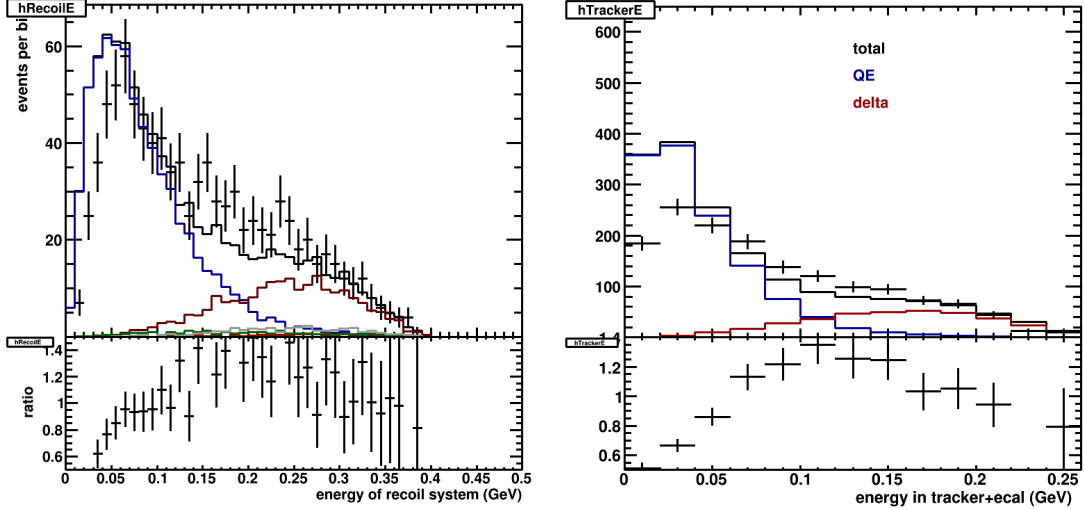
Figure 1.8: Recoil energy on the left and available tracker plus ecal energy on the right for the low q_3 sample.

These figures are now distributions from the reconstructed data and a full simulation (including detector effects) of the neutrino interaction model.

In looking at Figure 1.8, there are several different lines present in each figure. The first of which is the black crosses, which are points with statistical error bars, and represent a histogram of actual data events from the experiment. The next is the solid black line, this is all of the events in the simulation. For the simulated prediction, we know what interaction type was produced, so the rest of the colors are fractions of the simulation corresponding to different interaction types. Blue in this case is the Quasi elastic interaction and maroon is the Delta production portion of the simulation.

Looking back at the ratio plot in Figure 1.6, including the RPA and the MEC have the effect of bringing down the Quasi-elastic portion of the simulation and adding

events into the region between the QE and Delta regions. By looking at Figure 1.8 it seems that the simulation is over predicting events in the low portion of the QE and is under predicting the region in between the QE and the Delta. This is the exact thing that the addition of the RPA and MEC would correct.

Taking everything into account, it has still not been shown that the difference between the simulation and the data is due simply because of the errors in the reconstruction. When looking at the data and the MC (Monte Carlo Simulation) there are going to be errors brought about by many different effects. A large part of my thesis is devoted to examining these effects to see how they reduce our sensitivity to a need for the RPA and MEC effects.

Chapter 2

Design of the Detector and Simulation

This section describes how neutrinos are produced and how they are detected. It also highlights some of the quantities that are used to describe an event are measured. It will also go over how we get information from the simulation.

The experiment is broken down into two different areas, the data and the simulation. The data is the neutrino interactions that the detector is picking up and then being analyzed. The simulated data is an interaction model combined with a detector simulation to incorporate all of the smearing effects that are produced from the detector.

2.1 Data

When talking about the data there are three different areas that need to be discussed. The first is where exactly the neutrino is coming from. The next step is the detector and how it can measure the particles. The final step is getting the high level event quantities from the reconstruction software.

2.1.1 Producing Neutrinos

Figure 2.1 shows a diagram of what the Fermi lab beam paths are.
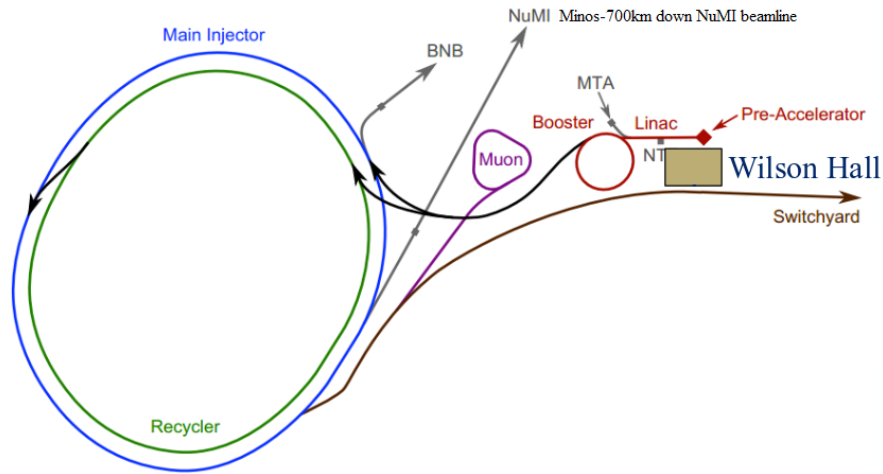
Figure 2.1: Fermi National lab beam diagram[4].

The start of the beam cycle is in the Pre-Accelerator which accelerates protons to have an energy of around 750 keV. The protons then go to the Linac and the booster which boosts the energy to 400MeV and then to 8GeV respectively. Finally they are dumped into the Main Injector which brings the energy up to 120 GeV.

The Main injector is a series of electromagnets and RF cavities that are used to accelerate the protons around the ring and to store them until a spill happens. A spill is when the main injector diverts roughly 30×10^{12} protons down a beamline. The beamline that MINERvA is on is the NuMI beam line, MINOS is also on this beam line with its detectors just after the MINERvA detector and 700 km away. For the NuMI beamline the spills are 10 μs long and occur every 1.667 seconds.

The protons go through a process which converts them into the muon neutrinos which are needed by MINERvA and MINOS. This production process is shown in Figure 2.2.

to Minerva

Muon Monitors

Target
Horns
Decay Pipe
Absorber
π^+
μ^+
ν_μ
10 m
30 m
675 m
Hadron Monitor
5 m
Rock
12 m
18 m

figure courtesy Ž. Pavlović
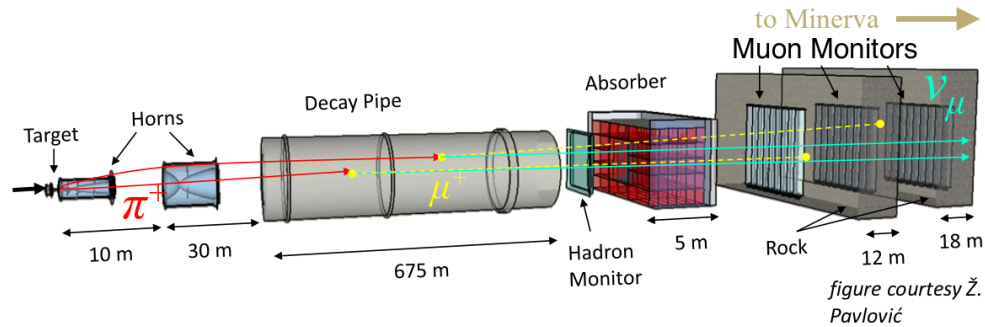
Figure 2.2: Path taken to obtain a beam of neutrinos. Image was taken from a talk given, used in many different presentations.

Looking at the above diagram, the first step is that the protons collide with a stationary graphite target in order to produce π^+. These particles are then sent through two horns in to focus them back to a beam. As particles are leaving the target they are diverging from the center of the beam line. A horn is essentially a large electromagnet whose field bends the π^+ back towards the beam axis which focuses them back into a beam heading towards the detector.

From the horns the π^+ are sent though a decay pipe which allows the particles to decay producing positive charged muons and muon neutrinos. There is a chance for a few other things to be produced, the rest of the diagram is a way to monitor those other particles and to absorb some of them[5]. The muon neutrinos are then on a collision course with the MINERvA detector.

Even though there are 30E12 protons being converted into 30E12 neutrinos in each spill, only about 6 interact in or near our detector. Neutrinos interact so rarely which makes them hard to study, but they can travel 700km through the earth to the MINOS and NOvA detectors in northern Minnesota enabling oscillation experiments possible due to the long baseline.

2.1.2 Detector

The MINERvA detector uses polystyrene scintillator strips[6] in order to track the products of a neutrino interaction. Scintillator material has the unique property that

when a charged particle moves through it, light is given off. This works by the charged particle depositing some of its energy in the material which is absorbed by a special organic molecule added to the polystyrene, and then when the molecule returns to its ground state, light is given off. This scintillator material is formed into triangular strips and then put together in order to form a plane of scintillator material. An example of a small section of two planes is shown in Figure 2.3.
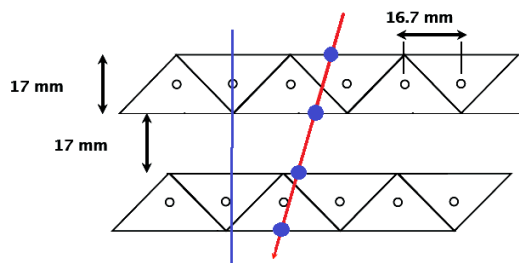
Figure 2.3: Combination of scintillator strips.

In Figure 2.3 there are two of these planes with the strips in the same direction. There is a 17 mm gap between them containing a plane whose strips run a different direction. The round hole in the middle of each strip is the optical fiber that is collecting the light and transmitting it to be read out by a photomultiplier tube, or PMT. A PMT is basically a device that takes light and turns it into an electrical signal. Each of the strips is also coated with a titanium dioxide coating that is highly reflective and opaque so that the light produced in one strip only goes into the single fiber. The amount of light produced is proportinal to the energy deposited in the scintillator. For particles that just track through, this is proportional to the path length it took through the active part of the strip.

The reason behind forming the strips into triangles is that it gives a better position resolution for a track. If for example the path in red is taken by a charged particle then you would know that the particle was between those scintillator strips. By looking at the relative magnitude of activity between the two strips in one plane you would even be able to tell if it went through one strip more than the other. If instead you were to form the strips as squares, the position of the particle would have a precision of the size of the square.

One plane combines 127 strips into the shape of an octagon two meters across. In order to track particles through the detector, planes were oriented as X, U and V planes. The X plane defines the horizontal coordinate system and is the reference plane of which U and V are rotated $\pm$ 60 degrees around. Then an X plane and either a U or V plane are stacked together in order to form a module. The size of each triangular strip is that its base measures 3.34 cm across and is 1.7 cm tall.

The MINERvA detector is not just made up of these planes of scintillator materials though. This in fact just makes up the active tracking region. The main part that is going to be examined throughout this analysis will be the active tracking region and the downstream electromagnetic calorimeter (ECAL). The whole detector is shown in Figure 2.4 below.
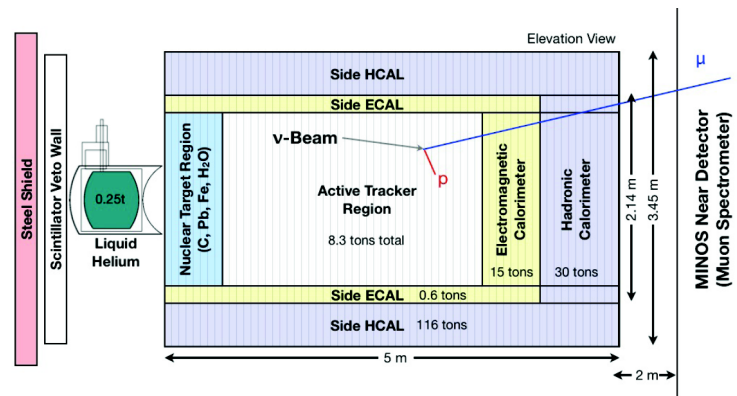
Figure 2.4: Side view of the MINERvA Detector.

In the front of the detector there is a steel shield which is there in order to shield the detector from any low energy particles coming down the beam line. The scintillator veto wall is another method which alerts us that another unrelated particle entered the detector.

The middle part is the active tracking region which is the best place to track the different particles that are being produced by the neutrino interaction. When we add up the energy deposited by hadrons to estimate the energy transfer, this is where most of it is found. The other areas are the ECAL and the HCAL, which are two additional

regions which the energy can be summed in. These would be the energies of particles that have escaped the active tracking region and are leaving the detector out of the back or the side.

The side calorimeters are better seen when looking at a cross section of the detector which is seen in Figure 2.5.
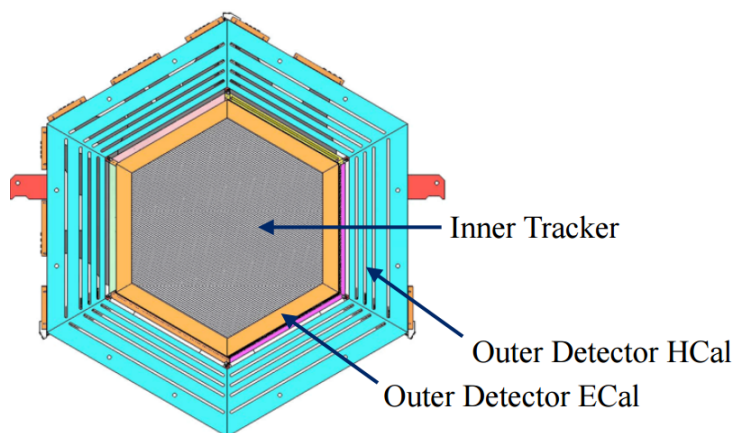
Figure 2.5: Cross section of the MINERvA Detector.

From this image it is clear to see that the active tracking region is contained within the ECAL and the HCAL. Therefor any particles leaving the detector or cosmics entering would pass through the side (Outer Detector) ECAL and HCAL. Likewise, particles passing out the back travel through the downstream ECAL and HCAL.

Once the neutrino interacts with the detector, information about the time energy of activity in each strip is read out and put into the reconstruction and also into a visualizing software, called Arachne[7]. Using Arachne the hits in every plane can be visually seen producing hit maps. Figure 2.6 shows an example of an event in Arachne.

The three windows in Figure 2.6 are the different orientations of the planes. Starting from left to right the order goes X, U and then V. Looking at the image, the one thing that is required for the analysis that we are doing is that a muon is present. A muon is easily seen in the detector since it goes straight through the detector with little to no change in direction. That is the long path in Figure 2.6.
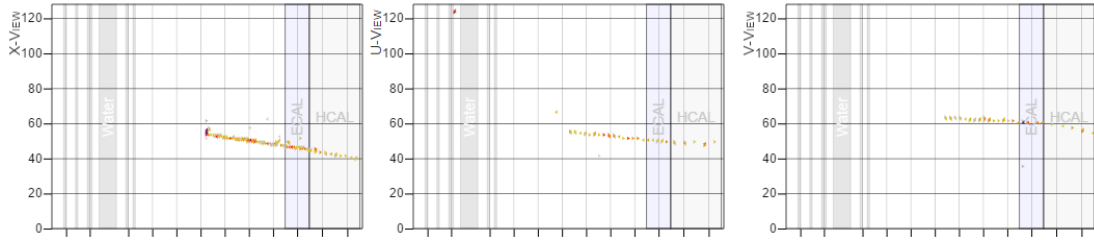
Figure 2.6: Arachne event display.

Another aspect of an is the activity at the vertex. The vertex is where the neutrino interaction took place. For the most part the muon path can be traced back to the beginning in order to identify the vertex. The different colors in the hit maps correspond to different energy values. The bluer the color the more energy deposited in that strip.

Also one thing to note here is that two planes that are adjacent in any of the windows are not actually right next to each other. For any given U or V plane there is at least one of the other V or U planes and two X planes. For example in-between two U planes the order would be XVX. More simply, in between two adjacent X planes there is only either a U plane or a V plane.

2.1.3   Reconstruction

In order to calculate the recoil energy or the total hadron energy transfered to the detector, the first step is that all of the energy on the muon path is removed. Next the remaining hits are multiplied by a passive material correction. After that all of the energy everywhere in the detector is added up. The last step is to apply a polynomial correction, based on the MC, to convert it to an unbiased estimate of the recoil energy.

The passive material correction accounts for the energy that is being deposited in between the planes. The energy can be deposited in the glue and the tape that holds the planes together or the planes of metal in between the scintillator planes in the ECAL and HCAL. This correction is the ratio of the amount of material the particle would see to the total material of the plane. This rescales the energy observed per hit

to be energy deposited per plane. Since this ratio would be different for the tracker, ECAL, and HCAL this has to be done before the energy is summed up.

A different hadron energy estimator is also used in this analysis, the "Tracker plus ECAL energy". It is calculated the same way except the energies are summed up only in the tracker region and the downstream ECAL. Instead of being an estimator for the total recoil or energy transfer, we are using it as an estimator for energy deposited by charged particles, especially protons and pions, which is referred to as "available energy". What this does not count is energy carried away by neutrons.

In order to calculate the muon angle, the first step is to create a best fit line of the data points on the muon path and from that extract the angle between that and the beam. A literal way is to fit X vs. Z points to get a slope, then do the same for the VZ and UZ. From these the three dimensional vector is obtained, which then needs to be further corrected because the Z axis of the detector is not identical to the beam direction. The actual method uses a sophisticated software algorithm called a Kalman Filter.

The muon energy is reconstructed from range or curvature using its path in the MINOS Near Detector. Then as estimate for the muon energy loss in the MINERvA detector from point of interaction to the exit point is added to the MINOS quantity to get the total.

The vertex energy is another quantity that is needed. The vertex energy is calculated be applying the passive material correction, and then add up the energy near the vertex. Specifically the energy is added into the vertex energy quantity if it is with-in 30 cm of the vertex and is not on the muon track.

2.2   Simulation

The simulation has two parts, the first of which is called GENIE[8]. It produces a list of particles and their momentum coming out of a neutrino-carbon interaction. From these values the different kinematic quantities like neutrino energy, q_3 and energy transfer can be calculated, this type of value will be referred to as the truth quantities.

This model is known to have shortcomings which this analysis is testing, but serves as a baseline model we can use to bootstrap to a measurement and a better model.

An additional value that can be calculated then is the True Visible energy. This value is the sum of all of the different energies of the charged hadrons which would be seen in the detector. Mostly it is the kinetic energy of the protons and charged pions. It also includes the total energy of neutral pions, and any electrons or photons that may have come out in the hadron system. The visible energy specifically excludes the kinetic energy of neutrons, which usually escape the detector and are unseen, making it not a reliable observable.

