

Benchmarking of neutrino energy reconstruction methods using electron-deuterium scattering data

by

Anjali Nambrath

Submitted to the Department of Physics
in partial fulfillment of the requirements for the degree of

Bachelor of Science in Physics

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2021

© Anjali Nambrath, MMXXI. All rights reserved.

The author hereby grants to MIT permission to reproduce and to
distribute publicly paper and electronic copies of this thesis document
in whole or in part in any medium now known or hereafter created.

Author

Department of Physics

May 14, 2021

Certified by

Or Hen

Assistant Professor of Physics

Thesis Supervisor

Accepted by

Deepto Chakrabarty

Associate Department Head, Department of Physics

Benchmarking of neutrino energy reconstruction methods using electron-deuterium scattering data

by

Anjali Nambrath

Submitted to the Department of Physics
on May 14, 2021, in partial fulfillment of the
requirements for the degree of
Bachelor of Science in Physics

Abstract

Neutrinos exist in one of three types or “flavors” (ν_e , ν_μ , or ν_τ) which oscillate from one to another when propagating through space. Understanding the detailed features of this phenomenon can provide new insight into the nature of our universe. Neutrinos oscillate as a function of their propagation distance divided by their energy (L/E), so experiments extract oscillation parameters by measuring the neutrino energy distribution at different locations. As accelerator-based oscillation experiments cannot directly measure E , their interpretation relies heavily on phenomenological models of neutrino-nucleus interactions to infer E . Here, we exploit the similarity of electron- and neutrino-nucleus interactions, and use electron-scattering data with known beam energies to test energy reconstruction methods and interaction models. In particular, we analyze 5 GeV electron-deuteron scattering data measured using the large-acceptance CLAS detector at Jefferson Lab. Despite deuterium having a simple nucleus that is relatively well-understood, we find that large nuclear effects can bias the electron beam-energy reconstruction. We thus develop a set of event selection cuts that strongly suppress nuclear effects and propose a viable path for precision measurements. We further compare our data with the theoretical predictions of the GENIE event generator, commonly used by most neutrino experiments based in the US, to validate simulation models used for interpreting lepton-nucleus scattering reactions.

Thesis Supervisor: Or Hen
Title: Assistant Professor of Physics

Acknowledgments

I would like to thank Professor Or Hen for this incredible research opportunity, as well as his support over the past four years. I would also like to thank Dr. Dien Nguyen for her lending her expertise to this project and for her constant guidance and patience. I'm grateful to Afroditi Papadopoulou for being generous with her time and advice, and Professor Lawrence Weinstein for his insights. Thanks to Professor Axel Schmidt and Dr. Adi Ashkenazi for their help and suggestions as well. Finally, I'd like to thank all the members of the Hen Lab; my time in the group has been a key part of my undergraduate experience.

Contents

1 Neutrino physics	19
1.1 The neutrino	19
1.2 Neutrino oscillation	21
1.2.1 Neutrino masses and mixing	21
1.2.2 Oscillation formalism	23
1.2.3 Current neutrino oscillation results	24
2 Neutrino energy reconstruction	29
2.1 Energy reconstruction methods	30
2.2 Improving neutrino event generators	33
2.3 Applying electron scattering data	36
2.3.1 The e4ν effort	38
2.3.2 Focus on electron-deuterium data	38
3 Energy reconstruction analysis on CLAS data	41
3.1 The CLAS detector	41
3.2 Event selection	42
3.2.1 Vertex determination	43
3.2.2 Weighting events	44
3.2.3 Fiducial cuts	45
3.2.4 Acceptance maps and correction	47
3.2.5 Background subtraction	51
3.2.6 Radiative photons	56

3.3	Results	62
4	Energy reconstruction analysis on GENIE simulation	71
4.1	About GENIE	71
4.2	Event selection	73
4.3	Energy reconstruction results	73
4.3.1	Background subtraction validation	75
4.3.2	Background processes	77
4.4	Comparison between data and GENIE	78
5	Conclusions	81

List of Figures

2-4	The impact on extracted oscillation parameters if the event generator does not capture the underlying physics. The grey regions show the 1σ , 2σ , and 3σ confidence regions for the extracted parameter values when the input data was produced with the event generator GIBUU, and also analyzed with GIBUU. The red, green, and blue lines show the same regions for analysis of the input GIBUU data with the GENIE event generator. The red dot indicates the input parameter values, and the black dot shows the extracted best fit parameters. Taken from [32].	35
2-5	Two-body current interactions. The upper left shows isobar configurations (IC), where the virtual photon is absorbed on a nucleon, which is excited to a Δ than then de-excites via $\Delta N \rightarrow NN$. In the upper right, meson exchange currents (MEC) show the virtual photon being absorbed on a meson exchanged between two nucleons, which are both knocked out. In the bottom left, the virtual photon is absorbed by one nucleon in a short-range correlated (SRC) pair, and its partner nucleon is also knocked out. In the bottom right, the virtual photon is absorbed by a nucleon, which rescatters via final state interactions (FSI). Figure from [13].	37
3-1	The CLAS detector and its main components.	42
3-2	The electron vertex and proton vertex. The liquid deuterium target is located at $z \sim 30$ cm, so the yellow band in the middle of the plot contains the $D(e, e'p)$ events.	44
3-3	The difference between the electron vertex and proton vertex, along with the applied cut. We select events where the difference between the two values is less than 1 cm.	45
3-4	Effect of fiducial cuts on the θ and ϕ distribution for protons. On the left is the distribution before cuts are applied, and on the right is the distribution after the cuts. The six sectors of the CLAS detector are clearly visible.	47

3-5	E_{cal} and E_{QE} energy reconstruction distributions for $D(e, e'p)X$ events. Here we have applied fiducial cuts to the electron and proton, but none of the other event selection steps.	48
3-6	The $\theta\text{-}\phi$ distribution for electrons. The red events are those that pass the fiducial and acceptance cuts (> 0.7), overlaid on top of the distribu- tion after just the fiducial cut. Events with electrons at high scattering angles θ are eliminated, as are events with very forward-angle electrons. This is due to the process by which the maps were generated, but have little effect on the final comparison in this study.	49
3-7	The $\theta\text{-}p_{e'}$ distribution for electrons. The red events are those that pass the fiducial and acceptance cuts (> 0.7), overlaid on top of the distribution after just the fiducial cut. Events with electrons at low momenta ($<\sim 800$ MeV/c) are eliminated. This is likely due to thresh- old effects in the CLAS simulation, but this has little impact on the final comparison in this study.	49
3-8	The reconstructed E_{cal} and E_{QE} energy distributions after applying the acceptance maps and correction. The size of the peak has been reduced as compared to Figure 3-5.	50
3-9	Energy reconstruction distributions for E_{cal} (left) and E_{QE} (right), af- ter the fiducial step and after the acceptance step. Note the substantial reduction in the peak and the shift of the left edge of the distributions. This is due to the acceptance cut significantly affecting the electron scattering phase space, and the rejection of low-momentum electrons.	51
3-10	The distribution of the angle between the electron and the photon, versus the difference of their azimuthal angles (counts are shown on a logarithmic scale). The box in red represents the cut used to select radiation photons.	52
3-11	Proton multiplicity. Most events have one proton, and a small number have two protons. Higher multiplicities are low enough to be ignored.	53

3-12 Pion multiplicity. Most events have between zero to three charged pions. Higher multiplicities can be ignored.	53
3-13 The diagrams for (a) π rotation and (b) π and proton rotation around the momentum transfer \vec{q} . By doing a random rotation repeatedly, we can calculate the probability of a multi-hadron state being erroneously detected as an event with a simpler topology.	54
3-14 Energy reconstruction distributions for events with detected pions. The black curve shows E_{cal} for events with one detected pion, the blue curve for two pions, and the red for three pions. The 5 GeV peak all but vanishes by the three pion case. These are major components of the background and the incorrectly reconstructed tail.	55
3-15 Energy reconstruction distributions for E_{cal} (left) and E_{QE} (right), after background subtraction. Note the substantial reduction in the size of the tail after the subtraction step. Nevertheless, a non-negligible tail remains in both cases.	56
3-16 The reconstructed E_{cal} and E_{QE} energy distributions after the background subtraction step. The background tail has been reduced. . . .	57
3-17 E_{miss} and p_{miss} for ν between 0 and 1.5 GeV. Both quantities are near zero, indicating that radiative photons are not present in this regime of the reconstructed energy distribution.	58
3-18 E_{miss} and p_{miss} for ν between 1.5 and 3.5 GeV. Both quantities are mostly near zero, indicating that radiative photons are not meaningfully present in this regime of the reconstructed energy distribution. There is however a tail where $E_{miss} \approx p_{miss}$, but this does not contain the majority of events.	58
3-19 E_{miss} and p_{miss} for ν between 3.5 and 4.5 GeV. Here, most events have $E_{miss} \approx p_{miss}$, suggesting that radiative photons are present in these events.	59

3-20 $\theta_{p_{miss}}$ and ν . Events with very low $\theta_{p_{miss}}$ are concentrated at high values of ν , particularly around $\nu = 4$. Thus suggests that particles traveling down the beam line are carrying away most of the energy missing from the reconstruction.	60
3-21 $\theta_{p_{miss}}$ and E_{cal} . Similarly to Figure 3-20, for near-zero values of $\theta_{p_{miss}}$ the energy is reconstructed very incorrectly. This suggests that radiative photons are contributing to the far end of the tail.	60
3-22 Energy reconstruction distributions for E_{cal} (left) and E_{QE} (right), after the $\theta_{p_{miss}}$ cut. Note the reduction in the tail, especially for lower E_{reco} values. This successfully eliminates radiative photon events that are irrelevant for this study.	61
3-23 The reconstructed E_{cal} and E_{QE} energy distributions after the $\theta_{p_{miss}}$ cut. The background tail has been reduced, and the shapes and sizes are fairly similar for both methods.	61
3-24 E_{cal} with a fit. The black curve is the (more finely-binned than before) E_{cal} distribution. In blue is the fit, which is composed of a Gaussian centered at the 5 GeV peak and a polynomial.	62
3-25 E_{QE} with a fit. The black curve is the (more finely-binned than before) E_{QE} distribution. In blue is the fit, which is composed of a Gaussian centered at the 5 GeV peak and a polynomial.	63
3-26 E_{cal} reconstruction distributions over different slices in the transverse missing momentum, p_T . For low p_T , most events are quasielastic and reconstructed properly. The background subtraction procedure reduced, but did not remove, events with higher p_T	64
3-27 E_{QE} reconstruction distributions over different slices in the transverse missing momentum, p_T . For low p_T , most events are quasielastic and reconstructed properly. As with E_{cal} , the background subtraction procedure reduced, but did not remove, events with higher p_T	65

3-28 Transverse variables, diagrammed for a neutrino event. The quantities \vec{p}_T^ℓ and \vec{p}_T^p are the transverse momenta of the scattered lepton and proton respectively. Their sum is \vec{p}_T (labeled on the right as δp_T), the transverse missing momentum. Also in blue on the right is ϕ_T (labeled $\delta\phi_T$), the measure of coplanarity between the scattered lepton and proton. The last variable, α_T (labeled $\delta\alpha_T$) is the angle of the recoil system projected onto the transverse plane.	66
3-29 Transverse missing momentum distribution. Events with low p_T were generally reconstructed correctly.	67
3-30 ϕ_T distribution. Most events have a small value for ϕ_T , indicating a high degree of coplanarity between the scattered electron and proton. The bump-like behavior is due to the six sectors of CLAS and the imperfect acceptance near their edges.	67
3-31 α_T distribution. If we had a purely quasielastic sample with minimal nuclear effects, this would be isotropic. This is not quite the case here.	68
3-32 2D plot of p_T and E_{cal} (logarithmic event counts). Most of the events that are reconstructed correctly ($E_{cal} \approx 5$) have $p_T < 0.2$ GeV/c. The edge on the left side of the plot is due to the θ_{miss} cut.	68
3-33 2D plot of ϕ_T and E_{cal} (logarithmic event counts). Most of the events that are reconstructed correctly ($E_{cal} \approx 5$) have $\phi_T < 30^\circ$	69
3-34 2D plot of α_T and E_{cal} (logarithmic event counts). The band across the top of the plot corresponds to events that correctly reconstructed. For these events, α_T is isotropic. However, for the events in the background (below the band), a similar isotropy is not observed. Most of these events have high values of α_T , indicating that the deuteron's neutron had a large backward recoil angle.	70
4-1 E_{cal} reconstruction for $D(e, e'p)$ events generated by GENIE using the G2018 configuration. A clear peak is visible at the beam energy of 5 GeV, but a substantial tail remains.	74

4-2	E_{QE} reconstruction distribution on GENIE events, using the deuterium-specific formula. A clear peak is visible, but as with the E_{cal} method, a large tail remains.	74
4-3	A comparison of the E_{cal} and E_{QE} methods of reconstruction on GENIE events. As with the CLAS data, there is a slight difference in the shape of the tails, but the sizes remain similar.	75
4-4	1p1 π subtraction validation. The black curve is the combined sample, with 1p0 π and 1p1 π events. The blue is the result of running the background subtraction process on the whole sample. The red is the true 1p0 π sample from the original file. The blue and the red are quite close, suggesting that the subtraction process is working well.	76
4-5	1p2 π subtraction validation. The black curve is the combined sample. The blue is the result of the background subtraction process. The red is the true 1p0 π . The blue and the red are very similar, validating the subtraction process.	76
4-6	1p3 π subtraction validation. The subtracted and true 1p0 π samples (blue and red) are similar, validating this subtraction method.	77
4-7	Breakdown of events in E_{cal} energy reconstruction distribution by interaction channel. The total distribution is shown in black. The quasielastic events are shown in blue, and almost entirely located in the correctly-reconstructed peak at 5 GeV. The events in red are resonant pion events, and these are generally reconstructed incorrectly. The green events are for deep-inelastic scattering interactions, also not correctly reconstructed.	78
4-8	Breakdown of events in E_{QE} energy reconstruction distribution by interaction channel. The total distribution is shown in black. The quasielastic events. shown in blue, are almost all in the 5 GeV peak. The resonant pion events (red) and deep-inelastic scattering events (green) are mostly not correctly reconstructed.	79

4-9 Comparison between E_{cal} distributions for CLAS data (blue) and GE-NIE events (red).	79
4-10 Comparison between E_{QE} distributions for CLAS data (blue) and GE-NIE events (red).	80

List of Tables

1.1 Three-neutrino oscillation parameters after a fit to global data. Taken from Ref [25].	27
--	----

Chapter 1

Neutrino physics

*I have done a terrible thing,
I have postulated a particle that
cannot be detected.*

Wolfgang Pauli

Neutrino oscillation, first observed in 1995, is one of the few phenomena through which it is possible to experimentally probe physics beyond the Standard Model. The study presented in this thesis deals with analysis techniques which have the potential to improve the precision of forthcoming neutrino oscillation experiments. This chapter presents a discussion of neutrino physics and neutrino oscillation.

1.1 The neutrino

The existence of the neutrino was posited by Wolfgang Pauli in 1930 to explain the continuous energy spectrum observed in nuclear beta decays, over twenty-five years before the particle was eventually detected. The continuous spectrum was suggestive of a three-body process, where the decay products are the emitted beta particle, the recoiling nucleus, and the proposed neutrino. The addition of the neutrino also solved the problem of energy conservation in beta decay measurements. This was formalized by Enrico Fermi in 1934 in his theory of beta decay, in which a neutron decays into

a proton, an electron, and the then-hypothetical neutrino.

The neutrino was eventually detected in 1956 in the Cowan-Reines experiment. The delay between Fermi’s theoretical work and the observation can be ascribed to the rarity of neutrino interactions. This, along with other properties of neutrinos, was confirmed in the decades following their initial discovery. Today we know neutrinos are extremely abundant, and produced from a variety of sources in the universe – decay processes within the Earth, solar fusion, atmospheric cosmic rays events, and extragalactic sources. Neutrinos interact with other particles via the weak force, but these interactions seldom occur. This is one of the difficulties inherent in neutrino detection.

The weak force produces neutrinos in one of three states called “flavors,” each corresponding to a different lepton. Thus the three types of neutrino are the electron neutrino (ν_e), muon neutrino (ν_μ), and tau neutrino (ν_τ). When a neutrino interacts with another particle, conservation of flavor dictates that the corresponding lepton is produced. Thus when an muon neutrino interacts with a neutron, for example, a muon and a proton are possible products. These types of interactions, called charged current interactions, are mediated by the electrically charged W^+ and W^- bosons. The boson induces either the absorption or emission of a charged lepton, and for each neutrino flavor there is only one such lepton with which these interactions occur. Neutrinos can also interact with other particles via neutral current interactions, mediated by the neutral Z boson. The Z boson-mediated interactions include, for example, the elastic scattering of neutrinos off of matter. [16]

Conservation of lepton number indicates that the products of a beta decay are an electron and an antineutrino, rather than an electron and a neutrino. As neutrinos are neutrally charged, it is not easy to determine how to distinguish between neutrinos and their antiparticles. Experiments studying the helicity of the neutrino uncovered the answer. All neutrinos are left-handed, with spins anti-parallel to their motion, and all antineutrinos are right-handed, with spins parallel to their motion. Therefore in total, there are three flavors of neutrinos, each with a corresponding antineutrino. [26]

1.2 Neutrino oscillation

Neutrinos were integrated into the Standard Model as massless particles. The Standard Model (SM), a set of quantum field theories developed over the 1960s and 1970s, provides a precise description of the fundamental particles and interactions that constitute ordinary matter in our universe. It explains the existence of the hundreds of known particles via an astoundingly simple underlying scheme. The existence of particle masses is explained by the couplings to the Higgs field in this framework; this was verified by the discovery of the Higgs boson [21]. The SM has passed numerous experimental tests, and many of its predictions have been tested to extremely high precision.

However, it is now abundantly clear that the SM has some limitations. For example, it does not offer a way to predict the values of particles' masses. Cosmological measurements also suggest that 95% of the universe is not made of SM particles. The SM also offers no explanations for the observed matter-antimatter asymmetry in the universe. Thus despite its successes, there remain a host of unexplained exceptions to SM physics, motivating the current rich research in phenomena beyond the Standard Model (BSM).

Another such exception was found via the discovery of neutrino oscillations, which indicated that neutrinos, contrary to all expectations, are not massless. Two groups, Super-Kamiokande [28] and SNO [9], independently discovered that neutrinos can change flavor – a discovery that consequently won the 2015 Nobel Prize. Specifically, these experiments found that a neutrino created with a particular flavor can later be measured to have a different flavor, with a probability that oscillates as the neutrino propagates through space (thus giving the phenomenon its name). This is completely unexpected behavior for massless particles.

1.2.1 Neutrino masses and mixing

To understand more deeply why oscillation is a consequence of neutrino mass, a discussion of neutrino mixing is necessary.

The SM can be extended to endow neutrinos with mass by introducing right-handed neutrino states. This is because fermion mass terms via coupling to the Higgs field are nonzero only if both chiral states exist. It is worth noting here that there are two possible mechanisms here. If the neutrino exists in left- and right-handed forms, it would be a Dirac fermion. As noted earlier though, it appears that all neutrinos are left-handed and all antineutrinos are right-handed. A right-handed neutrino has never been observed. This turns out to not be strictly necessary – the other possibility is that the right-handed neutrino is the antiparticle of the left-handed neutrino, as has been observed. This would make neutrinos Majorana fermions. However, a consequence of this is neutrinoless double beta decay, which has also not yet been observed. The question of whether neutrinos are Dirac or Majorana fermions is one of the biggest open questions in neutrino physics today. However, the difference between these is not terribly consequential for neutrino oscillation and will be ignored hereafter [38].

The existence of neutrino mass has some interesting consequences, as the mass eigenstates (ν_1, ν_2, ν_3) are not the same as the flavor eigenstates $(\nu_e, \nu_\mu, \nu_\tau)$. Rather, each flavor eigenstate is a superposition of the mass eigenstates. This is called neutrino mixing. Neutrino mixing is key to neutrino oscillation. As a neutrino of a certain flavor propagates through space, its component mass eigenstates evolve at different rates and pick up different phases. This leads to a nonzero probability that when the neutrino interacts, it is in a different flavor superposition than when it was produced.

