

# Validating nuclear models: electron data to interpret neutrino physics

Kevin D. Ewart<sup>1, a)</sup> and Minerba Betancourt<sup>1</sup>

*Fermi National Accelerator Laboratory, Batavia, Illinois 60510, USA*

(Dated: 11 August 2017)

Nuclear models in Monte Carlo event generators are crucial for interpreting neutrino scattering experiments, but are difficult to verify using only neutrino data due to incomplete data for some energy ranges. In particular, we seek to constrain the GENIE neutrino event generator’s models. To overcome the uncertainties inherent to neutrino experiments, it is useful to take advantage of more controllable electron scattering experiments to evaluate current models and place constraints on future models. In this preliminary study on the feasibility of this strategy, we evaluate variables sensitive to the choice of nuclear model for the GENIE default model and data from the Jefferson Labs CLAS experiment. We find some disagreement between them, likely explained by the use of a dated default and background known to be remaining in the CLAS data that is not accounted for.

## I. INTRODUCTION

Neutrino physics, and in particular the measurement of neutrino mass, is one of the few concrete leads for physics beyond the standard model. For this reason, neutrino physics has been identified as one of the Particle Physics Project Prioritization Panel’s (P5) science drivers for the next decade<sup>1</sup>.

However, neutrino experiments are challenged by nuclear kinematics and lower precision beam energy.<sup>2</sup> Neutrino beams are formed from the decay products of particles produced in beam-target collisions, and therefore it is very difficult to create a monoenergetic beam. Energy must be reconstructed on a per-neutrino basis from observable properties of the interaction, which inherently adds uncertainty, and means a more accurate nuclear model is absolutely necessary to compensate. Nuclear models must be used because it is impossible to know the exact configuration of the nucleus at the time of interaction— in particular, the momentum of the hit nucleon is never known. Additionally, secondary interactions within the nucleus can change the observed products, which can cause the reconstructed event to be misclassified. Thus data must be compared to Monte Carlo (MC) simulations, which, if given accurate nuclear models, reproduce the average behavior of the physical system. Because of this, thorough tests of the nuclear models being used in these simulations are necessary— if they are inaccurate, then incorrect conclusions can be drawn from the data.<sup>3,4</sup>

Verifying these models is very difficult for all of the above reasons. An idea is to verify the models experimentally using electrons. Electron experiments have the crucial advantage of a monoenergetic beam with small variance, and a thorough understanding of the properties of electrons. Thus, by comparing electron-nucleus scattering simulations to real-world electron data it is possible to test the accuracy of current Monte Carlo nuclear models in order to improve them to the needed accuracy

for modern neutrino experiments.<sup>5</sup> Additionally, different beam energy reconstruction methods can be tested for accuracy.

Our primary goal was to validate the Monte Carlo event generator used by the MINERvA group GENIE,<sup>6,7</sup> to ensure that it was self-consistent when generating electron and neutrino events, and that it was consistent with actual electron scattering data. We were particularly interested in the charged current quasi-elastic (CCQE) collision mode, diagrammed in figure I. CCQE interactions were selected because it is far easier to reconstruct the beam energy for these interactions, and since they involve whole-nucleon scattering they are sensitive to whole-nucleon intranuclear effects<sup>8</sup>.

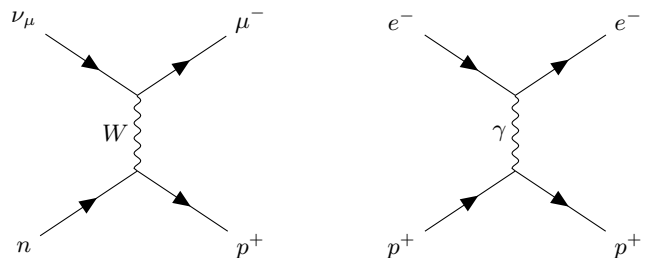


FIG. 1. The Feynman diagram for an idealized CCQE scattering event for a muon neutrino (left) and electron (right). In reality, the nucleon is generally part of an atomic nucleus, complicating the physics significantly, and the final proton may re-scatter before exiting the nucleus, leading to other products and hindering reconstruction.<sup>9</sup>

The rest of this paper is organized as follows: We will next discuss our methodology, then our results, and finally draw conclusions and explore next steps.