The second part of the simulation is called Geant4. This steps a simulated particle through the detector depositing energy in particular scintillator strips that mimic the real data as closely as possible. Once this is run, we also get hit maps for viewing in Arachne, just like for data. Once this happens these events are sent through the reconstruction process in order to get the measured value, just like the data. The Geant4 model is not only accurate but has known uncertainties.

In order to better represent the data, the true hit maps are laid over background noise seen by the detector. The background noise is produced by taking real data events with their real hit timing distributions.. This leads to the inclusion of different types of noise, for example after pulsing and delayed neutron activity.

These events are fed back into the reconstruction to recalculate the information including all of the potential sources of fluctuations, and even errors in the reconstruction along with the background noise. An example of this is the reconstruction saying that the vertex is further down the muon path than it really was, which would affect the muon energy, vertex energy, and everything else. In this way, the simulation is designed to capture all the sources of measurement fluctuation present in the real data.

To sum all of this up, for the data the different events are sent through the reconstruction in order to get the reconstructed information. The MC produces truth quantities for the initial interaction and all the activity that follows. This is sent back

through the reconstruction to get the reconstructed quantities. Later all of these will be combined in order to produce "truth" like data quantities.

2.3   Technical details on the data and simulation

The data is from the combination of the minerva1 and minerva13c data set, the two largest subsamples of neutrino-mode data. They integrate about two years of beam data and around $2x10^20$ protons on target. These data are compared to a combination of simulations of the minerva1 and minerva13c data periods, with about ten times the statistics as the actual data. The version of GENIE used to model neutrino interactions is 2.6.4. The version of Geant4 used for propagation of particles in the detector is v9.4p2. The set of calibration and reconstruction tools used for the data and simulation is named "Resurrection". We use the official processed data and simulation with no modifications, except for the systematic studies described herein.

Chapter 3

Manual Scanning of Events

If the MEC is included in the model then in the dip region, area between the Δ and the QE, its expected signature is two nucleons coming from the interaction. A large fraction of the time, both nucleons will be protons, which might be easy to identify. This means a selection can be made in order to see if a 2p2h signature is present. Using Arachne, many of these events can be looked at in order to see if the behavior in question can be easily seen in the event display.

The low q_3 sample was scanned over in order to see if the MEC was there. All of the low q_3 sample was chosen so that there were some in the dip region and some not. This means a total of 625 events were scanned for the data, along with 1100 events in the MC. The majority of the scanning was done by myself and John Demgen with additional help provided by Dr. Gran.

3.1   Scan Rules

Before scanning, a set of rules were developed in order to better identify the different properties for an event. The first is to define the different types of proton-like topologies. Our definitions specify four different types of hits: a vertex hit, a stub, a track and nothing.

A vertex hit is where a single strip at the vertex is over 10 MeV. If two strips that share a side and both have 10 MeV of energy, we could that as a single hit. If two

strips however have 10 MeV or more in energy and do not share a side, then that would be considered two hits. The number of vertex hits are incremented between the three different views, because its hard to apply the adjacently test between planes. An example of an event with a single vertex hit is shown below in Figure 3.1, in the X view as the purple triangle.
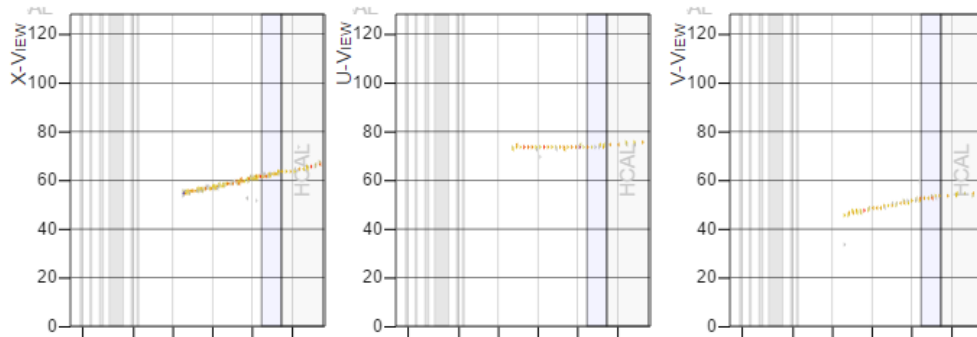
Figure 3.1: Event shown in Arachne with a single vertex hit.

A stub was defined as three or four unit lengths that line up and are coming from the vertex. A unit length is defined as the distance from one strip to the next in the horizontal or vertical direction. If there is a combination of the two the process was to count how many vertical and how many horizontal to determine the total length. For example of the stub went up two strips and over one strip the length would be three units. This counting simplified the scanning by not having the scanner apply a formula to calculate the actual distance. Figure 3.2 is an example of an event with a single stub.
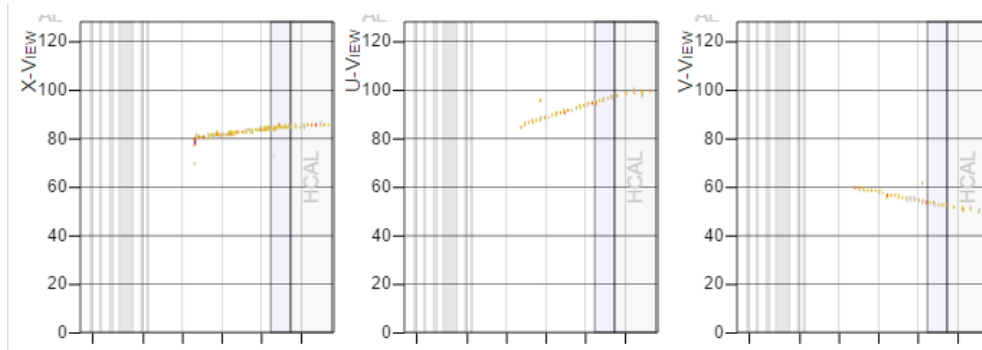
Figure 3.2: Event shown in Arachne with a single stub.

Unlike the vertex hit and the stub, the track seen in Figure 3.3 spans all three views and is quite long.
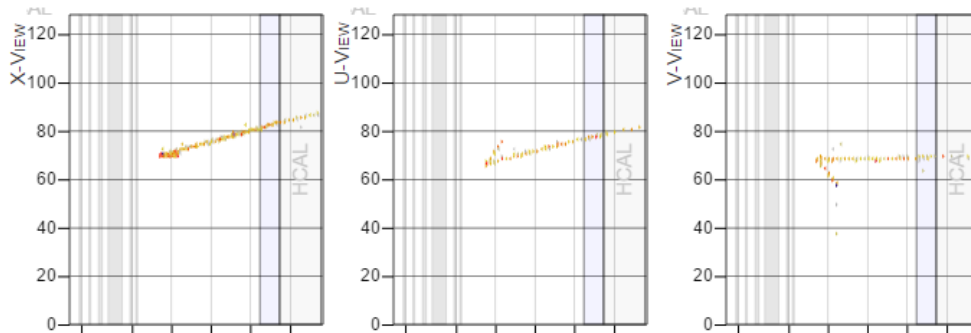
Figure 3.3: Event shown in Arachne with a single track.

Tracks are the easiest to identify, not just within a single view, but also across different views because of the length. The longer the track the more planes that it crosses which gives more hits in each of the views.

The last case, the "nothing" case is when an event does not have any of the other categories. There may be additional activity at the vertex, but it does not meet the criteria for a 10 MeV hit or a stub. The proton(s) have so little energy and/or deposit it in the passive material, or the final hadron state was all neutrons. The former should preferentially be very low energy transfer interactions.

3.2   Data Scanned Results

The events that were scanned for the Data were in the low q_3 cut region for the 2 to 5 GeV energy range in the MINERvA 1 data set. This results is about 625 events scanned for the data. All 625 data events were scanned twice.

Since the events were scanned twice, from the results an idea about how consistent the scanning can be. Out of the 625 events, 148 events where disagreed upon or about 23.7% of the time. Table 3.1 has the results for scanning for each run through.

Category	Run 1 scan results	Run 2 scan results	Resolved results	$\sigma_{Resolved}$
0 hadrons	24.16	22.56	23.52	1.7
1 hadron	46.24	48.8	47.36	2.0
2+ hadrons	29.44	27.68	28.48	1.8
not available	0.16	0.96	0.64	

Table 3.1: Data Scan results, percentage of events in each category.

The number of hadrons in an event was set equal to the sum of the number of tracks, vertex hits and stubs. For the not available case, this is when the event could not be found in Arachne, or something like a particle shower where there is no way to determine what was happening.

Even though 25 percent of the events had a discrepancy between the scanners, the percent for each category were only different by about two percent. What this says is that even though there was a discrepancy it is not one sided. A single scanner did not put more events in one bin over the other but, it was an even mixing between categories.

In both cases, about 47 percent of the events contained one hadron according to the scan. The other 53 percent where split fairly evenly between the zero hadron case and the more than two case, there is a slight lean towards more than two.

After all of the events where scanned, all of the conflicts where examined again by John and I. From this 38 events, or 6.1% where different from both of the original

scans. This means that only was there a disagreement but on further review it was found that both were wrong. The resulting resolved data is shown in the resolved results column in the table with the uncertainty (assuming binomial statistical fluctuations) in the last column.

Even though the categories only vary by about two percent and fall with in the corrected value and uncertainty, having roughly 25 percent of the events having a discrepancy is a cause for concern when designing future scanning efforts.

3.3   MC Scanned Results

For the simulation, there were 1100 events in total that where scanned from the low q_3 and in the energy range between 2 and 5 GeV. The scanned results can be compared to the truth information given by the simulation. This tells us how often one or two protons translate into the proton-like topologies we can see by eye. Table 3.2 below shows the number of hadrons from the MC scan along with the number of true charged hadrons according to the truth information.

Category	MC scan results	MC truth information	Matching percent
0 hadrons	34.82	12.73	35.25
1 hadron	43.55	60.55	75.77
2+ hadrons	21.45	26.55	30.11
not available	0.18	0.18	

Table 3.2: MC scan result of percentage of events in each sub group compared to the truth information on an event by event basis.

The numbers in the tables are percents with the total number of events present in this table being 1100. First comparing the more than two hadron case, the scan results and the truth information are about five percent difference with the truth populating this slightly higher. Those two numbers are fairly consistent with each other.

The bins that are less consistent are the zero and one hadron cases. For the zero hadron case the scan said there where about 22 percent more events here than the truth does. Similarly the truth says there should be 17 percent more one hadron events.

The explanation of why these numbers are so far different is because there are some paths that are just not distinguishable by a scanner. For the zero to one hadron case, these are typically low energy protons, meaning that the particles are near the nucleus and depositing less than 10 MeV of energy per strip. If this happens, according to the scan rules we write down nothing, which in turn puts more events in the zero hadron bin when in actuality the event could have a low energy hadron at the nucleus. Another way to say it, we have an imperfect efficiency for identifying the lowest energy hadrons.

This is not seen too much for the two hadron case because these are typically higher energies. The higher the energy the easier it is for a scanner to pluck out what is happening.

The last column is the percent of events that matched. The matching percent is defined as:

$$Matching\ percent = \frac{\#\ that\ matched}{Scan} \cdot 100\% \tag{3.1}$$

For the zero and two hadron case the matching percent are 35 and 30 percent. A physical description of this is that about a third of the time that the scanner says the event has this many hadrons, it agrees with the truth information. The most accurate was the one hadron grouping where the scanners matched the truth information 76 percent of the time. On average there is a 50 percent chance that the scanner correctly identified the number of hadrons in the final state.

Matching the truth information identically is not vital to get results comparing data to MC. Knowing how often and why we misclassify events is necessary for interpreting that comparison.

3.4   Data and MC Comparison

The main comparison of the scanning results is the comparison between the scanned data and the MC for the entire scanned sample. These results are shown below in Table 3.3.

Category	MC(1100 events)	σ_{MC}	Data(625 events)	σ_{Data}
0 hadrons	34.82	1.4	23.52	1.7
1 hadron	43.55	1.5	47.36	2
2+ hadrons	21.45	1.2	28.48	1.8
not available	0.18		0.64	

Table 3.3: Data and MC scan results, percentage of events in each category.

Starting at the zero hadron case, there are 10 percent more events in this group in the MC than in the Data. For the one hadron case these numbers are in agreement to within just over one sigma on both. Events with two or more hadrons are different by about seven percent, or about two and half sigma. These two numbers would also be in agreement with each other, just not as good as the one hadron case.

The MC does have over 11 percent more events in the zero hadron case than the data does. Since the error added straight together is only 3.1 percent, this is over three sigma away meaning that the number are inconsistent.

What this says is that something is wrong. We went into this scan with the working hypothesis that it could be the RPA or MEC effects. With the inclusion of the RPA, there would be fewer events in the zero and one hadron case. The MEC would be adding events into the two hadron case. Both of these would be raising the percent of two or more hadrons and lowering the percent of the other two. This could potentially increase the agreement between data and MC depending on the size of each of the effects.

If it were the MEC, then we would expect the strongest evidence of this in the dip region, see Figure 1.8. This subset of the scanned events is shown in Table 3.4.

Category	MC(372 events)	σ_{MC}	Data(306 events)	σ_{Data}
0 hadrons	11.29	1.6	5.88	1.3
1 hadron	54.84	2.6	57.19	2.8
2+ hadrons	33.60	2.4	36.60	2.8
not available	0.27		0.33	

Table 3.4: Data and MC scan results for the dip region.

In looking at the results, the one hadron case are very similar between the data and MC. With a difference of just over two percent and uncertainties at almost three, these numbers do agree with each other. The same thing can be said for the two hadron case, but with slightly more in the data.

In the zero hadron case there are almost twice as many events in the MC as there are in the data. These numbers are not consistent. With the inclusion of the RPA, which suppresses events in the low energy range which is populated with zero hadron events, the number in the MC would go down increasing the agreement. Importantly, these results do not suggest a strong population of two-proton topologies in this sample.

3.5   Additional Analysis

John Demgen has been pushing the analysis of the scan data, following up on the initial observation above. This briefly summarizes a few in-progress studies of the scanned events. The focus has been on diving further into the dip region in between the QE and delta. This is the region of the distribution which we are expecting to find the MEC.

He has done two things in the area. The first is the adaptation of an algorthim in order to measure the number of hits in the vertex region and then everywhere. The second was an analysis of proton energies for a one proton final state and a two proton final state.

3.5.1   Algorithm: Scanning for hadrons

In this section an existing MINERvA algorithm to look at energy upstream of the vertex was adopted to count the hits with energy greater than 10 MeV in the sample. This is a simplification of the vertex hit and stub finding procedure we did while scanning. Using an algorithm allows us to study the full sample of statistics, including all of the 10x MC simulation. The algorithm looked for hits above 10 MeV near the vertex and then everywhere in both the data and the MC. The results of the algorithm scanning and the manual scanning are shown below in Table 3.5.

# of hadrons	Algorithm VTX(MC)	Alg. VTX(D)	Alg. Whole(MC)	Alg. Whole(D)	Manual Total(MC)	Manual Total(D)
Totals	7342	749	7726	731	372	306
0	17.1	10.5	9.86	6.4	11.29	5.88
1	40.2	39	37.61	35.3	54.84	57.19
2+	42.7	50.5	52.52	58.3	33.6	36.6

Table 3.5: Manual and algorithm(alg) scan results for both the data(D) and the MC. Numbers for the hadrons are in percents of the totals. Algorithm searched both in only the vertex and everywhere in the tracker and ECAL for the events that populate the dip region.

Comparing first the algorithm search in the vertex for both the MC and the Data there are two differences. There are more two plus hadron events in the data than in the MC. On the other end, for the zero hadron case the MC has seven percent more events than the Data has. A potential explination of this is that even though the MC is trying to account for background noise, it is still not doing a perfect job and the background hits are causing more hadrons in the data. Another explination is a combination of RPA suppression reducing zero hadron events and MEC increasing two-hadron events.

When the algorithm went to search for the hits everywhere, the events could only move to more hadrons. We see this in that the zero and one hadron bins go down and the two hadron bin goes up. The same data vs. MC effects are seen here that are seen just in the vertex. There is more events in the MC with no hadrons and more in the data that have two or more.

Both of these algorithm scans match the data vs. MC trend seen in the manual scan. The absolute percentages naturally do not match, because they are using different definitions to count hadrons.

One potential cause for this is that the algorithm is searching for 10 MeV hits, but this is not well-correlated with the number of hadrons. The assumption is that every hit is potentially a different hadron when in reality one hadron could produce two or more hits with energy above 10 MeV. Also this algorithm could be picking up hadrons that went a ways from the nucleus and then reacted producing hits, in the manual scan this was not accounted for.

3.5.2 1 Protons energy Vs. 2 protons energy

John found that for the same total energy, one proton produced a similar distribution of 10 MeV hits as did two protons. He investigated the statistics of the two proton's energy and how they divided the energy. Then looked at the total energy compared to events with just a single proton. This analysis was done using the truth quantities in the MC. To start, for the two hadron case the fraction of energy in one proton to the total was plotted and is shown in Figure 3.4.
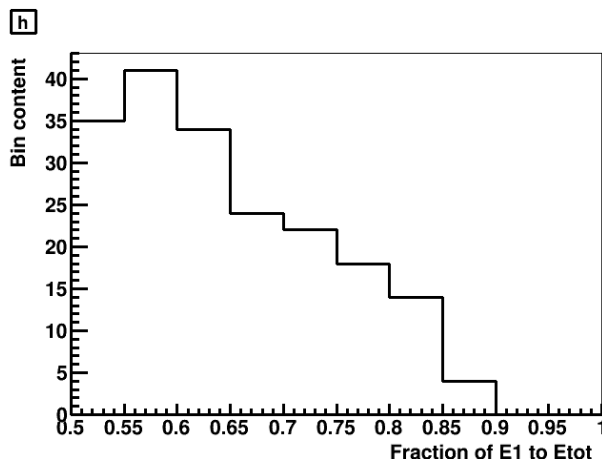
Figure 3.4: Fraction of energy that a single proton has compared to the total, for two-proton final states in the dip region.

The two proton events in this figure are being generated by the Delta and QE interaction models, not a MEC model. In looking at the figure, the distribution peaks in the 0.575 bin, which means that the most common thing to happen is that one proton has about 57.5 percent of the energy, and the other has the rest. The average of this distribution was also found to be 64.6 percent. This means that when two protons are produced the result is that they do not share the energy equally, one contains about 65 percent and the other has 35 percent.