The relationship between the mass and flavor eigenstates is described by a mixing matrix called the Pontecorvo-Maki-Nakagawa-Sakata (PMNS) matrix, as shown below.

$$\begin{pmatrix} \nu_e \\ \nu_\mu \\ \nu_\tau \end{pmatrix} = \begin{pmatrix} U_{e1} & U_{e2} & U_{e3} \\ U_{\mu 1} & U_{\mu 2} & U_{\mu 3} \\ U_{\tau 1} & U_{\tau 2} & U_{\tau 3} \end{pmatrix} \begin{pmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \end{pmatrix} \quad (1.1)$$

The vector on the left represents a neutrino expressed in its flavor-eigenstate basis, and the vector on the right represents a neutrino in its mass-eigenstate basis. The matrix on the right is the PMNS matrix. The PMNS matrix is usually parametrized with four

free parameters: three mixing angles (θ_{12} , θ_{13} , and θ_{23}) and a phase angle δ_{CP} related to charge-parity violation and the differences between neutrinos and antineutrinos. This parametrization is written as

$$U_{PMNS} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix} \begin{pmatrix} c_{13} & 0 & s_{13}e^{-i\delta_{CP}} \\ 0 & 1 & 0 \\ -s_{13}e^{i\delta_{CP}} & 0 & c_{13} \end{pmatrix} \begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad (1.2)$$

where $c_{ij} = \cos \theta_{ij}$ and $s_{ij} = \sin \theta_{ij}$ [29]. Having established neutrino mixing, it is now possible to better understand the formalism of neutrino oscillation.

1.2.2 Oscillation formalism

We can make use of some approximations to derive the general form for neutrino oscillation. After traveling a distance L or time t , a neutrino originally produced in flavor α can be described as follows.

$$|\nu_\alpha(t)\rangle = \sum_{j=1}^n U_{\alpha j}^* |\nu_j(t)\rangle \quad (1.3)$$

where $U_{\alpha j}$ is an element of the PMNS matrix. The neutrino can be approximated as a free particle and thus a plane wave: $|\nu_j(t)\rangle = e^{-iE_j t} |\nu_j(0)\rangle$. Since neutrinos are relativistic, we have (using natural units)

$$E_j = \sqrt{\mathbf{p}_j^2 + m_j^2} \approx \mathbf{p} + \frac{m_j^2}{2E} \approx \mathbf{p} + \frac{m_j^2}{2|\mathbf{p}|}, \quad (1.4)$$

where the last step comes from the ultrarelativistic approximation $\mathbf{p} \gg m$. We can also write

$$E_i(\mathbf{p}) - E_j(\mathbf{p}) \approx \frac{m_i^2 - m_j^2}{2|\mathbf{p}|} + \mathcal{O}(m^4) \quad (1.5)$$

Thus we can write the transition probability between ν_a and ν_β as

$$P_{\nu_a \rightarrow \nu_b} = |\langle \nu_b | \nu_a(t) \rangle|^2 \approx \sum_{i,j} U_{\alpha i}^* U_{\beta i} U_{\alpha j} U_{\beta j}^* \exp\left(-i \frac{\Delta m_{ji}^2 L}{2|\mathbf{p}|}\right) \quad (1.6)$$

where $\Delta m_{ij}^2 = m_i^2 - m_j^2$ and we have approximated the distance $L \approx t - t_0$.

Two-flavor oscillation is the simplest case of oscillation and is an illuminating example to see how the oscillation probability depends on other quantities. In this case, we look at the limited situation where only two flavor states are possible. In this case, there is only one mixing angle θ and only one mass difference Δm^2 . For the PMNS matrix, we have

$$U_{PMNS} = \begin{pmatrix} \cos \theta & \sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \quad (1.7)$$

The probability of a neutrino ν_α produced with energy E_ν oscillating into a neutrino ν_β after propagating a distance L is then given by

$$P_{\nu_\alpha \rightarrow \nu_\beta}(E, L) = \sin^2 2\theta \sin^2 \left(\frac{\Delta m^2 L}{4E_\nu} \right) \quad (1.8)$$

From this equation, we can see that the amplitude of the probability oscillations depends on the mixing angle θ , and the frequency of the oscillations depends on the quantities Δm^2 and L/E_ν . This approximation is quite helpful. In fact, it is used in discussion of atmospheric and solar neutrinos. Atmospheric oscillation involves the $\nu_\mu \rightarrow \nu_\tau$ channel, while solar oscillation deals with $\nu_e \rightarrow \nu_\mu$. It is possible to design experiments that are preferentially sensitive to these specific channels of neutrino oscillation. Current knowledge of the oscillation parameters (discussed in the following section and summarized in Table 1.1) suggest that θ_{13} is quite small, making this two-flavor approximation applicable.

In any case, it is clear that neutrino oscillation requires nonzero neutrino masses. Otherwise, the mass splittings Δm^2 are zero and oscillation does not occur.

1.2.3 Current neutrino oscillation results

In the expressions above for the oscillation probabilities, θ and Δm^2 are the governing parameters that we are interested in measuring. This leaves L and E_ν as tuneable parameters for oscillation experiments. In direct measurements of atmospheric or

solar neutrinos, it is impossible to control these values given that the neutrinos are generated by natural sources. Instead, we can use reactor-based or accelerator-based experiments to probe the same oscillation channels [41]. Reactor-based neutrino experiments can probe the solar oscillation regime. Neutrino beams generated at particle accelerator facilities offer the most control over the neutrino energy and distance traveled, and can be constructed to probe the atmospheric regime to give measurements of θ_{23} and Δm_{23}^2 .

These accelerator-based neutrino experiments consist of two detectors: a near detector, placed close to the beam source to measure the unoscillated neutrino flux, and a far detector, to measure the oscillated flux some distance away. The position of the far detector is chosen carefully to maximize the oscillation probability, so that

$$\frac{\Delta m^2 L}{4E} \approx \frac{\pi}{2}$$

The neutrino beams for these experiments are generated by colliding proton beams with solid targets, and filtering out all the decay products besides neutrinos. This produces a beam of muon neutrinos. In principle, the energy of the neutrinos can be controlled by changing the energy of the proton beam [23]. Thus the quantity L/E can be optimized to maximize the oscillation probability. This is a crucial step, given the rarity of neutrino interactions.

Measuring the neutrino flux at the far detector gives two handles with which we can observe oscillations. It is possible to measure neutrino disappearance, $P(\nu_\mu \rightarrow \nu_\mu)$, or neutrino appearance, $P(\nu_\mu \rightarrow \nu_e)$ in this case.

Previous oscillation experiments have obtained measurements of the mass splittings. Reactor experiments like Daya Bay [44] and Double Chooz [36] have given measurements for Δm_{12}^2 and Δm_{13}^2 , while accelerator experiments like T2K [3] and NOvA [6] have measured Δm_{23}^2 . However, since Δm_{ij} appears squared and as an argument of the sine squared, it is not possible to determine the sign of the splittings directly from oscillation measurements. Additionally, in the solar regime measured by reactor experiments, matter effects in the Sun constrain the product $\Delta m_{12}^2 \cos^2 \theta_{12}$

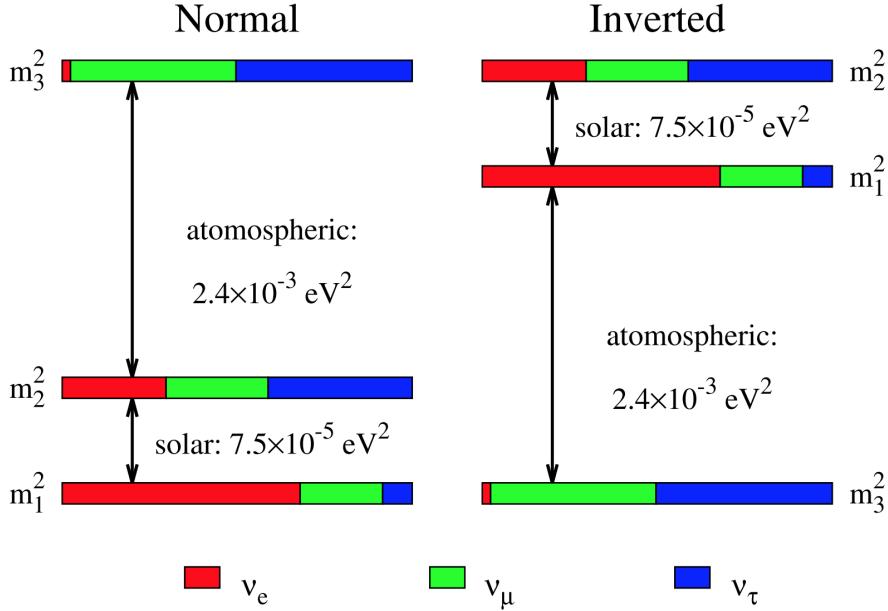


Figure 1-1: The two possible neutrino mass orderings, labeled with constraints on the mass splittings. The normal ordering has ν_1 as the lightest eigenstate and ν_3 as the heaviest, while the inverted ordering has ν_2 as the heaviest and ν_3 as the lightest. Figure taken from [10].

to be positive. Thus depending on the convention chosen to describe solar neutrino oscillations, matter effects either fix the sign of the mass splitting Δm_{12}^2 or the octant of the angle θ_{12} , where the former is positive by definition [18]. As a consequence, it is not clear what the mass hierarchy is; that is, the mass ordering of the eigenstates ν_1 , ν_2 , and ν_3 . The two possible orderings are shown in Figure 1-1.

Oscillation experiments have also resulted in measurements of the mixing angles and δ_{CP} . These values are based on global fits of all available relevant data. This fit was done in six-parameter space, to include the three mixing angles, the CP-violating phase, and two mass splittings. The two splittings are Δm_{12}^2 and $\Delta m_{3\ell}^2$, where the latter depends on the choice of mass hierarchy. Under the normal ordering, it is Δm_{32}^2 , and under the inverted ordering it is Δm_{13}^2 . Best fit parameters for all these values are presented in Table 1.1 for both mass hierarchies. The list of experiments used for this fit can be found in [25].

Experiments like the Deep Underground Neutrino Experiment (DUNE) will continue to make precise measurements of the oscillation parameters over the coming

	Normal ordering (best fit)		Inverted ordering ($\chi^2 = 2.7$)	
	bfp $\pm 1\sigma$	3σ range	bfp $\pm 1\sigma$	3σ range
$\sin^2 \theta_{12}$	$0.304^{+0.013}_{-0.012}$	$0.269 \rightarrow 0.343$	$0.304^{+0.013}_{-0.012}$	$0.269 \rightarrow 0.343$
$\theta_{12}/^\circ$	$33.44^{+0.78}_{-0.75}$	$31.27 \rightarrow 35.86$	$33.45^{+0.78}_{-0.75}$	$31.27 \rightarrow 35.87$
$\sin^2 \theta_{23}$	$0.570^{+0.018}_{-0.024}$	$0.407 \rightarrow 0.618$	$0.575^{+0.017}_{-0.021}$	$0.411 \rightarrow 0.621$
$\theta_{23}/^\circ$	$49.0^{+1.1}_{-1.4}$	$39.6 \rightarrow 51.8$	$49.3^{+1.0}_{-1.2}$	$39.9 \rightarrow 52.0$
$\sin^2 \theta_{13}$	$0.02221^{+0.00068}_{-0.00062}$	$0.02034 \rightarrow 0.02430$	$0.02240^{+0.00062}_{-0.00062}$	$0.02053 \rightarrow 0.02436$
$\theta_{13}/^\circ$	$8.57^{+0.13}_{-0.12}$	$8.20 \rightarrow 8.97$	$8.61^{+0.12}_{-0.12}$	$8.24 \rightarrow 8.98$
$\delta_{CP}/^\circ$	195^{+51}_{-25}	$107 \rightarrow 403$	286^{+27}_{-32}	$192 \rightarrow 360$
$\frac{\Delta m_{21}^2}{10^{-5}\text{eV}^2}$	$7.42^{+0.21}_{-0.20}$	$6.82 \rightarrow 8.04$	$7.42^{+0.21}_{-0.20}$	$6.82 \rightarrow 8.04$
$\frac{\Delta m_{3\ell}^2}{10^{-3}\text{eV}^2}$	$+2.514^{+0.028}_{-0.027}$	$+2.431 \rightarrow +2.598$	$-2.497^{+0.028}_{-0.028}$	$-2.583 \rightarrow -2.412$

Table 1.1: Three-neutrino oscillation parameters after a fit to global data. Taken from Ref [25].

decades [5]. DUNE has a fairly broad science program and will be conducting searches for supernova neutrinos and proton decay in addition to measuring oscillation parameters. These latter measurements, as a result of their projected precision, have the potential to answer open questions about the mass hierarchy and the existence of CP violation in the lepton sector. Other short-baseline oscillation experiments are conducting searches for sterile neutrinos, which are heavier species of neutrinos that have not yet been observed (although LSND [14] and MiniBooNE [8, 22] have reported some suggestive results). Accurately extracting parameter values from these upcoming experiments will be crucial to their success. The studies presented in this thesis investigate techniques currently used in oscillation data analysis, with the goal of improving their accuracy.

Chapter 2

Neutrino energy reconstruction

A lucky guess is never merely luck.

There is always some talent in it.

Jane Austen, *Emma*

The work in this thesis deals with new methods to approach the challenge of neutrino energy reconstruction. When a neutrino interacts with a nucleus in the detector volume, its partner lepton is produced. By detecting this lepton, a neutrino event is understood to have occurred. Thus the neutrino itself is never actually measured; its presence is registered by detecting the products of its interaction with a nucleus. However, in order to eventually calculate the oscillation parameters from the measured neutrino flux, it is necessary to obtain the energy of each incoming neutrino. Neutrino beams for accelerator oscillation experiments have a wide energy spread, and the neutrino flux as a function of energy is extracted from data using guidance from simulations of neutrino-nucleus scattering reactions. The incoming neutrino energies are not known *a priori*, and so the measurements of particles produced in neutrino-nucleus interactions are used to arrive at a guess for the neutrino's initial energy. This requires a robust understand of these nuclear interactions. All of this is, unsurprisingly, a complicated process which introduces large uncertainties in our measurements.

This chapter discusses the current approach for energy reconstruction. It also

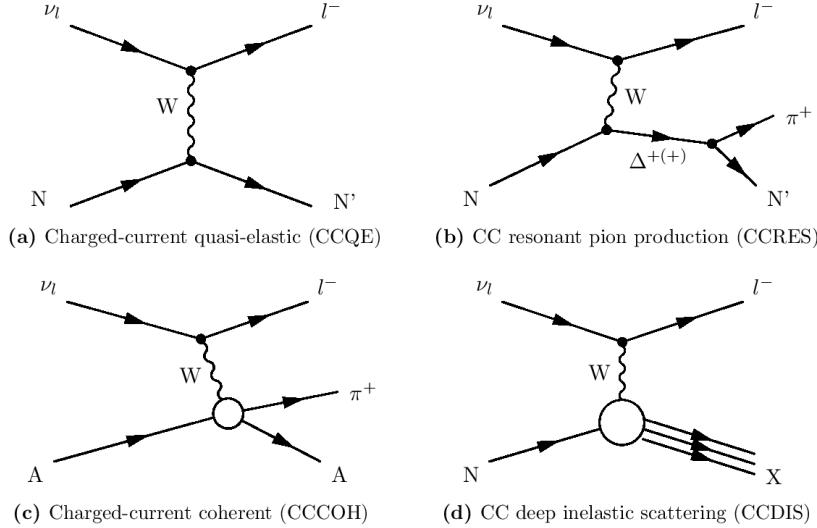


Figure 2-1: Different channels of charged-current neutrino-nucleus interactions. Quasielastic interactions (a) are the simplest of these, with no pions produced. Resonant pion production (b) and deep-inelastic scattering (d) become significant at higher energies, which will be probed by DUNE. Figure from [42].

introduces the motivation for applying electron scattering data, specifically with a deuterium target, to constrain and improve this process. Because deuterium is the simplest nucleus, we can in principle reconstruct the initial energy directly without the full complications of nuclear effects.

2.1 Energy reconstruction methods

Oscillation experiments must reconstruct neutrino energies event by event, on the basis of measurements of particles produced in neutrino interactions. There are two main methods of neutrino energy reconstruction. The first of these is the quasielastic method, employed by experiments like T2K. Here, the energy distribution of charged-current events (that is, events where the neutrino's partner lepton is produced) is calculated from the kinematics of the outgoing lepton. These quantities, such as the kinetic energy and emission angle, are measured by Cherenkov detectors filled with water or mineral oil. This technique is best applied to quasielastic (QE) events, whose Feynman diagram is shown in Figure 2-1(a).

In QE interactions, the incoming neutrino interacts with a nucleon to produce its

partner lepton and a different nucleon. For example, if a muon neutrino interacts with a neutron, a muon and a proton are produced. The quasielastic reconstruction technique is reasonable for experiments like T2K, which uses a neutrino beam with $E_\nu \sim 600 - 800$ MeV (specifically for the off-axis ND280 detector [1]), where QE events are dominant contribution to the total cross-section. Taking the example of an ejected muon and proton, the reconstructed energy is calculated with the expression shown below,

$$E_{rec}^{QE} = \frac{M_p^2 - m_\mu^2 + 2E_\mu(M_n - B) - (M_n - B)^2}{2(M_n - B) - E_\mu + p_\mu \cos \theta_\mu} \quad (2.1)$$

where M_p, M_n are the proton and neutron masses, B is the binding energy of the target nucleus, m_μ is the muon mass, E_μ, p_μ are the muon energy and momentum, and θ_μ is the angle between the muon and neutrino momentum vectors. Note that this calculation does not rely on knowledge of the knocked-out nucleon's momentum, because these particles are unobservable in Cherenkov detectors. This method assumes that no other nucleons were knocked out and that no other particles were produced [12]. More complicated methods are required to account for complex final states.

At higher energies above 1 GeV, inelastic processes become significant and eventually dominant. DUNE will operate in this energy range, with an energy spectrum between 0-10 GeV. As can be seen by looking at the cross-sections in this range in Figure 2-2, DUNE will be dominated by resonant pion production and deep-inelastic scattering events. The Feynman diagrams for these interactions are shown in Figure 2-1(b) and (d).

In this higher energy range, because multiple hadrons are produced, the quasielastic method fails to suffice. Additionally, the development of new detector technology enables better reconstruction methods. DUNE will employ state-of-the-art Liquid Argon Time Projection Chambers (LArTPC) detectors, which have already been used to take data for the MicroBooNE experiment. The detector volume is filled with liquid argon, and incoming neutrino interact with argon nuclei to produce charged particles. These particles ionize the liquid argon as they propagate, and an applied electric field causes the ionization electrons to drift towards the collection planes. The

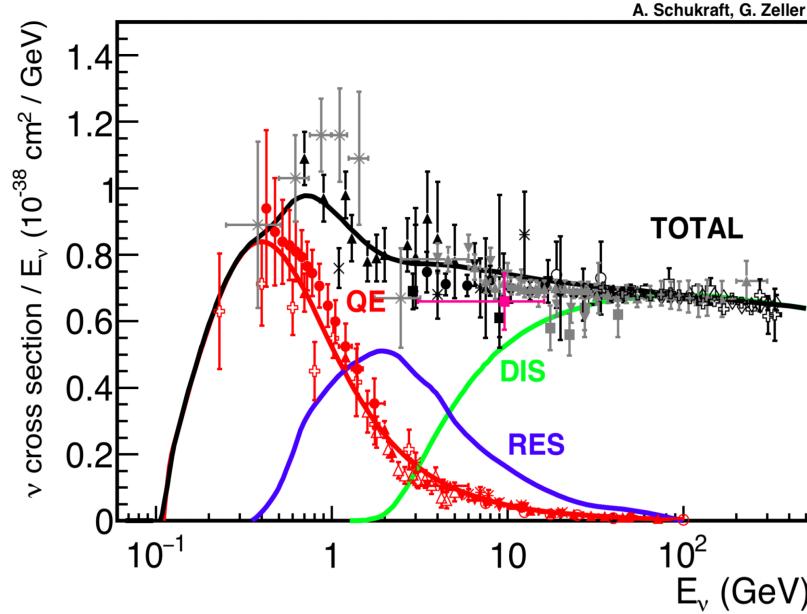


Figure 2-2: Neutrino scattering cross-sections by interaction channel. In the 0-10 GeV range relevant for DUNE, the QE, resonant, and deep-inelastic scattering channels are all important. Figure from [27].

detected charge can be used to calculate the deposited energy, and position information can be used to reconstruct particle tracks. LArTPC detectors make it possible to conduct fine-grained tracking of a number of interaction products. They produce bubble-chamber-like images of neutrino events, making it more straightforward to distinguish background events and allowing accurate measurements of complex final states [39].

