FERMILAB-THESIS-2022-20

Simultaneous Measurement of Visible Energy and Momentum Transfer in Anti-Electron Neutrino Interactions on Hydrocarbon

by

Sarah Henry

Submitted in Partial Fulfillment of the

Requirements for the Degree

Doctor of Philosophy in Physics

Supervised by Professor Kevin McFarland

Department of Physics and Astronomy
Arts, Sciences and Engineering
School of Arts and Sciences

University of Rochester

Rochester, New York

2022

To Harper Marie

"Go confidently in the direction of your dreams!" -Thoreau

Acknowledgments

This thesis would not have been possible without the support many people have shown me throughout this entire journey as a student. Among many offers, I accepted an athletic scholarship to attend and swim for Texas A&M University in College Station, TX. Little did I know that, with the incredible support from the academic and athletic programs, I would find myself on the path to pursuing a physics Ph.D. from the University of Rochester in Rochester, NY. A place I was not familiar with but grew to love.

I would first like to thank all my physics professors at Texas A&M University, with a special thank you to Dr. David Toback. I remember the day I mustered up the courage to walk into Dr. Toback's office. He willingly took me under his wing and sparked my initial interest in particle physics by allowing me to pick up projects in my spare time around athletic commitments. I especially have to thank Chris Barttelbort, my athletic-academic advisor, for helping me accomplish my highest academic goals in conjunction with my Olympic dreams and for preparing me for a world after swimming.

I would also like to thank everyone who has supported me at the University of Rochester.

Especially my first year class and first year office of B&L 373 for the late night study sessions and for the great community I found in a new city. Thank you to JJ (Elizabeth), Esteban, Sayat, and Andrew. Thank you to Connie, Laura and Linda, as well as the staff working so hard to make U of R a great place to be a student.

A huge thank you to Kevin for being a great advisor and mentor and for introducing me to road cycling(!). On that note, I also have to thank Tess for her kindness and compassion when she pulled an all-nighter with me in Highland hospital after I crashed my bike at 25mph...

I greatly enjoyed my time on MINERvA and learned so much from many brilliant people. I have fostered lifelong friendships and look forward to seeing what the future holds for everyone. To Steve for being a great mentor and Jeffrey, Aaron, Andrew, Dan, Rob, Marvin, and Gonzalo thank you for putting up with my contant questions throughout the years. A special thank you to Hang Su for being a great collaborator and for answering my many, many questions.

Finally, thank you to all my friends and family. To my Rochester friends, RBC, and Midtown Athletic Club for providing an amazing community. To my Dallas A&M friends for helping me make Dallas my home earlier than expected due to COVID. Thank you, Mom and Dad, for providing me many opportunities and encouraging me to challenge myself. Dad, thank you for instilling a sense of curiosity. And last but not least, thanks to my husband Omar for his encouragement and unwavering support.

Abstract

Precise knowledge of neutrino interaction cross sections is required for current and future accelerator-based neutrino oscillation experiments. Precision neutrino oscillation measurements require inference of the neutrino energy and flavor from the visible particles in the neutrino interaction in the detector. This inference is different for the true neutrino flavors measured, electron and muon neutrinos, and can be studied by observation of neutrino interactions in an experiment's near detector. However, anti-electron neutrinos make up only a few percent of an anti-muon neutrino beam and pose a challenge in making cross section measurements and predictions. The reported measurements were made using data from MINERvA, a neutrino-nucleus scattering experiment, with an anti-neutrino beam configuration of mean energy ~ 6 GeV.

This thesis provides two double-differential cross sections of anti-electron neutrino inclusive charged-current reactions using the kinematics of visible energy, three-momentum transfer, and transverse momentum. The analysis is carried out at low three-momentum transfer, making it sensitive to regions with multi-nucleon effects.

Contributors and Funding Sources

This work was supervised by a committee consisting of Advisor Kevin McFarland (Department of Physics), Professor Steve Manly (Department of Physics), Professor Kara Bren (Department of Chemistry), Professor Lynn Orr (Department of Physics), and chaired by Professor Dustin Trail (Department of Earth and Environmental Sciences).

The members of the MINERvA collaboration are listed in Appendix A. This work was supported by Dept of Energy grant DE-SC0008475 and the National Science Foundation grant DGE-1419118.

Contents

Acknowledgments **iii**

Abstract **v**

Constributors **vi**

List of Tables **xiv**

List of Figures **xxviii**

1 Introduction **1**

1.1 Neutrinos and the Standard Model 1

1.2 Brief History of the Neutrino 3

1.3 The Solar Neutrino Problem 6

1.4 Neutrino Oscillations 11

1.5 Neutrino Oscillation Theory 11

1.6 Current Oscillation Experiments 17

2 The NuMI Neutrino Beam **22**

2.1 Proton Production and Acceleration 22

2.2 NuMI Beamline 24
2.3 Neutrino Flux Simulation 30
2.3.1 Hadron Production 30
2.3.2 Neutrino Flux Constraint 31

3 The MINERvA Detector 38

3.1 The MINERνA Experiment 38
3.1.1 Active plane composition 40
3.1.2 Coordinate System 41
3.1.3 Module Assemblies and Nuclear Targets 43
3.1.4 Event Display 44
3.2 The MINOS Near Detector 45
3.3 Readout Electronics 46
3.3.1 Front-end Electronics 50
3.4 Calibration and Pedestal Subtraction 51
3.5 Sample Preparation 54
3.5.1 General Track and Calorimetry Reconstruction 54
3.5.2 Calorimetric and Spline Correction 55
3.5.3 Kinematic Variable Reconstruction 57
3.5.4 Available Energy Correction 60
3.6 Interaction Models and Simulation 65
3.6.1 Nuclear Model 67
3.6.2 Quasi-elastic scattering and the 2p2h Model 69
3.6.3 DIS and Resonant Scattering 71

3.6.4 Final State Interactions 74
3.7 The MINERvA Tune 74
3.7.1 2p2h and RPA Corrections 75
3.7.2 Non-resonant Pion Reweight 76
3.7.3 Low Q^2 Suppression 77
3.7.4 MINERvA Tune v1 and MINERvA Tune v2 78

4 $\bar{\nu}_e$ Low Recoil in Medium Energy 82

4.1 Characterization of Events in the MINERvA Detector 83
4.1.1 Minimum Ionizing Particles 83
4.1.2 Electrons and Photons 85
4.1.3 Electromagnetic Cascades 89
4.2 Electron Cone Candidate 93
4.2.1 Neutrino Type 93
4.2.2 Electron Cone Construction 94
4.3 Electromagnetic Shower Classification 97
4.3.1 Endpoint Energy Fraction 98
4.3.2 Mean dE/dx 100
4.3.3 Median Transverse Width 101
4.3.4 MVA Classifier 102
4.3.5 Fraction of Energy in non-MIP-like-clusters 104
4.4 Event-level Cleanup 105
4.5 Leakage Study 106

4.6 Separation of Events 108
4.6.1 Electron-Photon Discrimination: dE/dx 109

5 Charged Current Low Recoil $\bar{\nu}_e$ 117

5.1 Diffractive and Coherent Pion Production 117
5.1.1 Construction of Sidebands 122
5.1.2 Excess Sideband 127
5.2 Final Analysis Corrections 144
5.2.1 FHC Correction 145
5.2.2 Background Fitting 149
5.2.3 Post Background Tune 154

6 Double Differential Cross Section 168

6.1 Data and MC Sample 169
6.2 Unfolding 169
6.2.1 Warping Studies 175
6.3 Efficiency Correction 183
6.4 Cross Section Normalization 184
6.5 Double Differential Cross Section 186
6.5.1 Model Comparison 200

7 Systematic Uncertainties 206

7.1 Uncertainties 206
7.2 The Many Universe Method 207
7.3 Covariance Matrix 209

7.4 Systematic Uncertainty Summary 211
7.4.1 Flux Model Uncertainty 214
7.4.2 Interaction Model Uncertainty 216
7.4.3 MnvTune Uncertainties 223
7.4.4 Detector Model Uncertainties 228

8 Conclusion 236

8.1 Discussion 236
8.1.1 Excess of events at nearly zero E_{avail} 236
8.1.2 Model Comparison 238
8.1.3 Comparison with Published $\bar{\nu}_\mu$ Measurements 238
8.2 Ratio of $\bar{\nu}_\mu$ and $\bar{\nu}_e$ samples 246

Bibliography 254

A The MINERvA Collaboration 261

List of Tables

1.1 Given n generations of leptons, the corresponding number of mixing angles and complex phase factors. 13

3.1 Breakdown of the chemical composition of the MINERvA scintillator planes 41

3.2 Passive material corrections by sub detector 56

3.3 Additive weights for the reconstructed $E_{available}$ in bins of reconstructed $E_{available}$ 64

3.4 Table summary of Minerva Tune components for the "MnvGENIE-v1" tune. 81

3.5 Table summary of Minerva Tune components for the "MnvGENIE-v2" tune. 81

4.1 Critical energy of electrons Ec and radiation length X_0 for different materials. 85

4.2 Cuts and requirements used to form base sample used in the analysis . . . 116

5.1 Categorical breakdown of selection. 122

6.1 Protons on target for respective data and MC playlist production. Generic is all interactions (except NC and CC Diffractive π^0). The remainder of the playlists are separate MC productions that are integrated into the analysis: NC diffractive π^0 events, Valencia 2p2h modeled up to 2.0 GeV and a high statistical production of ν_e events created for efficiency corrections. 168

6.2 Definitions of the variables measured in the analysis. 169
6.3 List of warp models used in the unfolding study. 182
6.4 Cross section normalization values for the tracker region 186
6.5 Definitions for the signal component breakdown. 187
6.6 Parameters for the 2p2h reweighting function. The column "All" is used in the reference model while QE is used as the alternative model comparison. Table from Ref. [99]. 200