The idea was to look at the total energy in the two proton case compared to the energy in the same area with a single proton. The result is that in the dip region two protons on average shared about 61 MeV of energy and one proton had on average

61.4 MeV of energy. This means that event though two protons are produced, the energy being transfered is on average the same in the dip region, as expected.

Expanding that to include the whole sample results in protons sharing on average 51 MeV and a single proton having 40.8 MeV. This means that on average the interactions producing two protons happen higher in the q_3 q_0 plot than the single proton does. This makes sense, since the lower portion of the plot is the QE which most often has only only one hadron in its final state.

The working conclusion is that was drawn from this section is that this algorithm is less sensitive to the number of hadrons in the event that we thought, when applied to protons with only a few tens of MeV each. John Demgen is finalizing this analysis.

Chapter 4

Resolution

The main part of my work on this project starts by looking at the uncertainty and resolution of the recoil energy, muon energy and muon angle. These are the inputs to the calculation of q_3 and Q^2. The end result of this investigation will be the determination of the error band on the final data vs. MC comparison due to the uncertainties on the quantities.

To propagate the uncertainty to the end of the analysis, these values will be biased by an amount equal to the uncertainty. A bias in the value means that the actual value is reconstructed low, and the values should be shifted up by the bias. This can be caused by a problem in the calibration of any of the electronics or accounting for energy loss or inaccuracies in Geant4.

In contrast to a bias, the resolution is a smearing effect in the data. This value levels out peaks in the distributions. Why this is done is say if the resolution in a quantity is really poor, then a distribution with a very steep peak would be distorted. By applying a resolution style effect the peak will be lowered and all the dip regions would be filled in. The main difference between a bias and the resolution is that the resolution works in both directions rather than the one direction for the bias. However for this section the effects are going to only be studied as a bias. The true nature of resolution will be explored in the next chapter.

The three different parameters that are going to be examined in this method are going to be the Muon energy, the Muon angle and the recoil energy. Each of these were

examined at two different magnitudes corresponding to the bias uncertainty in the given values and the magnitude of the resolution of the measurement.

The main goal of this section is to demonstrate the effect a bias has in the calculation of q_3 and Q^2. In addition to a clear demonstration of the effect, we want to look if there are special effects different between q_3 and Q^2 selection that would cause us to change the design of the analysis.

4.1 Resolution and Bias Magnitudes

The resolution and bias values that are going to be used in this section are all shown in table 4.1.

Type	Recoil Energy	Muon Energy	Muon Angle
Bias	5%	2.6%	1 mrad
Resolution	50%	8%	10 & 23 mrad
Worse Resolution	55%	11%	

Table 4.1: Uncertainties and resolution magnitudes for the recoil energy, muon energy and muon angle.

For the biases the values are the standard MINERvA uncertainties on these values. Specifically these are one milliradian on the muon angle, 2.6 percent on the muon momentum and five percent on the recoil energy[6]. The recoil energy is slightly different from the standard MINERvA uncertainty, it was approximated using the information from test beam[9] and the standard MINERvA uncertainty. Starting with the hadronic energy scale from the test beam, it was inflated to account for additional effects in neutrino data such as muon removal and noise uncertainties.

The size and shape of the resolution can be illustrated for our specific samples. The ratio of the difference between the truth and reconstructed value to the true value was taken in order to determine the resolution. This was done using the simulated data since the real data does not have the truth information. Once the ratio is taken, the RMS of that distribution can be obtained and a histogram can be used to visualize the

shape of the distribution. RMS or the root mean square is just the square root of the average of the squares. This plot for the Muon energy is shown below in Figure 4.1.
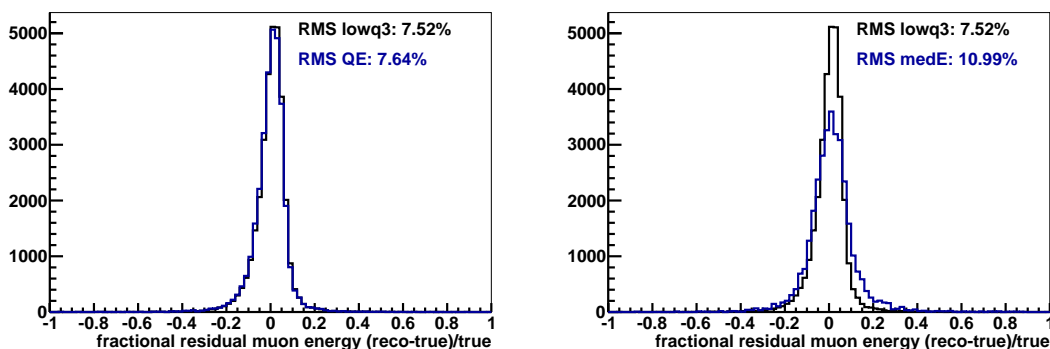
Figure 4.1: Fractional resolution for the muon energy for the low q_3 sample in both figures but the left plot contains the QE resolution in the low q_3 and the right plot contains the resolution for the medium energy in blue.

The RMS computed from the data for the low q_3 cut is 7.52 percent and is repeated in both plots. The effect of resolution will be illustrated using 7.52% rounded to 8%. The RMS in the medium energy sample (5 to 20 GeV) is 11 percent which can be used to estimate the uncertainty in the resolution. Though this is the predicted uncertainty in for a diferent sample, we are using it as a bound on how poorly the resolution might be modeled in our low q_3 sample. Degrading the resolution in the muon energy to 11% will be used in the later section.

Resolution for the recoil energy can be picked out from the plots shown in Figure 4.2.
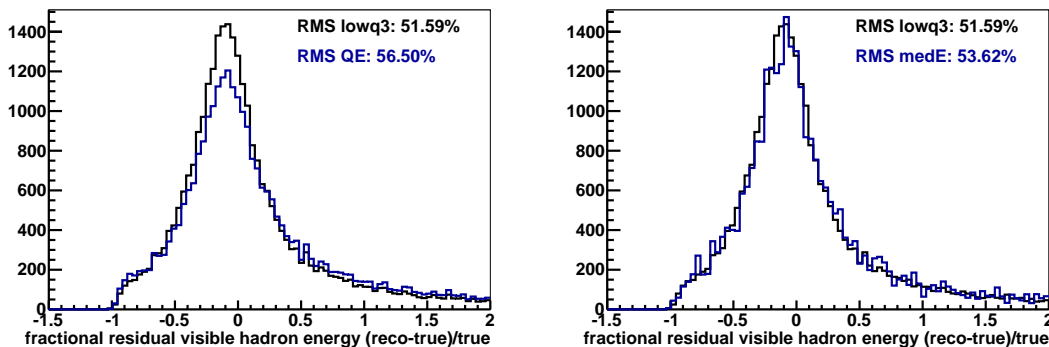
Figure 4.2: Fractional resolution for the Recoil energy for the low q_3 cut in both figures but the left plot contains the QE resolution in the low q_3 and the right plot contains the resolution for the medium energy.

Again looking at the RMS of these plots gives an idea about what the resolution is for the recoil energy. This quantity is constructed such that these are the fractional resolution in the hadron energy, or the energy transfer to the hadron system. Since the recoil energy is the energy seen in the detector, these two are often used interchangeably. However, because some fraction of the hadron system are energetic neutrons which are not reliably seen in the detector, the two quantities can refer to different things.

The resolution is not especially good nor is it Gaussian. It is a combination of detector effects when hadrons re-interact in the detector, plus fluctuations due to the intrinsic mix of neutrons, protons, and pions the model predicts are exiting the nucleus, plus effects of background activity incorporated into the hadronic system. The RMS values seen here are all just over 50 percent in size. The higher side of this effect is up to about 55 percent. This will also be used latter as a worse case type scenario. To evaluate the uncertainty we will degrade the resolution to 55%.

The angle of the Muon or the opening angle is the next one to be looked at and is seen below in Figure 4.3.
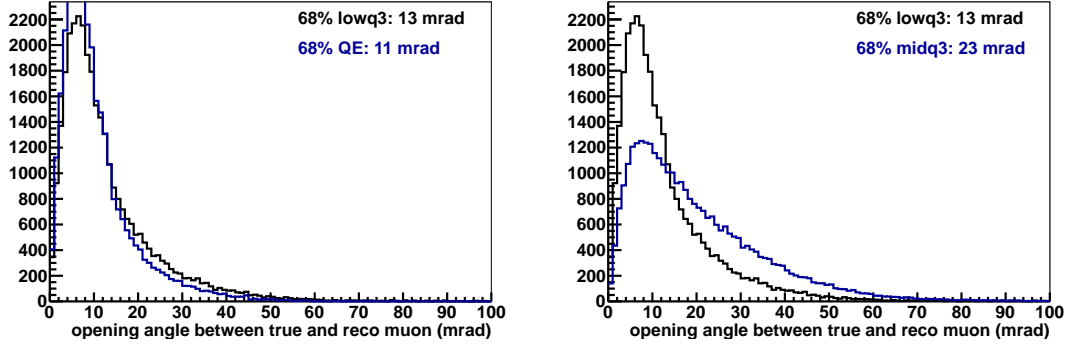
Figure 4.3: Resolution for the Muon angle for the low q_3 cut with the inclusion of the QE in the low cut on the left and the medium energy on the right.

This plot is slightly different from the rest. What is plotted in Figure 4.3 is simply the opening angle between the true muon momentum vector and the reconstructed vector. In the other two variables the resolution between cuts have been very similar. In this case however the resolutions are 10 milliradians in the low sample and 23 milliradians in the mid sample. The effect seen here is that the resolution in the muon angle depends on the energy of the hadron system, which can confuse the tracking algorithm.

4.2 The Sample

For this analysis the combined MINERvA 1 and 13c playlists were looked at. A playlist is a set of runs and sub runs that have similar characteristics like the configuration of the beam, calibrations and so forth. These playlists total 2/3 of MINERvA's low energy neutrino-mode data.

From the two playlists, several additional cuts are made in order to acquire the sample being tested.The first of the cuts is on the total energy. Specifically the energy that we want is going to be between 2 and 5 GeV. MINOS muon match is another cut. This ensures that the muon observed in MINERvA has its energy fully reconstructed by the MINOS near detector.

Another cut is the reduced fiducial cut. What this does is it makes the vertex of the

collision be in a specified region in the detector. This region is actually going to be the centered, upstream part of the tracker region. This region of the detector produced the cleanest results. At this time we are using a smaller fiducial region than a standard MINERvA analysis. Events near the edge of the detector might not have their hadron energy fully measured. A future analysis will redo this entire analysis with the larger, standard reduced fiducial volume and see how they compare.

The other cut that is going to be made divides the sample into regions based on Q^2 or q_3 and the analysis will focus on the lowest region in momentum transfer. This reduces the sample to a fraction of the total events. After making these cuts there are 1507 events in the low q_3 sample and 5310 events in the mid q_3 sample for the Data and 17426 in the low and 50364 in the mid for the MC. The MC was generated specifically to have approximately 10x the statistics as the data.

The difference in the number of events in the data and the MC brings about the need for a normalization. There are several different types of available, including normalizing be the number of protons delivered on the target (POT), subsample integral (or area) normalized and total sample integral (or area) normalized. All three types of normalizations apply a reweighting factor to the MC distribution to scale it back to a similar number as the Data. POT normalized is the ratio of the number of Protons on Target which is intrinsic to both the operation of the experiment and the design of the simulated sample's prediction of the flux neutrinos per second passing through the MINERvA detector.

Area normalization is taking the ratio of the number of events in a specific subsample of q_3 of the data to the MC, then reweighting the MC by that fraction, either event by event or analysis bin by analysis bin. The total sample area normalization is the ratio of the number of events in all q_3 for that energy region in the data to that in the MC. Through the other normalizations are also interesting, it is the total sample normalization that is going to be used throughout this analysis.

4.3   Smearing due to resolution sized biases

Now that the uncertainties and the resolution for the three different parameters have been found, the effect these have on the data and MC distributions can be examined. Shifts in these quantities due to resolution or bias will move events around in the q_0 q_3 space even to the point of moving them across the cut boundary to a different subsample. This section will quantify and visualize this movement in several ways, one per subsection.

Since the resolution effect is so much larger than the uncertainty, the effect can be shown for a resolution sized bias and it would be known that the effect is still there with the uncertainty just on a smaller scale. The process that was used, was to recreate Q^2 and q_3 from the reconstructed E_μ, θ_μ and E_{recoil} then one of those values is shifted according to the size of the resolution and everything is recalculated.

The first step is to calculate the muon's energy from its momentum p_μ, which is the relativistic relation between the energy, the mass and the momentum.

$$E_\mu = \sqrt{M_\mu^2 + P_\mu^2} \tag{4.1}$$

The muon's measured momentum or P_μ is read straight in from the MINERvA data structure and the muon's mass is 105.6583715 MeV/c^2. When the muon energy is tested it is going to be shifted by first shifting the muon's momentum instead and then recalculating the muon's energy. The next step is to calculate the neutrino's energy

$$E_\nu = q_0 + E_\mu \tag{4.2}$$

The value is a calorimetric sum of the total final state energies. In Equation 4.2 the recoil energy or q_0 is read in from the data structure tuple and the muon's energy was already recalculated by Equation 4.1. The recoil energy is another value that is going to be tested. Changing either the recoil energy or the muon's momentum changes E_ν.

At this point the value of the relativistic invariant, Q^2, can be calculated using Equation 4.3.

$$Q^2 = 2E_\nu(E_\mu - P_\mu cos(\theta_\mu)) - M_\mu^2 \tag{4.3}$$

In Equation 4.3 the only new variable is the muon angle which is read straight in from the tuple. This will be the third value that is going to be tested. From here the base equation for Q^2 can be rearranged in order to calculate the the three-momentum. The base equation is shown in the left side of Equation 4.4 and then rearranged for q_3 on the right hand side.

$$Q^2 = q_3^2 - q_0^2 \qquad => \qquad q_3 = \sqrt{Q^2 + q_0^2} \qquad (4.4)$$

We are not using it for the kinematic analysis, but this sequence of kinematics leads to

$$W^2 = M^2 + 2Mq_0 - Q^2 \qquad (4.5)$$

Where M is the mass of the struck nucleon. This expression uses the energy and momentum transfer basis and is actually used in Figure 1.1 instead of the equivalent hadronic system basis. Another note is that in the literature, the energy transfer is also variously called ν or ω, but here we will use the notation q_0.

4.3.1 Movement in the q_0-q_3 space

With a change in the recoil energy, muon energy or muon angle there will be some sort of motion in the q_0-q_3 space. The approach taken to look into the movement of events in the space was to divide up the space into squares that have each side being 0.1 GeV. Once the area was divided the before and after shift of the recoil energy and q_3 was found for a positive shift in the individual parameter. Then an average was taken of the events that fell in the lower energy corner of the box. Specifically the corner of interest would have dimensions of 0.06GeV on each side.

Recoil Energy

To start with the movement due to an upward bias in the recoil energy, the magnitude of the movement is different at different locations within this space. The movement of an event up and to the right is because a positive shift in recoil energy raises the three momentum pushing it to the right. Since the vertical is literally the recoil energy then the movement is up and to the right. Figure 4.4 shows how the events move around in the cuts due to a resolution sized shift.
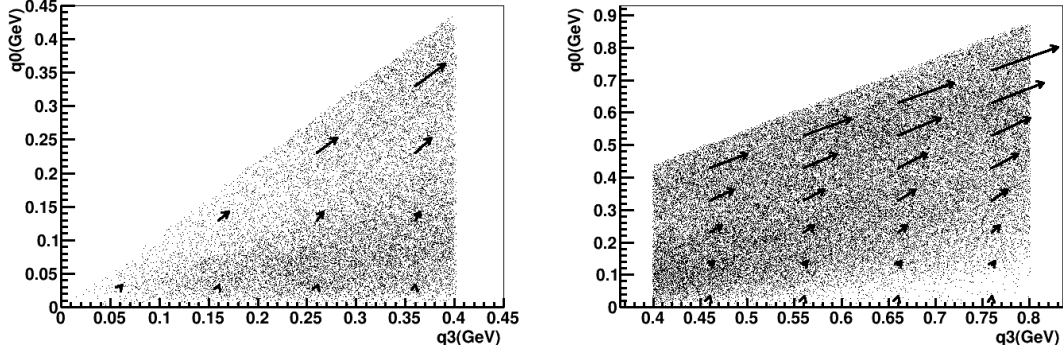
Figure 4.4: Movement due to a resolution sized shift in the recoil energy for the low q_3 on the left and the mid q_3 on the right.

The two key features here are when the recoil energy is equal to zero and along the diagonal. When the recoil energy is equal to zero there is no shift at all since the recoil energy is zero. Just higher than that the movement would be essentially vertical since the value of q_3 would be approximately equal to Q^2. Along the diagonal is when the recoil energy is equal to q_3. This means that when the recoil energy is biased by a certain amount, q_3 will be shifted the same amount. This moves the event directly along the diagonal.

These are still the resolution sized shifts in order to illustrate the shift with the larger magnitudes. In this case the size of the shifts where actually too big for the plot. Each of the arrows were scaled back to 20 percent of the full length, but the direction is still the same.

Looking back at Figure 1.7 to see what a possible selection in Q^2 looks like, with this type of shift in the recoil energy, the event can move a ways in the q_0-q_3 space but still not cross a cut line. Since the lines are bent over to the right the events seem to move along the cut lines rather than cross them. An event would need to be very close to the boundary if it was to cross over Q^2 line.

A cut in q_3 is vertical. This means that at higher energies the events are going to be able to be further away from the q_3 boundary but still cross it. What this is going to

do is shift a few of the QE interactions out, but a majority of the delta events will be pushed out of the sample.

Muon Energy

For the muon energy case, since the recoil energy is read straight in from the tuple and not affected by a bias in the muon energy, all the movement should be horizontal. The movement in the q_0 and q_3 space is shown below in Figure 4.5.

Figure 4.5: Movement due to a resolution sized shift in the muon energy for the low q_3 on the left and the mid q_3 on the right.

Unlike in the recoil energy case, these are the full sized movement vectors for a resolution sized bias in the muon energy. This means that on average these arrows are much smaller than those shown in Figure 4.4. In the recoil energy case the largest arrows occured along the diagonal, in this case though the largest arrows are when the recoil energy is equal to zero.

A bias in the muon energy or the muon angle affect the value of Q^2 directly and q^3 from that, see Equation 4.3 and 4.4. The highest effect that a change in Q^2 will have on q_3 is when the recoil energy is equal to zero, making Q^2 equal to q_3^2. This is the reason why the largest arrows in this case are at the lowest recoil energy values.