To take full advantage of these detectors and best account for inelastic events, a new reconstruction method is necessary. This is the calorimetric technique, which uses measurements of the energy deposited by all interaction products to reconstruct the initial neutrino energy. This method is already in use in experiments with calorimeter-based detectors, such as NOvA [7]. Clearly the success of the calorimetric technique depends on the ability to fully reconstruct the final state, but with advanced LArTPC detectors this will only improve. The expression for the calorimetric reconstruction for a QE event with an ejected muon and proton is shown below.

$$E_{rec}^{cal} = E_\mu + E_p - m_p + B \quad (2.2)$$

Nuclear effects can have a large impact here, as they may lead to a significant amount of missing energy. For example, a pion may be produced and reabsorbed by the spectator nucleus, but its energy will not be registered by the LArTPC.

In the studies presented in this thesis, we consider the calorimetric method. Upcoming oscillation experiments will be using heavier target materials like carbon, iron, and liquid argon, where the calorimetric method has much better performance than the standard quasielastic method. One aim of this work is to benchmark the performance of the calorimetric method using precise electron scattering data, as will be described later in this chapter.

2.2 Improving neutrino event generators

Another component of neutrino energy reconstruction relies heavily on phenomenological models of neutrino-nucleus interactions. At the far detector, by which point the neutrino beam has traveled some distance L , we expect some fraction of the neutrinos to have oscillated into different flavors. We can express the far detector flux as shown below,

$$\begin{aligned}\Phi_e(E, L) &\propto P_{\nu_\mu \rightarrow \nu_e}(E, L) \Phi_\mu(E, 0) \\ \Phi_\mu(E, L) &\propto [1 - P_{\nu_\mu \rightarrow \nu_e}(E, L)] \Phi_\mu(E, 0)\end{aligned}\tag{2.3}$$

where $P_{\nu_\alpha \rightarrow \nu_\beta}(E, L)$ is given in Equation 1.8. The proportionality here depends on the experiment. By measuring these fluxes and their energy dependence, we can extract the parameters that govern $\nu_\mu \rightarrow \nu_e$ oscillation.

This flux is calculated from the neutrino interaction rate with detector nuclei. This rate is what is actually measured, and can be expressed as

$$N_e(E_{rec}, L) \propto \sum_i \int \Phi_e(E, L) \sigma_i(E) f_{\sigma_i}(E, E_{rec}) dE\tag{2.4}$$

$N_e(E_{rec}, L)$ is the interaction rate at a given reconstructed energy E_{rec} and distance L from the beam source. The term $f_{\sigma_i}(E, E_{rec})$ is a smearing matrix that relates the

true (E) and reconstructed (E_{rec}) neutrino energies. If the energy is reconstructed perfectly, this is not necessary. But practically speaking, E_{rec} differs due to experimental effects (detector inefficiencies and background events) and nuclear interaction effects (nucleon motion, nucleon reinteraction, meson currents). Detector effects are minimizable by improving hardware and technology, but the nuclear effects are irreducible and must be accounted for with better phenomenological models. This is typically implemented using neutrino event generators.

To extract oscillation parameters accurately, we must be able to correctly extract $\Phi_e(E, L)$ from measurements of $N_e(E_{rec}, L)$. This depends on the accuracy of the theoretical models used by event generators to determine $f_{\sigma_i}(E, E_{rec})$. Current oscillation experiments report large systematics due to these factors [3, 6]. However, because there are no mono-energetic high-energy neutrino beams, it is difficult to test these models at specific neutrino energies.

This is problematic, as event generators rely on these models to simulate a complete set of interaction processes on a variety of target materials for arbitrary neutrino beam energies – an ambitious goal. Event generators usually rely on semiclassical models and ignore interference effects, which is already a limitation given that the nucleus is a quantum mechanical system. These models are also usually tuned to near-detector data, to be able to reproduce the unoscillated flux $\Phi(E, 0)$. This is helpful, but there is no guarantee that the models are correctly reproducing results for each value of E , or that the models will be correct at the far detector once oscillations dramatically change the observed fluxes. Figure 2-3 shows the difference between the projected energy spectra for DUNE, before and after oscillations.

As a result, incorrect modeling in the event generators is not unlikely. This can strongly bias the flux calculation in Equation 2.4, and thus skew the extracted oscillation parameters. This effect can be seen in Figure 2-4, where data has been simulated using one event generator (GIBUU) and analyzed using another (GENIE). Clearly, the measured parameter values depend on the choice of event generator, and it is not clear which is more accurate.

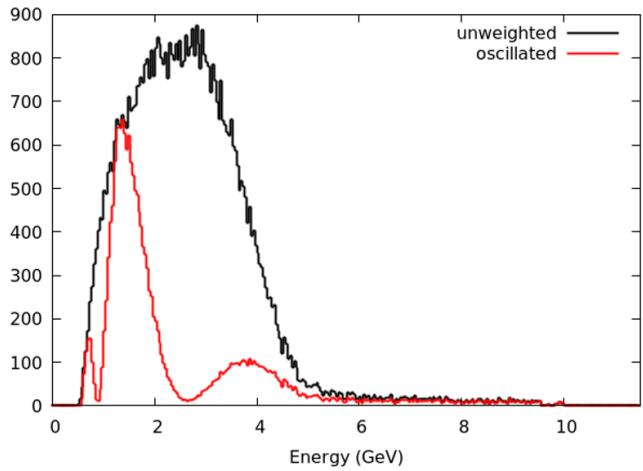


Figure 2-3: The expected incident energy spectra of the DUNE ν_μ beam at the near detector (before oscillations, shown in black) and at the far detector (after oscillations, shown in red).

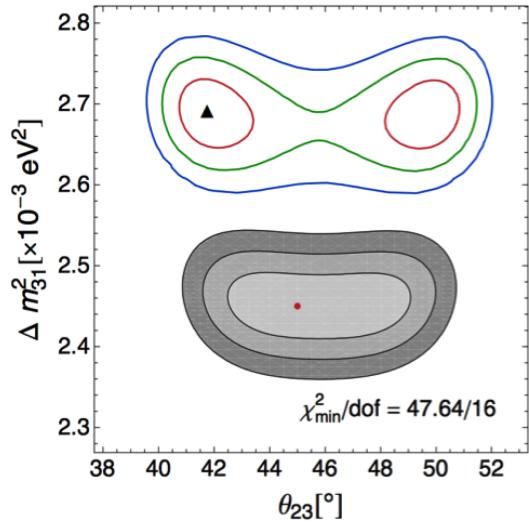


Figure 2-4: The impact on extracted oscillation parameters if the event generator does not capture the underlying physics. The grey regions show the 1σ , 2σ , and 3σ confidence regions for the extracted parameter values when the input data was produced with the event generator GIBUU, and also analyzed with GIBUU. The red, green, and blue lines show the same regions for analysis of the input GIBUU data with the GENIE event generator. The red dot indicates the input parameter values, and the black dot shows the extracted best fit parameters. Taken from [32].

2.3 Applying electron scattering data

It is important to understand and constrain the biases in energy reconstruction and event generators, in order to ensure the success of DUNE’s precision measurements.

Comparing neutrino event generators against mono-energetic electron scattering data can aid in this effort. Electrons and neutrinos do interact slightly differently with nuclei, but there are substantial similarities. This means we can use electron data to constrain parts of these event generator models, in fact in a complementary way to the current approach which uses short-baseline neutrino measurements. The advantage of doing this lies in the data – electron beams can be produced to precise, known energies. Therefore we can test energy reconstruction in specific regions of interest and identify where neutrino event generators (run in electron scattering mode [40]) succeed and where they fail.

Neutrinos, as described earlier, interact by exchanging W and Z bosons and interact via vector and axial currents. Note that both interaction types share vector currents. Currently, the axial component of neutrino interactions can be constrained using short-baseline neutrino data. Electron scattering data can help to constrain the vector response in neutrino interactions. Electron scattering also provides an analogue to neutrino scattering, and thus can be used to probe issues in nuclear modeling and serves as a useful testing ground for neutrino models [13, 34].

Electrons interact with nuclei by exchanging photons, and interact via longitudinal and vector currents. Electrons can interact with both one-body and two-body currents in the nucleus. In one-body currents, only one nucleon is involved in the interaction. QE knockout is an example of this. Two-body currents lead to two nucleons in the final state, as shown in the Feynman diagrams in Figure 2-5. For example, in isobar configurations (IC) the electron excites a nucleon to a Δ , which de-excites by interacting with another nucleon. In meson exchange currents (MEC), the virtual photon is absorbed on an exchanged meson leading to two-nucleon knockout. In short-range correlations (SRC), the electron knocks out one nucleon belonging to an SRC pair, and the correlated nucleon is also ejected. In final state interactions (FSI),

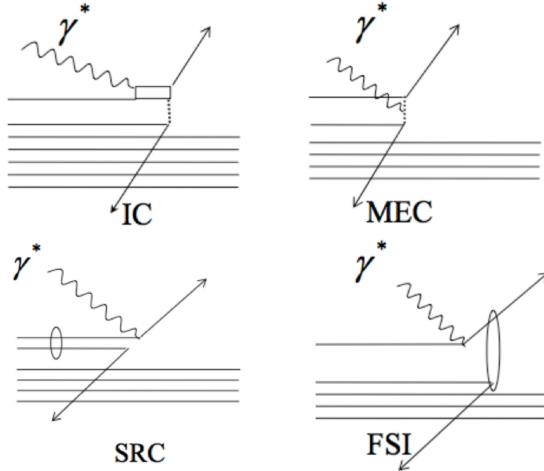


Figure 2-5: Two-body current interactions. The upper left shows isobar configurations (IC), where the virtual photon is absorbed on a nucleon, which is excited to a Δ than then de-excites via $\Delta N \rightarrow NN$. In the upper right, meson exchange currents (MEC) show the virtual photon being absorbed on a meson exchanged between two nucleons, which are both knocked out. In the bottom left, the virtual photon is absorbed by one nucleon in a short-range correlated (SRC) pair, and its partner nucleon is also knocked out. In the bottom right, the virtual photon is absorbed by a nucleon, which rescatters via final state interactions (FSI). Figure from [13].

the ejected nucleon rescatters from another nucleon.

All of these two-body current interactions lead to the same final state as the one-body interaction and thus interfere with each other. This is an effect which is not directly built into event generators, so using electron scattering data to constrain and tune neutrino models can incorporate these effects implicitly.

Because electron interactions are mediated by photons and neutrino charged-current interactions are mediated by the W boson, some correction must be made to account for the different propagators involved. Because the photon is massless, its propagator has a factor of $1/Q^2$, where Q^2 is the four-momentum transfer of the interaction. The W mass leads to a constant propagator. To correct for this, electron scattering events are weighted by a factor of Q^4 in order to be directly comparable to neutrino results [13].

Another crucial difference is the presence of radiative effects, which will be discussed in more detail in Chapter 3.

2.3.1 The $e4\nu$ effort

The first analysis under this electrons-for-neutrinos framework, conducted for the $e4\nu$ collaboration, used 1999 data from the CLAS detector (see Section 3.1) at the Thomas Jefferson National Accelerator Facility (JLab) on ${}^4\text{He}$, C, and Fe targets with mono-energetic electron beam at a variety of energies. The first analysis focused on QE events, by selecting events with one electron, one proton, and no charged pions ($1p0\pi$). That study found that most of the $1p0\pi$ events do not reconstruct to the correct energy, that there are 50% discrepancies in the percentage of correctly-reconstructed events between data and simulated event generator input, and that the event generator models failed to capture the shape of the low-energy tail for QE energy reconstruction [17].

These results indicate that there is tremendous room for improvement with current interaction models, and that this improvement is crucial as can lead to serious issues in the reconstruction of the far detector energy spectrum and neutrino flux.

2.3.2 Focus on electron-deuterium data

In this thesis, I investigate the effectiveness of energy reconstruction methods using electron scattering data with a view to understanding how well this process works. In particular, I focus on electron-deuterium scattering data. The choice of deuterium is because it is the simplest nucleus. As mentioned earlier, nuclear effects can have a substantial impact on the quality of the reconstruction. By looking at deuterium, these tangled nuclear effects can be eliminated, and in principle it is possible to obtain an upper bound on the performance of this method and benchmark its effectiveness. By studying the calorimetric technique in the electron scattering regime, we have finer knowledge of the data because the beam energy is known.

Theoretical work suggests that it ought to be possible to perfectly reconstruct events using a deuterium target [31, 43, 37]. In particular, because deuterium is a two-body system, it should be possible to kinematically reconstruct the beam energy for charged current interactions. This is because the kinematics of all of the particles

are in principle fully determined by energy-momentum conservation. The deuteron behaves in such a way that semi-inclusive measurements of two of the three particles in the final-state is actually exclusive, since there is no missing momentum or missing energy dependence, as is the case in general nuclei. This allows us, in principle, to reconstruct the incident beam energy with minor uncertainties. For the electron-scattering case, the reconstruction formula is shown below.

$$E_{rec}^D = \frac{m_D^2 + m_p^2 - m_n^2 - 2m_D(E_{e'} + E_p) + 2E_{e'}(E_p - p_p \cos \theta_{e'p})}{2(E_{e'}(1 - \cos \theta_{e'}) + E_p - p_p \cos \theta_p - m_D)} \quad (2.5)$$

Here, $E_{e'}$ is the energy of the scattered electron. We also define $\theta_{e'}$ as the angle between the scattered electron and the beam, θ_p as the angle between the proton and the beam, and $\theta_{e'p}$ as the angle between the scattered electron and the proton.

This thesis investigates the validity of this claim and verifies how well reconstruction with a deuterium target works. Thus we also consider the possible effectiveness of future deuterium-based neutrino detectors.

Chapter 3

Energy reconstruction analysis on CLAS data

Take nothing on its looks; take everything on evidence. There's no better rule.

Charles Dickens, *Great Expectations*

Electron-scattering data is plentiful; in particular, data from the Thomas Jefferson Accelerator Facility (referred to as Jefferson Lab) is well-suited for the purposes of this study. In this section, I describe the analysis and reconstruction process implemented on electron-deuterium scattering data.

3.1 The CLAS detector

Jefferson Lab contains a linear electron accelerator, which delivers beams to three experimental Halls. This study uses data from Hall B, which houses the CEBAF Large Acceptance Spectrometer (CLAS) [35]. The CLAS detector has a nearly 4π solid angle coverage, and can detect charged particles with polar angles 8° to 140° , and neutral particles from 8° to 75° . CLAS can accurately measure the momentum, time-of-flight, and trajectories of particles from a reaction target. This enables particle

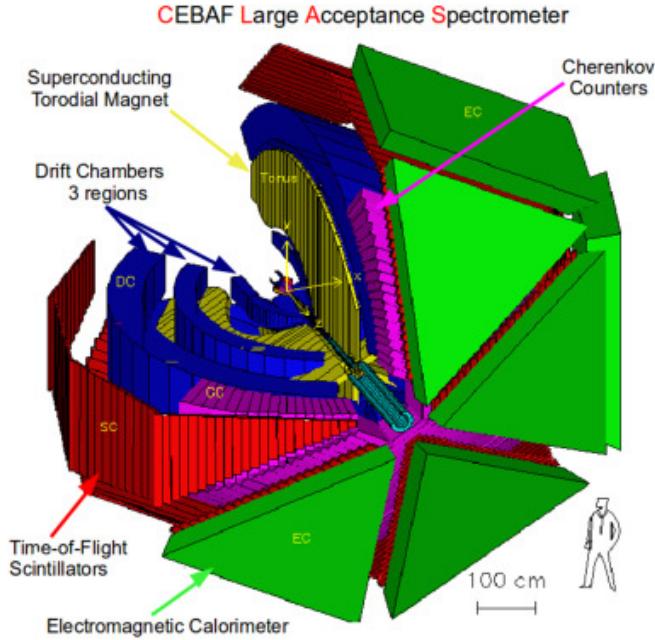


Figure 3-1: The CLAS detector and its main components.

identification using mass, charge, and momentum information [45]. Figure 3-1 shows the CLAS detector and its components.

The main torus magnet consists of six superconducting coils in a toroidal geometry, perpendicular to the electron beam, which split the CLAS detector into six parts (called sectors). The drift chambers are the first detector elements after particles scatter from the target. They can track particles and determine their momentum (if over 200 MeV/c) with a resolution of 0.5%. The forward electromagnetic shower calorimeter of CLAS has scattering angles from 8° to 45° with corresponding azimuthal coverage of the drift chambers. For higher angles, there is another calorimeter, the Large Angle Calorimeter. The next element is the Cherenkov counters, which are used to discriminate between electrons and pions. Finally, the time-of-flight system is a group of scintillators, used (along with the drift chambers) for particle identification. [45]

3.2 Event selection

The data for this study was collected by the CLAS detector for the *EG2* run, conducted in 2004. During this run, a liquid target and solid target were used simulta-

neously. The first target seen by the beam was the cryogenic liquid target located at a vertex of $z \sim 30$ cm, which was 2 cm long and contained either liquid hydrogen or deuterium. The second target was a solid target, which contained a variety of different materials (C, Al, Fe, Sn and Pb) over the course of the run. The relevant data for this study was collected with a beam energy of 5.014 GeV and the liquid deuterium target at $z \sim 30$ cm (density 0.1624 g/cm³) [30]. I used a sample of 11,128,235 events. My goal was to isolate quasielastic $D(e, e'p)$ scattering events, namely with one proton and no pions in the final state. This is the equivalent of the charged-current quasielastic events of interest in the neutrino picture.

Only particles that passed the certain momentum thresholds were considered: the pion threshold was 150 MeV/c, the photon threshold was 300 MeV/c, and the proton threshold was 300 MeV/c as well. These thresholds are comparable to those in neutrino tracking detectors [4], making these CLAS datasets particularly useful for developing constraints for neutrino experiments. Due to scattering from target walls and other backgrounds, some events had more than one proton detected. The proton from the $(e, e'p)$ interaction (referred to as the primary proton) was chosen to be the proton with the most momentum.

3.2.1 Vertex determination

Since particle identification was already complete for the data I used, the first step was of event selection was vertex determination. Figure 3-2 shows a plot of the electron vertex and the proton vertex. The yellow band in the middle of the plot, with z near 30, corresponds to the events of interest from the deuterium target. A secondary band is visible with a proton vertex near $z \sim -23$ cm, which corresponds to events from the solid target. Here, $z = 0$ corresponds to the center of the CLAS detector.

The electron vertex in the region of interest fell between -27 cm and -30 cm, as can be seen in the figure. In order to select events from this region, we apply a cut on the difference between the electron and proton vertex. The $(e, e'p)$ events were selected by cutting on events where the vertex of the primary proton was within 1 cm of the electron vertex. This ensures that the scattered electron and primary proton are

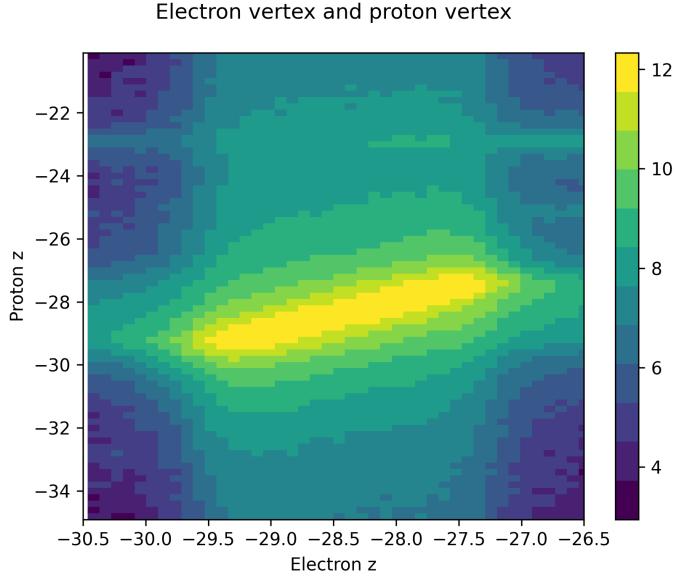


Figure 3-2: The electron vertex and proton vertex. The liquid deuterium target is located at $z \sim 30$ cm, so the yellow band in the middle of the plot contains the $D(e, e'p)$ events.

both generated at the target, and excludes accidental events with a proton produced outside the deuterium target. This cut is shown in Figure 3-3.

3.2.2 Weighting events

All events are weighted by Q^4 . This is the square of Q^2 , the squared four-momentum transfer, which is defined as

$$Q^2 = -q^2 = 4E_e E_{e'} \sin^2 \frac{\theta_e}{2} \quad (3.1)$$

where q is the momentum transfer vector, E_e is the initial electron energy, $E_{e'}$ is the scattered electron energy, and θ_e is the electron scattering angle.