7.1 Categorical breakdown of systematics uncertainties. 212
7.2 Systematic uncertainties for the GENIE universes based on $\pm\sigma$ two universe shifts. 216
7.3 Magnitude of systematic shifts for GENIE model universes. 219
7.4 Magnitude of systematic shifts for GENIE model universes. 220
7.5 Magnitude of systematic shifts for GENIE FSI universes. 221
7.6 Magnitude of systematic shifts for GENIE FSI universes. 222
7.7 Fractional uncertainties accounted for in the particle response systematic . 231

8.1 Comparison of the cross section for data and MC in the first bin of E_{avail} for all bins of q_3. The table entries are $10^{-39}\frac{cm^2}{GeV^2}$. 237
8.2 Comparison of the cross section for data and MC in the first bin of E_{avail} for all bins of p_T. The table entries are $10^{-39}\frac{cm^2}{GeV^2}$. 237

8.3 Comparison between the E_{avail} binning used for the low recoil ME $\bar{\nu}_e$ analysis (left) and the low recoil LE $\bar{\nu}_\mu$ analysis (right). The $\bar{\nu}_e$ binning is truncated at 0.5 GeV for the comparison to LE but the results are reported up to 1.2 GeV. 239
8.4 Comparison between the q_3 binning used for the low recoil ME $\bar{\nu}_e$ analysis (left) and the low recoil LE $\bar{\nu}_\mu$ analysis (right). 240
8.5 Correlation matrix ordering for both results. Note that the correlation matrix value for the data/MC peak is equivalent to the estimated peak at zero value. 243
8.6 Summary of results for ME $\bar{\nu}_e$. 244
8.7 Correlation matrix for ME $\bar{\nu}_e$ result for relevant bins. 244
8.8 Summary of results for LE $\bar{\nu}_\mu$. 245
8.9 Correlation matrix for the LE $\bar{\nu}_\mu$ result for relevant bins. 245

List of Figures

1.1 Feynman diagrams for neutrino electron scattering. All flavors of neutrinos are involved through neutral current type interactions (left) while only electron neutrinos interact through charged current interactions (right). . . 3
1.2 Cosmic ray interactions with the atmosphere. Figure from Ref. [13] . . . 9

1.3 Three neutrino oscillation in a vacuum for long (left) and shot (right) baselines beginning with initial neutrinos of ν_e (top), ν_μ (middle) and ν_τ (bottom). Figure from Ref. [19]. 18

2.1 The Fermilab accelerator complex. Figure from Ref. [24]. 23

2.2 Cross sectional view of the NuMI target for a proton traveling through the NuMI baffle to hit the target. Figure from Ref. [24]. 25

2.3 FHC horn configuration. Figure from Ref. [24]. 26

2.4 Neutrino flux for various species for forward horn current (FHC) for the MINERvA medium energy run. Figure from Ref.[26] 28

2.5 Neutrino flux for various species for reverse horn current (RHC) for the MINERvA medium energy run. Figure from Ref. [26] 29

2.6 Flux distribution with hadron production weight in ME. Figure from Ref. [31]. 32

2.7 Feynman diagram for a neutrino electron ($\nu + e \rightarrow \nu + e$) interaction . . . 33

2.8 Probability distribution for number of $\nu + e$ scatters, before and after constraining the *a priori* flux model. Figure from Ref. [34]. 34

2.9 Probability distribution of the ν_μ flux between 2 and 20 GeV, before and after constraining the *a priori* flux model. Figure from Ref. [34]. 35

2.10 The FHC (top) and RHC (bottom) fluxes (left) and uncertainties (right) as a function of E_ν for neutrino energies > 10 MeV. Slight reduction in fractional uncertainty can be seen as a result from IMD flux constraints. Figure from Ref. [36]. 36

2.11 Ratio of low - ν events with Data/MC before and after the muon energy scale shift. Figure from Ref. [39]. 37

3.1 Schematic of MINERvA. Figure from Ref. [40] 39

3.2 Cross section of the MINERvA scintillator strips. Figure from Ref. [40]. . 41

3.3 Orientation of MINERvA Planes, as defined by the strip orientation. Figure from Ref. [40]. 42

3.4 MINERvA tracker modules. The X and Y axes of the detector coordinate system are shown where the XY origin is located at the center. Looking downstream, the Z axis points in to the page. Figure from Ref. [40]. . . . 43

3.5 Charged current deep inelastic scattering MC event display in the MINERvA detector. Color is the amount of energy deposited in a scintillator strip. 44

3.6 Scintillator and steel places on the MINOS detector. Figure from Ref. [41] 46

3.7 Schematic of an overhead view of the MINOS upstream calorimeter and downstream muon spectrometer. Figure from Ref. [41] 47

3.8 Schematic construction of a photomultiplier tube. Figure from Ref. [42]. . 47

3.9 Optical PMT box prior to assembly. From left: the end cap, ribbon cable, PMT base board, the PMT (black rectangular box), woven fibers and the PMT circuit board. The blue cylinder is the steel housing. Figure from Ref. [40] 49

3.10 A pedestal distribution for an individual channel readout from one 16 μs gate. Figure from Ref. [40] 52

3.11 Number of hits as a function of charged measured. The first peak on the left is the pedestal (zero). The second peak represents 1 PE. Figure from Ref [40]. 53
3.12 E_{avail} migration before additive correction. 61
3.13 True E_{avail} in bins of Reco E_{avail}. The spike population at zero is a neutron population. 62
3.14 E_{avail} Additive correction calculated from the difference of the true available energy and reconstructed available energy. The correction is formed based on the mean of each distribution in bin of reconstructed energy. 63
3.15 Smooth function fit between additive reconstructed E_{avail} values. 65
3.16 E_{avail} migration matrix with additive correction to reconstructed available energy 66
3.17 Variety of neutrino interaction models contributing to the total charged current neutrino-nucleon cross section, from [47]. 67
3.18 RES scattering (left) and DIS scattering (right) Feynman diagrams. Figure from Ref. [45]. 72
3.19 Resonant scattering with a not-CCQE-like final state (top) and resonant scattering where pion is absorbed (bottom) creating CCQE-like final state. Figure from Ref. [45]. 73
3.20 The double-differential cross section $d^2\sigma/dE_{avail}dq_3$ compared to GENIE 2.8.4. Adding the RPA effect reduces the over-prediction in the first three q_3 bins. 2p2h modeled by Valencia is shown in solid, with RPA+2p2h corrections the best fit to the data. Discrepancies still remain. Figure from Ref. [55]. 76

3.21 The reconstructed available energy in bins of three-momentum transfer. The left panel refers to the MC excess of events and the right panel refers to the MC deficiency seen between the MC QE and MC Delta (RES) peaks. Figure from Ref. [55]. 78

3.22 Comparisons between the nominal and tuned MINERvA pion production models. Channels of $\nu_\mu CC1\pi^\pm$ (top left), $\bar{\nu}_\mu CC1\pi^\pm$ (top right), $\nu_\mu CC1\pi^0$ (bottom left) and $\bar{\nu}_\mu CC1\pi^0$ (bottom right) distributions are shown as a function of Q^2. the light green model (ANL/BNL) refers to the tuned GENIE model. The FrAbs shows the cross-section prediction with the nominal FrAbs model, and FrAbs+Q^2 with the suppression weight. Figure from Ref. [69]. 79

4.1 Mean energy loss rate for muons, pions and protons in various materials derived from the Bethe-Block equation. Figure from Ref. [73] 84

4.2 Fractional energy loss per radiation length in lead as a function of $e^\pm$ energy. The crossover of ionization and bremsstrahlung is referred to as the critical energy, E_c. Figure from Ref. [73] 86

4.3 Photon total cross section as a function of energy in carbon (top) and lead (bottom). Figure from Ref. [73]. $\sigma_{p.e}$ dominates at low photon energy while κ_{nuc} pair production in a nuclear field dominates at higher energy. . 88

4.4 Diagram of an electromagnetic shower for t radiation lengths. Figure from Ref. [78]. 90

4.5 A MC event display showing the interaction of a 12.23 GeV $\bar{\nu}_e$. The final state particles are protons, a π^-, and an 11.4 GeV positron which creates an electromagnetic shower in the detector. 92

4.6 Event display showing the signatures of different particles in the MINERvA detector. The triangles are the different energy deposits in the detector. . . 94

4.7 Fraction of true energy inside of electron candidate cone with different opening angles as measured in a simulated particle cannon. Figure from Ref. [79]. 95

4.8 An event display of the activity of different particles in the MINERvA detector. The black circles represent the purpose of the endpoint energy fraction, where we can see the different energy deposition profiles at the end of the track. The color represents amount of energy deposited in a scintillator strip. 96

4.9 A representation of the division of energy deposits into bins of 10 g/cm^2 of areal density. Figure from Ref. [81]. 98

4.10 Determining the median (blue) value of the energy deposits excluding the last bin (red). Figure from Ref. [81]. 99

4.11 Endpoint energy fraction as a function of calorimetrically reconstructed energy. Figure from Ref. [79]. Error bars on each point represent the smallest interval containing 68% of the distribution in that bin. 100

4.12 Mean dE/dx as a function of calorimetrically reconstructed energy. Error bars on each point represent the smallest interval containing 68% of the distribution in that bin. Figure from Ref. [79]. 101