At lower recoil energies a cut on Q^2 is fairly vertical which means that at lower energies an event moving over one of the cut lines should cross over both of the cut

boundaries. This is actually the same across the whole sample however. Since the motion is directly horizontal, if an event moves further than it is horizontally away from either cut it will migrate over the line.

With a bias in the recoil energy, the larger movement occured in the delta region moving more of them out than in the QE. For the muon energy it seems this is reverse. The movement in the delta region is very small, where in the QE it is larger.

Muon Angle

The movement in for the muon angle should be very similar to the muon energy, being that the movement in the space is directly to the right. Also since a shift in the muon angle effects Q^2 directly and then q_3, the largest arrows due to these shifts should be when the recoil energy is equal to zero. Figure 4.6 shows the movement in the space for the resolution sized smear in the muon angle.
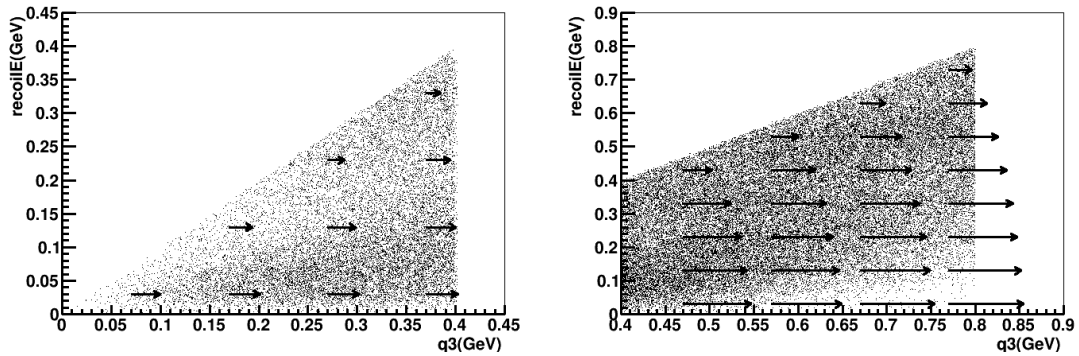
Figure 4.6: Movement due to a resolution sized shift in the muon angle for the low q_3 on the left and the mid q_3 on the right.

The trend in the movement in the recoil and q_3 space for a bias in the muon angle and the muon energy are almost the exact same. The movement of both are horizontal moving towards higher energies for a positive bias. It seems though that the size of the arrows for the muon angle are slightly larger than that of the muon energy, which is a mild surprise.

Similar to the muon energy bias, the larger movement occurs in the QE region of the distributions. One thing that is lacking in both of the plots is the direction of motion for events near the diagonal. Taking a limit on the arrow size as you approach the diagonal would suggest that they would go to zero. A bias in both the muon energy and angle is going to affect Q^2, the diagonal is when q_3 is equal to q_0 or equivalently when Q^2 is equal to zero. A change in something that is already practically zero would result in little to no movement in this space.

4.3.2   Percent Difference: first of three illustrations

Along with how the events move in the q_3-q_0 space, the percent difference of q_3 and Q^2 due to the biases can be calculated to quantify the size of the effect. The reason for doing this is that an event could move only a small amount in the triangular space but, the change in Q^2 could be large, these event could be in the middle to upper region of the plot when the q_3 has the majority of energy coming from the recoil energy.

What was done is similar to how the movement plots where created. First the code calculated Q^2 and q_3, then the biases where done one at a time and the code recalculated the values. The first set of values were saved as just the name and then the new shifted values were given the names RQ^2 and Rq_3. The percent difference from these values is then

$$PD_{Q^2} = \frac{RQ^2 - Q^2}{Q^2} \cdot 100 \qquad (4.6)$$

What the percent difference is going to say about the actual quantities is how much they are going to change due to these effects. In other words this is going to be getting at how stable Q^2 and q_3 are in terms of these quantities.

Percent difference of q_3 and Q^2 due to the recoil energy

From the resolution magnitude size section, the size of the shift due to the resolution in the recoil energy is 50%. The resulting percent difference plots for the resolution size scale in recoil energy are shown below in Figure 4.7 for both the low and mid q_3 selected regions.

Figure 4.7: Percent difference for a resolution sized shift in the recoil energy for the low q_3 on the left and the mid q_3 on the right. In both plots the black line is for q_3 and the blue line is for Q^2. In the plots the black line is for q_3 and the blue points are for Q^2.

Looking at both of these plots it is clear that over the whole data set q_3 is shifted more than Q^2. This is evident since the q3 line spreads out over higher percent differences than Q^2 does for both of the cuts in q_3. Another thing to note here is that the the the scale of the percent difference in q_3 is around 25 percent and is less than 10 percent for Q^2. The distribution for Q^2 is peaked at a central value with a modest tail. The q_3 distribution is much flatter in the low q_3 cut, with a small peak and pretty evenly distributed for the middle q_3 selection.

Percent difference of q_3 and Q^2 due to the muon energy

The same percent difference plots can be made when making an eight percent shift forwards in the muon energy. The result is shown below in Figure 4.8.
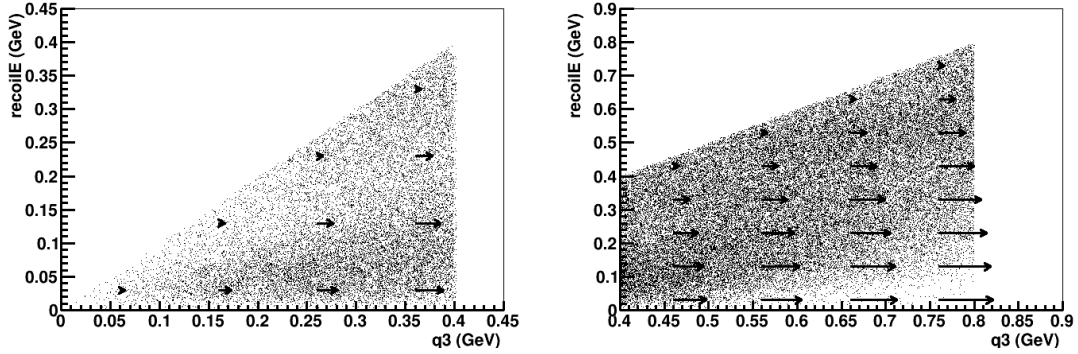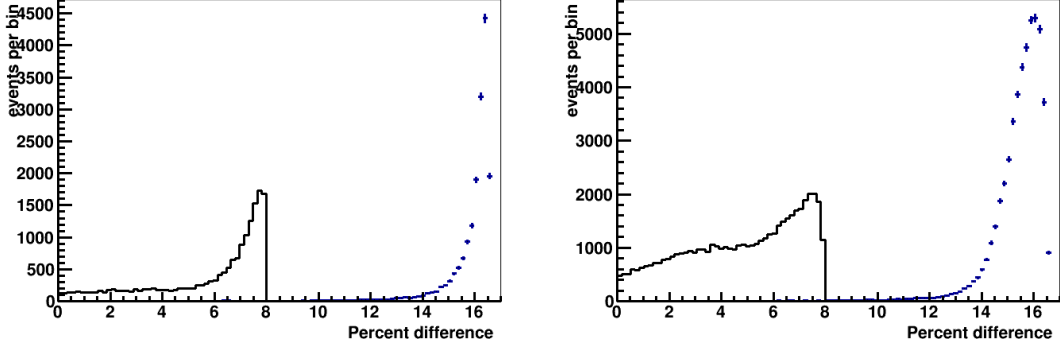
Figure 4.8: Percent difference for a resolution sized shift in the muon energy for the low q_3 on the left and the mid q_3 on the right. In the plots the black line is for q_3 and the blue points are for Q^2.

With a bias in the muon's energy the first thing to note is that the percent difference in Q^2 is larger than the percent difference in q_3. This is the opposite of what was seen in the case for the recoil energy. The size of the shift in q_3 is dispersed around five percent and for Q^2 is between 15 and 16 percent.

The big effect that the muon energy has on the distributions is at lower recoil energy. At lower recoil energy the values of q_3 and Q are approximately equal to each other. Assuming this is true, then a simple expression for the shift in Q^2 can be found in terms of shift in q_3.

$$(q_3)^2 \approx Q^2 \Rightarrow ((1+\epsilon) \cdot q_3)^2 = (1+\epsilon)^2 (q_3)^2 \approx (1+\epsilon)^2 Q^2 \tag{4.7}$$

In the expression the value ϵ is the percent change in decimal form. What this does is it says that the new value is one plus ϵ times the old one. An example of this is when there is an eight percent shift up in q_3, one plus ϵ would be equal to 1.08. 1.08 squared is just over 1.16 which means that the there should be roughly a 16 percent change in Q^2. Which is what is seen in the figure.

If the percent differences on average are summed, between the recoil energy and the muon energy, then the combination of the resolution shifts in recoil energy and muon energy results in the percent difference for q_3 being about 30 percent and Q^2 being about 25 percent.

Percent difference of q_3 and Q^2 due to the muon angle

Again the percent difference plots can be made for the shift in the muon angle. Unlike the recoil and muon energy, the muon angle is not a constant value across the low and mid q_3 cuts. Specifically for the low q_3 cut the size of the shift according to the resolution magnitudes is 10 milliradians and 23 milliradians for the mid q_3 cut. Since these are in milliradians, these will be true shifts not scales. The plot of the percent differences are shown below in Figure 4.9.
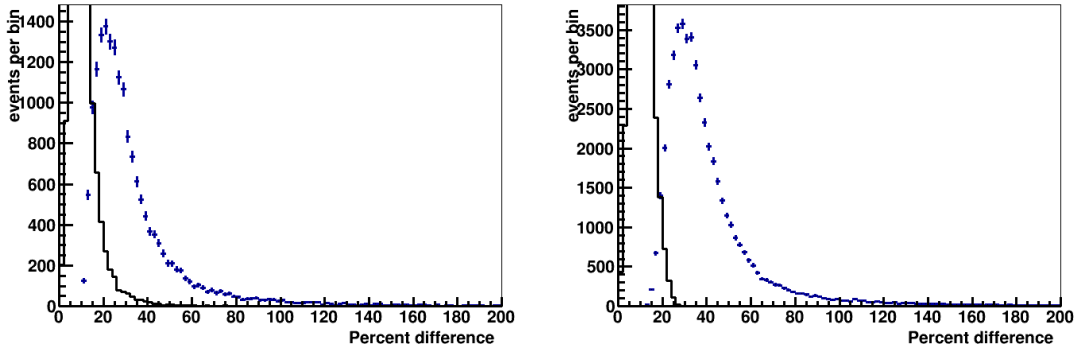
Figure 4.9: Percent difference for a resolution sized shift in the muon angle for the low q_3 on the left and the mid q_3 on the right. In the plots the black line is for q_3 and the blue points are for Q^2.

The smearing due to the muon angle produces a much greater percent difference than the muon energy and the recoil energy does. For the low q_3 sample the percent differences in q_3 is about 15 percent and about 30 percent for Q^2. The mid q_3 sample has about a 10 percent difference in q_3 and 40 percent in Q^2.

Part of this study was to understand if biases and resolutions caused problems when making cuts in either kinematic quantity q_3 or Q^2. In combination the outcome is neutral. The recoil energy potentially causes more problems for q_3 while muon angle and energy appear to cause more problems for Q^2.

4.3.3 Percent Difference in q_3 Vs. Percent Difference in Q^2

The comparison of the percent difference for both q_3 and Q^2 can be done for each data point. This would be interesting to see if there is a trend, that some subset of events are shifted more in q^3 and others more in Q^2. What can be done to address this is to plot the percent difference in q_3 on the x axis and the percent difference in Q^2 along the y axis. This generates the 2-D scatter plot which can be used to address this question.

Recoil Energy

Over both samples the percent difference in q_3 has a higher percent difference than Q^2. This means that the majority of the events will have a higher percent difference in q_3 so the events will mostly fall under a diagonal line which has a slope of one. This line would be if the percent difference in each quantity was the same. If all of the events fall below the diagonal then every event smears more in q_3 than in Q^2 due to the resolution size shift in the recoil energy. The 2D plot is shown below in Figure 4.10.
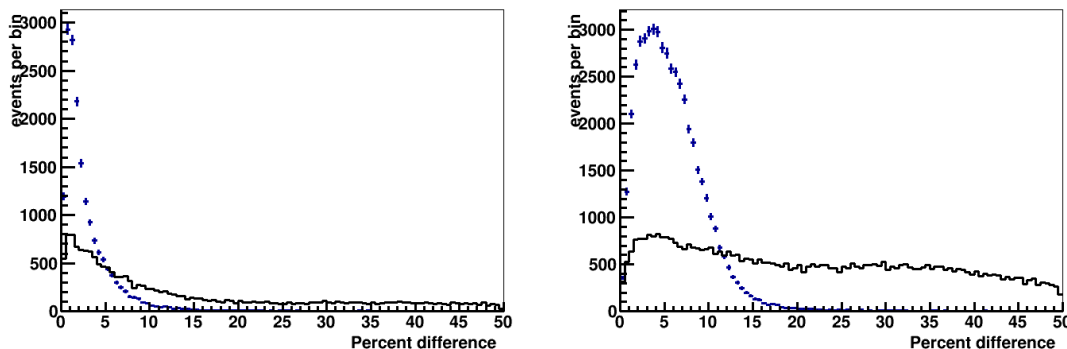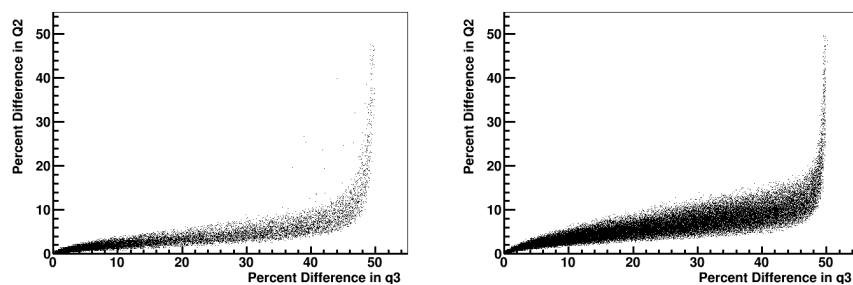

Figure 4.10: Comparison of the percent difference in q_3 and Q^2 for a resolution sized shift in the recoil energy for the low q_3 on the left and the mid q_3 on the right.

In both of these figures there is a steep tail up at about 48 percent difference in q_3. This tail is when the recoil energy is approximately equal to the q_3, or equivalently when Q^2 is approximately equal to 0. Also the majority of the events fall on the low side of the distribution with a lower distribution at the end. Which is what was seen in the percent difference plot.

Muon Energy

For the muon energy case, a resolution sized shift generated a percent difference of less than eight percent for both the low and mid cuts for q_3. Q^2 does not end up having significant statistics untill the percent difference of 10 percent. This means that ideally all of the events should be above the diagonal. The results are shown in Figure 4.11, below.
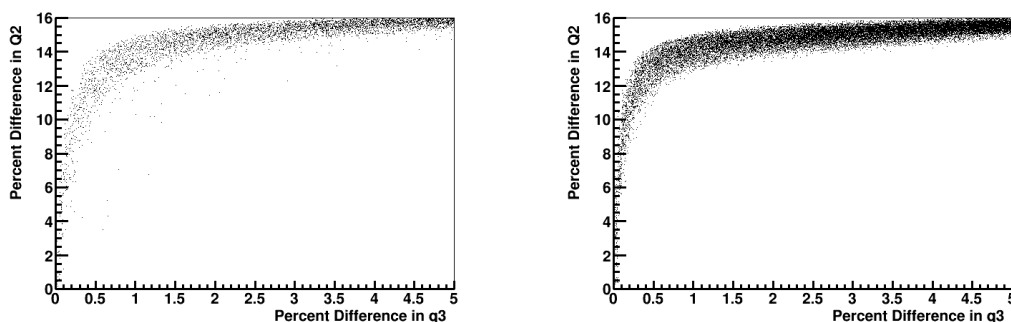
Figure 4.11: Comparison of the percent difference in q_3 and Q^2 for a resolution sized shift in the muon energy for the low q_3 on the left and the mid q_3 on the right.

The first thing to note is that there are actually events that have a percent difference less than 10 percent. The reason that they are not seen in Figure 4.8 is because there are just not enough of them to separate from the horizontal axis. Also another note here is that the diagonal line does not go from corner to corner as was seen in Figure 4.10. The reason why is because the percent difference in Q^2 is much higher than in q_3. All of the events in this case fall above the equilibrium line, which is opposite when the recoil energy was being tested.

Muon Angle

For the muon angle the majority of the events have a higher percent difference in Q^2 than in q_3. There is still a chance for some events that are in the reverse order though. The resulting 2D plot is shown below in Figure 4.12.
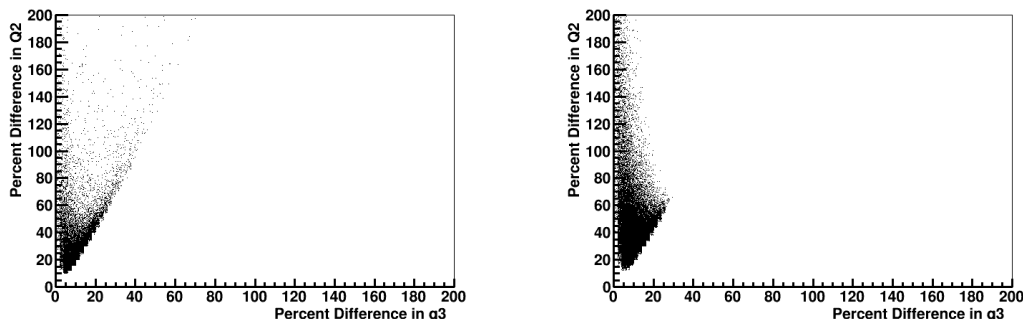
Figure 4.12: Comparison of the percent difference in q_3 and Q^2 for a resolution sized shift in the muon angle for the low q_3 on the left and the mid q_3 on the right.

Again it is seen in the figure that for any given event the percent difference in Q^2 is greater than the percent difference in q_3 because all of the events are above the diagonal. In fact the percent difference in Q^2 is more than twice the percent difference in q_3. unlike the other two, where the percent difference increased together, this one has a complicated structure relative to the diagonal.