## II. METHODOLOGY

We decided to perform this analysis by comparing several MC data sets and one set of electron scattering data. The MC data is either simulated electron scattering or simulated muon neutrino scattering. Muon neu-

<sup>a)</sup> Also at Department of Physics and Astronomy, Rice University.

trinos were chosen because the long-term goal of this project is to constrain the nuclear models used by the MINERvA collaboration specifically, and MINERvA focuses on muon neutrinos. In this paper, “neutrino” is to be taken to mean “muon neutrino” if unspecified. All MC data were generated by others (see acknowledgments), and the electron scattering data were obtained from JLab’s CLAS experiment. The MC data were, unless otherwise noted, generated using the GENIE trunk version circa July 2017 and processed with ROOT version 6.08<sup>10</sup>. The CLAS detector is a large-acceptance detector designed for electron scattering experiments in Jefferson Labs Hall B. It features large solid-angle acceptance and high momentum resolution<sup>11</sup>.

### A. Neutrino/electron comparison

The MC data for comparing electron and muon neutrino scattering were generated with GENIE running in both electron mode and neutrino mode at 4.461 GeV, as well as in muon neutrino mode with the flux observed in the MINERvA detector. The MINERvA flux MC data were generated with GENIE 2.12.6. These data sets were cut so that only CCQE events were considered. Data was binned into one-dimensional histograms using ROOT, with binning parameters chosen for low uncertainty for the monoenergetic neutrinos, the smallest data set. All relevant variables were then compared visually to determine qualitative differences between electron and neutrino scattering and assess applicability of the electron data to neutrino data.

Our major focus was on beam energy reconstruction. For both the Monte Carlo simulations and the CLAS data (considered in the next section), beam energy is precisely known. Therefore, we can determine with equal precision the resolution of different energy reconstruction formulas by comparing the reconstructed energy with the known beam energy. We used two formulas. The first was the quasi-elastic reconstruction (QE hypothesis), which assumes the event was CCQE and uses only kinematic information from the scattered lepton:

$$E_{\text{QE}} = \frac{m_n^2 - (m_p - \epsilon)^2 - m_{\text{lepton}}^2 + 2(m_p - \epsilon)E_{\text{lepton}}}{2(m_p - \epsilon - E_{\text{lepton}} + p_{\text{lepton}} \cos \theta_{\text{lepton}})} \quad (1)$$

where  $m_n$  is the neutron mass,  $m_p$  is the proton mass,  $\epsilon$  is the nucleon binding energy,  $m_{\text{lepton}}$  is either the electron mass or muon mass depending on context,  $p_{\text{lepton}}$  is the momentum of the primary lepton, and  $\theta_{\text{lepton}}$  is the polar angle of the lepton. Here “primary lepton” can be either a muon or an electron, and is the lepton produced by the CCQE interaction diagrammed in figure I. Of course, for MC data the assumption that the event is CCQE can be and was strictly enforced.

The other formula was the calorimetric energy reconstruction, which uses both the scattered lepton and pro-

duced proton:

$$E_{\text{cal}} = E_{\text{lepton}} + T_p + \epsilon \quad (2)$$

where  $T_p$  is the kinetic energy of the produced proton (assumed to be the proton with the highest kinetic energy at the end of the interaction) and others are as before. This method assumes that most of the interaction energy is carried away by the scattered lepton and a single proton, which is generally only valid for CCQE interactions. Since both of these reconstructions assume CCQE events, we performed the QE analysis cut on all data, which has the criteria that there is exactly one proton with momentum greater than threshold, no other charged particles with greater than threshold, and four-momentum transfer  $Q^2 > 1$  GeV, where the threshold momentum is set to 0.3 GeV. Here the  $Q^2 > 1$  GeV cut is not physically motivated, but is put in to be able to generate electron MC events in a reasonable amount of time. Note that for MC  $Q^2$  is reported directly by GENIE. These criteria select the proton product in figure I while allowing for low-momentum additional products, and generally exclude resonant and deep inelastic events since these produce multiple high-momentum products. This cut is required for using eqn. 2 to be able to select a primary proton, and is used for other analysis as well since it will be eventually applied to real-world data where the true event type is unknown. We do not consider neutral products due to detector capabilities. In addition, if we were comparing electron data to neutrino data, we weighted the electron data by the inverse of the Mott cross-section. This was necessary because electrons interact primarily through massless photon exchange, whereas neutrinos interact through massive W or Z boson exchange. The massive propagators change the scattering cross-section to be approximately constant, so it is sufficient to divide out the Mott cross-section for electron scattering to render the data comparable.<sup>5</sup>

We were also interested in the energy transfer and momentum transfer, called  $q_0$  and  $q_3$  respectively. The energy transfer is defined as  $q_0 = E_{\text{incoming lepton}} - E_{\text{outgoing lepton}}$ , and the momentum transfer is defined as  $q_3 = \sqrt{Q^2 + q_0^2}$  where  $Q^2$  is the four-momentum transfer squared. The energy transfer provides a general probe of the hadronic interactions by measuring how much energy is available to them, while the momentum transfer provides more of a perspective on the primary lepton.

### B. MC/data comparison

We then sought to compare electron MC data with the electron scattering data from CLAS. Data of both kinds was binned into one-dimensional histograms using ROOT, with binning chosen to have sufficiently high statistics for the CLAS data.

The MC data were then weighted per-event by the probability of CLAS actually detecting the primary lep-

ton and proton based on their momenta and angles, thus simulating the detector output and rendering the MC data comparable to the CLAS data. This probability mapping is known for the detector and was obtained from JLab, and depends on the polar angle, azimuthal angle, and momentum of the scattered electron and produced proton.

In order to investigate CCQE events, we had to perform cuts on the data. Monte Carlo data was tagged with its channel, including CCQE, but the CLAS data must cut on other parameters to approximate this. Specifically, we required the Bjorken  $|x - 1| < 0.2$  and the invariant mass  $W < 2.0$  GeV to perform this selection, and the cut was applied to both MC and CLAS for comparability. CLAS data were reported with the primary lepton and proton already selected. We therefore applied the QE analysis cut to the MC and CLAS data to select the primary proton in MC. With these cuts it is valid to compare the MC data generated with only CCQE events to the CLAS data.

Since we desired to assess the validity of nuclear models, we needed to use variables sensitive to nuclear interactions. We used the variables determined by X-G. Lu et al. to have the necessary sensitivity:  $\delta p_T$ ,  $\delta \alpha_T$  and  $\delta \phi_T$ .<sup>12</sup> These are defined in figure 2: If we take  $\vec{q} = -p_{\text{lepton},T}$ , then  $\delta p_T = |\vec{q} - p_{p,T}|$  where  $p_{p,T}$  is the transverse momentum vector of the proton,  $\delta \alpha_T$  is the angle between  $\vec{q}$  and  $(\vec{q} - p_{p,T})$ , and  $\delta \phi_T$  is the angle between  $\vec{p}$  and  $p_{p,T}$ .