4.13 Median transverse width as a function of calorimetrically reconstructed energy. Figure from Ref. [79]. 103
4.14 The ξ_{kNN} values (left) from the kNN classifier and the corresponding efficiency-purity curve (right). Figure from Ref. [79]. 104
4.15 Distribution of Non-MIP-cluster fraction. 105
4.16 Calculation of leakage correction from electron particle cannon 108
4.17 Difference of leakage energy and corrected leakage energy from electron particle cannon 109
4.18 Difference of cone leakage and corrected leakage in bins of true electron energy 110
4.19 Difference of cone leakage and corrected leakage in bins of true electron theta 111
4.20 Difference of cone leakage and corrected leakage in bins of transverse momentum 112
4.21 Diagram of the dE/dx calculation. The clusters inside of the electron cone candidate are projected onto the cone axis. The dE/dx is calculated inside of a 100 mm block sliding in 25 mm increments. There is a 25 mm exclusion radius around the vertex. 113
4.22 Optimization of the minimum dE/dx (MeV/cm) requirement used to isolate signal electron events. e is the electron particle cannon, gamma is for the photon particle cannon 114
4.23 Minimum dE/dx distribution (MeV/cm), showing the signal optimization . 115

5.1 Feynman diagrams for CC coherent pion production (left) and NC coherent pion production (right) 120

5.2 Feynman diagrams for CC diffractive pion production (left) and NC diffractive pion production (right) 120
5.3 Mean dE/dx distribution used to determine sideband definitions 123
5.4 $\psi * E_e$ for events of dE/dx $>$ 2.4 MeV/cm 125
5.5 Representation of signal and sideband regions with respective cuts. 126
5.6 $\psi * E_e$ distribution in the Signal region of Front dE/dx $<$ 2.4 MeV/cm . . 127
5.7 E_{avail} vs p_T distribution for events in the π^0 sideband defined by dE/dx $>$ 2.4 MeV/cm and $\psi * E_e > 0.5$ GeV. 128
5.8 E_{avail} vs q_3 distribution for events in the π^0 sideband defined by dE/dx $>$ 2.4 MeV/cm and $\psi * E_e > 0.5$ GeV. 128
5.9 Vertex z position for dE/dx $>$ 2.4 MeV/cm and $\psi * E_e < 0.5$ GeV 130
5.10 Vertex $R^2 = x^2 + y^2$ for dE/dx $>$ 2.4 MeV/cm and $\psi * E_e < 0.5$ GeV . . . 131
5.11 Inline Upstream Energy diagram 132
5.12 Energy inside a 7.5 degree cone pointing upstream. 132
5.13 Energy outside of electron candidate cone and vertex region 133
5.14 Shapes of particle cannon samples in median shower width 134
5.15 Transverse shower asymmetry shape distributions of ad hoc sample compared to the data excess 136
5.16 Extra energy ratio, ψ, where the excess is best described by the shape of the coherent π^0 event 138

5.17 A schematic showing the calculation of the transverse and longitudinal distance of clusters. The green triangles represent the clusters of energy resultant from a neutrino interaction. The red cone represents the electron candidate cone formed around the optimized axis. The clusters outside of the cone are assigned a longitudinal (red dotted line) and transverse (blue dotted line) distance with respect to the cone axis. 139
5.18 Transverse distance of extra energy clusters with respect to cone axis . . . 140
5.19 Longitudinal distance of extra energy clusters with respect to cone axis . . 140
5.20 $E_e\theta^2(GeV \text{ x } radian^2)$ 141
5.21 Inline Upstream Energy in bins of E_e in the excess sideband, NC Diffractive has a scale factor of 12. 143
5.22 Representation of signal and sideband regions with respective cuts. The excess sideband has been subdivided into low and high upstream inline energy. 144
5.23 Ratio of RHC/FHC ν_e in True Neutrino Energy 146
5.24 Neutrino energy estimator vs true neutrino energy 147
5.25 The scaled FHC sample providing an electron neutrino (blue) prediction for the RHC electron neutrino background. 149
5.26 A breakdown of the RHC sample into electron neutrino and anti-electron neutrino contributions. 150

5.27 A comparison of the predicted electron neutrino contribution from the scaled FHC sample (left) to the RHC electron neutrino background prediction (right) for a randomly chosen bin of $0.3 < q_3 < 0.4$. The number of events (y - axis) is comparable between the two samples. 151

5.28 Scale factor values for signal (top left) and π^0 backgrounds (top right) as a function of p_T, NC Diffractive π^0 background (bottom left) and NC Coherent π^0 backgrounds (bottom right) as a function of E_e as calculated by a global fit 153

5.29 Vertex z position in the Excess sideband. 154

5.30 Vertex $R^2 = x^2 + y^2$ in the Excess sideband. 155

5.31 Energy inside a 7.5 deg cone pointing upstream in the Excess sideband. . 155

5.32 Energy outside of electron candidate cone and vertex region in the Excess sideband. 156

5.33 Transverse shower asymmetry in the Excess sideband. 156

5.34 Transverse distance of extra energy clusters with respect to cone axis . . . 157

5.35 Longitudinal Difference (m) 157

5.36 $E_e\theta^2$ 158

5.37 Signal distributions (dE/dx $<$ 2.4 MeV/cm) Pre and Post background tuning. These figures do not have the Signal scaled factor applied. 160

5.38 Pre and Post background tuning in the NC Coherent π^0 sideband (dE/dx $>$ 2.4 MeV/cm, $\psi * E_e < 0.5$ GeV, Inline upstream energy $<$ 10 MeV). . . . 161

5.39 Pre and Post background tuning in the NC diffractive π^0 sideband (dE/dx $>$ 2.4 MeV/cm, $\psi * E_e < 0.5$ GeV, Inline upstream energy $>$ 10 MeV). . . 162

5.40 Pre and Post background tuning in the π^0 sideband (dE/dx $>$ 2.4 MeV/cm, $\psi * E_e > 0.5$ GeV). 163
5.41 Signal distributions (dE/dx $<$ 2.4 MeV/cm) Pre and Post background tuning. These figures do not have the Signal scaled factor applied. 164
5.42 Pre and Post background tuning in the NC Coherent π^0 sideband (dE/dx $>$ 2.4 MeV/cm, $\psi * E_e < 0.5$ GeV, Inline upstream energy $<$ 10 MeV). . . . 165
5.43 Pre and Post background tuning in the NC diffractive π^0 sideband (dE/dx $>$ 2.4 MeV/cm, $\psi * E_e < 0.5$ GeV, Inline upstream energy $>$ 10 MeV). . . 166
5.44 Pre and Post background tuning in the π^0 sideband (dE/dx $>$ 2.4 MeV/cm, $\psi * E_e > 0.5$ GeV). 167

6.1 Migration matrix for E_{avail} vs p_T distribution. The x axis is reconstructed bins and the y is for true bin. 173
6.2 Migration matrix for E_{avail} vs q_3 distribution. The x axis is reconstructed bins and the y axis is for true bins. 174
6.3 Migration matrix for E_{avail} vs q_3 distribution. The x axis is reconstructed bins and they axis is for true bins. 175
6.4 Migration matrix for E_{avail} vs q_3 distribution. The x axis is reconstructed bins and the y axis is for true bins. 176
6.5 Migration matrix for E_{avail} vs q_3 distribution. The x axis is reconstructed bins and the y axis is for true bins. 177
6.6 Warping study using the RPA model. 178
6.7 Warping study using the SuSA model. 179
6.8 Warping study using the Low Q2 model. 180

6.9 MINERvA 2p2h Tune Variant: nn pair. 181
6.10 Efficiency for E_{avail} vs q_3 distribution. The x axis is reconstructed bins and the y axis is for true bins. 184
6.11 Efficiency for E_{avail} vs q_3 distribution. 185
6.12 Cross section result in E_{avail} vs q_3. The blue regions in each q_3 bin are kinematically forbidden. 188
6.13 Cross section result in E_{avail} vs q_3. The y axis is on a log scale truncated at $10^{-}2$ to enable a better view of the tail end of the cross section. 189
6.14 Data/MC cross section ratio for E_{avail} vs q_3 distribution. 190
6.15 Cross section result for E_{avail} vs p_T distribution. 191
6.16 Cross section result for E_{avail} vs p_T distribution shown on a log y scale. . 192
6.17 Data/MC cross section ratio for E_{avail} vs p_T distribution. 193
6.18 A decomposition of the cross section result into contributing interaction types in E_{avail} vs q_3. 194
6.19 A decomposition of the cross section result into contributing interaction types in E_{avail} vs q_3 on a y log scale. 195
6.20 Data/MC ratio including fractional contribution of signal processes shown in E_{avail} vs q_3. 196
6.21 A decomposition of the cross section result into contributing interaction types in E_{avail} vs p_T. 197
6.22 A decomposition of the cross section result into contributing interaction types in E_{avail} vs p_T on a y log scale. 198

6.23 Data MC ratio including fractional contribution of signal processes shown in E_{avail} vs p_T. 199
6.24 Reconstructed available energy for neutrino events (left). Reconstructed available energy for anti-neutrino events (right). The contribution of 2p2h events is in yellow. The green circle represents the regions where simulations underestimate data. Figure from Ref. [102]. 201
6.25 Comparison of cross section result to alternative models in E_{avail} vs q_3. . . 202
6.26 Ratio of cross section and alternative model MC to data in E_{avail} vs q_3. . . 203
6.27 Comparison of cross section result to alternative models in E_{avail} vs q_3. . . 204
6.28 Ratio of cross section and alternative model MC to data in E_{avail} vs p_t. . . 205

7.1 All systematic uncertainties fall under one of the categories. 212
7.2 Total cross section error summary broken down into four major subgroups. 213
7.3 Cross section error summary showing flux contribution to total uncertainty. 215
7.4 Cross section error summary showing GENIE FSI and GENIE interaction model uncertainty contribution. 224
7.5 A decomposition of the cross section systematic errors broken into leading GENIE systematic error contributors for E_{avail} vs p_T. The leading systematic uncertainty is GENIE_MaRES (light blue), GENIE_MvRES (green) and GENIE_NormCCRES (yellow). 225
7.6 A decomposition of the cross section systematic error broken into leading GENIE systematic error contributors for E_{avail} vs q_3. The leading systematic uncertainty is GENIE_MaRES (light blue), GENIE_MvRES (green) and GENIE_NormCCRES (yellow). 226