All of these 2D plots are what were expected coming from the histograms of the percent differences. Given a bias in one of the variables, either q_3 or Q^2 had a larger percent difference uninamously. For the muon energy and angle case the percent difference in Q^2 was greater than the percent difference in q_3. The recoil energy was vice-versa.

4.3.4 Migration of events between q_3 cuts

A final study is to visualize and quantify how many events potentially cross a q3 cut boundary. The analysis presented in a later chapter will correct for this effect, and systematic biases in q_3 will become a systematic uncertainty in that correction. This really matters because if the large movement in one of the cut parameters is far from the cut region then it would not be as significant as if there are medium size movements close to the cut region. One way to examine this is to plot the unbiased value on the x axis and the biased value on the y axis. This can be done for both q_3 and Q^2.

Recoil Energy

Previously it has been seen that for a resolution sized smearing in the recoil energy, q_3 was shifted more than Q^2 was. First the plot of q_3 was made for this case. The result is Figure 4.13.

Figure 4.13: Unbiased vs regular q_3 for a resolution sized shift in the recoil energy for the low q_3 on the left and the mid q_3 on the right.

In these two plots any points above the 0.4 GeV in the low cut and above 0.8 GeV cut in the mid are events that are going to be shifted into the next cut after the bias. Reading from the figure, events with a three momentum transfer greater than 0.3 GeV in the low and greater than 0.6 GeV in the mid can migrate up into the next sample.

Another thing to note is that as the value of q_3 increases, the width of the band is also increasing. The reason for this is that for a given value of q_3, the triangle nature of the triangle plot seen in Figure 1.1, there is a wide range of what the recoil energy is. The recoil energy can vary between zero and the value of q_3. This results in the higher percent shifts in q_3 correspond to the higher values of the recoil energy in for that value of q_3.

Making the same plot for a similair cut in Q^2 results in Figure 4.14.

Figure 4.14: Unbiased vs regular Q^2 for a resolution sized shift in the recoil energy for the low Q^2 on the left and the mid Q^2 on the right.

The cut regions for the low Q^2 cut is from 0 to 0.16 GeV and 0.16 to 0.64 GeV for the mid cut. Similarly any event that falls above the maximum cut would be washed into the next cut region. Unlike in the q_3 cut though, the width of the bands are fairly consistent throughout both of the cut regions, and the widths are much smaller in this case.

Looking at these two sets of images there should be many more events migrating between cuts for a cut in q_3 compared to Q^2. This can be quantified by comparing the percent of events that migrate for each cut. Table 4.2 shows those results for both the data and the simulation for the low and mid cuts for both Q^2 and q_3.

Data Type	Data Value	Cut	migrating event %	# of Events before Migration
Data	q_3	Low	35.4	1507
Data	Q^2	Low	8.3	3945
Data	q_3	Mid	36.1	5310
Data	Q^2	Mid	5.8	7034
MC	q_3	Low	26.8	17426
MC	Q^2	Low	7.0	42084
MC	q_3	Mid	35.9	50364
MC	Q^2	Mid	6.3	70571

Table 4.2: Migrating events in percents for a resolution sized bias in the recoil energy.

Due to a resolution sized bias in the recoil energy nearly 36 percent of the events migrate for the cuts in q_3. For the cut in Q^2 there is more along the lines of five percent of events migrating. Why this is hapening was explained previously through the movement plots. A bias in recoil energy essentially moves events along Q^2 cut lines.

This is for RMS resolution sized shifts. The real modeled resolution includes both smaller and larger shifts in both directions, so it illustrates the effect on a single event, but not how the resolution affects the whole sample. Likewise, the shift due to a 5% bias is one tenth that of the resolution, and would approximately have an order of magnitude less effect.

Muon Energy

For the muon energy the percent difference was larger in Q^2 than in q_3. For the muon energy the reverse pattern should be seen here than what was seen for the recoil energy. Meaning that there should be more events migrating in a cut in Q^2 than in q_3. First the plot for q_3 is shown in Figure 4.15.
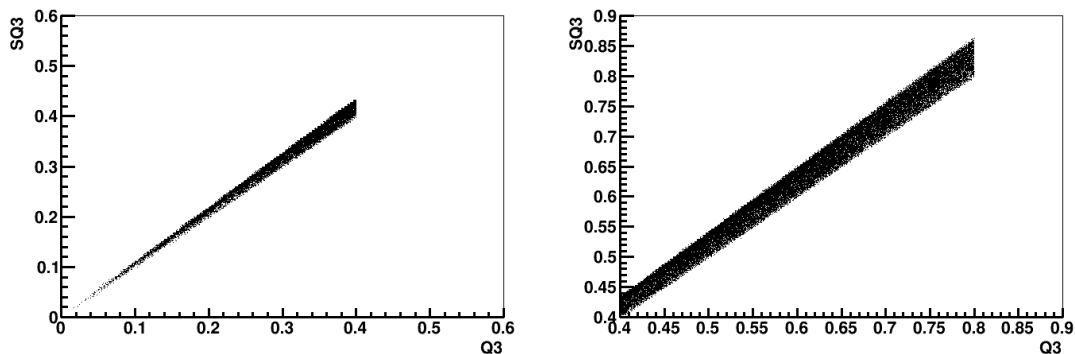
Figure 4.15: Unbiased vs regular q_3 for a resolution sized shift in the muon energy for the low q_3 on the left and the mid q_3 on the right.

Again there is a widening of the of the band as q_3 is larger. The width is a lot smaller than was seen with the recoil energy. Making the same plots for the Q^2, Figure 4.16 is created.
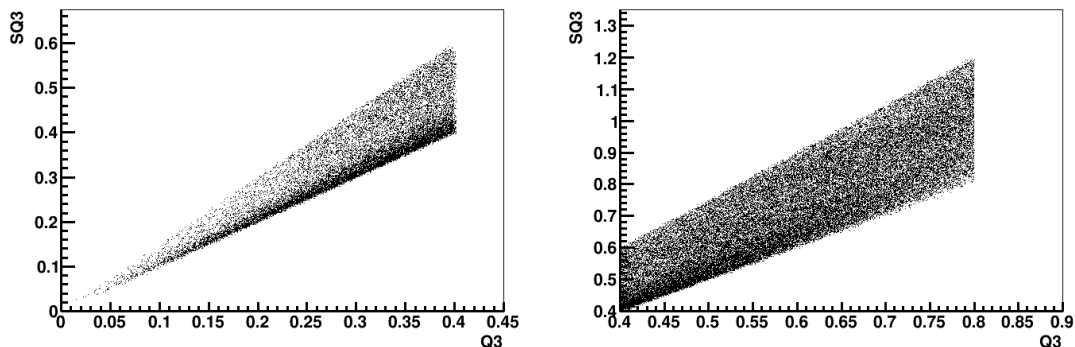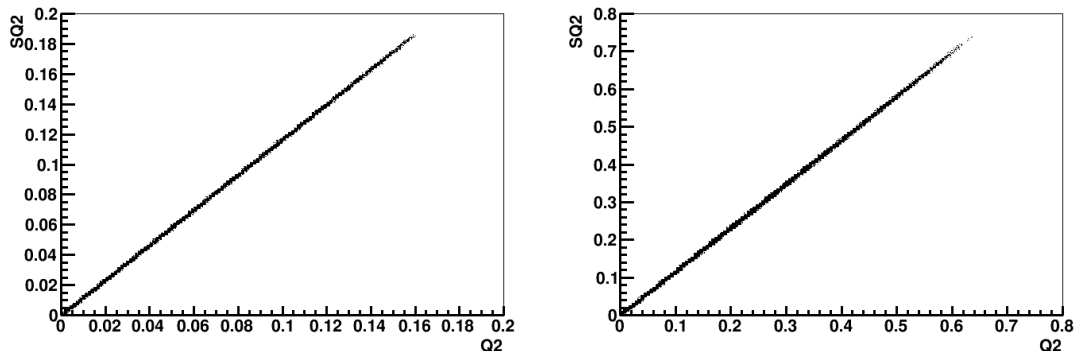
Figure 4.16: Unbiased vs regular Q^2 for a resolution sized shift in the muon energy for the low Q^2 on the left and the mid Q^2 on the right.

Very similar to what was seen for the recoil energy case. At this point what is going to be expected is that there will be more events that migrate in q_3 than there will be in Q^2. Again looking at the same table that is shown in Table 4.2 for the muon's energy results in Table 4.3.

Data Type	Data Value	Cut	migrating event %	# of Events before Migration
Data	q_3	Low	14.9	1507
Data	Q^2	Low	15.2	3945
Data	q_3	Mid	7.3	5310
Data	Q^2	Mid	8.2	7034
MC	q_3	Low	12.1	17426
MC	Q^2	Low	12.7	42084
MC	q_3	Mid	7.5	50364
MC	Q^2	Mid	8.4	70571

Table 4.3: Migrating events in percents for a bias in the muon's energy.

The results seen in Table 4.3 are not what was expected from looking at the plots but is what was expected when going back to the percent differences. There are more events that are shifted in Q^2 than there is in q_3. The explanation for this is that even though the width of the bands seen in Figures 4.16 are still smaller than in q_3 the slope of the base line is no longer one. This can be seen by looking at the y value of

the bands at a Q^2 of 0.16 GeV. The value is 0.18 GeV, this means that given a single value of Q^2 all of the values are shifted by a very similar amount.

Muon Angle

For the muon angle, a similar thing should be seen when compared to the muon energy. The percent difference in Q^2 is much higher so more events should migrate. The plot for q_3 is shown below in Figure 4.17.
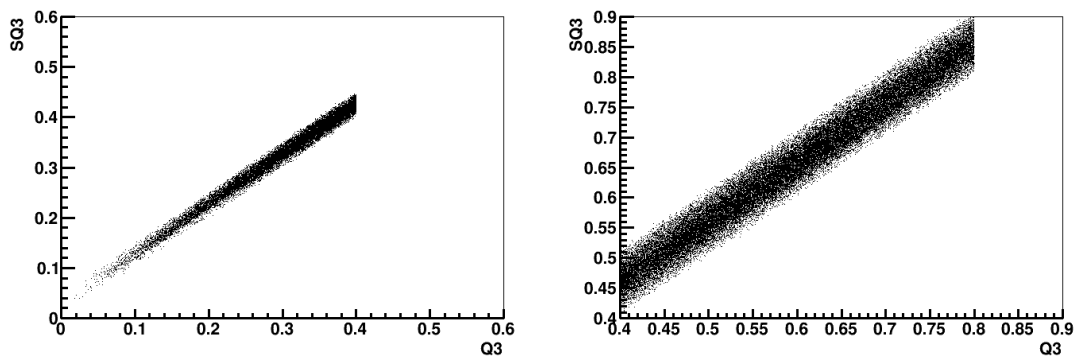
Figure 4.17: Unbiased vs regular q_3 for a resolution sized shift in the muon angle for the low q_3 on the left and the mid q_3 on the right.

Again there is a similar pattern in the low cut, how it goes up and widens out towards the top. The mid q_3 cut however stays at a pretty consistent width. This could be because the bias in this case is a set amount rather than a percentage. The width of this band looks to fall right in between the recoil energy and the muon energy. Next the Q^2 equivalent was made, which can be seen in Figure 4.18.
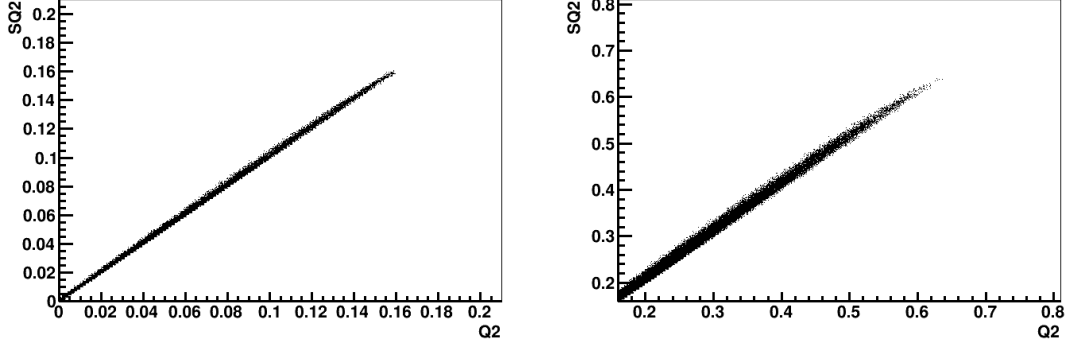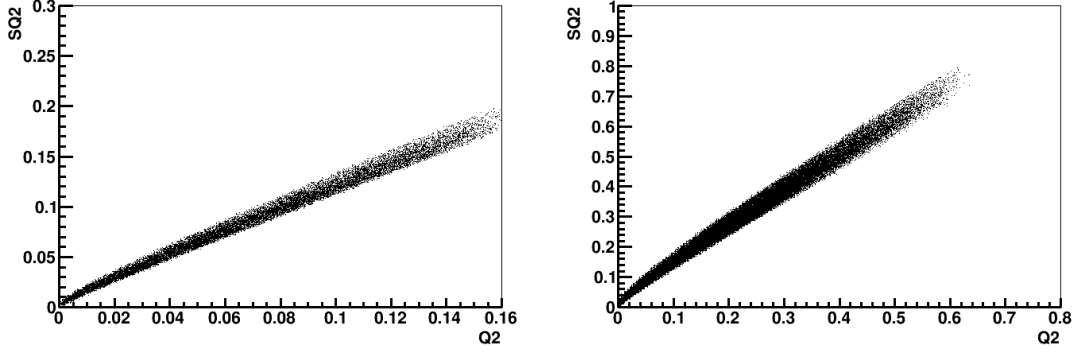
Figure 4.18: Unbiased vs regular Q^2 for a resolution sized shift in the muon angle for the low Q^2 on the left and the mid Q^2 on the right.

The width of these bands in Q^2 are the widest out of the three variables and are also sloped up like what was seen in the muon energy case. With the combination of both of those factors there should be a substantial amount of migration in the Q^2 cuts. However with the width of the band for the q_3 cut, there would be a substantial amount of migration there as well. Table 4.4 is the actual numbers for the migration due to a bias in the muon angle.

Data Type	Data Value	Cut	migrating event %	# of Events before Migration
Data	q_3	Low	19.11	1507
Data	Q^2	Low	17.2	3945
Data	q_3	Mid	13.5	5310
Data	Q^2	Mid	12.6	7034
MC	q_3	Low	15.4	17426
MC	Q^2	Low	14.7	42084
MC	q_3	Mid	13.5	50364
MC	Q^2	Mid	13.1	70571

Table 4.4: Migrating events in percents for a resolution sized bias in the muon's angle.

These are all pretty close q_3 is migrating only a few more events to above the cut than Q^2 does. For the muon angle and muon energy, the migrations are fairly close but with one having a higher migration in q_3 and the other in Q^2. Yet the recoil energy has a very high migration in q_3 and a lower migration in Q^2.

To this point it has been shown that there is a higher percent difference in Q^2 yet there are more migrations in q_3. The cause for this discrepancies is the design of the cut boundaries. That in the q_3 and recoil 2D area, the cut lines for q_3 are vertical lines where the Q^2 lines are bending over to the right.

When the recoil energy is being biased the movement is up and to the right, where the movement in this space for the muon energy and angle are straight to the right. At lower recoil energy this is not a big difference as a cut in q_3 and Q^2 are roughly the same. At higher recoil energies though, an event has to still be shifted roughly the same amount to cross the q_3 cut line but would need to move much further to cross the Q^2 cut line since the movement is up and to higher energies. When using a q_3 selection for the analysis later, the most significant movement across the selection boundary will be for the Delta events at higher energy transfer.

Chapter 5

Migration within the recoil spectrum

Another systematic study that was done was to shift and or scale the recoil energy and the vertex energy in the simulation. The reason for doing this was to see if a bias in the energy quantity could produce the discrepancies seen between the data and simulation. This was an initial exploration of the technique that becomes the full uncertainty analysis in the next chapter. I was looking for what direction and what magnitude could produce a better agreement to the data.

The idea is to explore known uncertainties and see if they cause the discrepancy in Figure 1.8. The scales used in the previous section would account for anything that could be causing the energy readout to be lower than what it actually is. For example if you learned the passive material correction is slightly off in the MC, you would correct it by applying an additional scale. A shift in the value would be because something is wrong by a constant amount of energy. For example if energy in the vertex is being added into the muon path where it should be the hadron's energy, or subtraction of background noise is wrong, you would add or subtract a constant amount from the MC.

This chapter is also my opportunity to look at all four subsamples in q_3 at once. The entire sample that passes the initial cuts was then divided into four different regions of q_3. These four samples are shown below in Figure 5.1.
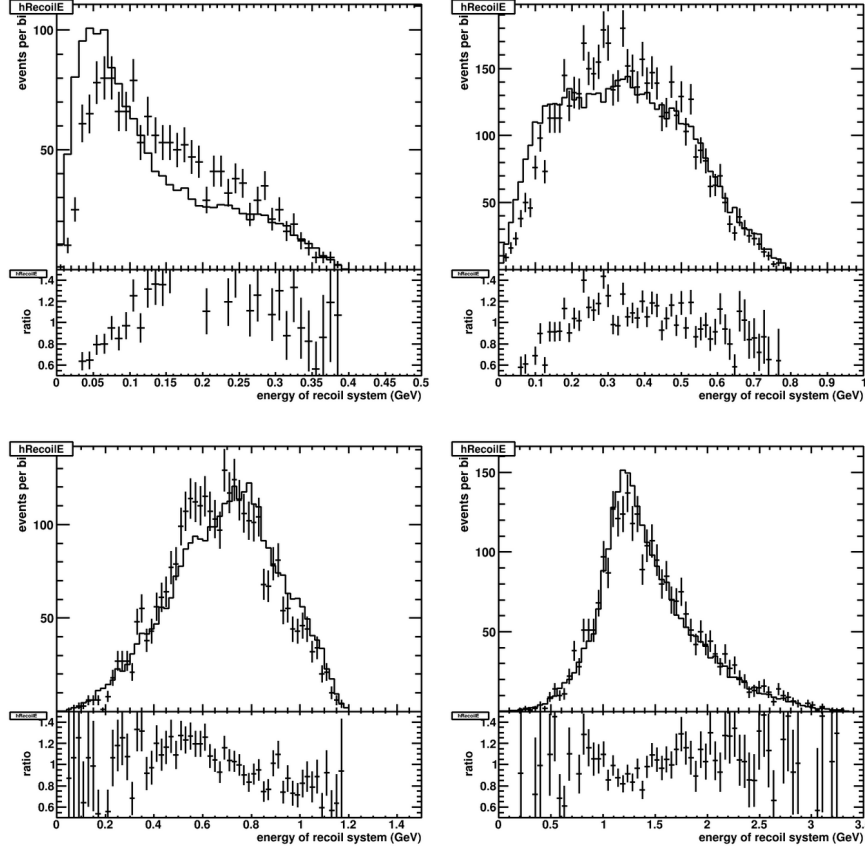
Figure 5.1: Recoil energy distributions over the different q_3 regions. The top left, top right, bottom left and bottom right plots are the low, mid, high and very high q_3 selections in that order.