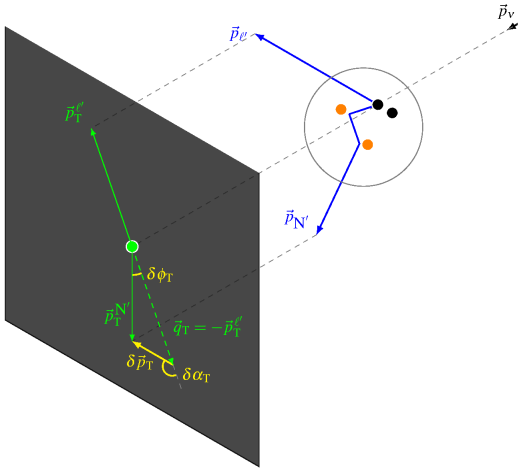


FIG. 2. Definitions of  $\delta p_T$ ,  $\delta \alpha_T$  and  $\delta \phi_T$  as used by X-G. Lu et al.<sup>12</sup> and in this paper. Diagram created by X-G. Lu.

A  $\chi^2$  goodness of fit test was then performed between the monoenergetic electron MC data and the CLAS data to evaluate model validity.

### III. RESULTS

#### A. Neutrino/electron comparison

Comparisons of kinematic variables between electron and neutrino MC after the QE analysis cut (but *before* the QE selection cut) are shown in figure 3. Both data sets are monoenergetic at 4.461 GeV. While the shapes match in general, there are still significant differences in the fall-offs of the tails that need further investigation. Perhaps the most likely explanation is that the “primary lepton” is an electron for electron scattering and a muon for the neutrino scattering, and that we need to do more to compensate for that. Regardless, the close match in the number of protons produced is reassuring. Note that the proton data is for only the “primary” proton of each event, defined to be the final state proton with the greatest momentum.

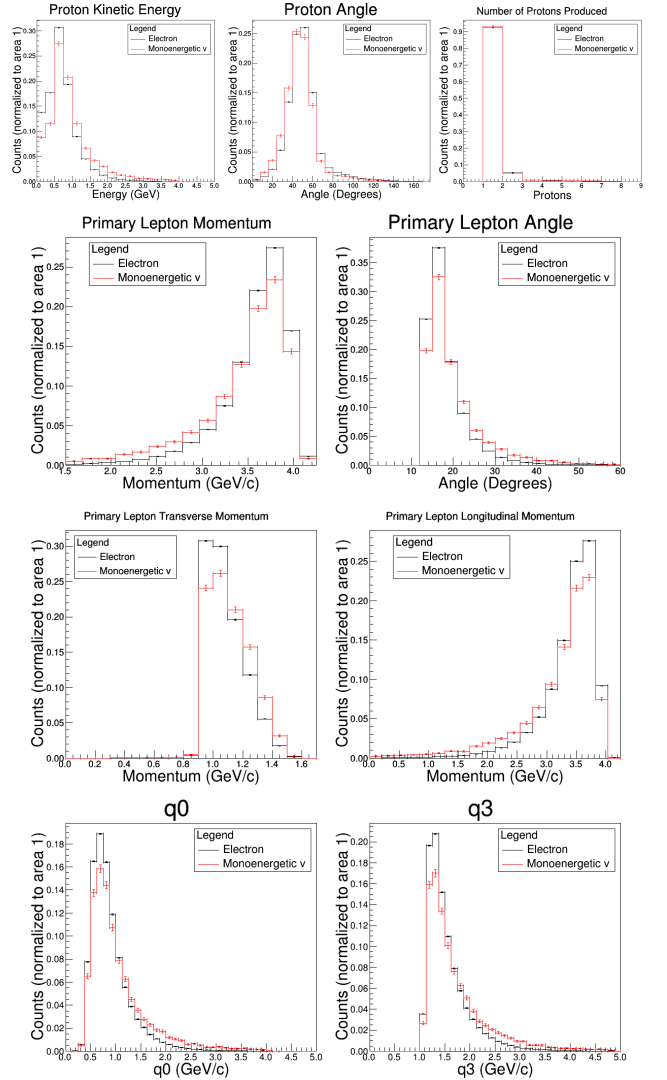


FIG. 3. Proton and lepton kinematics for monoenergetic neutrinos and electrons at 4.461 GeV,  $Q^2 > 1$ , QE events only.