7.7 Cross section error summary showing GENIE FSI and GENIE interaction model uncertainty contribution. 227
7.8 Cross section error summary showing MnvTune uncertainties. 229
7.9 Distribution used to derive leakage systematic uncertainty value. 233
7.10 Cross section error summary looking at the leading detector response uncertainties. 234
7.11 Systematic uncertainties for GEANT. 235

8.1 $d^2\sigma/dE_{avail}dq_3$ cross section per nucleon compared to the model with RPA and tune 2p2h components. Figure from Ref. [107] 241
8.2 The $\bar{\nu}_e$ cross section result presented in this thesis truncated at 0.8 GeV on logy scale in q_3 for comparison to the LE result. 242
8.3 Neutrino energy estimator of electron energy plus available energy on the x-axis vs true neutrino energy on the y-axis. 247
8.4 MC comparison of the $\bar{\nu}_e$ (red) and the flux weighted $\bar{\nu}_e$ (green) samples in true neutrino energy. 248
8.5 Comparison plot between the true efficiency of the flux weighted $\bar{\nu}_\mu$ events (green) and the unweighted $\bar{\nu}_e$ events (red). 250
8.6 Comparison plot between the reconstructed efficiency of the flux weighted $\bar{\nu}_\mu$ events (green) and the unweighted $\bar{\nu}_e$ events (red). 251

8.7 The flux weighted and reconstructed efficiency corrected $\bar{\nu}_\mu$ sample compared for MC (light green) and to the reconstructed efficiency corrected $\bar{\nu}_e$ sample for MC (light red) and data on a logy scale. The data has not yet been studied in detail and no comments can be made about its agreement to the MC. 252

8.8 A MC ratio of the reconstructed efficiency corrected $\bar{\nu}_e$ sample to the the flux weighted and reconstructed efficiency corrected $\bar{\nu}_\mu$ sample. 253

1. Introduction

1.1 Neutrinos and the Standard Model

The Standard Model of Particle Physics is the theory explaining the fundamental constituents that make up all particles and interactions of matter. The Standard Model contains six quarks denoted u, d, t, b, s, and c. Quarks cannot be free in nature and so they form two-quark groups (mesons) and three-quark groups (baryons) called hadrons. The most stable hadrons are protons (uud) and neutrons (udd). While hadrons are composed of constituent quarks, leptons are particles that are unable to be broken down into smaller units of matter. The Standard Model has three negatively charged leptons, which are described ass having lepton flavors called the electron (e), the muon (μ) and the tau (τ). Each of the charged leptons is accompanied by a neutral lepton partner called the neutrino, denoted by their lepton flavor as ν_e, ν_μ and ν_τ. Quarks and leptons have an equivalent antiparticle denoted either by a bar ($\bar{\nu}_e$) in the case of quarks and neutrinos or by a superscript of $+$ referring to the positive charge.

The Standard Model encompasses three of the four known fundamental forces referred to as the electromagnetic, strong and weak forces. The exchange of a boson mediates

each interaction. The electromagnetic force, the interaction between electrically charged particles, is mediated by the exchange of a photon (γ). The strong force, responsible for the confinement of quarks to a nucleus, is mediated by the exchange of a gluon (g). The weak force, responsible for the decay of unstable subatomic particles, is mediated by the exchanging of a charged W or neutral Z boson.

Neutrinos, the focus of this analysis, are neutral and interact via the weak force, limiting interactions to a very short effective range on the order of 10^{-18} meters. Weak interactions occur in two separate channels through either the exchange of a charged W boson (charged current interaction) or a neutral Z boson (neutral current interaction). Figure 1.1 shows first-order Feynman diagrams for neutrino scattering on electrons for the two separate channels.

Although the Standard Model of particle physics is generally successful in describing particles and their interactions, important unanswered questions remain. Even with the recent addition of the Higgs boson, the mechanism by which particles receive mass, the Standard Model still has glaring deficiencies. The Standard Model does not explain the matter-antimatter asymmetry observed in the universe, nor does it encompass the fourth fundamental force known as gravity. It also does not provide a dark matter candidate and there are gaps in its prediction of neutrino phenomena. One of the Standard Model predictions is that neutrinos are massless. However, experiments have found that neutrinos are able to oscillate between flavors. Neutrino oscillation is a phenomenon in which a neutrino created as a specific lepton flavor can later be measured to have a different lepton flavor, an occurrence only possible if the neutrino has some small non-negligible mass. The remainder of this chapter will explain the historical significance of the neutrino as well as

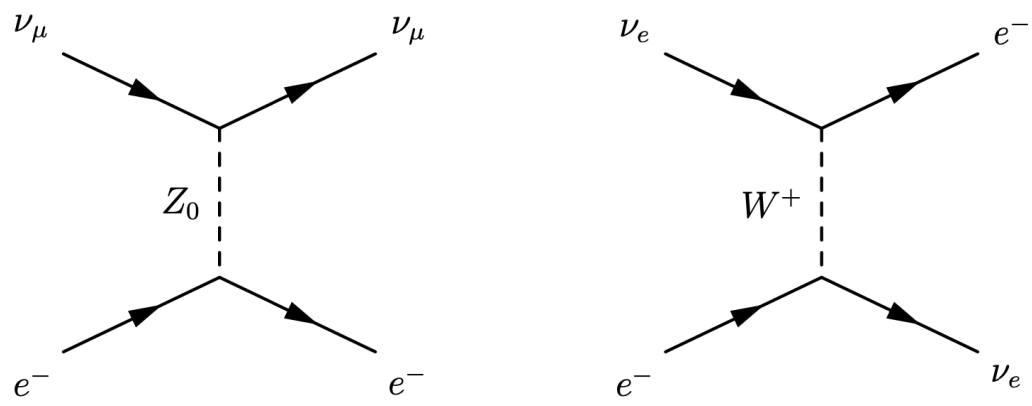


Figure 1.1: Feynman diagrams for neutrino electron scattering. All flavors of neutrinos are involved through neutral current type interactions (left) while only electron neutrinos interact through charged current interactions (right).

the mechanisms by which neutrinos oscillate.

1.2 Brief History of the Neutrino

The scientific history of the neutrino began in 1930 when Wolfgang Pauli proposed a lightweight neutral particle as a solution allowing for the conservation of momentum in beta decay. These particles (originally known as neutrons) would carry away energy allowing energy, momentum, and spin to be conserved but would be able to escape direct experimental detection. Although Pauli's hypothetical particle addressed a standing issue in beta decay, he was concerned about the implication, famously writing in a letter for the Tübingen convention: "I have done a terrible thing, I have postulated a particle that cannot be detected" [1]. Despite Pauli's concerns about experimental verification, in 1932 James Chadwick proved the existence of the first neutral particle, known now as the neutron [2].

The discovery confirmed the existence of particles devoid of charge, but the neutron was much too massive to be Pauli's hypothetical lightweight particle. Enrico Fermi would go on to popularize the name neutrino meaning "the little neutral one" in Italian to describe Pauli's hypothetical particle referencing its small mass [3]. In 1934, Fermi delivered his own theory of beta decay

$$n^0 \rightarrow p^+ + e^- + \bar{\nu}_e \tag{1.1}$$

proposing that particles can change their identity using a new force that is extremely weak compared to the only existing known force at the time, electromagnetism. Termed Fermi theory, this theory of beta decay required the neutrino to be chargeless, massless and was the precursor to the theory of weak interactions [4].

The first attempt at experimental detection of the theorized neutrino occurred in 1939 by Horace Crane, who used a neutrino capture method. The method was to expose a bag of salt (NaCl) to a radioactive material (neutrino source ν) and observe the process of neutrino absorption

$$\nu + {}^{35}Cl \rightarrow {}^{35}S + e^+ \tag{1.2}$$

by detecting sulfur produced through inverse beta decay of chlorine. Crane was the first to establish an upper limit of 10^{-30} cm^2 for the cross section of neutrino capture by chlorine-35 [5].

In 1956, the experiment now known as the Cowan-Reines experiment was the first to prove the existence of the neutrino. Cowan-Reines used a nuclear reactor as a source of

neutrinos providing a neutrino flux far greater than any other radioactive source. In the experiment, a cadmium (Cd) doped water target was set between two liquid scintillator detectors. An anti-neutrino produced by the nuclear reactor through beta decay would have a signature in the Cd appearing as

$$\bar{\nu}_e + p \rightarrow n^0 + e^+ \tag{1.3}$$

where the positron (e^+) proceeds to interact with an electron, annihilate and immediately produce two energetic photons, also known as gammas. Cadmium allows for a high level of discrimination because it absorbs the thermal neutron (n^0) produced and simultaneously releases a gamma ray, allowing for detection in the liquid scintillator. With the proton annihilation and neutron capture, the coincidence between the gammas rays is used as the distinct signature of the anti-neutrino event. The anti-neutrino discovered by Cowan and Reines become known as the anti-electron neutrino, $\bar{\nu}_e$ [6].

In 1962 Lederman, Schwartz, and Steinberger won the 1988 Nobel prize for their discovery of the muon neutrino, allowing for the existence of more than one type of neutrino. In 1975 a third type of lepton, called the tau, was discovered at the Stanford Linear Accelerator Center and with it came the expectation for the third flavor of neutrino. However, the tau neutrino would not be experimentally detected until 2000 by the DONUT (Direct Observation of the Nu Tau) collaboration [7].