The q_3 selection regions are 0-0.4(top left), 0.4-0.8(top right), 0.8-1.2(bottom left), and greater than 1.2(bottom right). Looking at these four plots it is clear that the simulation and the data do not agree perfectly. For example in the low q_3 cut there is a surplus of the simulation at low recoil energy but an underestimate at about 0.2 GeV, same as was illustrated in the introduction. The test here is if a shift in the recoil energy in either direction by 10MeV or a scale of $\pm 10\%$ would create a better agreement between the data and the simulation.

The 10 percent scale applied to the values was a reasonable value to use because of the uncertainty and the resolution values. The uncertainty in the recoil energy is five

percent and the resolution is 50 percent. This 10 percent is twice the uncertainty, the size of a two standard deviation effect, and is big enough that we will clearly see the effect in the plot.

One motivation for choosing the 10 MeV as an absolute shift parameter is because that is what each bin width is. Looking at the distributions to the eye it seems that if a few of the events were moved by one bin, then the agreement would be better. 10 MeV is also a reasonable number in terms of the uncertainty. It is similar in size to the uncertainty in how much energy is lost unbinding nucleons, how well we model muon energy that leak into the reconstructed hadron system and the energy deposits from stray neutron activity.

Changing the recoil energy is going to have two effects in these distributions. The first is that it will move events around in the sample, shifting the contents of some bins up or down. Also changing the recoil energy as seen previously will cause events to migrate between the different selections.

The two different effects were tested individually. This analysis was run through by first making the selection, and then applying the shifts and scales. It was also done by applying the shift and scale before any of the calculations so that both effects were seen. I am going to present the results for the later combined shift, but the conclusions were the same for both of these styles.

From here multiple plots are tiled in order to quickly compare the results between all of the shift values and scale values. The best comparison would be to look at the ratio part of the plots and see if any shift or scale values made the distribution of ratio's flatter. Each row in the tile plots has a constant shift value, going from top to bottom the shift values are -10, -5, 0, 5 and 10 MeV. The same thing can be said for the columns and the scale values. Starting from the left and going right the columns have a scale value of -10, -5, 0, 5 and 10 percent.

An example of one of the tile plots is shown below in Figure 5.2.
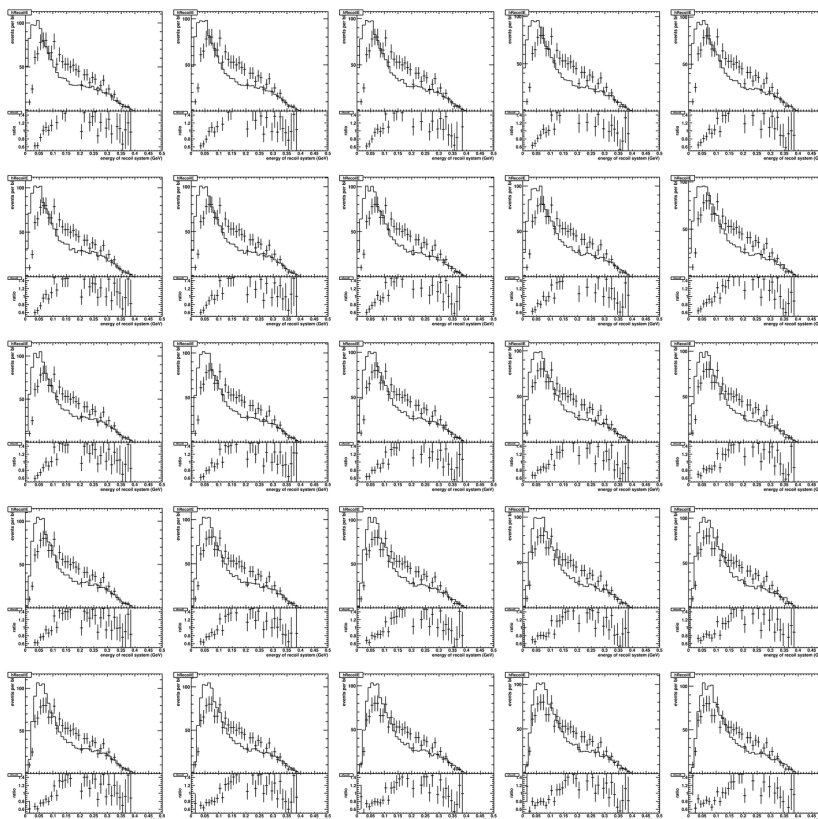
Figure 5.2: Tile plot for the low q_3 selection. Shift values for the rows starting at the top and going down are -10, -5, 0, 5 and 10 MeV. Scale values for the columns starting at the left and going right are -10, -5, 0, 5 and 10 percent.

In looking at Figure 5.2, the plot with a shift and scale value of 10 produced a noticeably flatter ratio distribution at the lowest energies, the very left part of each plot. The higher energy region is only changed a little. That means that at these values the agreement improved. The plot with shift and scale values of -10 seemed to have slightly made the agreement worse, but in this case it was actually easier to tell looking at the actual distribution. The peak in the shifted or scaled simulation is clearly quite low.

This type of plot was made for all four different q_3 selections. Once all four were made three different sets of shift and scale events were selected in order to draw a conclusion from this section. These three picked values along with the default were put side by

side to create Figure 5.3.
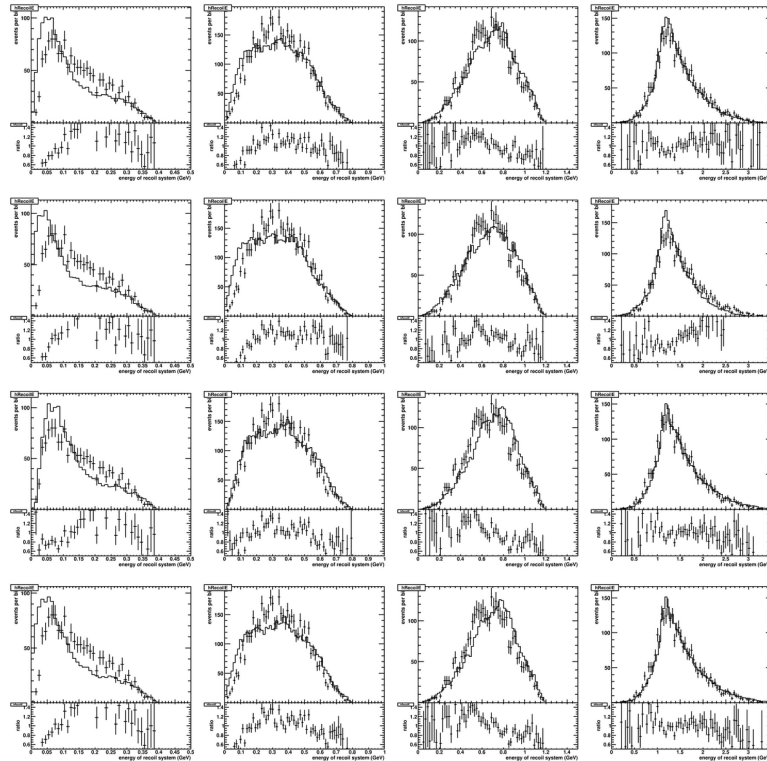
Figure 5.3: Select values for the shift and scale parameters. Shift and scale parameters are (0,0), (-10,-10), (10,10) and (-10,10) starting at the top row and working down. Columns are the low, mid, high and very high q_3 cuts going from the left most column to the right.

Each column of the above plot is a selected q_3 region with low q_3 on the left and very high q_3 on the right. Each row of the image has a different shift and scale parameter. Starting at the top the shift and scale parameters are 0MeV shift 0 percent scale, next is the -10 MeV and -10 percent scale, then it is 10MeV and 10 Percent scale and the bottom row is -10MeV and 10 percent scale. The comparison was done by eye, no fit parameter was used to determine the best fit.

To the eye it looks as if the second row and the fourth row did not make an improvement to the agreement between the data and the simulation. However the

third row does indeed show an improvement across all four cuts. The agreement is improved, but still far from perfect. This suggests that a systematic adjustment to make the MC hadron energy scale higher would help, but might not be enough.

Even though a slight improvement was seen here in the agreement between the data and simulation for the recoil energy when it was shifted before the reconstruction, was not the case for a quantity called the vertex energy. This is adding up just the energy in the tracker region near the vertex. It is the same quantity used in the MINERvA publications[10][11]. In my analysis, certain shifts or scales produced better agreement in the low q_3 sample, but made the agreement visibly worse in the other samples. A known but mild problem with muon energy leaking into the hadron energy estimate would affect the recoil energy but not the vertex energy, and could account for the two results observed in this study.

This study, and the realization that the recoil energy suffers from non-hadron backgrounds, caused us to change the quantity that is being examined. The tracker + ECAL energy is a comparable quantity to the recoil energy, but does not include energy in any of the hadronic calorimeters or upstream of the tracker region. This eliminates large regions that might contain background energy that may be porrly modeled. It includes more energy than the vertex energy does, but less than the total of everything seen in the detector.. This value is going to be used in a similar analysis in the next section.

In the context of our particular measurement, these effects will be turned into a systematic error band on the available tracker plus ECAL energy. With a systematic error band, the default model and alternate models for the low momentum transfer reaction can be evaluated to see which of them best describe the data.

Chapter 6

Analysis

The results from the resolution and bias chapter can be rolled into an analysis to generate error bands on the Monte Carlo. All of the biases have been done on the MC which we can use to generate an error band on the results. Through a process called unfolding, these error bands will be propagated onto a truth-like data quantity, which is a precursor to a measured cross-section.

The reconstruction of events is not perfect. An event could have a different energy from what is actually reconstructed, which is what was quantified with the resolution studies. A process called unfolding is a way to account for the smearing of the detection and reconstruction process. This produces a quantity for the data that is the equivalent of the MC truth information; its what the data would look like if there was no resolution smearing. Figure 6.1 shows the reconstructed MC and data and the resulting unfolded data and MC. The tracker+ECAL energy is being unfolded to the available charged particle energy. Quantitatively, it is the sum of the Proton kinetic energy, charged pion kinetic energy, neutral pion total energy, gamma ray photon energy and the electron energy. Only the first three have significant contributions to the total.
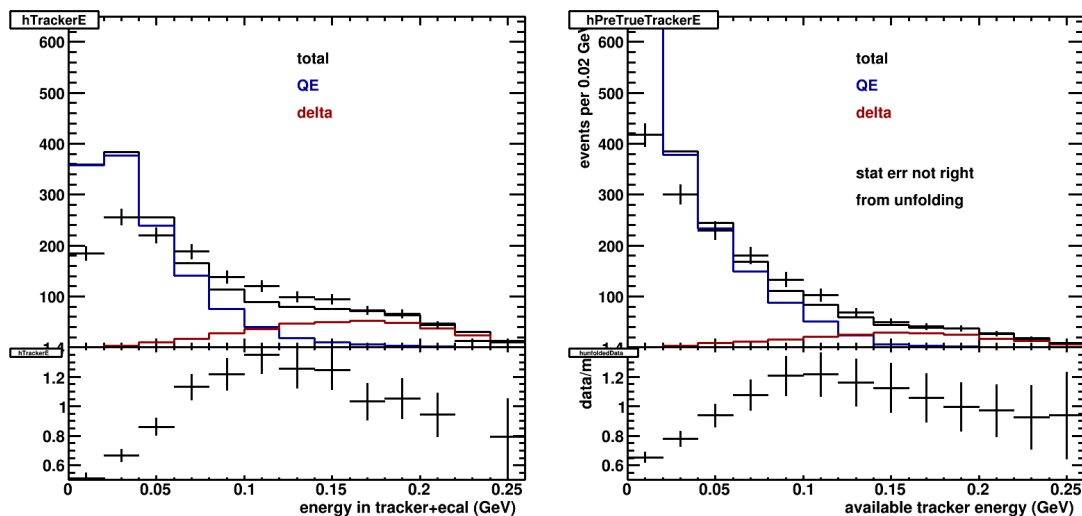
Figure 6.1: Reconstructed available tracker + ECAL energy for the data and MC shown on the left and the unfolded or truth data and MC on the right.

In both Figures, the data are the points with error bars and the MC is the solid black line. The blue and red lines that appear here are the contribution of the QE and delta respectively to the total MC.

The left hand plot is the reconstructed data and the MC. This was the amount of energy in the tracking region plus that in the ECAL with the passive material correction applied to it and a factor of 1.12 to produce an unbiased estimator. A reminder, the main details that are being looked at are the MC's over estimate of the number of events in the low energy bins and the lack of events in the dip region between the Delta and QE.

The right hand plot is the unfolded Data and MC. For the MC this is the sum of the true energies of the charged particles seen in the detector. The data was produced by applying the unfolding matrix to undo the predicted smearing effects and an acceptance correction to undo the movement of events in or out of the cut region.

Here is how the migration matrix is made. For the MC the Truth information and the reconstructed information can be used in order to calculate a matrix that essentially

shows how events are truthfully in one bin but reconstructed in others. The real and simulated smearing occurs on an event-by-event basis, but the migration matrix is constructed in bins of reconstructed and truth quantities.

The matrix can then be inverted in order to find the opposite effect, this is the unfolding or migration matrix. The inverse of the matrix looks at the reconstructed data and puts the events back into "Truth" bins. Using this on the data is part of what is used to produce the right side plot in Figure 6.1. The migration matrix for that example is shown below in Figure 6.2.
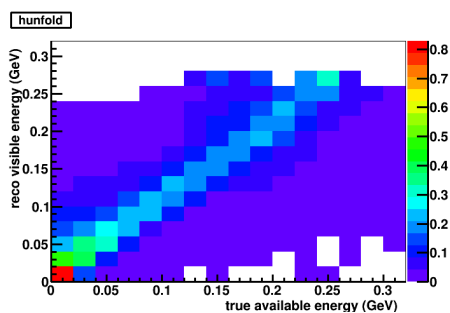
Figure 6.2: Unfolding Matrix for the transformation shown in Figure 6.1.

Take a single row in the matrix, each value in that row on the plot is the percent of the total events that have a true energy in that spot. Since there is a strong tendency towards the diagonal then the events have the highest chance to be reconstructed to the same spot in the unfolding histogram as the bin of the truth histogram from which they came. Since each row is the percent of events that have that truth value, the sum of the values in a row should be one.

The other effect that is present going from the left to right in Figure 6.1 is is the acceptance correction. What this does is it examines the effect of events migrating both into and out of this q_3 selection. As seen in the previous section the resolution smearing of the recoil energy for example could move the events into or out of the sample.

For this situation what happens is that the reconstruction of q_3 puts an event in or out of the sample when the truth information determines it must be out or in the sample.

The ratio can be found of the events in a bin for the true MC with the truth q_3 versus the reconstructed q_3. It creates a scale for each bin which accounts for events being reconstructed into or out of the low q_3 cut incorrectly. Typically tens of percent of events are affected, these corrections for the low q_3 cut are shown in Figure 6.3
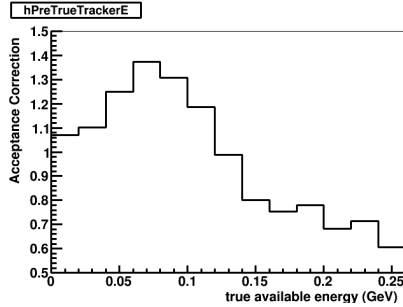
Figure 6.3: Acceptance correction for the unfolding process that was used in creating Figure 6.1.

When looking at the above figure, this means there should be more events in all of the bins less than 0.12 GeV than are put there by the reconstruction. Above 0.12 GeV, there are more events due to a cut in the reconstruction than there should be. This acceptance correction is applied to the truth like quantity bin-by-bin, even though the effect itself is event-by-event.

Once the reconstructed plot is created, the unfolding process starts by applying the migration matrix to the data. Then the acceptance matrix is applied to the distribution. This process turned the data distribution in the left plot in Figure 6.1 to the unfolded data shown in the right plot.

6.1   Flow diagram for this analysis

At this point the uncertainty biases previously discussed can be applied to the MC to see how they translate through the unfolding process and distort the final result. The uncertainty in the recoil energy, muon energy, and angle are about five percent, 2.6 percent and about one milliradian respectively. In order to better understand how

these shifts propagate through the code, a flow chart was created and is shown below
in Figure 6.4
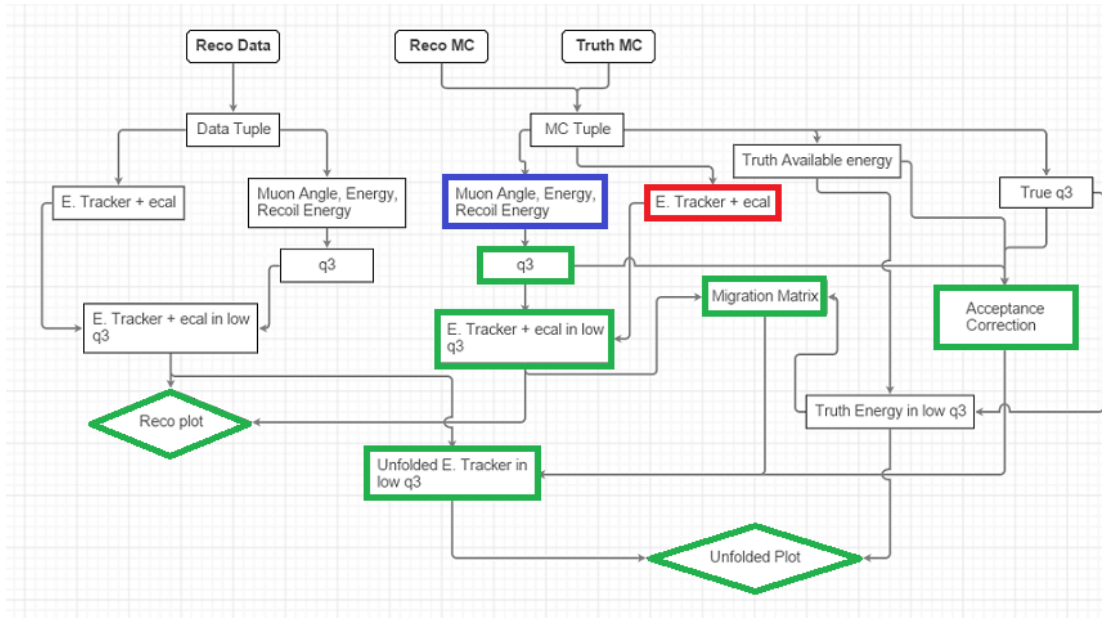
Figure 6.4: Flow chart to describe how the code works along with highlighting areas of
interest.