This scaling factor is applied to account for the different propagators involved in electron interactions as opposed to neutrino interactions. Because the photon is massless, its propagator has a factor of $1/Q^2$. This means that the cross section of electron interactions is $\sim 1/Q^4$. By contrast, in the neutrino case, the large W and Z masses lead to effectively constant propagators: there is a factor of $1/M^2$ in the

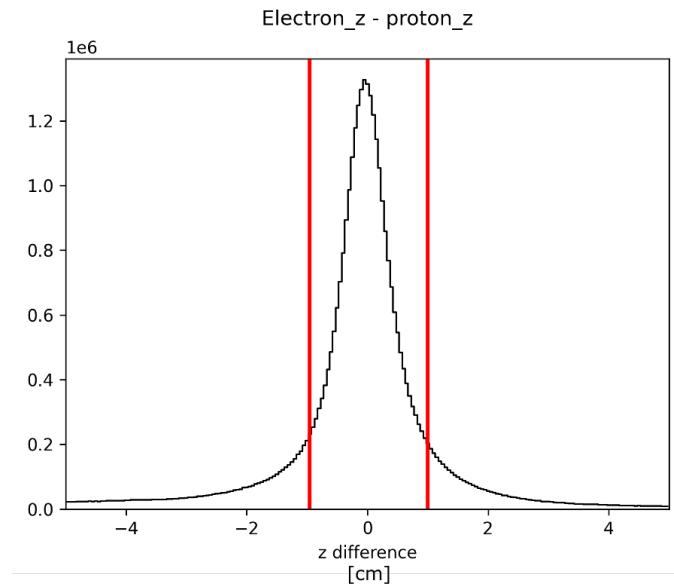


Figure 3-3: The difference between the electron vertex and proton vertex, along with the applied cut. We select events where the difference between the two values is less than 1 cm.

propagator, which becomes a factor of $1/M^4$ in the cross-section. To directly compare the electron scattering results to neutrino results, as is the goal here, I weight the electron scattering events by Q^4 to account for the different cross-sections. This makes electron events “look like” neutrino events [13]. The effect is that for a given incident energy, electron scattering is concentrated at lower momentum transfer than neutrino scattering.

3.2.3 Fiducial cuts

The components of the CLAS detector, like the drift chambers, Cherenkov counters, and calorimeter, have inefficient zones. In these regions, there might be distortions in the information collected. For example, at the edges of the Cherenkov counters, there might be a decrease in the efficiency of light collection. There are various such inefficiency zones. Fiducial cuts are applied to define detection zones where the efficiency is constant, and reproducible by simulations of the CLAS detector. Different fiducial cuts are used for different particles and magnetic fields. This reduces the statistics of the data but improves its quality.

I applied the *EG2* fiducial cuts to the detected electron and primary proton. Fiducial cuts were also necessary for pions, in order to implement the background subtraction procedure described in Section 3.2.5. The cuts are generated by inspecting the acceptance in different regions of the detector, and performing a fit on where the acceptance is near 100%. Regions where the acceptance was uniform (and the detector components were not known to be malfunctioning) were determined and used to generate the fiducial cuts, which are thus functions of particle angle and momentum. These cuts defines the edges of the angle and momentum distributions where the acceptance is full. The cuts I used are defined as follows.

The fiducial cuts on θ and ϕ are given by

$$\begin{aligned} \text{if } \theta < \theta_{\min} \text{ or } \theta > \theta_{\max} &\implies \phi = 60(\text{Sector} - 1) \\ \text{else } &\implies \phi = 60(\text{Sector} - 1) + (\text{Side})a \left(1 - \frac{1}{\frac{\theta - \theta_{\min}}{b} + 1}\right) \end{aligned} \quad (3.2)$$

Here, θ_{\min}, a, b are free parameters, θ_{\max} has a default value for each particle, and Side (± 1) switches for the left and right halves of each sector. The parameters θ_{\min}, a, b are determined as a function of the momentum p as shown below, using parameters p_0, \dots, p_5 obtained through fit procedures described in [45, 24].

$$\begin{aligned} a &= p_0 + p_1 e^{p_1(p-p_3)} \\ b &= p_0 + p_1(p)e^{p_2(p-p_3)^2} \\ \theta_{\min} &= p_0 + \frac{p_1}{p^2} + p_2 p + \frac{p_3}{p} + p_4 e^{p_5 p} \end{aligned} \quad (3.3)$$

An example of the effect of the fiducial cuts is shown in the case of protons in Figure 3-4. The figure shows the proton scattering phase-space (θ and ϕ) before and after the cuts. After applying the cuts, the six sectors of CLAS are clearly visible. Events near the edges of the sectors have been rejected due to inefficiencies in those regions. Areas within the sectors have also been rejected, likely due to faulty components and dead zones.

We can reconstruct the beam energy using the two techniques specified in Equa-

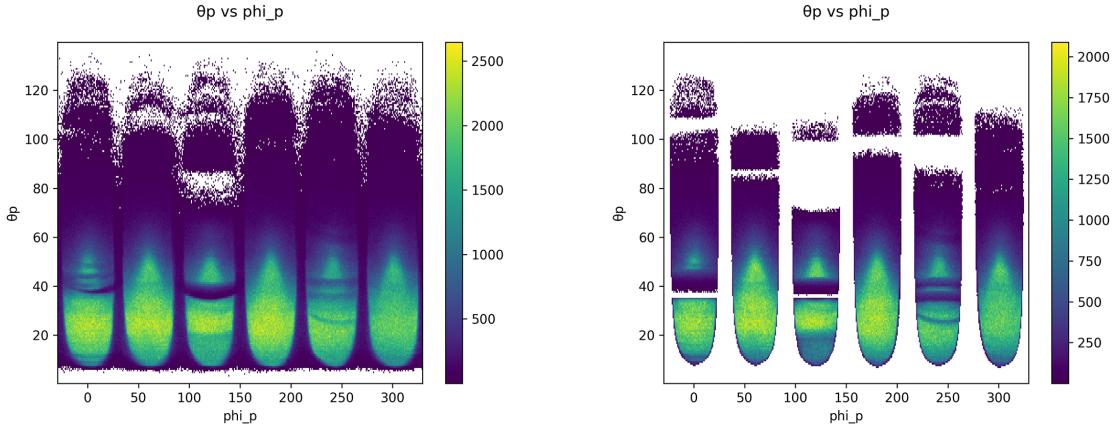


Figure 3-4: Effect of fiducial cuts on the θ and ϕ distribution for protons. On the left is the distribution before cuts are applied, and on the right is the distribution after the cuts. The six sectors of the CLAS detector are clearly visible.

tions 2.2 and 2.5. In both cases, we are using information about the proton. The former is the standard calorimetric reconstruction, and the latter is a quasielastic method specifically for deuterium. Figure 3-5 shows the results. Both methods result in a clearly defined peak near 5 GeV, which is the correct beam energy. However, there is a substantial tail extending to the left of the peak. In both cases, the size of the tail is fairly similar although the shape differs somewhat.

3.2.4 Acceptance maps and correction

In addition to the fiducial cuts, I applied a cut based on the *EG2* acceptance maps. These maps serve as a double check for the fiducial cuts. The CLAS acceptance maps are a combination of geometrical acceptance, detection and track reconstruction efficiencies, and event selection efficiencies. They are calculated using simulations. By comparing simulated events before and after detection with CLAS, we can extract the acceptance. This is defined by generating events in θ -momentum bins and calculating the ratio between N_{acc} , the number of “accepted” events from each bin that pass through CLAS, and N_{gen} , the number of generated events in that bin. $N_{\text{acc}}/N_{\text{gen}}$ can be thought of as the total detection efficiency or the detection probability.

We require that the acceptance for both the scattered electron and the primary

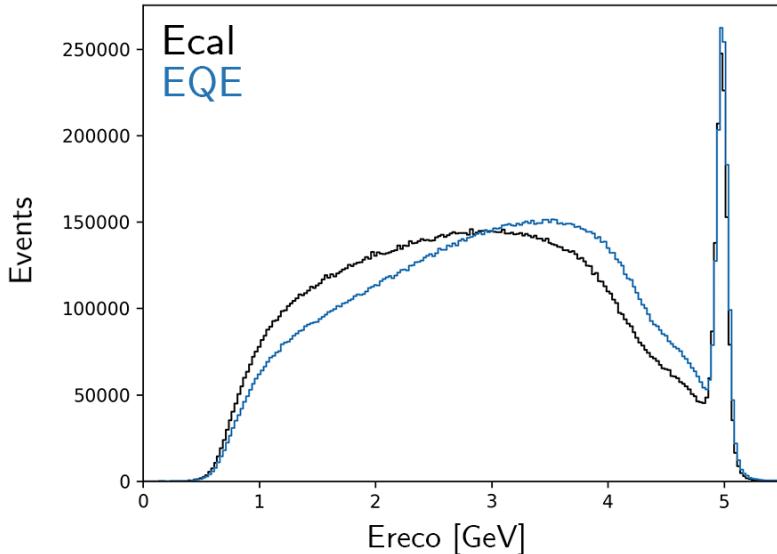


Figure 3-5: E_{cal} and E_{QE} energy reconstruction distributions for $D(e, e'p)X$ events. Here we have applied fiducial cuts to the electron and proton, but none of the other event selection steps.

proton be greater than 0.7. This ideally serves to validate the fiducial cuts and ensure that only events from high-efficiency parts of the detector are used for the analysis. In the case of the proton acceptance maps, applying this cut makes little difference after the fiducial cut. However, this does have significant impacts on the electron distributions.

Figure 3-6 shows the θ - ϕ distribution for electrons, and Figure 3-7 shows the θ - $p_{e'}$ distribution. The acceptance maps were generated with evenly-sized bins in $\cos\theta$, which leads to resolution problems for forward-angle electrons where $\cos\theta$ is close to 1. This is visible in Figure 3-6, where electrons at forward angles are rejected. In the same plot, we can see that electrons at high scattering angle θ are also rejected. These correspond to low-momentum electrons, visible in Figure 3-7. There is an effective momentum threshold in the acceptance maps, likely due to threshold effects in the simulation software used to generate these maps (for example track reconstruction might degrade at low momenta, or the trigger in the simulation might not recognize these events). Thus the main effect of the acceptance cut is to pare down the electron distribution. Although the effect at forward angles appears to be detrimental, since

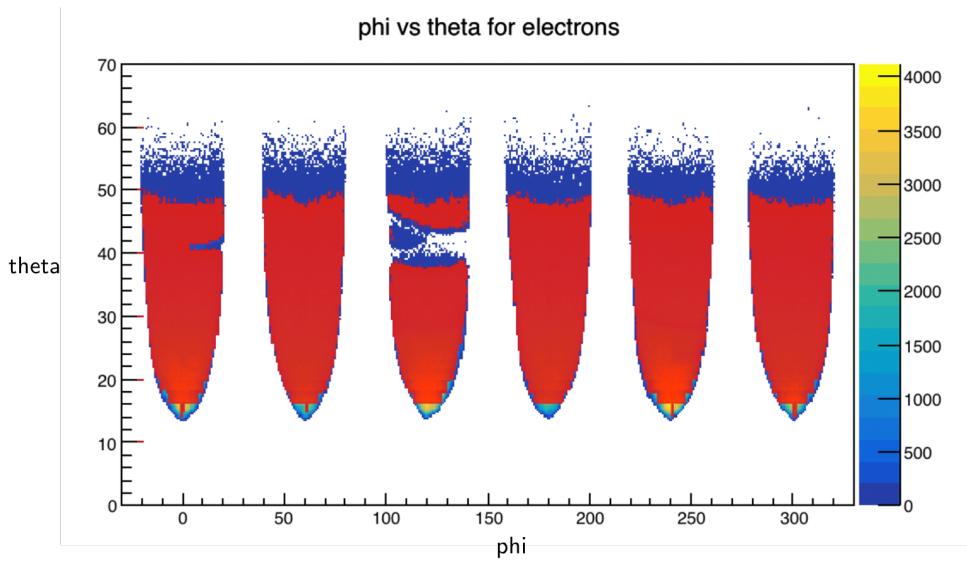


Figure 3-6: The θ - ϕ distribution for electrons. The red events are those that pass the fiducial and acceptance cuts (> 0.7), overlaid on top of the distribution after just the fiducial cut. Events with electrons at high scattering angles θ are eliminated, as are events with very forward-angle electrons. This is due to the process by which the maps were generated, but have little effect on the final comparison in this study.

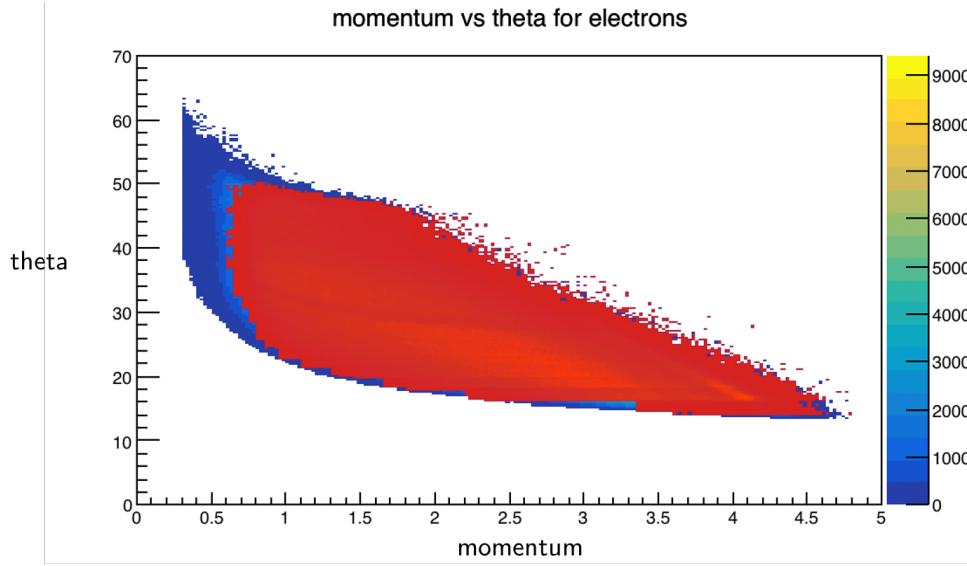


Figure 3-7: The θ - $p_{e'}$ distribution for electrons. The red events are those that pass the fiducial and acceptance cuts (> 0.7), overlaid on top of the distribution after just the fiducial cut. Events with electrons at low momenta ($<\sim 800$ MeV/c) are eliminated. This is likely due to threshold effects in the CLAS simulation, but this has little impact on the final comparison in this study.

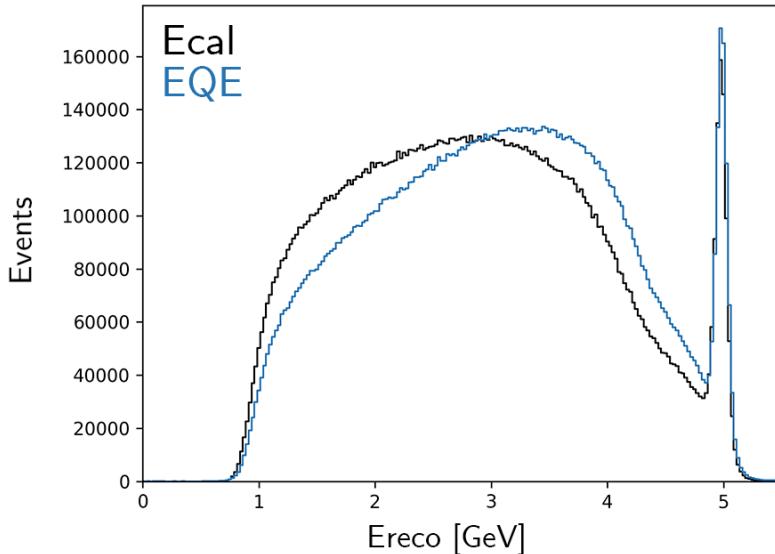


Figure 3-8: The reconstructed E_{cal} and E_{QE} energy distributions after applying the acceptance maps and correction. The size of the peak has been reduced as compared to Figure 3-5.

ultimately we are interested in a comparison of the reconstruction between data and simulation, it is not critical as long as the same procedure is applied in both cases.

The second step with the acceptance maps is applying an acceptance correction. In order to account for different efficiencies in different regions of the detector, we reweight the events that pass the acceptance and fiducial cuts by the reciprocal of the acceptance value. Each event was weighted by a factor of $(a_e a_p)^{-1}$, where a_e is the acceptance of the electron and a_p is the acceptance of the proton. This corrects for the fact that certain types of events might be underrepresented in the data due to lower efficiencies in the detector. This ends up increasing the event counts somewhat, because the acceptance values are always less than 1, so this correction only upweights events.

We can again inspect the reconstruction after these steps. Figure 3-8 shows the results. We still have a clearly defined peak near 5 GeV, which is the correct beam energy. Both cases still have a large tail, and the acceptance maps have actually reduced the relative size of the peak, resulting in a tail of comparable size. This effect can be seen more clearly in Figure 3-9. This figure shows the change in the energy

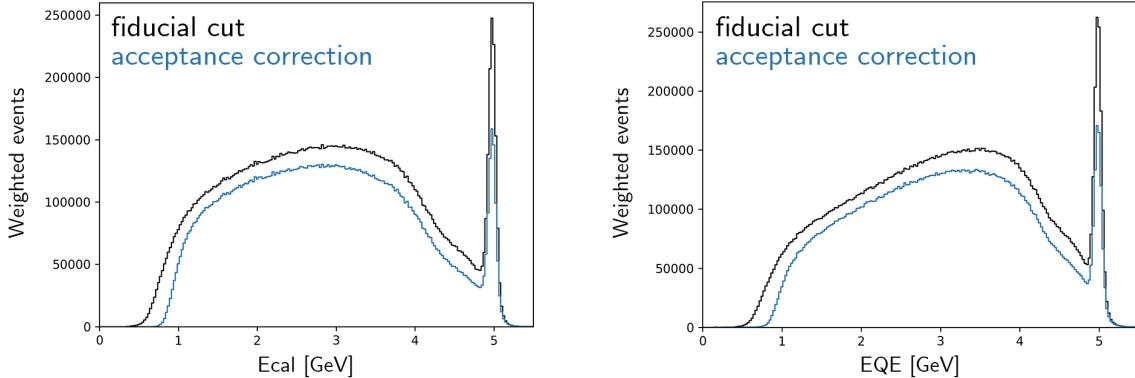


Figure 3-9: Energy reconstruction distributions for E_{cal} (left) and E_{QE} (right), after the fiducial step and after the acceptance step. Note the substantial reduction in the peak and the shift of the left edge of the distributions. This is due to the acceptance cut significantly affecting the electron scattering phase space, and the rejection of low-momentum electrons.

reconstruction distributions for both E_{cal} and E_{QE} after the fiducial cut and after the acceptance step. The reduction in the peak is clear. Note also that there is a shift in the left edge of the distribution; this is due to the rejection of low-momentum electrons.

3.2.5 Background subtraction

The fiducial and acceptance cuts, along with the acceptance correction, account for detector effects and leave us with a sample of valid electron-deuterium scattering events. The next step is to select $(e, e'p)$ events.

We want to obtain a sample where the final state consists of one proton and no pions or photons. There are various other non-quasielastic interactions that lead to multi-hadron final states, which we wish to ignore. However, it is not sufficient to use the naive method of simply applying a cut on events where no pions are detected. This is because of gaps and inefficiencies, as the CLAS detector does not have perfect acceptance. Undetected particles, either below the CLAS momentum thresholds or in the dead zones of the detector, can fly under the radar and contaminate the $1p0\pi$ sample. To obtain a true $1p0\pi$ sample, we must implement a data-driven background subtraction procedure to account for any significant background processes with un-

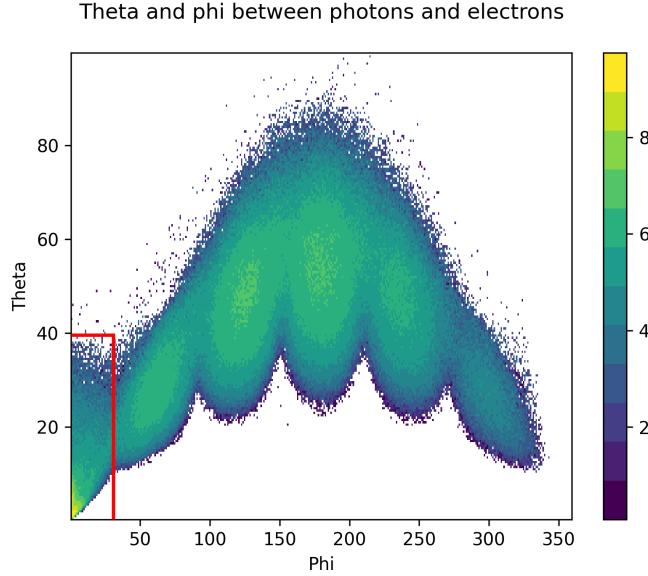


Figure 3-10: The distribution of the angle between the electron and the photon, versus the difference of their azimuthal angles (counts are shown on a logarithmic scale). The box in red represents the cut used to select radiation photons.

detected pions or photons.