1.3 The Solar Neutrino Problem

In the late 1960s, astrophysicists Raymond Davis and John Bahcall spearheaded an experiment in the Homestake Gold Mine of South Dakota (referred to as the Homestake experiment) to test whether converting hydrogen nuclei to helium nuclei was the primary source of energy in the sun [8]. The thermonuclear reaction chains taking place in the sun generate solar energy and the Standard Solar Model [9] predicts that over 98 percent of this energy is generated from the proton-proton (pp) chain conversion of four protons into helium

$$4p \rightarrow^{4} He + 2e^{+} + 2\nu_e. \tag{1.4}$$

Bahcall theorized the number of produced neutrinos with different energies from the sun and predicted how many neutrinos would interact with Cl atoms in a 380 cubic meter tank. The tank, filled with 610 tons of C_2Cl_4, was set 4,900 meters underground to shield it from unwanted cosmic rays. The signature for neutrino capture on chlorine is given by

$$\nu_e +^{37} Cl \rightarrow e^{-} +^{37} Ar. \tag{1.5}$$

The radioactive isotopes of ^{37}Ar were extracted by a helium purge and placed in a shielded counter where Auger electrons, electrons produced at very low energy emitted by radionuclides, were produced at 2.8 KeV, giving distinct pulses with a fast rise. Based on the predicted flux of neutrinos produced from boron-8 beta decay, it was theorized that the argon count should be 1.5 Ar atoms per day. The experiment took continuous data for a 20

year period and eventually the experiment published a result of 0.49 ± 0.03 Ar atoms per day. This observational rate was $\frac{1}{3}$ the theoretical prediction [10] [11]. This discrepancy, known as the Solar Neutrino Problem, persisted for the next 30 years.

Many subsequent experiments would go on to confirm the Solar neutrino deficit reported by the Homestake experiment. It was eventually realized that neutrinos were even more complicated than initially thought. In 1957, Bruno Pontecorvo postulated that the three flavors of neutrinos could oscillate back and forth between one another as the neutrino propagates through space. However, to oscillate, at least one of the neutrino flavors requires some small but non-zero mass. Finally, in 1998, the first confirmation of neutrino and anti-neutrino oscillations was reported by the Super-Kamiokande (Super-K) atmospheric neutrino experiment and the Sudbury Neutrino Observatory (SNO) solar neutrino experiment.

Super-K was designed to study high energy neutrinos produced as secondary particles from cosmic rays. Cosmic rays, originally discovered by Victor Hess in 1911, are found to contain energetic protons emitted in every direction. When cosmic rays hit the earth's upper atmosphere they interact with atoms and produce a shower of hadrons [12]. The hadrons, mostly pions and kaons, have very short lifetimes and will immediately decay into muons and neutrinos. The muons will again decay and produce electrons and more neutrinos, referred to as atmospheric neutrinos [13]. The decay chain, found in Figure 1.2

is written as

$$\pi^+ \rightarrow \mu^+ + \nu_\mu$$

$$\mu^+ \rightarrow e^+ + \nu_e + \bar{\nu}_\mu$$

$$\pi^- \rightarrow \mu^- + \bar{\nu}_\mu$$

$$\mu^- \rightarrow e^- + \nu_\mu \bar{\nu}_e.$$

The atmospheric neutrinos can transit through the earth and interact in underground detectors designed to detect the neutrinos scattering off of nuclei. Super-K, a 50,000 metric ton water Cherenkov device (a particle detector based on the Cherenkov effect where particles travel faster than light in a medium), first provided conclusive evidence for the flavor oscillation of atmospheric neutrinos. The predicted neutrino flux on the earth's surface depends on the location of the experiment and the direction of the incoming cosmic rays. Neutrinos passing through the earth before entering the detector can interact with the rock in the earth, producing some high energy muons, known as upward-going muons. Muons produced from neutrino interactions above the detector are indistinguishable from the constant shower of muons created from cosmic rays. However, cosmic ray muons do not have enough energy to traverse the rock in the earth; therefore, upward-going muons provide the key signature of a neutrino interaction. Results are typically shown in the form of a double ratio, R, which is the ratio of neutrino fluxes through observation at the detector and theoretical simulation based on the neutrino flux model predictions.

$$R = \frac{(\frac{\nu_\mu}{\nu_e})_{observed}}{(\frac{\nu_\mu}{\nu_e})_{theoretical}} \tag{1.6}$$

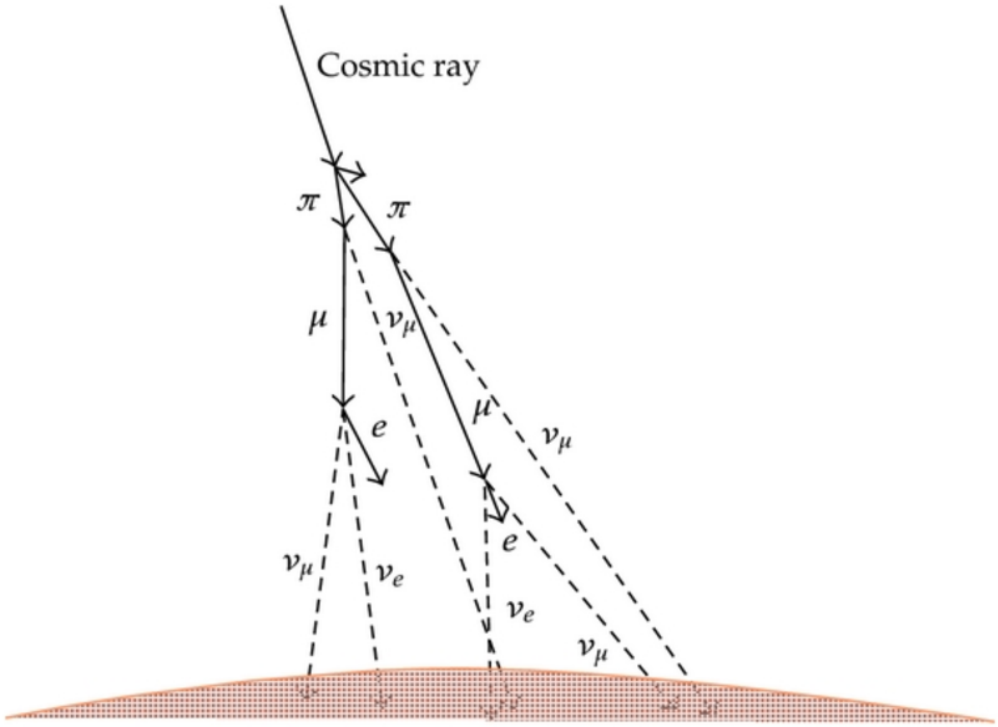


Figure 1.2: Cosmic ray interactions with the atmosphere. Figure from Ref. [13]

In the ratio from Equation 1.6, if the data observed and the theoretical prediction are equivalent then the ratio cancels to about 1, however the Super-K ratio result was R = 0.675±0.085. The Super-K neutrino data showed strong evidence for $\nu_\mu \rightarrow \nu_\tau$ oscillation [14], also known as ν_μ disappearance. The discovery of atmospheric neutrino oscillations was announced in 1998 by Super-K.

Built in 1984, SNO was a neutrino observatory designed to detect solar neutrinos with a heavy water Cherenkov detector looking to measure $\nu_e \rightarrow \nu_\mu$ and $\nu_e \rightarrow \nu_\tau$ oscillations

through charged current and neutral current neutrino interactions.

$$\nu_e + {}^2H \rightarrow p + p + e^- \tag{1.7}$$

$$\nu_x + {}^2H \rightarrow n + p + \nu_x \tag{1.8}$$

The charged current reaction seen in Equation 1.7 occurs when a neutrino converts the neutron in a deuteron (${}^2H$) to a proton. The neutrino is absorbed and an electron is produced and detected. Solar neutrinos have energy smaller than the mass of a μ and a τ, limiting the interactions to only electron neutrinos. After the electron interacts in the detector, it carries off most of the energy, leaving a resultant proton with barely enough energy to be detected. In the case of a neutral current reaction seen in Equation 1.8, the neutrino breaks up the heavy water into its constituent proton and neutron and continues traveling through the detector with slightly less energy. The low energy neutron is captured by a deuterium nucleus. The capture produces a photon that can be detected. All three neutrinos can interact in SNO's detector in this manner. The total flux of all neutrino flavors measured by SNO agreed with the theoretical prediction for neutrino oscillations [15].

The combined efforts from Super-K and SNO, building on the results of previous experiments, concluded that the Solar Neutrino problem was a consequence of neutrino oscillations. One of the consequences of neutrino oscillations is that neutrinos must have mass, directly contradicting the Standard Model of Particle Physics' requirement that neutrinos are massless. Additional experiments like those that look at neutrinos from nuclear reactors ($\bar{\nu}$ disappearance) and neutrinos produced by particle accelerators ($\overset{(-)}{\nu}_\mu$ disappearance, $\bar{\nu}_\mu \rightarrow \bar{\nu}_e$, and $\nu_\mu \rightarrow \nu_\tau$) among others, altogether help us determine the

actual parameters that govern the oscillations.

1.4 Neutrino Oscillations

Each neutrino flavor eigenstate is a coherent sum of mass eigenstates. As the coherent eigenstate propagates through space, it changes its flavor composition. Therefore, a neutrino beginning its journey as an electron neutrino will become some mixture of an electron, muon, and tau neutrino after traveling some distance. The rate of mixing is governed by quantum mechanical phases, each unique to the three neutrino mass states. The mixing can be written as a unitary transformation relating the flavor and mass eigenstates. The mixing of flavor states to form mass states was first observed in the quark sector; however, the following section will describe the analogous mathematical case of lepton mixing.