We now consider MC data for a neutrino beam with the flux profile of the MINERvA experiment as well—the previous variables depend heavily on incoming energy, and so the MINERvA flux data had to be excluded. Figure 4 shows the resolution  $\text{res} = (E_{\text{true}} - E_{\text{reco}})/E_{\text{true}}$  for reconstructed energies of the electrons and neutrinos. We find that performance is roughly equivalent between data sets for  $E_{\text{QE}}$ , though it is skewed for the MINERvA flux data. The calorimetric reconstruction is biased slightly high for the monoenergetic neutrinos (likely the result of a overestimated binding energy  $\epsilon$ ), but is in general very good. Additionally, we see that the peak of the calorimetric reconstruction is much sharper than that of the QE hypothesis, but it has an asymmetric tail towards reconstructing low that consists of a significant fraction of events ( $\sim 20\%$ ). The origin of this tail is multi-nucleon knockout that distributes momentum relatively evenly between multiple nucleons or that produces a high-momentum neutron, but with a single proton (not necessarily the true primary) coincidentally still above threshold. Since  $E_{\text{cal}}$  considers only the primary lepton and a single proton, this results in a large amount of missing energy. While  $E_{\text{QE}}$  is less sharp, it is in general symmetrically distributed for the monoenergetic samples, and would be appropriate when proton data is not available or when a symmetric reconstruction is required.

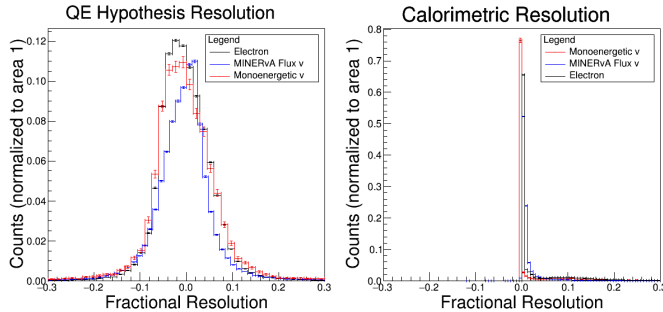


FIG. 4. Monte Carlo monoenergetic muon neutrino, monoenergetic electron, and MINERvA flux neutrino energy reconstructions  $E_{\text{QE}}$  and  $E_{\text{cal}}$ . Data is under the QE analysis cut (see text), since  $E_{\text{cal}}$  is ill-defined otherwise.

## B. MC vs CLAS data

We compared  $q_0$  and  $q_3$  between MC and CLAS data to check that they agree on better understood properties as well as nuclear effect. As we see in figure 5, there is different tail behavior in both  $q_0$  and  $q_3$ , likely indicating a flaw in the model. However, going back to figure 4 we see that the data appears to be closer to the falloff shown by the neutrinos. This might indicate an issue with how GENIE generates electron scattering events, and warrants further exploration. There are also major shape discrepancies between hA and data in  $\delta p_T$ , both in peak

location and fall-off rate. There is closer agreement in  $\delta\alpha$  and  $\delta\phi$ , but still sufficient differences to be concerning. It should be noted that the jump at  $\phi = 0$  is a known issue with the hA model, which has since been fixed in newer models. The differences for all are sufficiently small that we can expect them to be fixed or tuned out in future models.