In the case of photons, there are two processes at play. Radiative photons are produced by bremsstrahlung radiation when the electron beam interacts with the target nucleus. A pair of photons is also produced by the decay of a neutral pion. Radiative photons will be treated separately (described in Section 3.2.6); for the background subtraction, we are interested in the non-radiative π^0 photons because their presence is indicative of a multi-hadron final state. To distinguish between these photons for the purposes of the background subtraction, we consider the distribution of the angle between the electron and the photon, plotted as a function of the difference between the electron and photon azimuthal angles ϕ . This latter quantity should always be small for radiative photons. Figure 3-10 shows this plot for the data, along with the cuts on θ, ϕ used to exclude radiative photons. The remaining non-radiative photons can serve as proxies for the π^0 in the subtraction process.

The first step of the background procedure is the naive $1p0\pi$ cut. Then we deal with multi-hadron states. The proton and pion multiplicity plots in Figures 3-11 and 3-12 show the number of protons and charged pions produced in each event. From

these plots, it is visible that the main background processes are those with between 1 and 3 pions or 2 protons. The background is dominated by the $1p1\pi$, $1p2\pi$, $1p3\pi$, and $2pN\pi$ channels, where N accounts for the pion multiplicity. In order to account for events from these channels that might appear in the data as $1p0\pi$ events, we use a pion subtraction procedure.

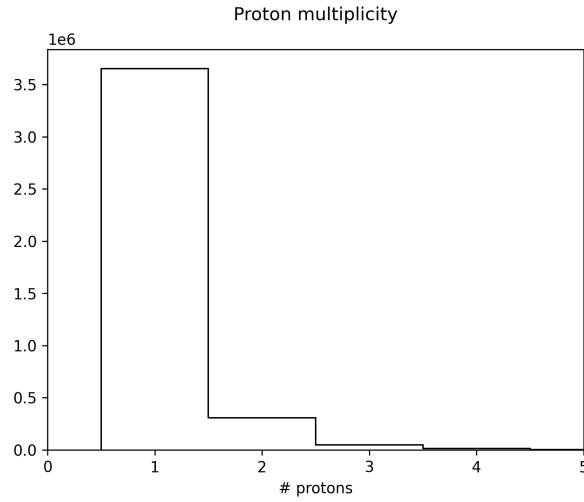


Figure 3-11: Proton multiplicity. Most events have one proton, and a small number have two protons. Higher multiplicities are low enough to be ignored.

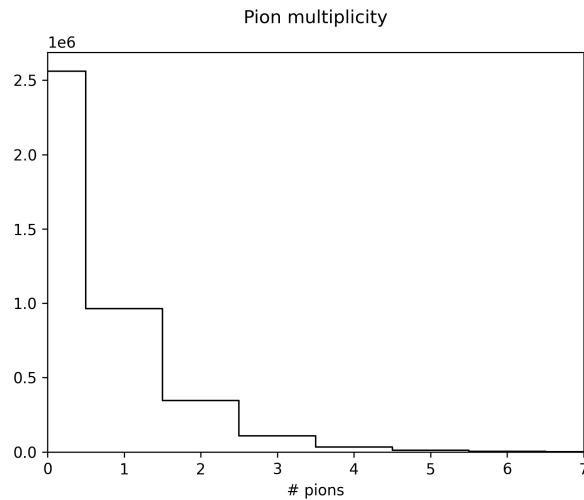


Figure 3-12: Pion multiplicity. Most events have between zero to three charged pions. Higher multiplicities can be ignored.

The subtraction scheme is as follows. After the naive $1p0\pi$ cut, we consider

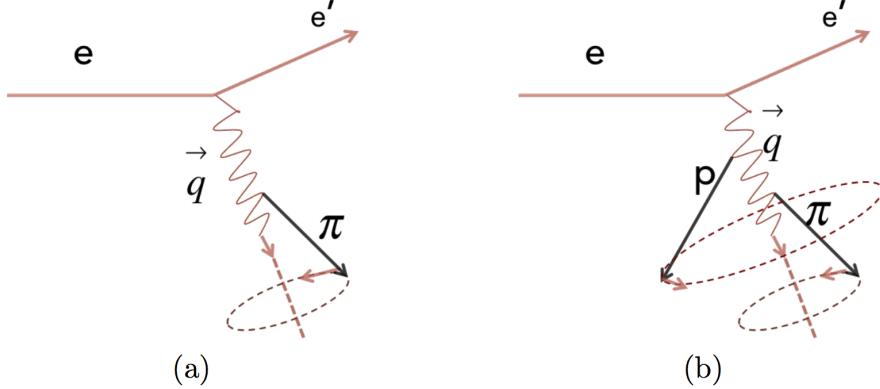


Figure 3-13: The diagrams for (a) π rotation and (b) π and proton rotation around the momentum transfer \vec{q} . By doing a random rotation repeatedly, we can calculate the probability of a multi-hadron state being erroneously detected as an event with a simpler topology.

events with detected $(e, e'\pi)$ topologies. We rotate the π momentum vector around the direction of the three-momentum transfer \vec{q} , by a randomly generated angle ϕ between 0 and 2π . Rotating around \vec{q} preserves the cross section of the reaction. Figure 3-13 shows a graphic representation of this process.

By doing this random rotation repeatedly, we can find the probability of the π being inside the acceptance region, using the pion acceptance maps. This probability is given by $P(1\pi) = N_{\text{Detected}}/N_{\text{Rotations}}$, where N_{Detected} is the number of rotations for which the rotated pion was within the acceptance region, and $N_{\text{Rotations}}$ is the total number of rotations performed. From this probability, we can estimate the number of *undetected* events with one pion, $W = -(1 - P(1\pi))/P(1\pi)$. These are events that might have erroneously passed the naive $1p0\pi$ cut. The negative sign in front means that this contribution should be subtracted from the naive $1p0\pi$ sample. Thus for each event with an electron and one detected pion, there is a single negative contribution to the event sample that has passed the pion rejection cut.

We can continue for higher multiplicities. Events with two pions might appear as 1π or 0π events. Thus for each event with two pions, we rotate both around \vec{q} to estimate the contribution to both the 0π and 1π events. This reduces the number of true 0π events, as well the true 1π events. The reduction in 1π events also reduces the

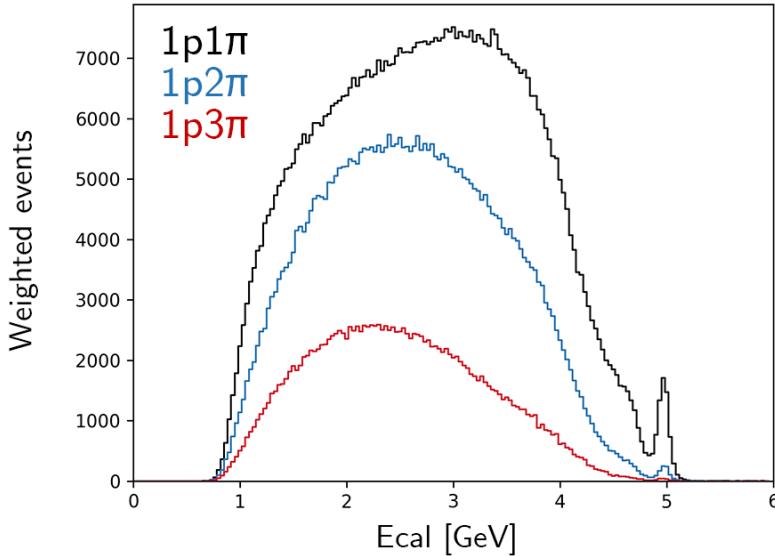


Figure 3-14: Energy reconstruction distributions for events with detected pions. The black curve shows E_{cal} for events with one detected pion, the blue curve for two pions, and the red for three pions. The 5 GeV peak all but vanishes by the three pion case. These are major components of the background and the incorrectly reconstructed tail.

subtraction of 1π events from the 0π sample, so this actually contributes an addition to the true $1p0\pi$ sample. For each pion in a detected 2π event, we can calculate the probability of it not being detected, which is the probability of erroneously measuring a 1π event. We can add a contribution, given by the product of these probability weights, to the $1p0\pi$ event sample that passes the pion rejection cut.

Similarly, events with three pions can appear as 3π , 2π , 1π , or 0π events. We can rotate these pions to estimate the number of 2π , 1π , and 0π events. Similarly, we can repeat this procedure for events with two protons and higher pion multiplicities.

The advantage of this method is that it is entirely data-driven, based on reconstructed events with higher hadron multiplicities. This subtraction also substantially reduces the tail in the energy reconstruction distributions, as these multi-hadron final states contribute substantially to the incorrectly reconstructed background events. Figure 3-14 shows the E_{cal} energy reconstruction distributions for events with detected pions. These events are mostly reconstructed incorrectly, and thus subtracting them from the distribution results in a cleaner sample with better reconstruction.

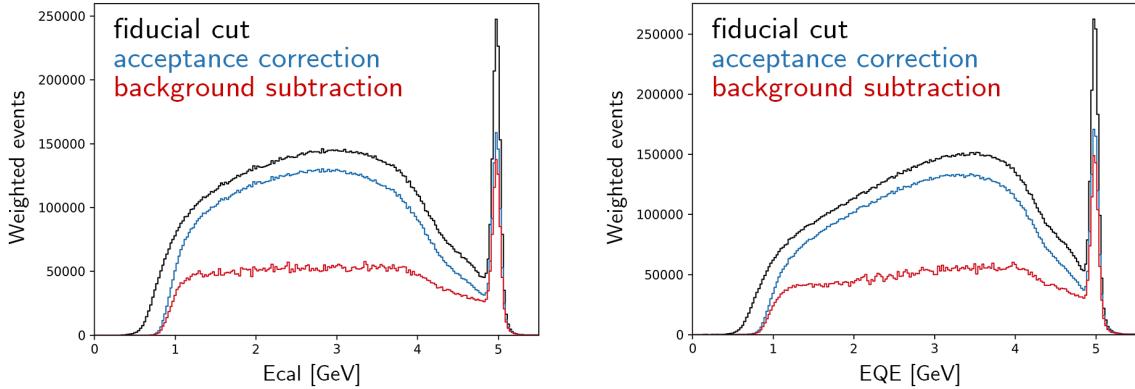


Figure 3-15: Energy reconstruction distributions for E_{cal} (left) and E_{QE} (right), after background subtraction. Note the substantial reduction in the size of the tail after the subtraction step. Nevertheless, a non-negligible tail remains in both cases.

The effect of the background subtraction on the reconstruction tail is shown in Figure 3-15. The size of the peak is not affected much, as would be expected. The tail is significantly reduced in both cases, and the shape is relatively similar between the two reconstruction techniques, as can be seen in Figure 3-16. However, a number of events remain that correspond to energy values that are off from the true beam energy. These are not quasielastic events, as we will observe in Chapter 4 when we study simulated events.

3.2.6 Radiative photons

Radiative photons occur via bremsstrahlung when the electron beam interacts with the target nucleus. We are not interested in events with radiative photons, as this is a process that is not relevant to neutrino interactions. In order to reject these events, we apply a cut on the angle of the missing momentum vector.

The events reconstructed to extremely low energies are most likely those where radiative photons were produced, because these photons carry away a large amount of energy. To see this, we can look at the events with ν between 3.5 and 4.5 GeV, where we define $\nu = E_{beam} - E_{e'}$ (the difference between the beam energy and the scattered electron energy). These events correspond to the events between 0.5-1.5 GeV in the reconstructed energy from data in Figure 3-16. To understand why the

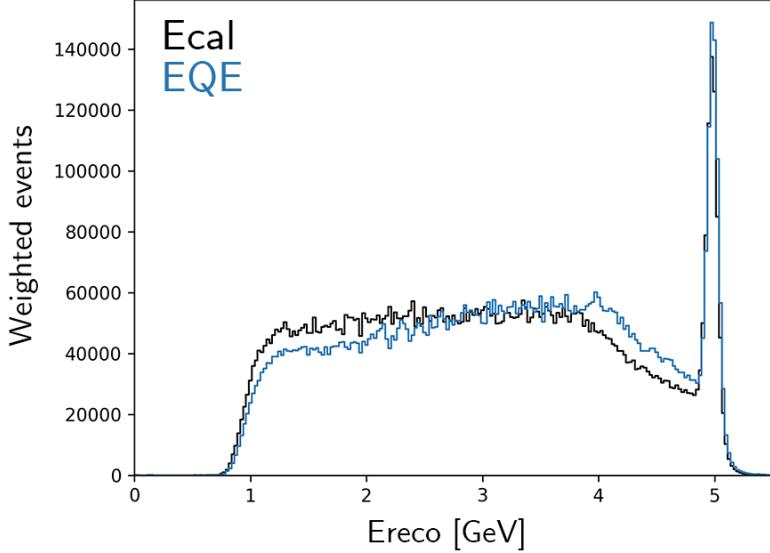


Figure 3-16: The reconstructed E_{cal} and E_{QE} energy distributions after the background subtraction step. The background tail has been reduced.

scattered electron energy is so low, we can look at E_{miss} and p_{miss} for these events.

We define these variables as follows:

$$\vec{p}_{miss} = \vec{q} - \vec{p}_p \quad E_{miss} = \nu - T_b - T_p \quad (3.4)$$

where T_b is the kinetic energy of the recoiling neutron and T_p is the kinetic energy of the proton.

If the low-energy events is due to photons, for ν between 3.5 GeV and 4.5 GeV we should see $E_{miss} \approx p_{miss}$. We should also see that E_{miss} and p_{miss} are both close to 0 for other values of ν . For ν between 0.5 and 1.5 GeV, we see in Figure 3-17 that E_{miss}, p_{miss} are both near 0. For ν between 1.5 and 3.5 GeV, in Figure 3-18, we still have most events with E_{miss}, p_{miss} near 0, but we see a tail where $E_{miss} \approx p_{miss}$. Finally, in the region of interest where ν is between 3.5 and 4.5 GeV, we see in Figure 3-19 that most events have $E_{miss} \approx p_{miss}$.

As further confirmation, we can look at the relationship between $\theta_{p_{miss}}$ and ν , and also $\theta_{p_{miss}}$ and E_{cal} , the reconstructed energy. We see in Figures 3-20 and 3-21 that $\theta_{p_{miss}}$ ranges from 0 to 180° . In Figure 3-20, we see at $\nu \approx 4$ GeV that $\theta_{p_{miss}}$ is

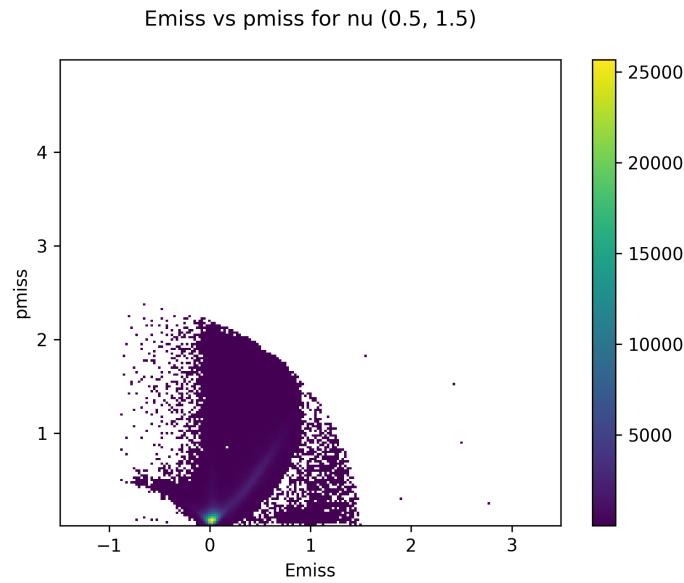


Figure 3-17: E_{miss} and p_{miss} for ν between 0 and 1.5 GeV. Both quantities are near zero, indicating that radiative photons are not present in this regime of the reconstructed energy distribution.

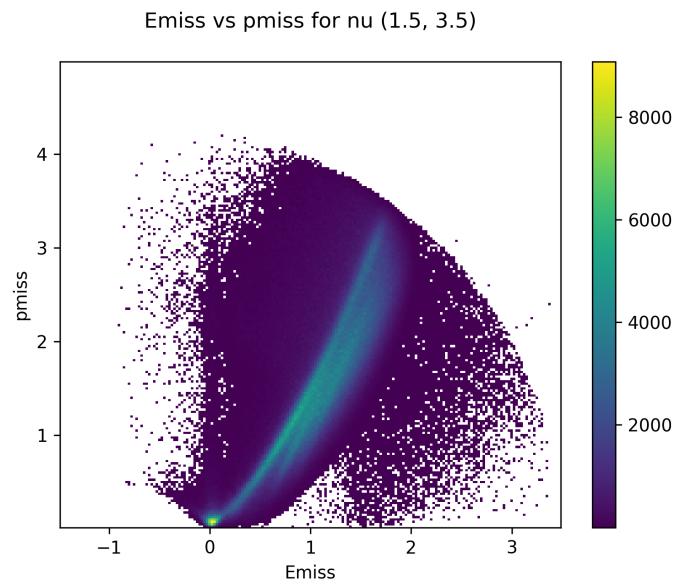


Figure 3-18: E_{miss} and p_{miss} for ν between 1.5 and 3.5 GeV. Both quantities are mostly near zero, indicating that radiative photons are not meaningfully present in this regime of the reconstructed energy distribution. There is however a tail where $E_{miss} \approx p_{miss}$, but this does not contain the majority of events.

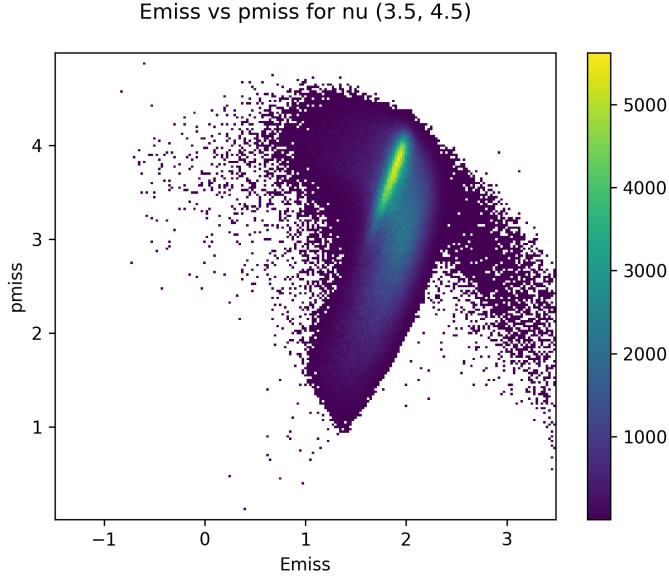


Figure 3-19: E_{miss} and p_{miss} for ν between 3.5 and 4.5 GeV. Here, most events have $E_{miss} \approx p_{miss}$, suggesting that radiative photons are present in these events.

very close to 0, i.e. along the beam axis. This is what we would expect for radiative photons – because they travel so close to the beam line, they are not detected by CLAS. In Figure 3-21, we see that these events with $\theta_{p_{miss}}$ nearly 0 occur for $E_{cal} \approx 1$ GeV, which corresponds to the left end of the distributions in Figure 3-16.

All of this suggests that the excess visible there is likely due to radiative photons, and thus is not indicative of any failing in the energy reconstruction process itself. In fact, we can apply a cut on $\theta_{p_{miss}}$ to remove these radiative events from the data and thus eliminate the peak. Because radiative photons travel down the beam line, we can make a fairly narrow cut on events with $\theta_{p_{miss}} < 3^\circ$.

Applying this cut substantially reduces the size of the incorrectly-reconstructed tail, as can be seen in Figure 3-22. This cut removes a significant number of events on the left end of the tail, corresponding to the events with the largest difference between the true energy and the reconstructed value. Figure 3-23 shows a comparison of the E_{cal} and E_{QE} methods after this cut, which is the final step in the event selection process.