1.5 Neutrino Oscillation Theory

The theory of neutrino oscillations between three flavor eigenstates results from the weak coupling neutrino flavor states $|\nu_\alpha\rangle(\alpha = e, \mu, \tau)$ being superpositions of the neutrino mass eigenstates $|\nu_i\rangle(i = 1, 2, 3)$,

$$|\nu_\alpha\rangle = \sum_{i=1}^{3} U^\dagger_{\alpha i}|\nu_i >$$
$$|\nu_i\rangle = \sum_{i=1}^{3} U_{\alpha i}|\nu_\alpha >$$

for $U^{\dagger}_{\alpha i}$, $U_{\alpha i}$ the elements of the 3×3 PMNS (Pontecorvo-Maki-Nakagawa-Sakata) mixing matrix. For three neutrino flavors α refers to e, μ, τ and the index i refers to the mass eigenstates 1, 2, or 3. Properties of the PMNS matrix can be derived based on the mathematics of a unitary matrix.

In linear algebra, a complex square matrix is unitary if its conjugate transpose is also its inverse, or $UU^{\dagger}$ = $U^{\dagger}U$ = I. Consequently, transforming from the flavor eigenbasis to the mass eigenbasis and back returns to the original state. A unitary matrix also meets orthogonality requirements $< U_i, U_j >= \sum_k \bar{U}_{ki} U_{kj} = \delta_{ij}$. In the case $i = j$, $\sum_{k=1}^{n} |U_{ki}|^2 = 1$, implying that for $i = 1, 2....n$ there are n independent real conditions. In the case $i \neq j$ there are $\frac{n(n-1)}{2}$ real elements below the matrix diagonal and $\frac{n(n-1)}{2}$ elements above the matrix diagonal. This gives the total number of independent real conditions as, $n + \frac{1}{2}n(n-1) + \frac{1}{2}n(n-1) = n^2$ unitary constraints. Consequently, the total number of independent real parameters in the PMNS matrix is equal to the total number of real numbers minus the total number of independent real conditions, $2n^2 - n^2 = n^2$. A general n×n unitary matrix depends on n^2 real parameters. However, in the PMNS matrix not all of these parameters have physical meaning. There are $2n$-1 lepton phases that are not physically meaningful and may be removed by phase re-definitions [16]. The number of free parameters is actually the n^2 real parameters minus the $2n - 1$ phase freedom.

$$= n^2 - (2n - 1)$$
$$= (n - 1)^2$$

Of these $\frac{n(n-1)}{2}$ are mixing angles and $\frac{(n-1)(n-2)}{2}$ are complex phases. For the Standard

Model there are three generations of leptons (n = 3), giving three mixing angles and one complex phase as seen in Table 1.1.

n(families)	Total independent params. $(n-1)^2$	Real Mixing Angles $\frac{1}{2}n(n-1)$	Complex Phase Factors $\frac{1}{2}$(n-1)(n-2)
2	1	1	0
3	4	3	1
4	9	6	3

Table 1.1: Given n generations of leptons, the corresponding number of mixing angles and complex phase factors.

Mass eigenstates, $|\nu_i\rangle$, can be expanded as plane wave solutions to the time-dependent Schrödinger equation as

$$|\nu_i(t)\rangle = e^{-i(E_i t - \vec{p}_i \cdots \vec{x})}|\nu_i\rangle,$$

for $\vec{p}_i$ and $\vec{x}$ the momentum and position four-vectors of the neutrino and the mass state energy E_i of the eigenstate given by the special relativity relation

$$E^2 = \vec{p}^{\,2} + m^2.$$

Though the precise mass of the neutrino remains unknown, it is known that the neutrino masses are very small (upper bound of 1.1 eV [17]). Therefore, neutrinos are typically considered ultra-relativistic ($E^2 \gg m^2$). In this relativistic regime the magnitude of the

momentum mass state from Einstein's energy equation is

$$
\begin{aligned}
|\vec{p}_j| &= \sqrt{E_j^2 + m_j^2} \\
|\vec{p}_j| &= E_j \sqrt{1 + \frac{m_j^2}{E_j^2}} \\
|\vec{p}_j| &\approx E_j (1 - \frac{m_j^2}{2E_j^2}) \\
|\vec{p}_j| &\approx E_j - \frac{m_j^2}{2E_j}.
\end{aligned}
$$

For the plane wave propagation in the relativistic limit ($t \approx L$, for L the distance the neutrino travels)[1] written as

$$
\begin{aligned}
e^{-i(E_j t - \vec{p}_j \cdot \vec{x})} &\approx e^{-i(E_j - |\vec{p}|)L} \\
&\approx e^{-i[E_j - (E_j - \frac{m_j^2}{2E_j})]L} \\
&\approx e^{-i\frac{m_j^2}{2E_j}L}
\end{aligned}
$$

the propagation of the neutrino flavor state in terms of the mass eigenstate evolves as

$$
|\nu_\alpha(t \approx L)\rangle = \sum_{j=1}^{3} U_{\alpha j}^* e^{-i\frac{m_j^2}{2E_j}L} |\nu_j >
$$

assuming a fixed energy state $E \equiv E_j$. Each mass state will propagate at different rates if

[1] We are using the "natural unit" convention, where t=L/c for c=1. $\hbar$ is also assumed to be 1.

the masses among the three neutrino flavors are nonequivalent and consequently, a neutrino beginning as a pure state of a certain flavor can oscillate into a mixture of flavor eigenstates as it propagates through space. The probability of a neutrino created in a pure flavor state $|\nu_\alpha\rangle$ interacting as a flavor state $|\nu_\beta\rangle$ after propagating some distance L in a vacuum is given by

$$\begin{aligned}
P(\nu_\alpha \to \nu_\beta) = |\left\langle \nu_\alpha | \nu_\beta(L) \right\rangle|^2 &= |(\sum_j \left\langle \nu_j | U_{\beta j}) (\sum_j U^*_{\alpha j} | \nu_j(L) \right\rangle)|^2 \\
&= |(\sum_j \left\langle \nu_j | U_{\beta j}) \sum_j U^*_{\alpha j} e^{-i\frac{m_j^2}{2E_j}L} | \nu_j \right\rangle)|^2 \\
&= |\sum_j U_{\beta j} U^*_{\alpha j} e^{-i\frac{m_j^2}{2E_j}L}|^2 \\
&= \delta_{\alpha\beta} - 4\sum_{j>k} \Re(U^*_{\alpha j} U_{\beta j} U_{\alpha k} U^*_{\beta k}) \sin^2(\frac{\Delta m^2_{jk} L}{4E}) \\
&+ 2\sum_{j>k} \Im(U^*_{\alpha j} U_{\beta j} U_{\alpha k} U^*_{\beta k}) \sin^2(\frac{\Delta m^2_{jk} L}{2E})
\end{aligned}$$

for $\Delta m^2_{jk} = m_j^2 - m_k^2$ and a fixed energy state [18].

It can be helpful to reduce the complexity of the three-flavor mixing by using the example of the probability of oscillation in the case of a two-flavor mixing. The probability that a neutrino begins in an initial flavor state α and oscillates into flavor state β seen in Equation 1.9.

$$P_{\alpha\to\beta} = sin^2(2\theta_{\alpha\beta}) sin^2(\frac{\Delta m^2_{\alpha\beta} L}{4E}) \tag{1.9}$$

The simplified two-flavor equation gives better insight into the case of neutrino oscillation of $\alpha \to \beta$. Neither the mixing angle nor the mass scale can be directly measured because

the squared nature of $\Delta m^2_{\alpha\beta}$, referred to as the "mass squared-splitting", indicates that only the difference of masses squared can be determined. Additionally, the oscillation must be observed at several different L/E values to provide separation between the $\Delta m^2_{\alpha\beta}$ and $\theta_{\alpha\beta}$ terms. The oscillation also depends on the distance the neutrinos travel L and their energy E each being controllable experimental parameters used to maximize the probability of oscillation.

The reality of the three generations in the neutrino sector is a rather complex three-dimensional mixing matrix parameterized by three mixing angles and three phases.

$$U = \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix} \begin{pmatrix} c_{13} & 0 & s_{13}e^{-i\delta}_{CP} \\ 0 & 1 & 0 \\ -s_{13}e^{-i\delta}_{CP} & 0 & c_{13} \end{pmatrix} \begin{pmatrix} c_{21} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} e^{i\eta_1} & 0 & 0 \\ 0 & e^{i\eta_2} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

for $s_{ij} = \sin\theta_{ij}$ and $c_{ij} = \cos\theta_{ij}$ the three mixing angles, the δ_{CP} phase, relating it to the breaking of symmetry under charge-parity operations and $\eta_{1,2}$ the Majorana phases. A non-zero δ_{CP} would lead to a difference between the oscillation rates of neutrinos and anti-neutrinos and $\eta_{1,2}$ is only non-zero if neutrinos are Majorana particles (their own antiparticle). The Majorana phases are not detectable in neutrino oscillations, reducing the parameterization to three mixing angles and one phase factor. The complete PMNS matrix

may be expressed as

$$U = \begin{bmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{bmatrix}.$$

Oscillation parameters are inferred from fits extracted from the measurements of two flavors of ν_μ and ν_e. Visualization of three neutrino oscillations in a vacuum is shown in Figure 1.3. Note the dependency of the L/E terms on the oscillations.

1.6 Current Oscillation Experiments

The current strategy for studying neutrino oscillations is to build two detectors: one at a distance of small L/E, referred to as the near detector, and another some distance downstream at L, referred to as the far detector. Neutrinos are first observed interacting in the near detector, where their energy and flavor distributions are measured. The neutrino beam is then sent to the far detector a distance L away and again measured for the energy and flavor, allowing the neutrinos the chance to oscillate to a different flavor.

There are major efforts to measure neutrino oscillations, including long baseline experiments, T2K (Tokai to Kamioka) [20] and NOvA (NuMI Off-axis ν_e Appearance) [21], and a third, the Deep Underground Neutrino Experiment (DUNE) [22], currently under construction and set to run in the near future. T2K, located in Japan, studies the oscillation of muon (anti)neutrinos into electron (anti)neutrinos at a baseline distance of about 300 km. NOvA, located in the US, studies $\nu_\mu \to \nu_e$ oscillations at a baseline around 800 km.