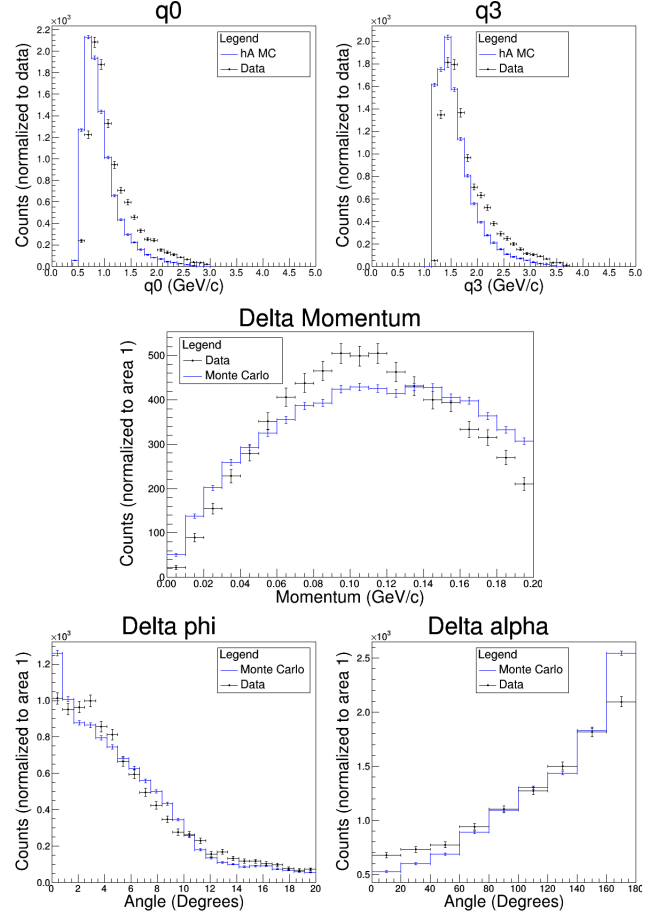


FIG. 5. Plots of  $q_0$ ,  $q_3$ ,  $\delta p_T$ ,  $\delta\phi_T$ , and  $\delta\alpha_T$  for the GENIE default hA nuclear model compared to CLAS data.

## IV. DISCUSSION AND NEXT STEPS

While not an exact match, electron and neutrino simulations are sufficiently comparable in the measured variables to warrant further investigation into this idea of using electron data to constrain the nuclear model parameters for neutrino experiments.

For comparing the MC to CLAS data, we find overall poor agreement in  $\delta p_T$ , but this is unsurprising because of the old FSI model we used for these comparisons. This FSI uses an effective model, in which hadronic intranuclear rescattering cross sections increase with the nuclear size according to  $A^{2/3}$  scaling.<sup>6,7</sup> It is expected that a model with a particle cascade treatment will have better

agreement with data. Agreement in  $\delta\phi_T$  and  $\delta\alpha_T$  is also poor, with the  $\chi^2$  test giving extremely low p-values.

This first study shows some disagreement between the hA model and the data. Part of the disagreement is due to the resonance background contribution in the data that is not present in the MC, as well as the use of an old MC model. Future work will test more models, especially more modern ones, and perform background subtraction for all known backgrounds.

## V. ACKNOWLEDGMENTS

We would like to acknowledge Afroditi Papadopoulou for generating the monoenergetic MC data, Dr. Daniel Ruterbories for generating the MINERvA flux MC data, and Mariana Khachatryan for the CLAS data. We would like to thank Joshua Ellis for his publicly available Feynman diagram creation package<sup>9</sup>. This work was supported in part by the U.S. Department of Energy, Office of Science, Office of Workforce Development for Teachers and Scientists (WDTs) under the Science Undergraduate Laboratory Internships Program (SULI).