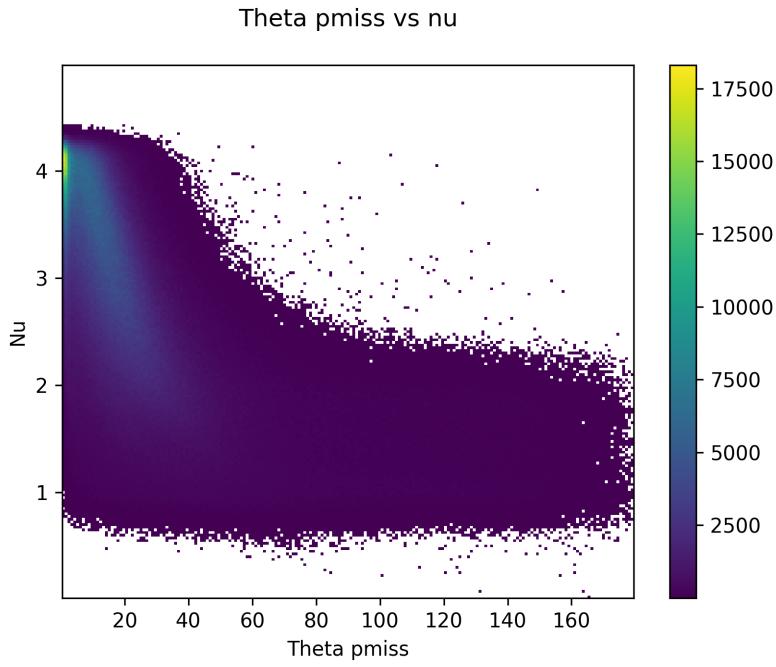


Figure 3-20: $\theta_{p_{miss}}$ and ν . Events with very low $\theta_{p_{miss}}$ are concentrated at high values of ν , particularly around $\nu = 4$. Thus suggests that particles traveling down the beam line are carrying away most of the energy missing from the reconstruction.

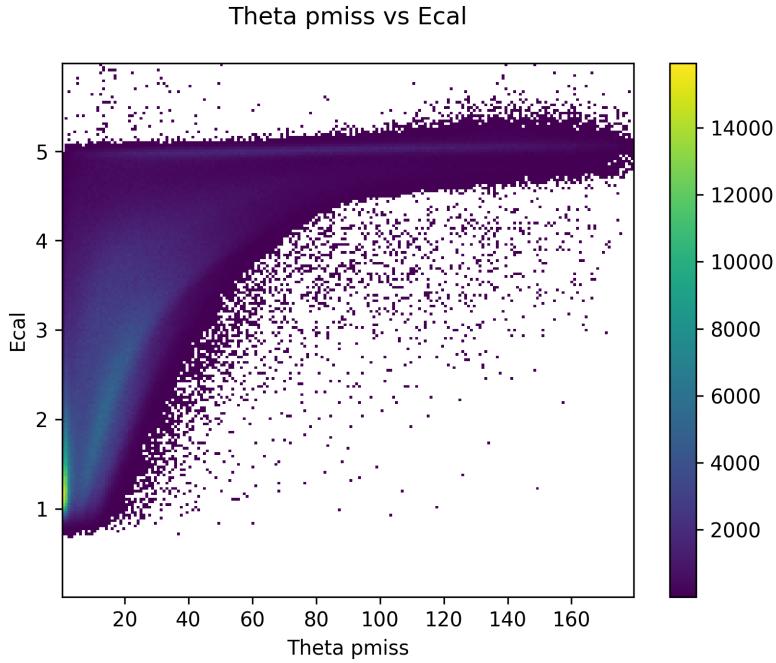


Figure 3-21: $\theta_{p_{miss}}$ and E_{cal} . Similarly to Figure 3-20, for near-zero values of $\theta_{p_{miss}}$ the energy is reconstructed very incorrectly. This suggests that radiative photons are contributing to the far end of the tail.

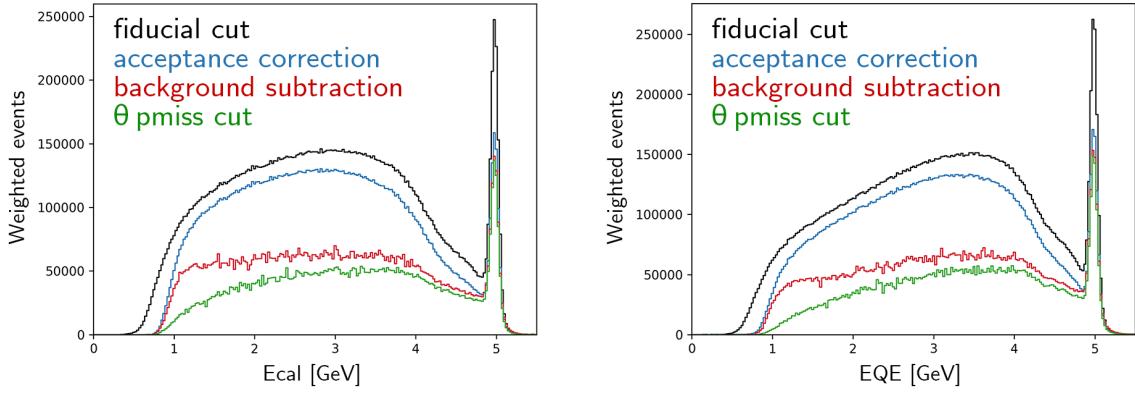


Figure 3-22: Energy reconstruction distributions for E_{cal} (left) and E_{QE} (right), after the $\theta_{p_{miss}}$ cut. Note the reduction in the tail, especially for lower E_{reco} values. This successfully eliminates radiative photon events that are irrelevant for this study.

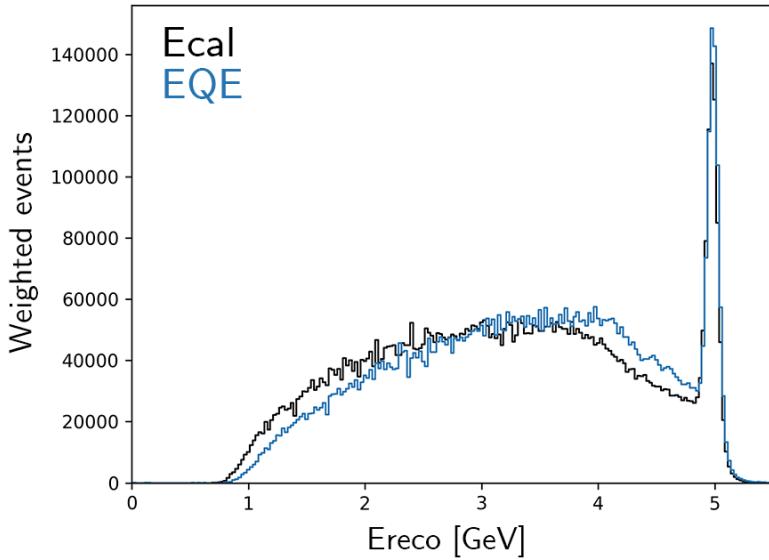


Figure 3-23: The reconstructed E_{cal} and E_{QE} energy distributions after the $\theta_{p_{miss}}$ cut. The background tail has been reduced, and the shapes and sizes are fairly similar for both methods.

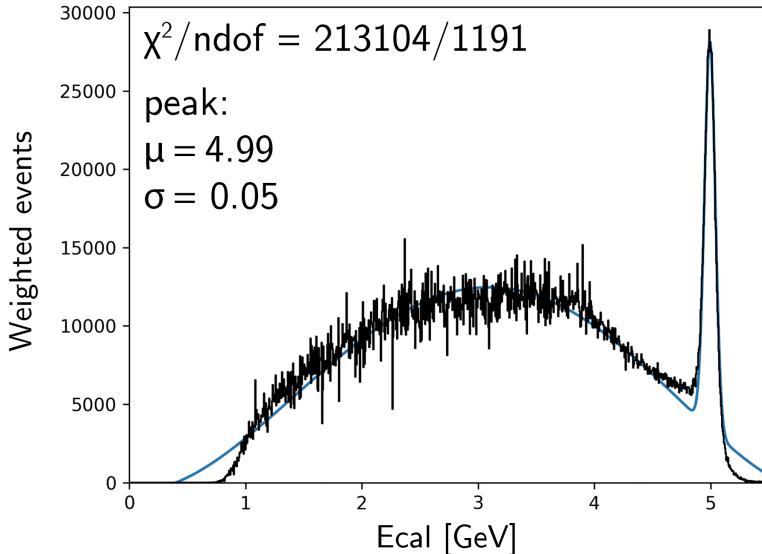


Figure 3-24: E_{cal} with a fit. The black curve is the (more finely-binned than before) E_{cal} distribution. In blue is the fit, which is composed of a Gaussian centered at the 5 GeV peak and a polynomial.

3.3 Results

The reconstruction plot in Figure 3-23 has a sharp peak, using both techniques. Both techniques have tails to the left of the peak, although the shapes are slightly different. This is likely due to pions below the momentum threshold. In both cases, the peaks do have similar widths, but the E_{QE} peak is narrower than the E_{cal} peak. We can see this by performing a fit to the E_{reco} distributions in both cases. I fit the distributions with the sum of a Gaussian distribution and a fifth-degree polynomial, where the former accounts for the peak at 5 GeV and the latter suffices to describe the tail. The results of the fits are shown in Figures 3-24 and 3-25 (these plots are more finely-binned than the distributions shown previously). For the E_{cal} distribution, the width of the Gaussian peak is 0.0463 ± 0.0007 , and it is centered at 4.9943 ± 0.0007 . Additionally, 90.2% of events lie below 3σ of the peak in this case. For the E_{QE} distribution, the width of the Gaussian peak is 0.0440 ± 0.0006 , and it is centered at 5.0015 ± 0.0006 GeV. Here, 89.3% of events lie below 3σ of the peak. Thus the E_{QE} appears to be a better description of the true energy, but only marginally.

Figure 3-26 shows E_{cal} as a function of p_T , the transverse missing momentum.

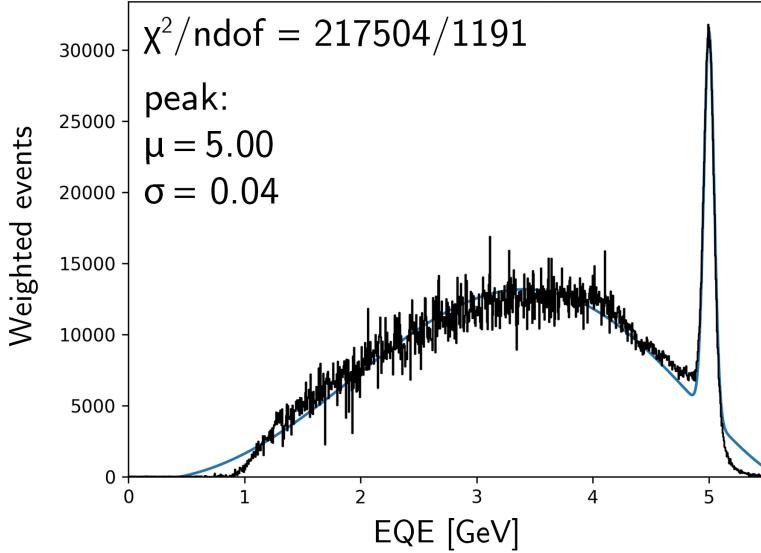


Figure 3-25: E_{QE} with a fit. The black curve is the (more finely-binned than before) E_{QE} distribution. In blue is the fit, which is composed of a Gaussian centered at the 5 GeV peak and a polynomial.

This is simply the sum of the transverse momentum of the scattered electron and scattered proton. Taking z to be the direction of the beam line, $\vec{p}_{T,e'} = (p_{x,e}, p_{y,e}, 0)$ and $\vec{p}_{T,p} = (p_{x,p}, p_{y,p}, 0)$. Then, simply, $\vec{p}_T = \vec{p}_{T,e'} + \vec{p}_{T,p}$. Since the electron beam carries all its momentum in the z direction, and the proton is assumed to be at rest in the nucleus for the quasielastic model, the transverse momentum is a proxy for understanding the electron-nucleus interaction. Without nuclear effects, it should be 0. It is sensitive to the Fermi momentum of nucleons, and the nuclear interaction energy.

The transverse missing momentum thus points to the presence of non-quasielastic processes. If the nucleons in the $D(e, e'p)$ reaction carried some transverse momentum, or there was a more complicated final-state topology, we expect to see higher a p_T value for those events. In the figure, the quasielastic events of interest appear as a distinct peak in the distribution for small values of p_T , with E_{cal} close to the beam energy. The subtraction process has reduced the non-quasielastic background at higher values of p_T , but some events remain. The majority of the quasielastic events occur with $p_T < 200$ MeV, which is the expected result for a quasielastic knockout of

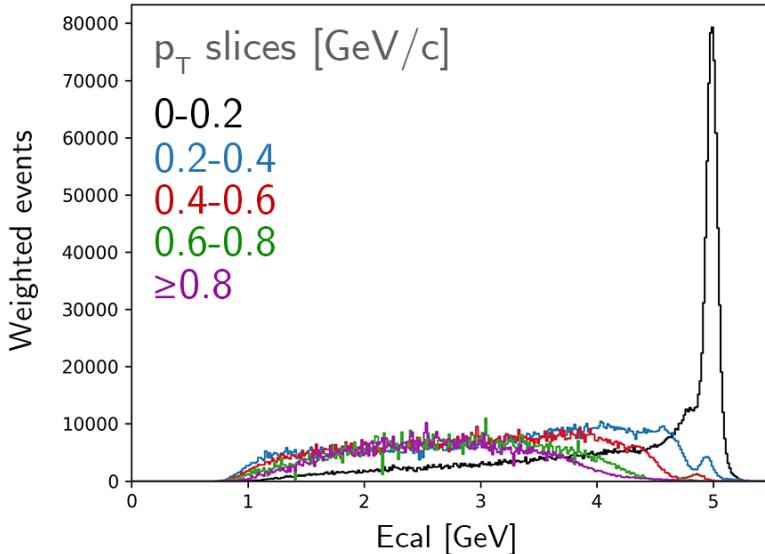


Figure 3-26: E_{cal} reconstruction distributions over different slices in the transverse missing momentum, p_T . For low p_T , most events are quasielastic and reconstructed properly. The background subtraction procedure reduced, but did not remove, events with higher p_T .

a mean field proton. Figure 3-27 shows the same plot for the deuterium-specific E_{QE} reconstruction. The main features are the same.

Transverse kinematic variables like p_T are interesting to consider, because they do not depend on the beam energy. Other quantities commonly used for event selection in electron-scattering data like Q^2 , the invariant mass W , and Bjorken- x x_B depend on precise knowledge of the beam energy. This makes them poor tools for neutrino analyses, because the energy of the incoming neutrino is not known. Transverse variables can be calculated without this information, and thus can be helpful tools for further selection of quasielastic events. In addition, we consider ϕ_T and α_T , defined as below [2, 20, 33].

$$\phi_T = \arccos \frac{-\vec{p}_{T,e'} \cdot \vec{p}_{T,p}}{|\vec{p}_{T,e'}| |\vec{p}_{T,p}|} \quad (3.5)$$

$$\alpha_T = \arccos \frac{-\vec{p}_{T,e'} \cdot \vec{p}_T}{|\vec{p}_{T,e'}| |\vec{p}_T|} \quad (3.6)$$

Physically, ϕ_T is a measure of the coplanarity between the scattered electron and

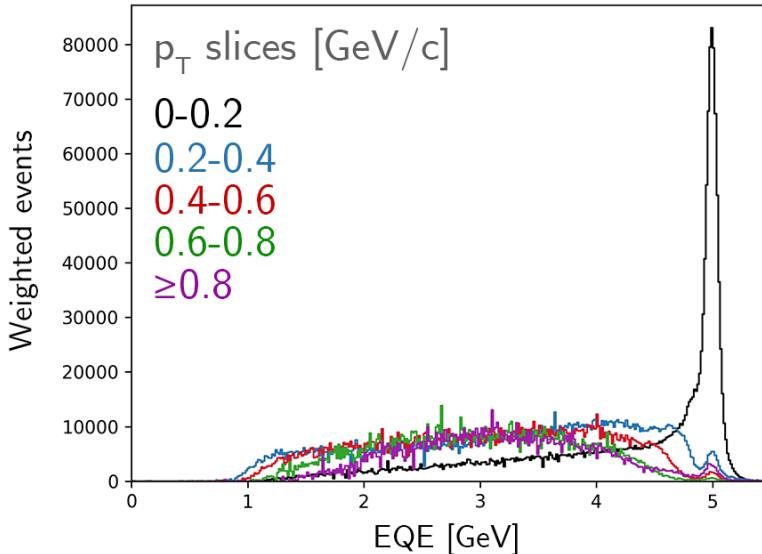


Figure 3-27: E_{QE} reconstruction distributions over different slices in the transverse missing momentum, p_T . For low p_T , most events are quasielastic and reconstructed properly. As with E_{cal} , the background subtraction procedure reduced, but did not remove, events with higher p_T .

proton. It is the angle between the plane containing the beam and the scattered electron, and the plane containing the beam and the scattered proton. Because it can be calculated from p_T measurements, which do not depend on E_{beam} , it is also independent of the beam energy. The other quantity, α_T is the projection of the angle of the recoil system (in this case, the neutron in deuterium) onto the transverse plane. If the neutron is not involved in the interaction, it should have the same momentum and angle that it did before the interaction, and α_T should thus be isotropic. These transverse variables facilitate the possible identification of Fermi motion of the initial state nucleon, final state-interactions of the nucleons in the nucleus, and multi-nucleon interactions.

Figure 3-28 shows these quantities labeled on a diagram for a neutrino event (the same as for the present case, just with a difference in labeling).

Figures 3-29, 3-30, and 3-31 show plots of these transverse variables. Note that a large number of events (Figure 3-29) have high values for p_T ; these generally belong to the background. In Figure 3-30, we see that most events have a low value for ϕ_T , indicating a high degree of coplanarity between the scattered electron and proton. In

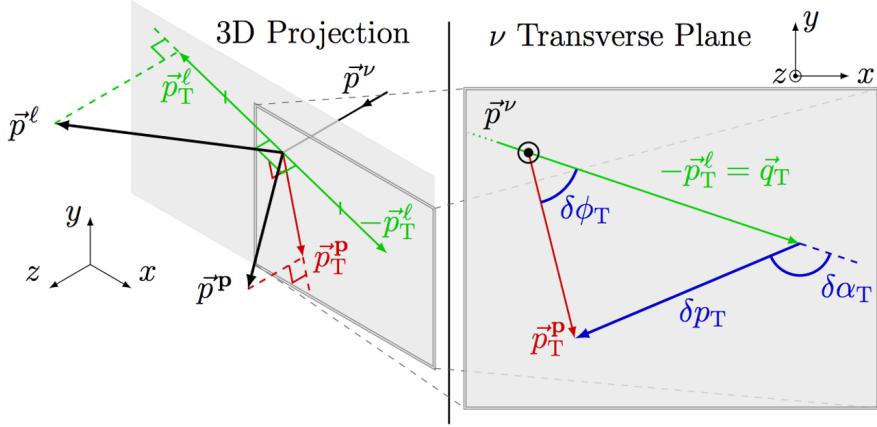


Figure 3-28: Transverse variables, diagrammed for a neutrino event. The quantities \vec{p}_T^ℓ and \vec{p}_T^p are the transverse momenta of the scattered lepton and proton respectively. Their sum is \vec{p}_T (labeled on the right as δp_T), the transverse missing momentum. Also in blue on the right is ϕ_T (labeled $\delta\phi_T$), the measure of coplanarity between the scattered lepton and proton. The last variable, α_T (labeled $\delta\alpha_T$) is the angle of the recoil system projected onto the transverse plane.

Figure 3-31, we see that α_T is not isotropic, suggesting that nuclear effects do play a role in the events that remain after the selection process used here.

Figures 3-32, 3-33, and 3-34 show two-dimensional plots of the transverse variables and the E_{cal} reconstruction values. Here we can see that events with low p_T and ϕ_T were generally reconstructed correctly. From the α_T plot in Figure 3-34, we see that correctly reconstructed events had isotropic α_T values. From these plots, it is possible to suggest threshold values for these transverse variables to further refine the quasielastic sample and obtain a better reconstruction result. However, because neutrino experiments often have reduced statistics relative to electron scattering data, I do not propose any specific values for these cuts – it might be more efficient for neutrino studies to make more generous cuts than this data suggests.

This concludes the portion of this study using electron-scattering data. We have found that it is not straightforward to eliminate the tail of the energy reconstruction distribution, and that nuclear effects seem to be playing a significant role in the incorrectly reconstructed events. In the next chapter, we look at simulated electron-scattering data using GENIE to understand the processes that lead to this tail.