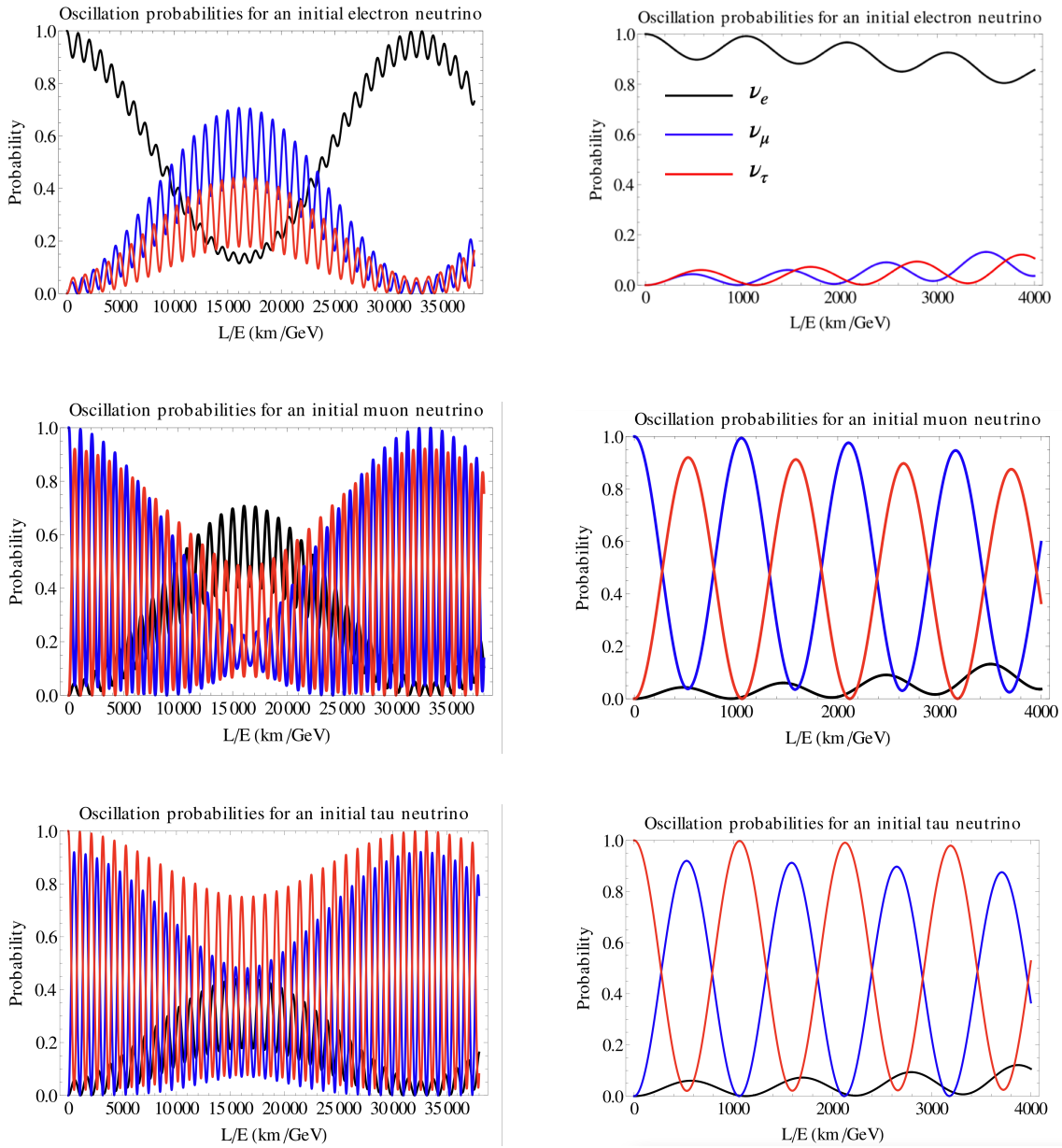


Figure 1.3: Three neutrino oscillation in a vacuum for long (left) and shot (right) baselines beginning with initial neutrinos of ν_e (top), ν_μ (middle) and ν_τ (bottom). Figure from Ref. [19].

To understand the goal of these experiments more precisely, consider a neutrino beam composed of ν_μ. An *appearance* experiment would begin with the ν_μ neutrino beam, measuring its composition in the near detector then counting the number of ν_e interactions seen in the far detector. In practice, identifying a ν_e interaction can be complicated and the algorithms used to reconstruct interactions are imperfect. The algorithm might mistake some type of interaction for a ν_e interaction, incorrectly counting it towards the total. Nuclear effects also make it difficult to predict the exact number of ν_e interactions and can be highly model dependent, an important topic discussed in Chapter 3. As a result, oscillation experiments typically apply corrections to the muon neutrino cross-section based on theoretical arguments and obtain a prediction for the electron neutrino cross-section. These oscillation experiments have a similar construction beginning with an intense beam of ν_μ fired into the near detector. The ν_e sample population is on the order of a few parts per thousand of the total neutrino flux, making for a challenging measurement.[2] The results presented in this thesis can serve as a blueprint for $\bar{\nu}_e$ measurements pertinent to current and future oscillation experiments.

There are many unresolved questions regarding neutrinos. Physicists look to measure the parameter δ_{CP} in which a non-zero value would indicate that neutrinos and anti-neutrinos undergo oscillation at different rates. Such a discovery could significantly impact the question of why the universe contains more matter than antimatter. Additionally, the signs of Δm^2 would define the ordering of masses of the neutrinos. The determination of each of these parameters relies on high-precision measurements of several different energy-dependent flavor transition probabilities. For example, an experiment beginning with a

[2]This number is slightly better for DUNE's on-axis near detector [23]

beam of muon neutrinos and observing the oscillation into electron neutrinos takes into account the oscillation probability term seen in Equation 1.10

$$\frac{\sin 2\theta_{12} \sin 2\theta_{23}}{2 sin\theta_{13}} \sin \frac{\Delta m_{21}^2 L}{4E} \sin^2 2\theta_{13} \sin^2 \frac{\Delta m_{31}^2 L}{4E} \sin \delta_{CP} \tag{1.10}$$

for θ_{12}, θ_{23}, and θ_{13} the three mixing angles, Δm_{21}^2, Δm_{31}^2 the mass splittings and δ_{CP} the charge-parity violating phase. The oscillation probability is heavily reliant on all factors involving the θ_{13} parameter. θ_{13} is small, about 9°, and with the wealth of data taken by oscillation experiments it is quite difficult to find small variations in such a noisy data set. Additionally, seeking out the oscillation signatures in large detectors also relies heavily on predicting the interaction properties of the neutrinos and their signatures in the respective detectors. In theory, neutrino interactions with a single nucleon can be understood and predicted well. However, the complexity of the nuclear environment introduces large uncertainties on any results that require a model to predict the kinematic picture of particles in a neutrino interaction. Neutrino oscillation experiments are required to precisely model and measure these interactions to identify the exact parameters of oscillation. This leads to the current work attempting to understand and quantify the interaction probabilities of these neutrinos on different nuclei.

Analyzing cross section distributions for neutrino interactions on different nuclear targets allows for comparisons among various models providing important information at a high level of precision. The MINERvA experiment (Main Injector Neutrino ExpeRiment to study v-A), the focus of this analysis, is one of the experiments making precise cross section measurements of neutrino-nucleus scattering, providing insight into neutrino-nucleus interaction models and reduction of systematic uncertainties pertinent to oscillation

experiments.

2. The NuMI Neutrino Beam

The Neutrinos at the Main Injector (NuMI) neutrino beam is a high-intensity neutrino beam mainly composed of muon neutrinos and anti-neutrinos located at the Fermi National Accelerator Observatory (Fermilab) in Batavia, Illinois. Fermilab's high intensity NuMI beam delivers a broad energy spectrum centered at 6 GeV. This chapter describes the process that generates the protons and the beam of neutrinos, and the energy spectrum of the produced neutrinos, with more detail found in Ref. [24].

2.1 Proton Production and Acceleration

Each stage required to produce neutrinos at the Fermilab accelerator complex is detailed in the sequence below. The accelerator complex has four main components: an H^- source, a Linac, the Booster and the Main Injection synchrotrons.

NuMI's proton source begins with H^- ions accelerated by a Cockroft-Walton accelerator to 750 keV. A single gap radio frequency (RF) cavity divides each 15 Hz beam pulse into $\approx$ 200 MHz bunches in preparation for injection into the Linac. The Linac is a linear accelerator 150 meters in length, with the first component of the Linac being a series of Alvarez drift-tubes that accelerate the H^- ion beam to 116 MeV. The second component is

a linear accelerator that accelerates the H^- ions to 400 MeV. At the end of the Linac, an RF cavity debunches the ion beam in preparation for the Booster synchrotron.

Upon injection into the Booster synchrotron [25], the ion beam is passed from the Linac through a carbon foil which strips the electrons from the ions creating the proton source. Dipole magnets steer the protons into the Booster until the Booster accumulates $\sim 3 \times 10^{12}$ protons. The protons will then undergo multiple orbits around the 150 m diameter synchrotron where they will be accelerated until they reach the Booster's target energy of 8 GeV. The 8 GeV protons are extracted from the Booster by kicker magnets for injection into the Main Injector. Each extraction is referred to as a *batch* which cycles at 15 Hz. A bird's eye view of the accelerator complex is shown in Figure 2.1.