- <sup>1</sup>S. R. et al., HEPAP Subcommittee (2014).
- <sup>2</sup>L. Alvarez-Ruso *et al.*, “NuSTEC White Paper: Status and Challenges of Neutrino-Nucleus Scattering,” (2017), arXiv:1706.03621 [hep-ph].
- <sup>3</sup>P. Coloma and P. Huber, “Impact of nuclear effects on the extraction of neutrino oscillation parameters,” *Phys. Rev. Lett.* **111**, 221802 (2013).
- <sup>4</sup>P. A. Rodrigues *et al.* (MINERvA Collaboration), “Identification of nuclear effects in neutrino-carbon interactions at low three-momentum transfer,” *Phys. Rev. Lett.* **116**, 071802 (2016).
- <sup>5</sup>F. H. et al., “Electrons for neutrinos, addressing critical neutrino-nucleus issues,”.
- <sup>6</sup>C. Andreopoulos *et al.*, “The GENIE Neutrino Monte Carlo Generator,” *Nucl. Instrum. Meth.* **A614**, 87–104 (2010), arXiv:0905.2517 [hep-ph].
- <sup>7</sup>C. Andreopoulos, C. Barry, S. Dytman, H. Gallagher, T. Golan, R. Hatcher, G. Perdue, and J. Yarba, “The GENIE Neutrino Monte Carlo Generator: Physics and User Manual,” (2015), arXiv:1510.05494 [hep-ph].
- <sup>8</sup>K. Saraswat, P. Shukla, V. Kumar, and V. Singh, “Charged current quasi elastic scattering of muon neutrino with nuclei,” (2016), arXiv:1606.06916 [hep-ph].
- <sup>9</sup>J. Ellis, “TikZ-Feynman: Feynman diagrams with TikZ,” *Comput. Phys. Commun.* **210**, 103–123 (2017), arXiv:1601.05437 [hep-ph].
- <sup>10</sup>R. Brun and F. Rademakers, “ROOT: An object oriented data analysis framework,” *New computing techniques in physics research V. Proceedings, 5th International Workshop, AIHENP ’96, Lausanne, Switzerland, September 2-6, 1996*, *Nucl. Instrum. Meth.* **A389**, 81–86 (1997).
- <sup>11</sup>B. Mecking *et al.*, “The cebaf large acceptance spectrometer (clas),” *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **503**, 513 – 553 (2003).
- <sup>12</sup>X.-G. Lu, L. Pickering, S. Dolan, G. Barr, D. Coplowe, Y. Uchida, D. Wark, M. O. Wascko, A. Weber, and T. Yuan, “Measurement of nuclear effects in neutrino interactions with minimal dependence on neutrino energy,” *Phys. Rev. C* **94**, 015503 (2016).

## VI. APPENDIX

We plotted  $Q^2$  against  $q_0$  or primary lepton angle as shown in figure 6 for both electron MC weighted by CLAS acceptance and MINERvA flux MC (CCQE muon neutrino only) unweighted.

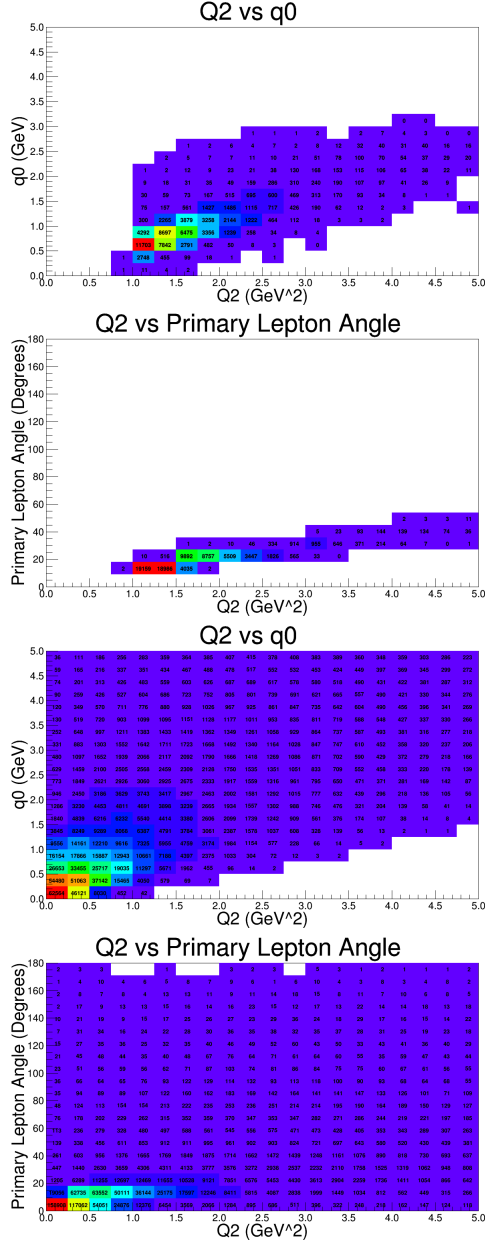


FIG. 6.  $Q^2$  vs  $q_0$  or lepton angle for electron MC (top) and MINERvA flux MC (bottom)