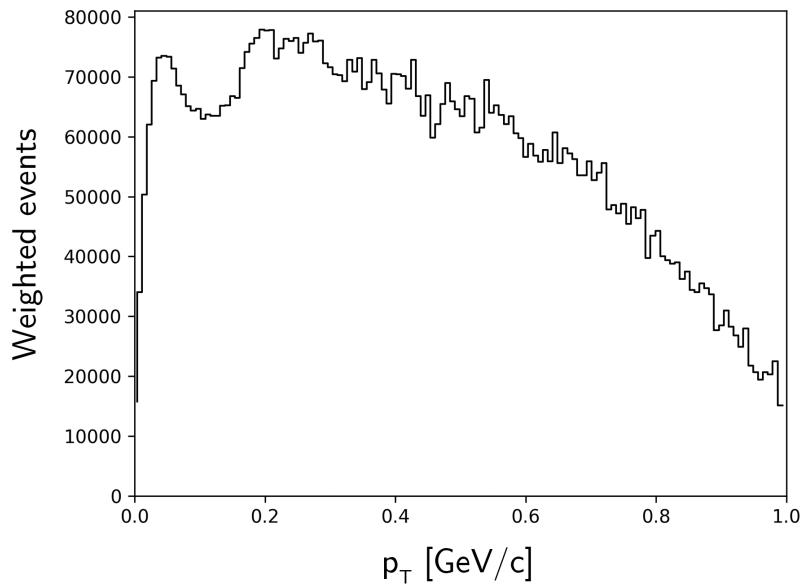


Figure 3-29: Transverse missing momentum distribution. Events with low p_T were generally reconstructed correctly.

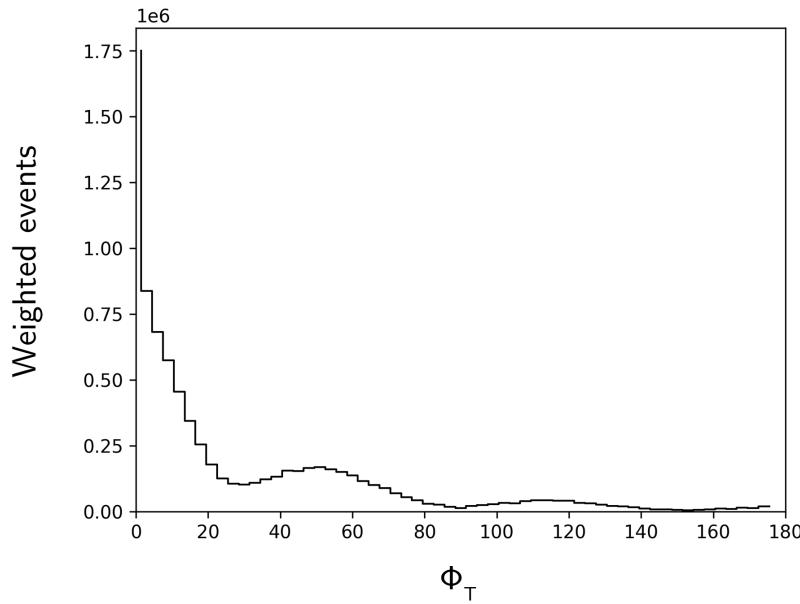


Figure 3-30: ϕ_T distribution. Most events have a small value for ϕ_T , indicating a high degree of coplanarity between the scattered electron and proton. The bump-like behavior is due to the six sectors of CLAS and the imperfect acceptance near their edges.

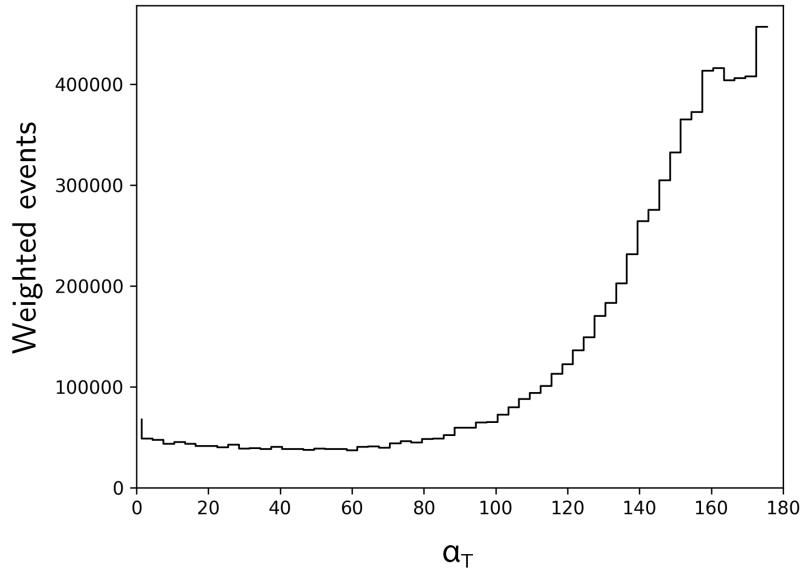


Figure 3-31: α_T distribution. If we had a purely quasielastic sample with minimal nuclear effects, this would be isotropic. This is not quite the case here.

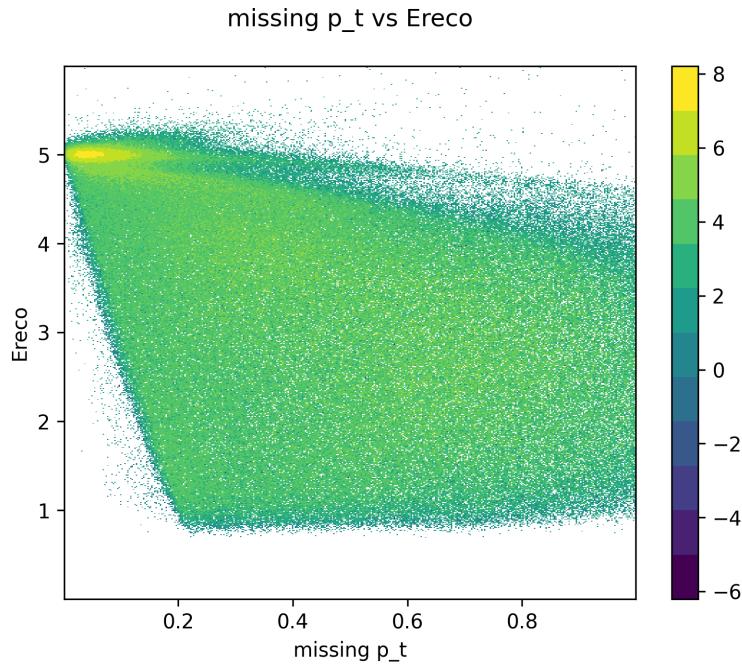


Figure 3-32: 2D plot of p_T and E_{cal} (logarithmic event counts). Most of the events that are reconstructed correctly ($E_{cal} \approx 5$) have $p_T < 0.2$ GeV/c. The edge on the left side of the plot is due to the $\theta_{p_{miss}}$ cut.

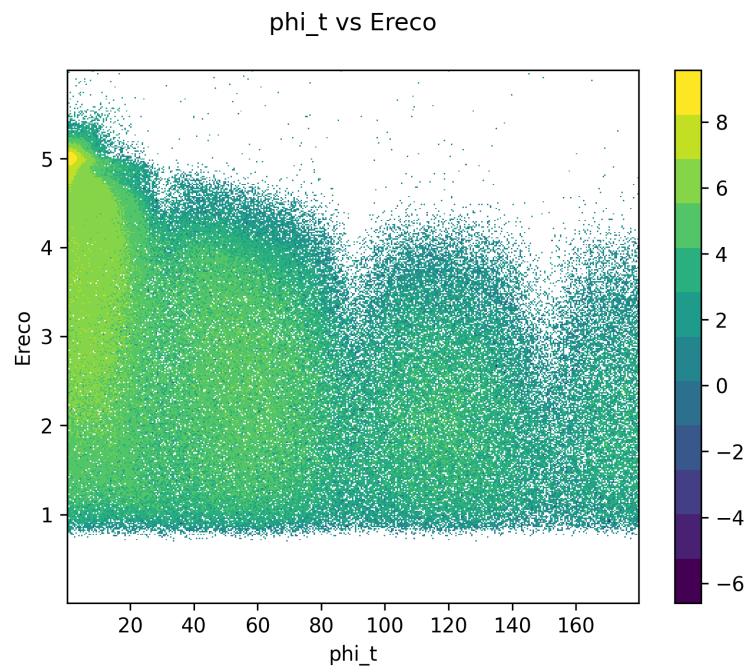


Figure 3-33: 2D plot of ϕ_T and E_{cal} (logarithmic event counts). Most of the events that are reconstructed correctly ($E_{cal} \approx 5$) have $\phi_T < 30^\circ$.

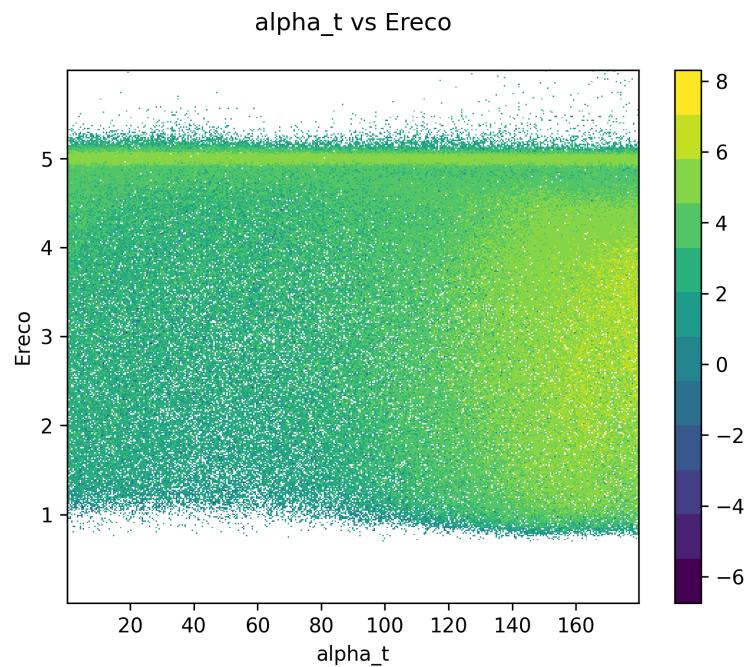


Figure 3-34: 2D plot of α_T and E_{cal} (logarithmic event counts). The band across the top of the plot corresponds to events that correctly reconstructed. For these events, α_T is isotropic. However, for the events in the background (below the band), a similar isotropy is not observed. Most of these events have high values of α_T , indicating that the deuteron's neutron had a large backward recoil angle.

Chapter 4

Energy reconstruction analysis on GENIE simulation

*Humankind cannot bear
very much reality.*

T. S. Eliot, Four Quartets

Because the energy reconstruction is not perfect, we can simulate it using neutrino event generators like GENIE [11] in order to better understand the underlying neutrino-nucleus interactions. In this section, we repeat the event selection and reconstruction machinery on simulated events, and study the results to learn more about the reconstruction process and obtain a better understanding with respect to the underlying interactions.

4.1 About GENIE

GENIE is one of the most popular neutrino event generators used for neutrino experiments, particularly those at Fermilab. It is a ROOT-based software, first released in 2007. It implements Monte Carlo simulations of neutrino interaction models, and plays a crucial role in the analysis and interpretation of oscillation experiments.

For this study, I have used *e*-GENIE, which is an electron-scattering version of

GENIE [40]. Wherever possible, *e*-GENIE uses the same code and the same models as the standard neutrino version. It was constructed by setting the axial part of neutrino-nucleus interaction to zero (since electrons interact via vector currents only), and using the appropriate new cross-sections. It has been benchmarked against electron scattering data.

GENIE samples an initial energy E for each event using a probability density of the form

$$P(E) \propto \Phi(E) \sum_i \sigma_i(E) \quad (4.1)$$

where Φ is the incident flux and the sum runs over the cross sections σ_i for each available interaction mode. These interaction modes include quasielastic interactions, resonant pion production, deep-inelastic scattering, and meson exchange currents. All but the latter are relevant for this study. GENIE does not include interference between the amplitudes of different interaction channels [40].

The models used for this study use the GENIE implementation of the local Fermi gas model for the nuclear ground state. Here, the Fermi momentum at a given radial position depends on the local nuclear density. To incorporate this dependence, GENIE samples an interaction location inside the nucleus according to the nuclear density. Then it draws an initial momentum for the struck nucleon from a Fermi distribution using the local Fermi momentum at that location.

For this study, we use the GENIE models from the `G18_10a_02_11a` configuration of GENIE v3.0.6, which I will refer to as G2018. These models were tuned on bubble chamber charged-current data. Resonant pion production is modeled with the Berger-Sehgal model [15], and deep-inelastic scattering with Bodek and Yang [19].

The *e*-GENIE G2018 configuration models final state interactions (FSIs) using the IntraNuclear Cascade model. This is based on an empirical data-driven method that makes use of pion and nucleon cross-sections with nuclei as a function of their kinetic energies. The simulated *e*-GENIE data used here was run with radiative corrections as well (these are not officially implemented in GENIE yet). There are a variety of radiative processes in electron-scattering – incoming and outgoing electrons can

radiate photons, which change the kinematics of the event. The radiative corrections implemented in this version of GENIE were validated via comparison with Jefferson Lab data [40].

4.2 Event selection

The same event selection machinery was used for the GENIE events as for the CLAS data events. The fiducial and acceptance maps accounted for detector effects, and the pion subtraction and $\theta_{p_{miss}}$ cuts were implemented as before. The latter cut has minimal impact on the event selection process, so its effect is not explicitly shown in the figures in this section.

In order to get slightly more realistic results, I added some smearing to the scattered electron and proton momenta. Using the values generated by GENIE, I resampled these quantities for each event using a Gaussian distribution centered at the GENIE value, with a width of 10 MeV for the proton and 5 MeV for the electron.

4.3 Energy reconstruction results

Figure 4-1 shows the calorimetric reconstruction on GENIE events that pass the event selection process. The black curve shows the initial sample after the fiducial cuts, the blue shows the effect of the acceptance cuts and correction, and the red is the result after background subtraction (and the $\theta_{p_{miss}}$ cut, whose effect is negligible). A clear peak is evident at the beam energy (5 GeV), but a substantial tail remains – as it did with the CLAS data. Figure 4-2 shows the same sequence with the deuterium-specific E_{QE} reconstruction. The main features remain the same. Figure 4-3 shows a comparison of the two methods. Both methods result in similar peaks and tails of similar sizes, although the shapes slightly differ (as was the case with the CLAS data).

Note that there is a gap of sorts between the 5 GeV peak in the reconstruction and the tail. This is due to known issues in GENIE with simulations of off-shell pion

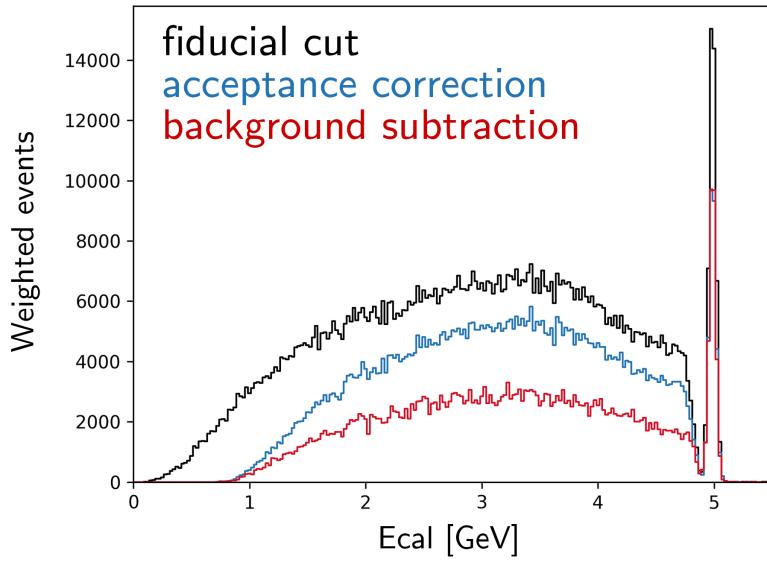


Figure 4-1: E_{cal} reconstruction for $D(e, e'p)$ events generated by GENIE using the G2018 configuration. A clear peak is visible at the beam energy of 5 GeV, but a substantial tail remains.

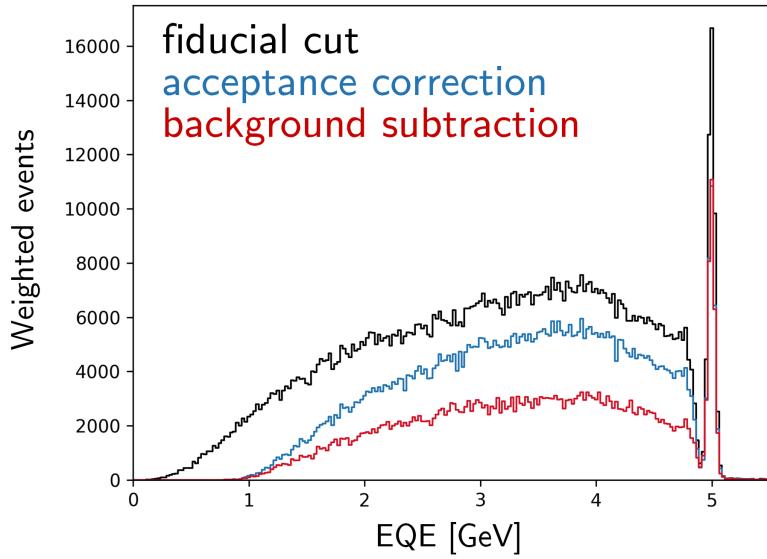


Figure 4-2: E_{QE} reconstruction distribution on GENIE events, using the deuterium-specific formula. A clear peak is visible, but as with the E_{cal} method, a large tail remains.

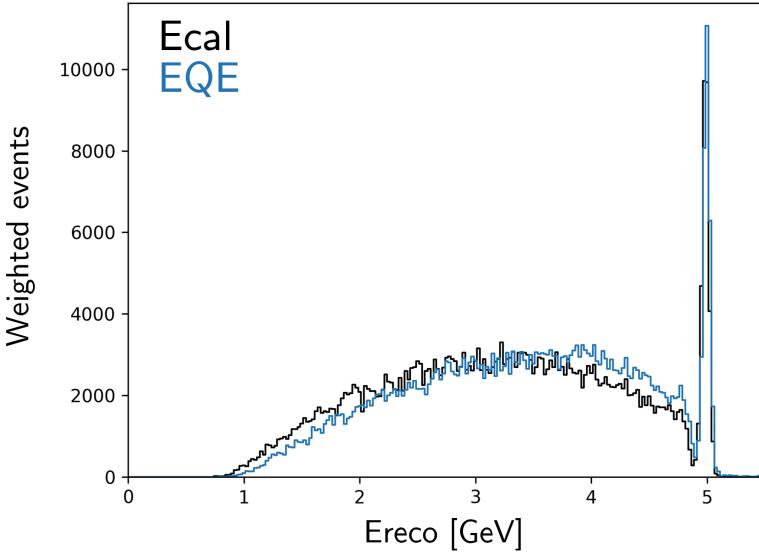


Figure 4-3: A comparison of the E_{cal} and E_{QE} methods of reconstruction on GENIE events. As with the CLAS data, there is a slight difference in the shape of the tails, but the sizes remain similar.

interactions [11]. This leads to a gap in events that are reconstructed to nearly the beam energy.

4.3.1 Background subtraction validation

The GENIE events can also be used to validate the background subtraction procedure through a series of closure tests. I generated subset samples with specific topologies to test the method described earlier. These samples consisted of $1p0\pi$ events and specific higher-multiplicity events (for example, $1pN\pi$ events). These were not necessarily true $1pn\pi$ events, though – using information about the CLAS acceptance, I selected events that would have been detected with these topologies. It was then possible to run the specific $1pn\pi$ channel of the subtraction procedure to observe how the subtracted $1p0\pi$ sample compared to the true $1p0\pi$ sample.

Figures 4-4, 4-5, and 4-6 show the results of this test on samples with $1p0\pi$ events and $1p1\pi$, $1p2\pi$, and $1p3\pi$ events respectively. The black curve in all three plots is the initial sample, with the $1p0\pi$ events and the multi-hadron events of choice. The blue curve is the result of running the background subtraction procedure on this

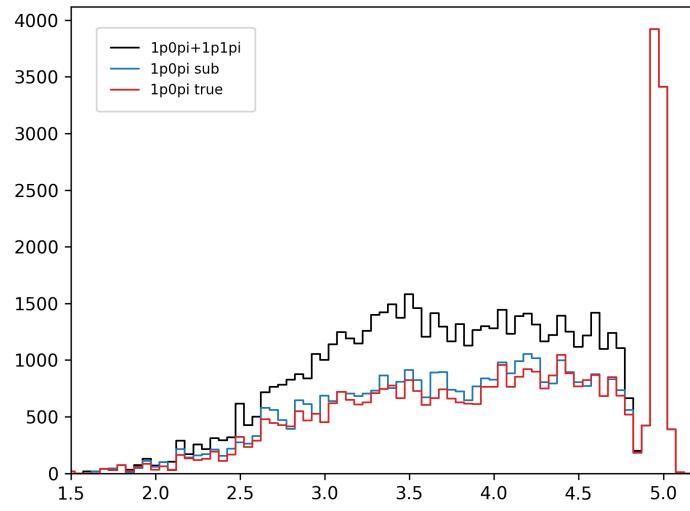


Figure 4-4: $1p1\pi$ subtraction validation. The black curve is the combined sample, with $1p0\pi$ and $1p1\pi$ events. The blue is the result of running the background subtraction process on the whole sample. The red is the true $1p0\pi$ sample from the original file. The blue and the red are quite close, suggesting that the subtraction process is working well.