To explain the flow diagram lets first start with the data side. From the reconstructed
data a file called a tuple is created. This holds all of the information that is needed for
the selection, analysis, and the plots. From here two groups are chosen. The first is
the muon angle, muon energy and recoil energy; using these a value for q_3 can be
calculated. The value on q_3 is used to select the tip of the kinematic triangle. The low
q_3 subset events have their tracker + ECAL energy analyzed; that is the Energy in
tracker + ECAL value, that is the second path from the data tuple, in order to
produce the data line in the reco plot.

Jumping over to the MC now the first difference is that not only is the simulated reco
information available, but so is the truth information. Using the reconstructed
quantities the same thing is done that was done for the data. With both of those
combined the reconstructed plot is created to compare them to each other.

On the right side of the diagram, the migration matrix is formed from the truth available energy in the low q_3 and the energy in tracker + ECAL in the low q_3. Also at this point the truth q_3, the reco q_3 and the truth available energy are all combined to form the acceptance correction. Both of these are then used on the data in the low q_3 region to unfold the data.

The unfolded data represent our measured value of nature's true distribution in available energy. We can compare the default model we have been using all along, and expect to see qualitatively the same disagreement we saw with the reconstructed quantity comparison. In addition, we can compare any other model, even one like Nieves MEC model that has not been simulated to the reconstructed quantity stage.

In the flow chart there are certain parts of the process that are colored. The first one is the blue box that contains the MC recoil energy, muon energy and muon angle. It is at this step where the different systematic uncertainty effects are applied. Everything that comes in line after that is going to be affected and will be recalculated with the modified, shifted inputs. All of these boxes are colored green. The three changes primarily affect which events are in the low q_3 sample, and have a secondary effect on other quantities.

The only other box is the red box. Towards the end of this section, a short analysis will be done in showing what applying a shift or resolution effect to the available tracker energy and ECAL would have on the final distributions. These effects will happen in this red box, and propagated through all of the green boxes, but do not change the q_3 calculation or which events are in the low q_3 sample.

6.2   Acceptance and Background effects due to a bias in the MC

Changing the recoil energy, muon energy and angle effects the low q_3 sample, the migration matrix and the acceptance matrix. This means even though the data events are unchanged, the unfolded data distribution is distorted. The effect that these changes have on the migration matrix though are not significant when compared to

the effect on the acceptance matrix. Applying the biases in question, the main effect being examined is the change to the acceptance correction.

Applying the uncertainty biases in the recoil energy, muon energy and muon angle will be done the exact same way as previously in in the biasing chapter. In short, before any of the reconstruction is done the individual parameters will be biased forwards by there respective uncertainty. This will only be done in the MC and not the data. The uncertainties were applied one at a time in order to see the effects that each has on the distributions shown in Figure 6.1. The results for the bias are shown in Figure 6.5.
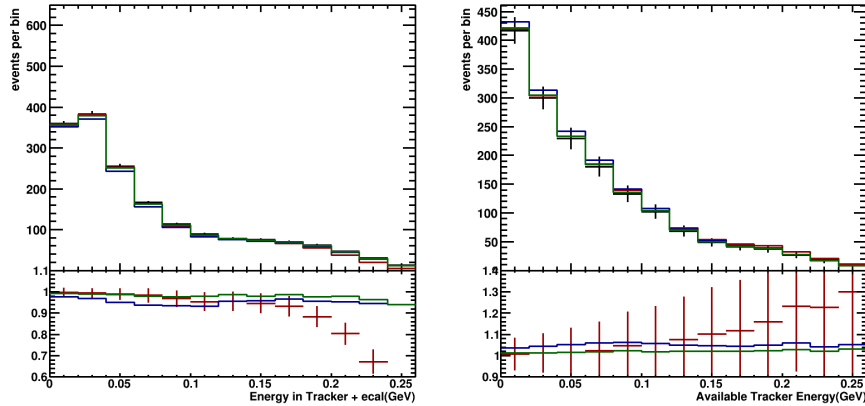
Figure 6.5: Reconstructed MC on the left and unfolded Data on the right for the uncertainty bias forwards in the recoil energy, muon energy and muon angle.

The plots are of the reconstructed MC information on the left and the unfolded data on the right. The different color lines correspond to the different biases being applied. Specifically the black line is the default case, red is with a bias in the recoil energy, blue is the muon energy and green is for the muon angle. The effects are small, so the lower portion of each plot shows the ratio of the shifted samples to the default, which highlights the magnitude of the distortion.

Looking first at the plot of the MC, every single bin is lower than the default case. Since all of the values are being biased up and the Energy in Tracker + ECAL

horizontal axis is independent of the three variables this is not surprising. All that is being seen with these shifts is a migration of events out of the sample, which means that each bin is either equal to or lower than the default case.

When looking at the Muon energy and angle there is not a huge effect, all bins were lowered by no more than five percent. There is a slight decrease in the ratio at the higher energy values. Since this is the low q_3 sample with the bias being forwards, there are only events moving out of the low q_3 region. For the muon energy and angle it is consistent across all bins with slightly more percentage of events being shifted out at higher energies.

For the recoil energy, at higher energies there is a high percentage of events moving out of the low q_3 cut. Referring back to Figure 4.4, events in the upper corner of the low q_3 region right at the cut boundary are being heavily migrated out of the sample. Due to these bias effects, the delta portion of the events are especially being moved out of the sample.

Now if we consider the plot of the unfolded data we see the opposite effect. Every single bin is higher than the default case. The reason for this is that the migration out shown in the reconstructed plot is simultaneously changing the predicted acceptance correction, which will return that higher fraction of events to the unfolded distribution. Again the effects in the unfolded data are less than five percent except for the recoil energy bias at higher energies.

From the plots the difference between the adjusted data and the default can be found for every bin. The difference represents the error in the default case due to that adjustment. Taken together, they are the systematic uncertainty of the reconstructed MC prediction and the unfolded measurement. The first step was breaking down the differences into a table and then summing the differences in quadrature. The results for the MC are shown in the Table 6.1 and the unfolded data in the Table 6.2.

Bin Low	Bin High	Default	$\Delta_{E_{Recoil}}$	Δ_{E_μ}	Δ_{θ_μ}	$\sum_i(\Delta_i^2)$
0.00	0.02	359.31	0.873	8.22	3.02	8.80
0.02	0.04	383.44	1.79	12.35	4.01	13.11
0.04	0.06	255.29	2.69	12.59	3.67	13.38
0.06	0.08	166.11	2.85	10.31	3.48	11.25
0.08	0.10	113.29	3.68	7.46	2.72	8.75
0.10	0.12	88.82	4.08	6.07	1.90	7.56
0.12	0.14	78.89	3.57	3.56	1.05	5.15
0.14	0.16	75.36	4.09	3.14	1.54	5.38
0.16	0.18	70.76	4.80	2.45	0.97	5.48
0.18	0.20	62.69	7.31	2.72	1.44	7.93
0.20	0.22	46.68	9.20	2.17	0.95	9.50
0.22	0.24	30.21	9.91	1.63	1.14	10.11
0.24	0.26	12.58	7.64	0.76	0.77	7.72

Table 6.1: Breakdown of the difference per bin in the reconstructed MC for a bias forward in the recoil energy, muon energy and the recoil energy.

Bin Low	Bin High	Default	$|\Delta_{E_{Recoil}}|$	$|\Delta_{E_\mu}|$	$|\Delta_{\theta_\mu}|$	$|\sum_i(\Delta_i^2)|$
0.00	0.02	417.18	2.85	14.61	4.80	15.64
0.02	0.04	300.27	3.67	13.42	4.15	14.52
0.04	0.06	299.39	3.51	12.18	3.72	13.21
0.06	0.08	180.38	4.28	10.98	3.31	12.24
0.08	0.10	133.20	6.07	8.37	2.90	10.75
0.10	0.12	102.31	5.71	5.71	1.87	8.29
0.12	0.14	68.42	5.19	3.42	1.33	6.36
0.14	0.16	49.10	4.96	2.31	1.01	5.56
0.16	0.18	41.05	4.84	1.81	0.82	5.23
0.18	0.20	37.29	5.95	1.79	0.85	6.27
0.20	0.22	26.47	6.15	1.56	0.77	6.39
0.22	0.24	16.85	3.82	0.71	0.36	3.90
0.24	0.26	8.30	2.49	0.44	0.26	2.54

Table 6.2: Breakdown of the difference per bin in the unfolded Data for a bias forward in the recoil energy, muon energy and the recoil energy.

The columns in the table are as follows. The first and second columns define what bin is being looked at. Number three is the number of events in the bin for the default case. Four, five and six are the absolute value of the difference between the shifted plot and the nominal. Finally the very last column is the sum of columns four, five and six in quadrature in order to give an idea of what the error would be due to all three of these effects. To put it in other terms, the uncertainty on the measurement in the first bin is $\frac{15.64}{417.18}$ or 3.7 percent.

6.3 Acceptance and Background effects due to resolution effects.

To this point, resolution effects have been illustrated by biasing quantities by resolution sized amounts. The true effect that a resolution has is in both directions with a varying magnitude and will move events into and out of the low q_3 sample. The way the smearing is going to happen is it will tend to take locations with high concentrations and move them to lower concentrations. If a resolution was applied to a peaked distribution what would happen is the peak would flatten out, decreasing the peak and raising the edges.

The simulation already takes into account the predicted resolutions shown in the previous section, but we need an estimate for the uncertainty, how badly mis-modeled could the simulation be. The simulation already predicts that nearby kinematic samples have a worse resolution. This section looks at what happens if the resolution in our sample is much worse than what is being put into the reconstruction.

The muon angle requires special care because it is a 2D smearing of the polar and azimuth angles. The resolution is known to be 10mrads, but the main point is that this is the resolution in the muon's direction not the opening angle between the muon direction and the beam direction which has been called the muon angle throughout the paper. This means that this 10 mrads can be in any azimuth direction. The smearing that is done with this follows the procedure of applying a Gaussian with a sigma of 10mrads to change the opening angle or polar angle relative to the original muon direction. Then the new vector is spun around the old one in azimuth by an angle pulled from a uniform distribution between zero and two pi. This effectively smears the angle in a 3D space. Then the angle with respect to the beam direction is recalculated from this new vector.

In order to degrade the resolution in the recoil and muon energies, a value was drawn from a Gaussian distribution centered around 1.0 and serves as a scaling factor for these quantities. For the recoil energy the resolution is 50 percent, but for the QE part in the low q_3 is 56 percent. The resolution in the recoil energy was degraded by about five percent which is done by applying the gaussian with a sigma of 0.25. Similarly the resolution for the muon energy is 8% but for the low q_3 in the medium energy range the resolution is 11% . In order to degrade the resolution from 8% to 11% the sigma value in the gaussian was set to 0.075.

The study in this section is also looking primarily at the acceptance correction part of the unfolding process. The results of applying the 2-D resolution smearing in the muon angle and degrading the resolution in the recoil and muon energies are shown below in Figure 6.6.
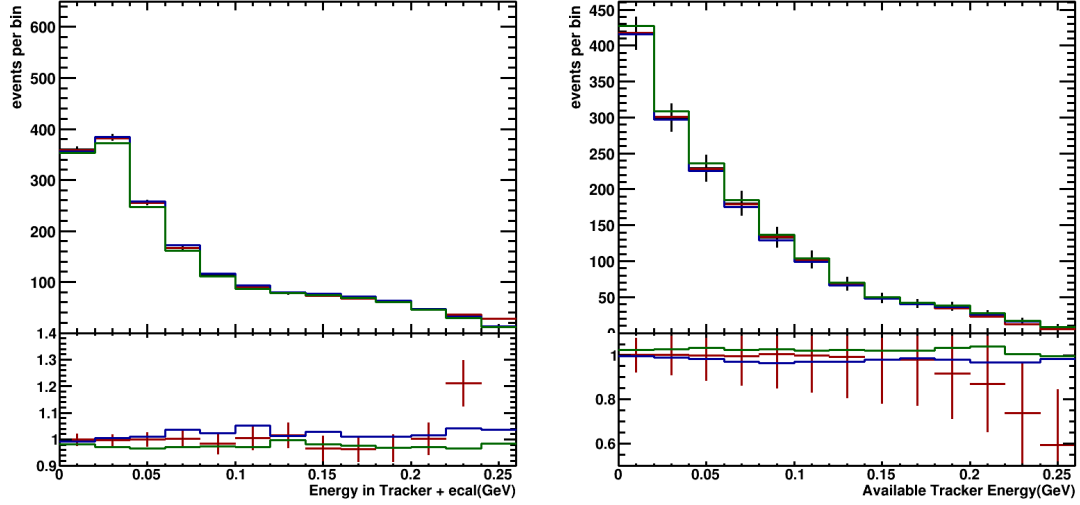
Figure 6.6: Reconstructed MC on the left and unfolded Data on the right after degrading the resolution one at a time for the recoil energy(red), muon energy(blue) and muon angle(green).

Again starting with the reconstructed MC plot, the effects seen are still less than five percent except for the recoil energy case at higher charged particle energies. Since this is a true degraded resolution events can be smeared both into and out of the low q_3 sample. The muon angle still is tending to move events out of the sample, which can be seen as the ratio is always less than one. The muon energy however tends to move events into the sample. With a degraded resolution in the recoil energy, both are seen. Some of the bins have events moving out of them while some of the events have events moving in. The biggest effect seen is events moving into the sample from the mid q_3 delta region.

The unfolded plot is everything that was seen in the reconstructed plot except inverted, because the primary effect is on the MC acceptance correction. After the unfolding there are more events with a change in the muon angle, fewer events in the muon energy and the recoil energy has a little of both. The bins in the recoil case that have more events in them are very hard to see, but if you look at the actual plot, not the ratio, in the low energy bins there is a small slice of red above the black line.

Again the difference per bin for these three effects can be put into a table in order to combine the results. For the MC Table 6.3 has the information for the resolution effects and Table 6.4 has the data numbers.

Bin	Bin	Default	$\Delta_{E_{Recoil}}$	Δ_{E_μ}	Δ_{θ_μ}	$\sum_i(\Delta_i^2)$
0.00	0.02	359.31	0.45	3.07	6.78	7.46
0.02	0.04	383.44	1.60	1.54	10.78	11.01
0.04	0.06	255.29	0.06	2.62	8.71	9.09
0.06	0.08	166.11	0.43	6.20	4.82	7.86
0.08	0.10	113.29	1.82	2.76	2.93	4.42
0.10	0.12	88.82	0.45	4.70	2.47	5.33
0.12	0.14	78.89	1.23	1.08	0.24	1.65
0.14	0.16	75.36	2.56	2.22	1.39	3.67
0.16	0.18	70.76	2.63	0.73	1.73	3.23
0.18	0.20	62.69	2.02	0.65	2.03	2.94
0.20	0.22	46.68	0.08	0.67	1.39	1.55
0.22	0.24	30.21	6.34	1.28	1.07	6.56
0.24	0.26	12.58	15.51	0.46	0.19	15.52

Table 6.3: Breakdown of the difference per bin between the nominal and the degraded resolution type for the recoil energy, muon energy and muon angle in the reconstructed MC.

Bin	Bin	Default	$|\Delta_{E_{Recoil}}|$	$|\Delta_{E_\mu}|$	$|\Delta_{\theta_\mu}|$	$|\sum_i(\Delta_i^2)|$
0.00	0.02	417.18	0.10	1.51	9.82	9.94
0.02	0.04	300.27	0.35	3.00	8.46	8.98
0.04	0.06	299.39	0.56	3.80	7.23	8.19
0.06	0.08	180.38	1.02	5.11	4.33	6.77
0.08	0.10	133.20	0.69	4.65	3.57	5.90
0.10	0.12	102.31	0.18	3.12	2.05	3.74
0.12	0.14	68.42	0.49	1.95	1.60	2.57
0.14	0.16	49.10	1.00	1.07	1.05	1.80
0.16	0.18	41.05	0.80	0.59	0.82	1.29
0.18	0.20	37.29	3.10	0.73	1.18	3.39
0.20	0.22	26.47	3.44	0.84	1.06	3.70
0.22	0.24	16.85	4.41	0.54	0.08	4.44
0.24	0.26	8.30	3.36	0.13	0.04	3.36

Table 6.4: Breakdown of the difference per bin between the nominal and the degraded resolution for the recoil energy, muon energy and the muon angle in the unfolded Data.

6.4 Total acceptance and background corrections

The four different tables now contain all of the information that is needed in order to find the error produced from all six effects. This is going to be done by summing in quadrature the two sets of three for both the data and the simulation. The results are shown in Table 6.5 for the MC and the Table 6.6 for the Data.

Bin Low	Bin High	$\sum_i(\Delta^2_{bias})$	$\sum_i(\Delta^2_{res})$	$\sum_i(\Delta^2_{tot})$
0.00	0.02	8.80	7.46	11.54
0.02	0.04	13.11	11.01	17.12
0.04	0.06	13.38	9.09	16.18
0.06	0.08	11.25	7.86	13.72
0.08	0.10	8.75	4.42	9.80
0.10	0.12	7.56	5.33	9.25
0.12	0.14	5.15	1.65	5.41
0.14	0.16	5.38	3.67	6.51
0.16	0.18	5.48	3.23	6.36
0.18	0.20	7.93	2.94	8.46
0.20	0.22	9.50	1.55	9.62
0.22	0.24	10.11	6.56	12.05
0.24	0.26	7.72	15.52	17.33

Table 6.5: Total sums of the differences per bin between the nominal and the shifted values in the reconstructed MC for all 6 effects due to the biases and degraded resolution in the recoil energy, muon energy and muon angle.