We also made plots illustrating the affects of a partial QE analysis cut on the MINERvA flux MC, shown in figure 7 below. This partial cut's only difference from the QE analysis cut is that the  $Q^2 > 1$  condition is not used, so that the below plots have the full range of  $Q^2$  in both curves. The data used was only CCQE events from

muon neutrinos, and has not been weighted.

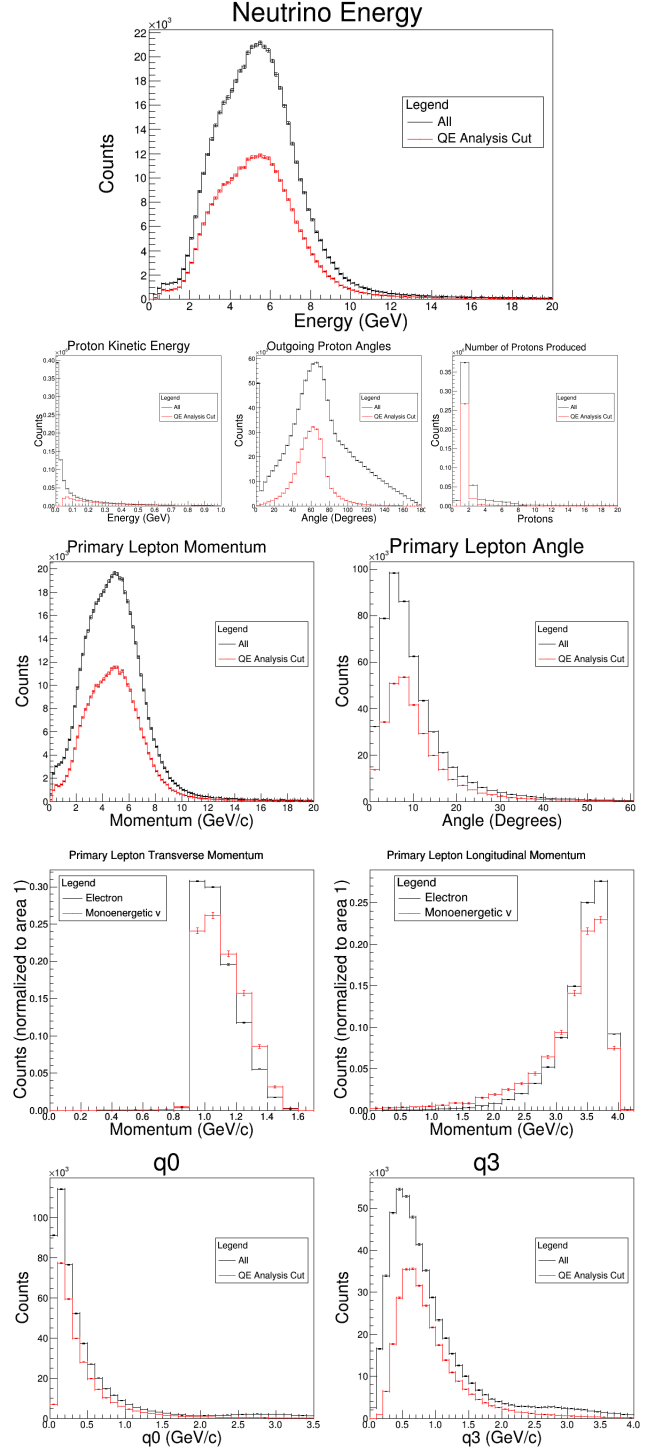


FIG. 7. Effects of the QE analysis cut on CCQE muon neutrino events from the MINERvA flux MC data.