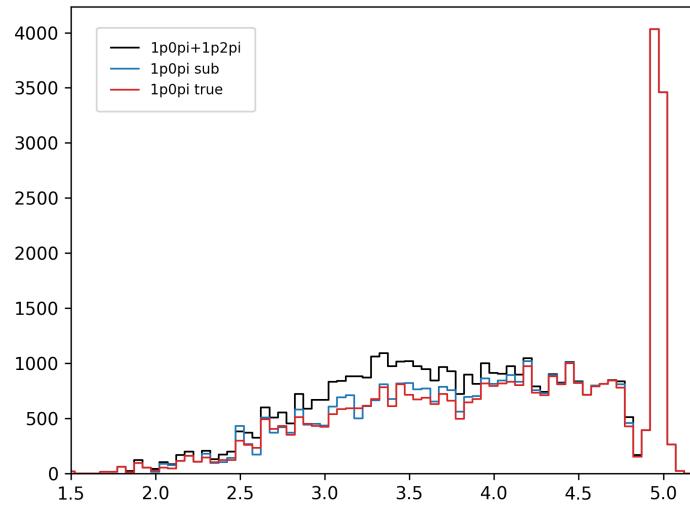


Figure 4-5: $1p2\pi$ subtraction validation. The black curve is the combined sample. The blue is the result of the background subtraction process. The red is the true $1p0\pi$. The blue and the red are very similar, validating the subtraction process.

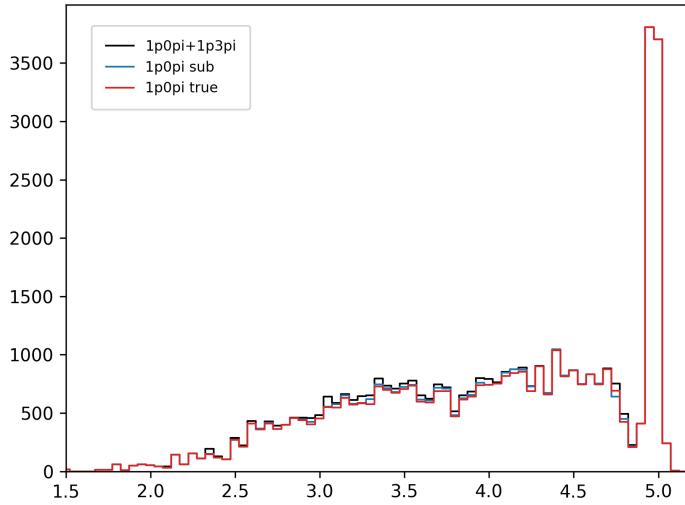


Figure 4-6: $1p3\pi$ subtraction validation. The subtracted and true $1p0\pi$ samples (blue and red) are similar, validating this subtraction method.

sample. The red is the true $1p0\pi$ sample. In all cases, the blue and the red are quite similar, suggesting that this method is working well. They will not be identical, given the randomized nature of the subtraction step. This lends confidence to this technique and suggests that we have mostly removed these multi-hadron events from the background.

4.3.2 Background processes

One of the advantages of studying the GENIE events is that we can look at the distribution of events in the reconstruction distribution tail. This is shown in Figure 4-7 for the calorimetric method, and Figure 4-8 for the quasielastic method. The quasielastic events, shown in blue in both figures, are almost entirely reconstructed correctly. These distributions are concentrated in the 5 GeV peak. The resonant pion and deep-inelastic scattering events, in red and green respectively, are generally not reconstructed properly. This is unsurprising, given that these events produce other particles whose properties are not considered in the two reconstruction techniques being studied here.

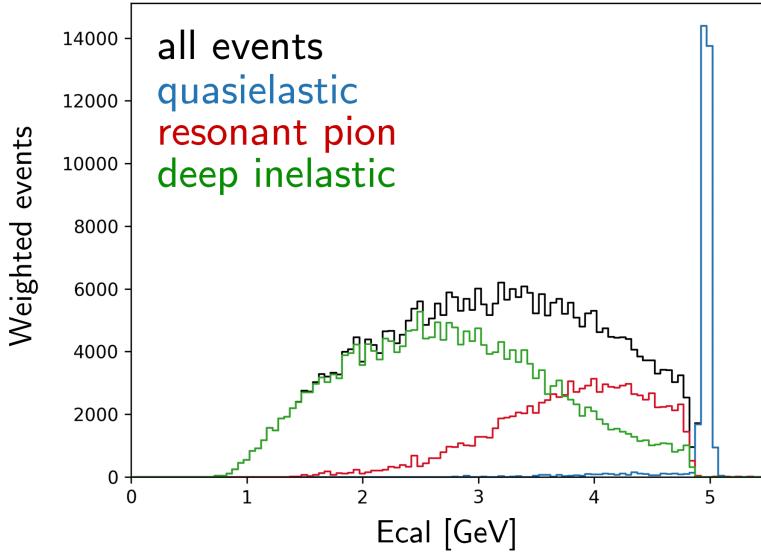


Figure 4-7: Breakdown of events in E_{cal} energy reconstruction distribution by interaction channel. The total distribution is shown in black. The quasielastic events are shown in blue, and almost entirely located in the correctly-reconstructed peak at 5 GeV. The events in red are resonant pion events, and these are generally reconstructed incorrectly. The green events are for deep-inelastic scattering interactions, also not correctly reconstructed.

However, this does suggest that particles below the CLAS thresholds are driving the large amount of background in these distributions, since the results from the previous section suggest that the background process for detected particles is adequately handling the issue of uneven acceptance in the CLAS detector.

4.4 Comparison between data and GENIE

We can compare the results of these two energy reconstruction techniques on CLAS data and GENIE events. In order to accurately compare these, we normalize the distributions using a simple area-normalization technique. The results for the E_{cal} method are shown in Figure 4-9, and for the deuterium-specific quasielastic technique in Figure 4-10.

In both cases, the GENIE peak is narrower than the peak in data. We see the gap between the peak and the tail in the GENIE distribution, which is not visible in the data distribution. The shapes of the backgrounds tails are fairly similar. There

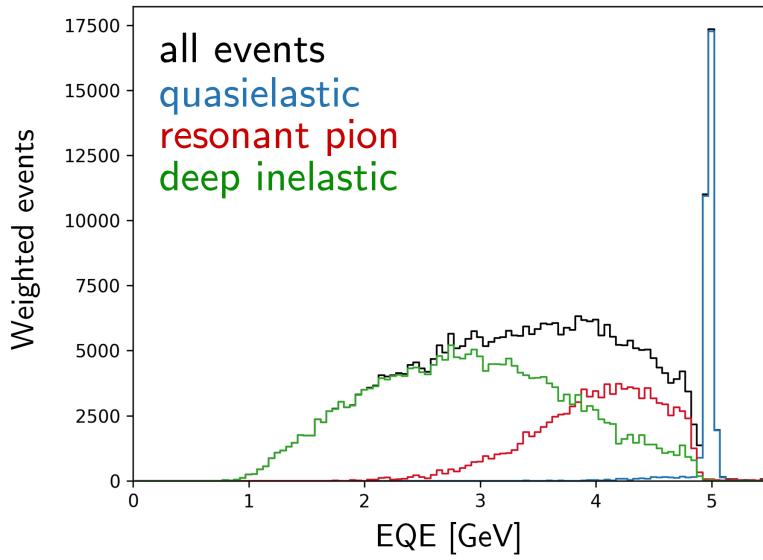


Figure 4-8: Breakdown of events in E_{QE} energy reconstruction distribution by interaction channel. The total distribution is shown in black. The quasielastic events, shown in blue, are almost all in the 5 GeV peak. The resonant pion events (red) and deep-inelastic scattering events (green) are mostly not correctly reconstructed.

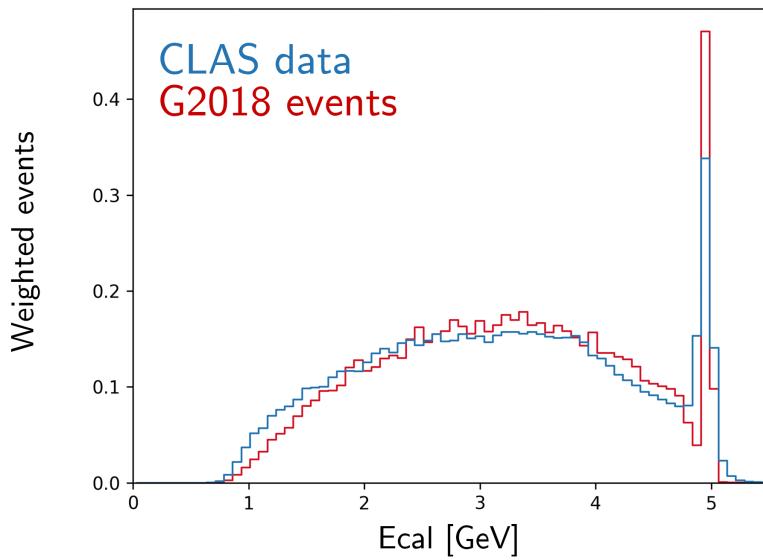


Figure 4-9: Comparison between E_{cal} distributions for CLAS data (blue) and GENIE events (red).

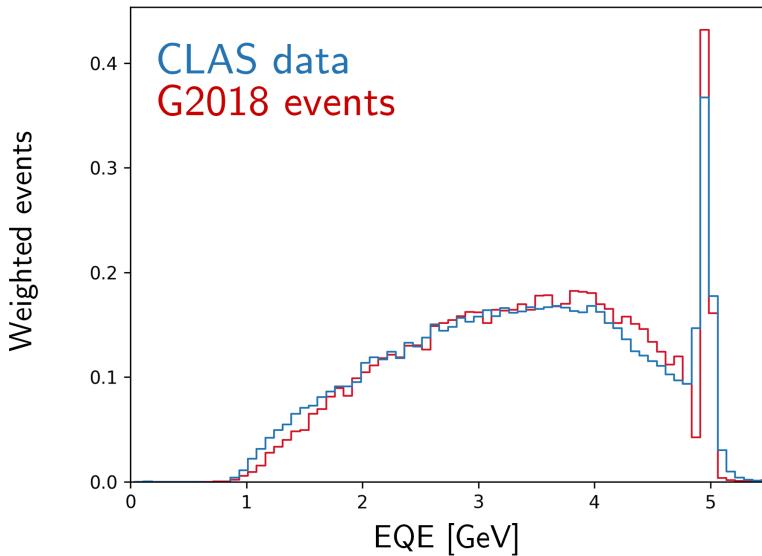


Figure 4-10: Comparison between E_{QE} distributions for CLAS data (blue) and GENIE events (red).

appears to be a shift towards higher reconstruction values in the GENIE sample. The level of agreement is somewhat surprising, given that e -GENIE is under active development as are GENIE's deuterium models. Nevertheless, this is encouraging for energy reconstruction pipelines, since the event generators results seem to be not entirely dissimilar from reality.

Chapter 5

Conclusions

*In real life, I assure you, there is
no such thing as algebra.*

Fran Lebowitz

Although theoretical studies suggested that the energy reconstruction process could work perfectly with a deuterium target, this study has shown that when tested on both real and simulated events, things are not so straightforward. There is a background of resonant and deep-inelastic events that result in a fairly substantial tail of incorrectly-reconstructed energy values. This is likely attributable to particles below the CLAS momentum thresholds.

Of the energy reconstruction methods considered, the E_{QE} method designed for deuterium events was shown to be fairly similar to the standard E_{cal} method that is in use for LArTPC-based oscillation experiments now. However, the deuterium-specific technique does seem to be marginally better. The E_{QE} peak is narrower than the E_{cal} peak, and fewer events are contained in the E_{QE} tail than the E_{cal} tail. We found a relatively high level of agreement between the GENIE simulation models and the CLAS data, which is also a validation of current *e*-GENIE development.

While this study seems to suggest that there is an irreducible background in the energy reconstruction distributions, even for a simple target like deuterium, the amount of background could potentially be smaller in neutrino experiments. The

momentum thresholds for detectability in neutrino experiments are generally lower, and it might be easier in practice to identify and ignore multi-hadron final states. Thus although a perfect energy reconstruction is unlikely to be completely achievable, future deuterium- or heavy-water-based neutrino detectors might have potential.

The next step of this work is to more deeply investigate the nature of the events that did not reconstruct correctly. Although we suspect that the CLAS thresholds are the main reason for the large tail, further work could investigate this expectation. It would also be possible to use the GENIE simulations to study the particles produced in these background events beyond pions, to understand what is carrying away the undetected energy. This would strengthen our ability to offer concrete insights for neutrino oscillation experiments. The reconstruction process is one of the main challenges for the next generation of oscillation experiments, and this study offers some clarity and directions for new work in addressing this problem.

Bibliography

- [1] K. Abe et al. T2K neutrino flux prediction. *Phys. Rev. D*, 87, 2013.
- [2] K. Abe et al. Characterization of nuclear effects in muon-neutrino scattering on hydrocarbon with a measurement of final-state kinematics and correlations in charged-current pionless interactions at T2K. *Phys. Rev. D*, 98, 2018.
- [3] K. Abe et al. T2K measurements of muon neutrino and antineutrino disappearance using 3.13×10^{21} protons on target. *Phys. Rev. D*, 103, 2021.
- [4] P. Abratenko et al. First measurement of differential charged current quasielasticlike ν_μ -argon scattering cross sections with the MicroBooNE detector. *Phys. Rev. Lett.*, 125, 2020.
- [5] R. Acciarri et al. Long-baseline neutrino facility (LBNF) and deep underground neutrino experiment (DUNE). 2016.
- [6] M. A. Acero et al. First measurement of neutrino oscillation parameters using neutrinos and antineutrinos by NOvA. *Phys. Rev. Lett.*, 123, 2019.
- [7] M.A. Acero et al. New constraints on oscillation parameters from ν_e appearance and ν_μ disappearance in the NOvA experiment. *Phys. Rev. D*, 98, 2018.
- [8] A. A. Aguilar-Arevalo et al. Significant excess of electronlike events in the Mini-BooNE short-baseline neutrino experiment. *Phys. Rev. Lett.*, 121, 2018.
- [9] Q.R. Ahmad et al. Direct evidence for neutrino flavor transformation from neutral current interactions in the Sudbury Neutrino Observatory. *Phys. Rev. Lett.*, 89, 2002.
- [10] F. An et al. Neutrino physics with JUNO. *J. Phys. G: Nucl. Part. Phys.*, 43(3), 2016.
- [11] C. Andreopoulos et al. The GENIE neutrino Monte Carlo generator. 2015.
- [12] A. M. Ankowski et al. Comparison of the calorimetric and kinematic methods of neutrino energy reconstruction in disappearance experiments. *Phys. Rev. D*, 92, 2015.
- [13] A. Ashkenazi et al. Electrons for neutrinos: Addressing critical neutrino-nucleus issues. Run group proposal, Jefferson Lab PAC 45, 2017.

- [14] C. Athanassopoulos et al. Candidate events in a search for $\bar{\nu}_\mu \rightarrow \bar{\nu}_e$ oscillations. *Phys. Rev. Lett.*, 75, 1995.
- [15] C. Berger and L. M. Sehgal. Lepton mass effects in single pion production by neutrinos. *Phys. Rev. D*, 76, 2007.
- [16] J. Beringer et al. Review of particle physics. *Phys. Rev. D*, 86:010001, Jul 2012.
- [17] M. Betancourt et al. Using electron scattering to constrain neutrinos. Letter of interest, Snowmass 2021, 2020.
- [18] M. Blennow and A.Y. Smirnov. Neutrino propagation in matter. *Adv. High Energy Phys.*, 2013, 2013.
- [19] A. Bodek and U.K. Yang. Higher twist, ξ_W scaling, and effective LO PDFs for lepton scattering in the few GeV region. *J. Phys. G: Nucl. Part. Phys.*, 29, 2003.
- [20] T. Cai et al. Nucleon binding energy and transverse momentum imbalance in neutrino-nucleus reactions. *Phys. Rev. D*, 101, 2020.
- [21] S. Chatrchyan et al. Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Phys. Lett. B*, 716, 2012.
- [22] J. M. Conrad, W. C. Louis, and M. H. Shaevitz. The LSND and MiniBooNE oscillation searches at high δm^2 . *Annual Review of Nuclear and Particle Science*, 63, 2013.
- [23] U. Dore, P. Loverre, and L. Ludovici. History of accelerator neutrino beams. *Eur. Phys. J. H*, 44, 2019.
- [24] L. El Fassi et al. Search for the onset of color transparency via ρ^0 electroproduction on nuclei. Approved CLAS analysis note, 2012.
- [25] I. Esteban, M.C. Gonzalez-Garcia, M. Maltoni, T. Schwetz, and A. Zhou. The fate of hints: updated global analysis of three-flavor neutrino oscillations. *J. High Energ. Phys.*, (178), 2020.
- [26] G. Fantini et al. The formalism of neutrino oscillations: An introduction. Chapter preprint, Gran Sasso Science Institute, February 2018.
- [27] J.A. Formaggio and G. P. Zeller. From eV to EeV: Neutrino cross sections across energy scales. *Rev. Mod. Phys.*, 84, 2012.
- [28] Y. Fukuda et al. Evidence for oscillation of atmospheric neutrinos. *Phys. Rev. Lett.*, 81, 1998.
- [29] I. Gil-Botella. Neutrino physics. *CERN Yellow Report*, pages 157 – 206, 2013.
- [30] O. Hen et al. Probing pp-SRC in ^{12}C , ^{27}Al , ^{56}Fe , and ^{208}Pb using the $A(e, e'p)$ and $A(e, e'pp)$ reactions. Approved CLAS analysis note, 2012.

- [31] S. Jeschonnek, J.W. Van Orden, and T.W. Donnelly. Neutral-current neutrino scattering from the deuteron. *Phys. Rev. C*, 101, 2020.
- [32] M. Khachatryan. *Validation of neutrino energy estimation using electron scattering data*. PhD thesis, 2019.
- [33] X.-G. Lu et al. Measurement of nuclear effects in neutrino interactions with minimal dependence on neutrino energy. *Phys. Rev. C*, 94, 2016.
- [34] K. Mahn et al. Electron scattering and neutrino programs. Letter of interest, Snowmass 2021, 2020.
- [35] B. A. Mecking et al. The CEBAF Large Acceptance Spectrometer (CLAS). *Nucl. Instrum. Meth. A*, 503, 2003.
- [36] A. Minotti et al. Latest results from Double Chooz. *Phys. Part. Nuclei*, 48, 2017.
- [37] O. Moreno, T.W. Donnelly, J.W. Van Orden, and W.P. Ford. Coincidence charged-current neutrino-induced deuteron disintegration. *Phys. Rev. D*, 92, 2015.
- [38] P. B. Pal. Dirac, Majorana, and Weyl fermions. *American Journal of Physics*, 79:485, 2011.
- [39] A. Papadopoulou. Neutrino energy reconstruction methods using electron scattering data. *J. Phys.: Conf. Ser.*, 1056, 2018.
- [40] A. Papadopoulou et al. Inclusive electron scattering and the GENIE neutrino event generator. *Phys. Rev. D*, 2021.
- [41] M. Pavin. *Measurements of hadron yields from the T2K replica target in the NA61/SHINE experiment for neutrino flux prediction in T2K*. PhD thesis, 2017.
- [42] E. Scantamburlo. Measurement of the water to scintillator charged-current cross-section ratio for muon neutrinos at the T2K near detector. 2017.
- [43] J.W. Van Orden, T.W. Donnelly, and O. Moreno. Coincidence charged-current neutrino-induced deuteron disintegration for ${}^2\text{H}_2 {}^{16}\text{O}$. *Phys. Rev. D*, 96, 2017.
- [44] V. Vorobel et al. Latest results from Daya Bay. *J. Phys.: Conf. Ser.*, 873, 2017.
- [45] L. Zana. *Search for the onset of color transparency through ρ^0 electroproduction on nuclei*. PhD thesis, 2010.