Bin Low	Bin High	$\sum_i(\Delta^2_{bias})$	$\sum_i(\Delta^2_{res})$	$\sum_i(\Delta^2_{tot})$
0.00	0.02	15.64	9.94	18.53
0.02	0.04	14.52	8.98	17.08
0.04	0.06	13.21	8.19	15.54
0.06	0.08	12.24	6.77	13.99
0.08	0.10	10.75	5.90	12.26
0.10	0.12	8.29	3.74	9.09
0.12	0.14	6.36	2.57	6.86
0.14	0.16	5.56	1.80	5.85
0.16	0.18	5.23	1.29	5.39
0.18	0.20	6.27	3.39	7.13
0.20	0.22	6.39	3.70	7.39
0.22	0.24	3.90	4.44	5.91
0.24	0.26	2.54	3.36	4.21

Table 6.6: Total sums of the differences per bin between the nominal and the shifted values, in the unfolded data for all 6 effects due to the biases and degraded resolution in the recoil energy, muon energy and muon angle.

The last column can then be substituted in for the error bars on the reconstructed MC and the unfolded data. With an uncertainty estimate, we can finally begin to interpret the disagreement between the data and the model. Is it significant or is it likely to be explained simply by energy and resolution uncertainties. The resulting figures are shown below in Figure 6.7.
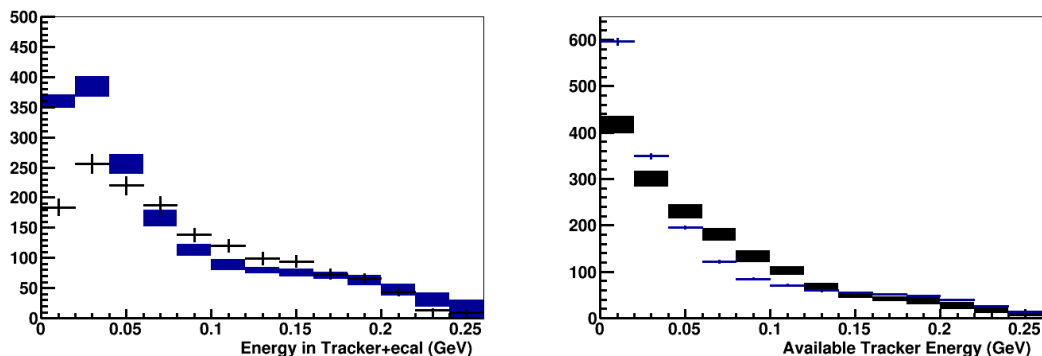
Figure 6.7: Reconstructed MC on the left and unfolded Data on the right with error bars produced from the effects of biasing and degrading the resolution in the recoil energy, muon energy and muon angle.

In both of the plots the blue is the MC and the black is the data. In both of the plots there is a huge problem at medium to low energies in this sample. A conclusion can be quickly drawn from these that something is missing. Even at the edge of this set of systematic uncertainties, the model does not describe the data.

6.5 Bias and resolution effects in the energy observable and migration matrix

Looking back at the flow diagram, if we change the quantity on the horizontal axis of the plot, the energy in the tracker + ECAL, then the MC events will shift within the plot. If this is changed then the migration matrix will change along with it, but the acceptance correction will not change. This means that with a change in the tracker energy + ECAL, what is going to be seen is only a change in the migration matrix.

The recoil energy shifted in the previous subsection is closely connected to the energy in the tracker quantity on the horizontal axis, the latter is the former less the neutron energy and with a different scaling factor. This means what can be done is a bias can be taken in both of these values and then the combined result in order to see the whole effect that this bias has. Figure 6.8 shows the result for a bias in these two values.
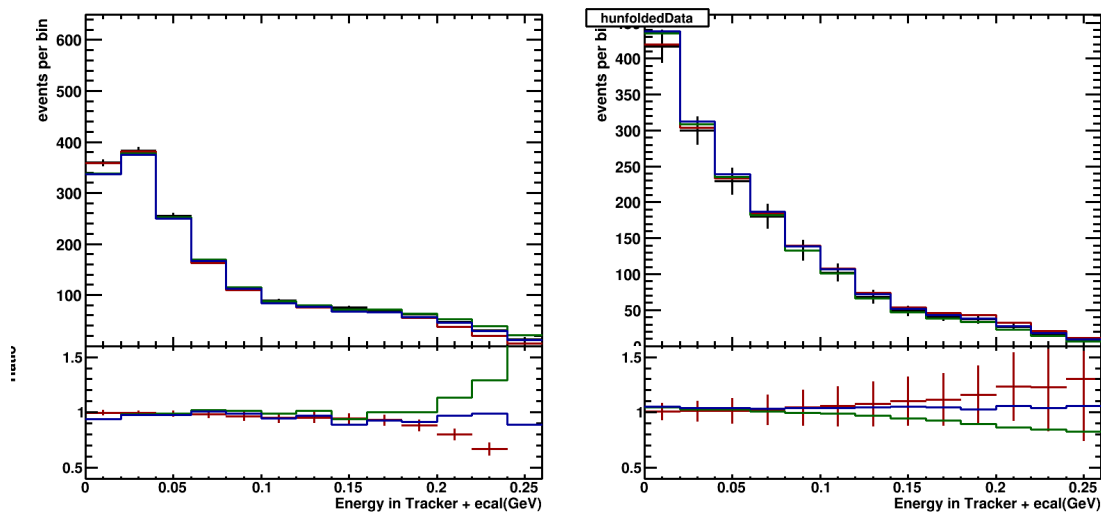
Figure 6.8: Bias forwards by five percent in the recoil energy, available tracker + ECAL energy and then both combined for the low q_3 cut. The colors correspond to the bias effect in the recoil energy for red, energy in tracker plus ecal for the green and combined for blue.

Looking at just the reco plot first, the big thing is that at higher energies the recoil energy moves events out of the low q_3 sample, biasing the Tracker + ECAL energy moves events from low to high energy such that the combination cancels out at the right hand edge of the plot. In the very first bin however the recoil energy does not produce a difference, where the a bias in the Tracker plus ECAL energy does.

The bias values for the recoil energy used to create Figure 6.7 are replaced by these differences. What this does, is it more accurately represents the correlated effect that a bias in the recoil energy has if the uncertainty is coming at the very beginning and effecting the energy in tracker value as well.

Moving along with the analysis strategy, the resolution in both of these values can be degraded the same way as the recoil energy was done previously. The result is shown in Figure 6.9.
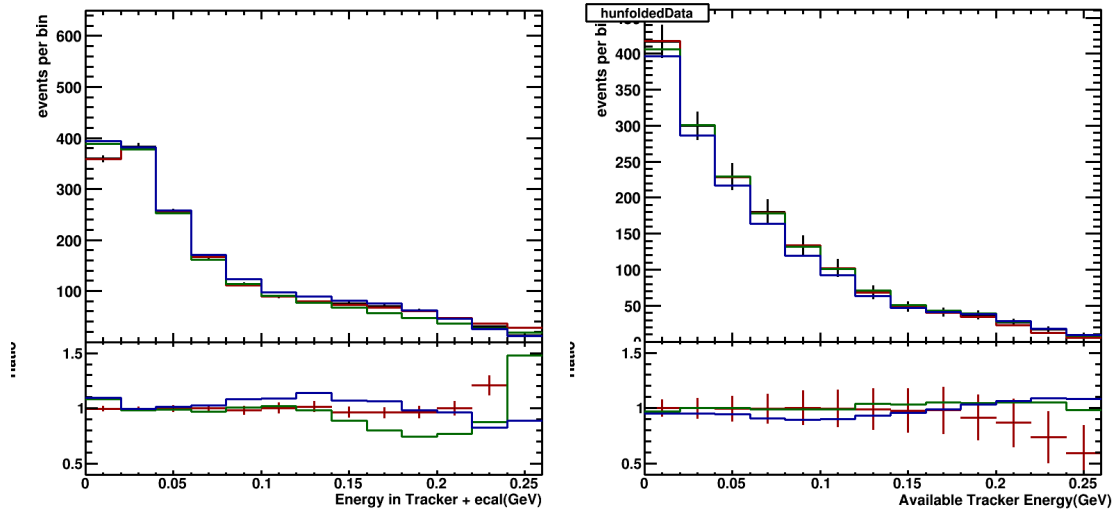
Figure 6.9: Degrading the resolution in the recoil energy, available tracker + ECAL energy and then both combined by five percent for the low q_3 cut. The left hand plot is for the reco and unfolded on the right. The colors correspond to the bias effect in the recoil energy for red, energy in tracker plus ecal for the green and combined for blue.

Looking at the reco plot for the first bin and the middle region of the plot the combined combination of the recoil energy and the tracker is higher than the recoil energy. This means that the error in the bins is higher.

Again the resolution quantity used to make Figure 6.7 can be replaced with the difference per bin seen from the combination of the recoil and tracker energy. The result of replacing both the bias and resolution numbers in Figure 6.7 produces Figure 6.10.
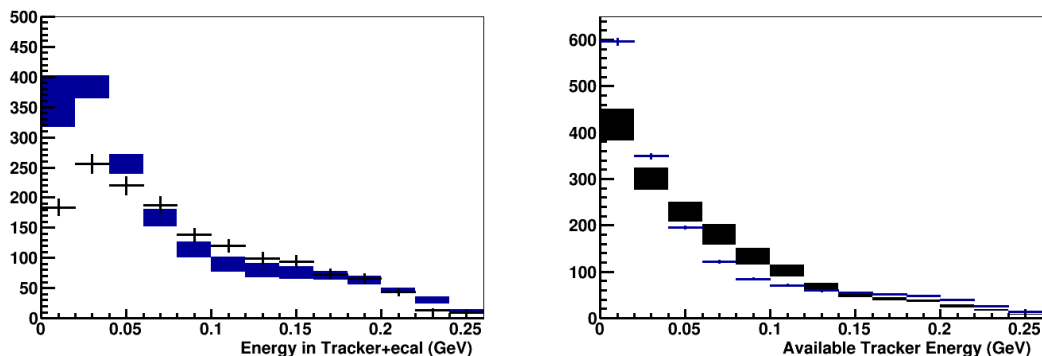
Figure 6.10: Reconstructed MC on the left and unfolded Data on the right with error bars that mimic Figure 7 but have the information for the recoil energy replaced by the information of the recoil and tracker energy values combined.

Comparing this to Figure 6.7, reveals that the error bars are increased at the lower and mid energies. At higher energies however, the error bars are essentially reduced to zero.

The take home message from this is the larger error bars means the significance of the discrepancy is not as large. But the simulation still doesn't describe the data at high significance.

6.6   Results of adding in RPA and MEC

Two of the proposed changes to the default model are to add a RPA suppression to the QE and to add a MEC (2p2h) process. The RPA will suppress the QE simulation at low energies and the MEC will enhance the rate especially in the middle of this sample. Separate from the physics they describe, there is a difference in how they are implemented in this analysis. The RPA effect can be applied as a reweighting of fully simulated quasi-elastic events, and can be illustrated for both the reco and the unfolded distributions. The MEC events have not been run through the detector simulation and are not like the existing events in the simulation. But the model does report the available charged particle energy, so they can be added to the predicted event rate and compared to the unfolded data distribution.

Using the model with the addition of the RPA and MEC, the error band was recomputed using the same methods as the samples without RPA. The resulting error band plots are shown for the reco plot in Figure 6.11 and the truth plot in Figure 6.12 along side the model without these effects.
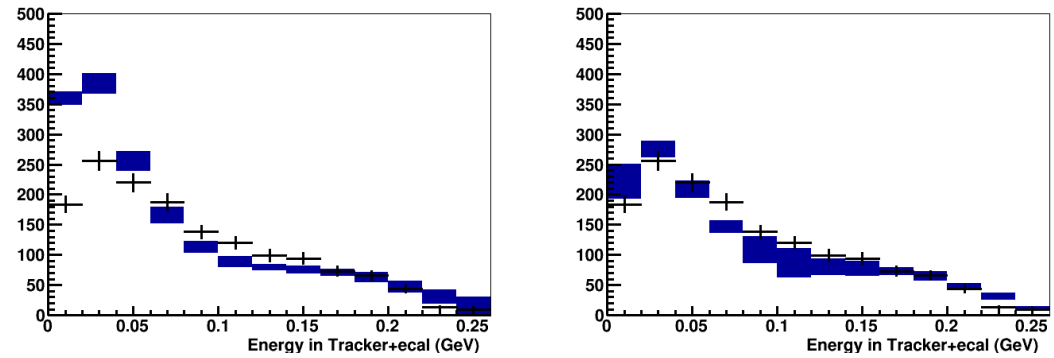

Figure 6.11: Reconstructed plots of the Data and MC for the model without(left) and with(right)the addition of the RPA, but not the MEC effect, including error bars produced from the biasing and degrading resolution in the recoil energy, muon energy and muon angle.
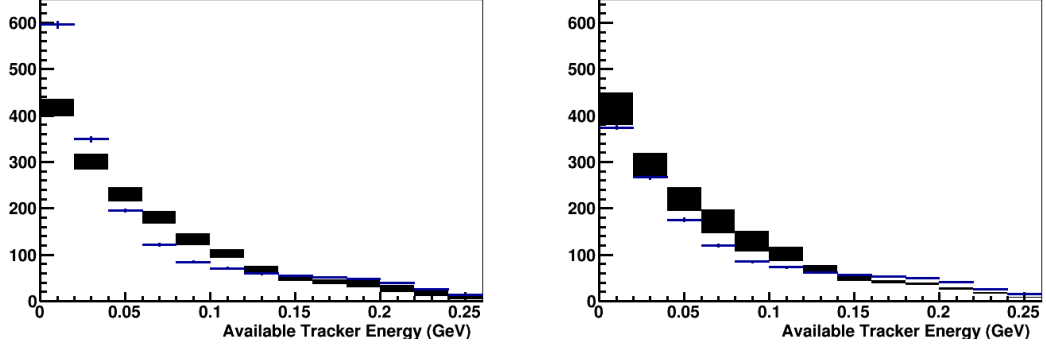
Figure 6.12: Truth plots of the Data and MC for the model without(left) and with(right)the addition of the RPA and MEC effect including error bars produced from the biasing and degrading resolution in the recoil energy, muon energy and muon angle.

First, the magnitude of the RPA effect takes care of almost all the discrepancy in the lowest two bins. However, in the reco plot in the dip region between 0.1 and 0.15 GeV the under prediction of the rate persists. The shape of the model is very different than the data.

Going now to the unfolded plots, Figure 6.12, the discrepancy at the lower energies is still being corrected for through the inclusion of the RPA. The middle region where there was a discrepancy in the reco plot now better describes the data when the MEC is put in.

6.7 χ^2 calculation

In order to determine which agreement is better a goodness of fit parameter can be calculated. For this case the goodness of fit parameter is going to be χ^2 which is defined by:

$$\chi^2 = \sum_n \frac{(D_n - S_n)^2}{\sigma_n^2} \tag{6.1}$$

χ^2 is essentially the ratio of the difference between the data and MC and the combined uncertainty and then the whole thing squared. This is done bin to bin and then

summed over the entire sample. This was done for the unfolded distributions for the default and all three cases for which the error bars where produced. The results are shown in Table 6.7.

Default	Uncorrelated	Correlated	RPA and MEC
156.43	86.77	70.12	36.00

Table 6.7: χ^2 parameter for the default data, with the addition of the uncorrelated systematic error bar, the correlated one and the full error bar with the inclusion of the RPA and MEC.

Since these samples are normalized to a larger sample. The two samples where renormalized in order to produce the best comparison between the data and MC.

When calculating these values the uncertainty was taken to be the systematic uncertainty, statistical uncertainty and the uncertainty on the true MC summed in quadrature. For the default case there is no systematic uncertainty to be looked at so it is only the statistical and the uncertainty on the MC. The statistical uncertainty is the square root of the number of events in the bin.

These numbers back up the results seen in the previous section. With the addition of the systematic uncertainties the agreement is improved, then the uncorrelated error bars increase the agreement further. Finally when the RPA and the MEC are added in the agreement between the data and the MC is the best.

Chapter 7

Conclusion and Discussion

With the inclusion of the RPA and MEC the general agreement between the data and the MC is improved.

The MEC events in particular are predicted to have a large fraction with two proton final states, compared to QE and Delta interactions. Scanning events at the kinematics where these should stand out yields an ambiguous result.

When the scanned results for the MC and the data where compared in the dip region, there was a surplus of events in the zero hadron category for the MC. With the inclusion of the RPA there is a supression of the cross section at low energies which in turn reduces the number of zero hadron events in the MC.

A study was also done with the energy in tracker + ECAL. The known systematic uncertainties in muon energy, muon angle and hadronic energy were applied to the MC. Through an unfolding process these uncertainties where then propagated to the unfolded data distribution, a precursor to a cross section measurement.

These uncertainties are small enough that the default MC clearly does not agree with the data. When the MC was modified to include the RPA and MEC, the agreement between data and the MC are just slightly outside of the systematic error band.

References

[1] Juan Nieves. Approximate methods for nuclei (ii) - part 1 of 3. Presentation at NuSTEC, October 26, 2014.

[2] F. Sanchez R. Gran, J. Nieves and M. J. Vicente Vacas. Neutrino-nucleus quasielastic and 2p2h interactions up to 10 gev. *arXiv.org*, arXiv:1307.8105 [hep-ph].

[3] J. Nieves, J. E. Amaro, and M. Valverde. Inclusive quasielastic charged-current neutrino-nucleus reactions. *Phys. Rev. C*, 70:055503, Nov 2004.

[4] Carrie McGivern. The story of a detector, and its daq. Presentation at MINERvA 101, slide number 3, June 16, 2014.

[5] S. Kopp. Proceedings of the 2005 particle accelerator conference. Knoxville, TN arxiv:[physics]0508001; A.G. Abramov et al., Nucl. Instr. and Meth. A485 (2002) 209.

[6] Design, calibration, and performance of tthe minerva detector. *arXiv.org*, arXiv:1305.5199 [physics.ins-det].

[7] Arachne - A web-based event viewer for MINERvA N. Tagg et al. [MINERvA Collaboration]. arXiv:1111.5315 [hep-ex] Nucl. Instrum. Meth. 676, 44 (2012).

[8] C. Andreopoulos et al. The GENIE Neutrino Monte Carlo Generator. *Nucl. Instrum. Meth.*, A614:87–104, 2010, 0905.2517.

[9] Nuclear Inst. and Methods in Physics Research, A 789 (2015), pp. 28-42.

[10] L. Fields et al. Measurement of muon antineutrino quasielastic scattering on a hydrocarbon target at $e_\nu \approx 3.5gev$. *Phys. Rev. Lett.*, 111:022501, Jul 2013.

[11] L. Fields et al. Measurement of muon neutrino quasielastic scattering on a hydro-carbon target at $e_\nu \approx 3.5 gev$. *Phys. Rev. Lett.*, 111:022502, Jul 2013.