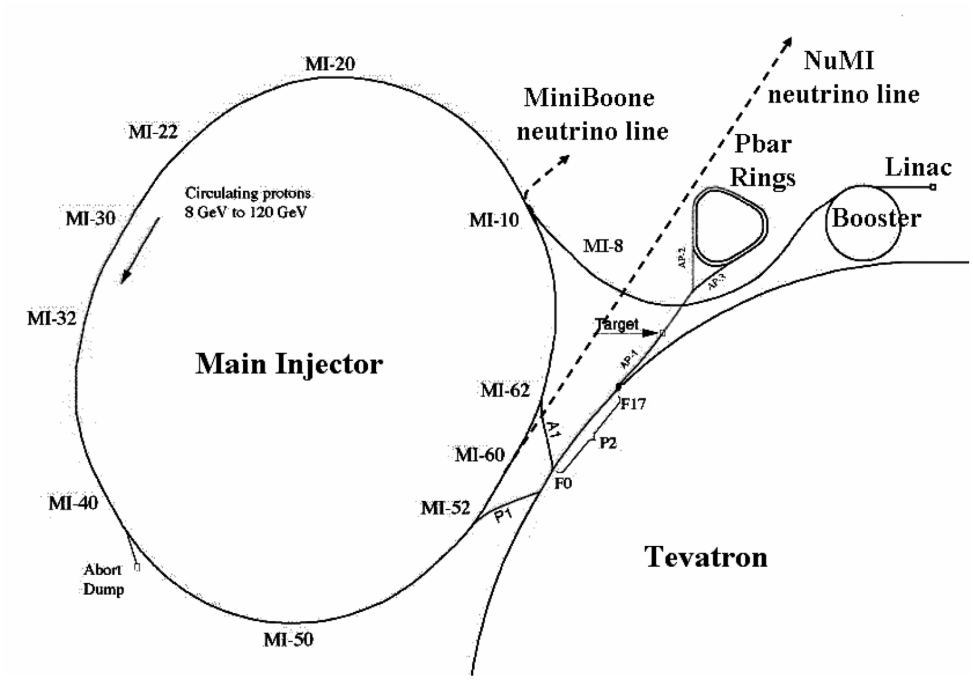


Figure 2.1: The Fermilab accelerator complex. Figure from Ref. [24].

Multiple proton batches are sent from the Booster to the Main Injector (3.3 km in

circumference), where they are accelerated to 120 GeV. The protons are extracted by kicker magnets, where each extraction is referred to as a *spill*. Each spill contains 35 $\times 10^{12}$ protons at 120 GeV. The spills are extracted at a rate of 0.45 Hz and each spill is 10 μs long. The extracted spills create a pulsed proton beam aimed at the NuMI target, located in the Target Hall.

2.2 NuMI Beamline

The proton beam spills from the Main Injector first pass through a graphite baffle which collimates the beam to 1 mm in diameter before interacting with carbon nuclei in a target region in the NuMI target hall. The proton beam is angled 58 mrad downward to hit the MINOS far detector located in Minnesota.

The NuMI target (Figure 2.2) contains 47 graphite fins 20 mm in length and 6.4 $\times$ 15 mm in rectangular cross section. The target length is 95 cm, corresponding to $\sim$ 2 nuclear interaction lengths. Water pipes run along the fin edges to keep the target cooled. The interactions of the proton beam on the carbon nuclei in the graphite target produce hadrons consisting of charged pions ($p + C \rightarrow \pi^{\pm}$), charged kaons ($p + C \rightarrow K^{\pm}$), neutrons and low energy protons. The secondary particles, mainly $\pi^{\pm}$ produced by the incident protons, are then focused by a system of two magnetic horns.

A magnetic horn is a toroidal aluminum conductor with a parabolic inner cavity. A pulsed current in each horn generates a magnetic field where the direction of the current determines which charged particles are focused and defocused. When the horns are focusing positively charged particles, the resulting beam is primarily composed of ν_{μ} and is referred

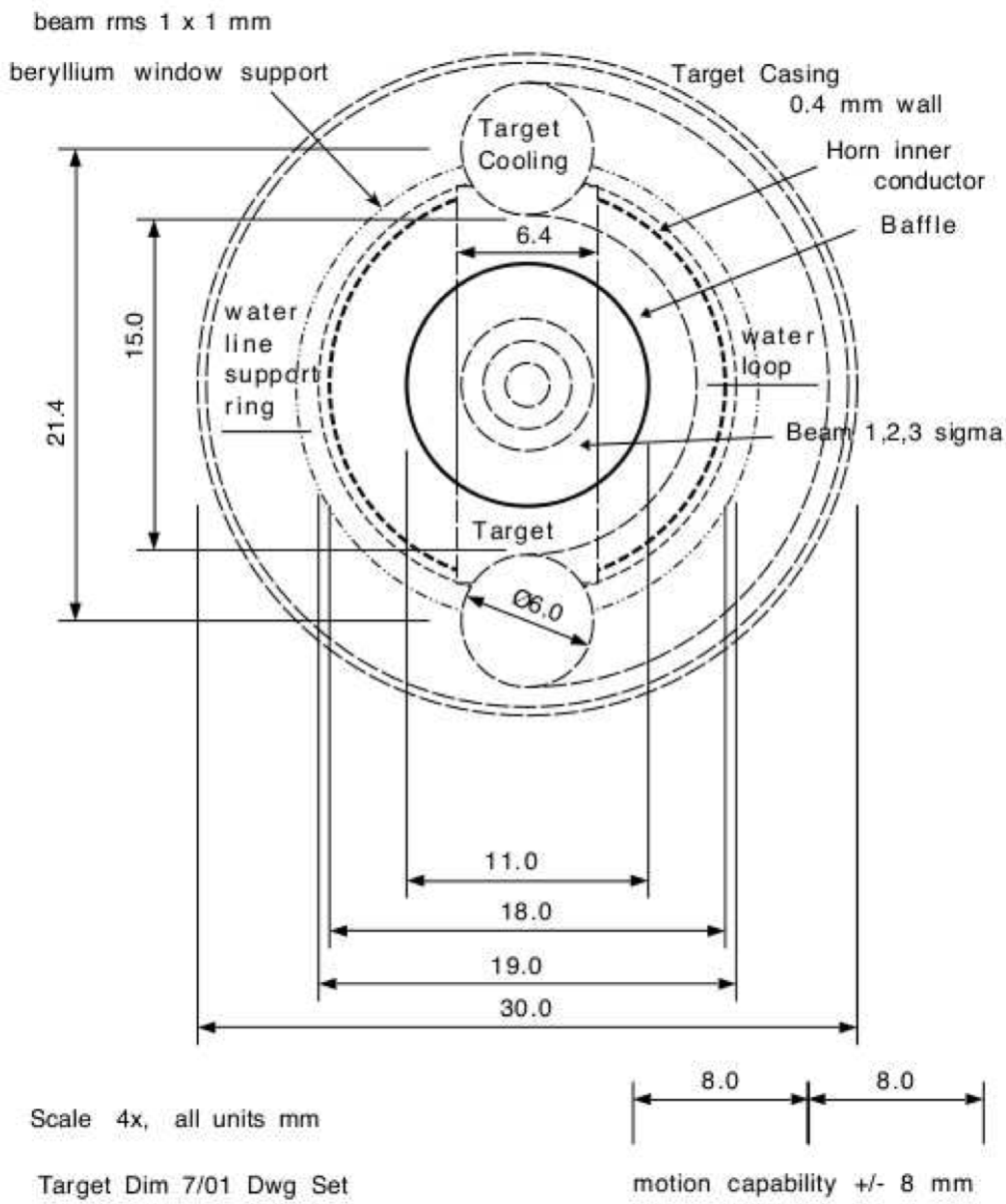


Figure 2.2: Cross sectional view of the NuMI target for a proton traveling through the NuMI baffle to hit the target. Figure from Ref. [24].

to as a forward horn current (FHC) configuration, seen in Figure 2.3. When the horns are focusing negatively charged particles, the beam is primarily composed of $\bar{\nu}_\mu$ and is referred to as a reverse horn current (RHC) configuration. The relative position of the target and horns as well as the magnitude of the horn current allows for sculpting of the energy spectra for the focused particles and, by extension, of the neutrino beam. The horns direct the

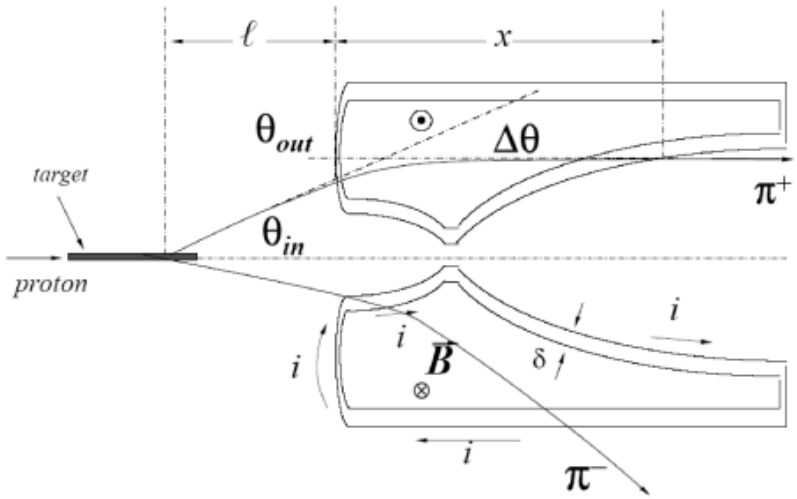


Figure 2.3: FHC horn configuration. Figure from Ref. [24].

particles, mostly pions and some charged kaons, down a 675 m long decay pipe where they will decay and create the neutrino beam. The main channels of the pion decay (99% branching ratio) are from $\pi^+ \to \mu^+ + \nu_\mu$ and $\pi^- \to \mu^- + \bar{\nu}_\mu$. However, the charged kaons have decay modes that produce electron neutrinos

$$K^+ \to \pi^0 + e^+ + \nu_e \tag{2.1}$$

$$K^- \to e^- + \bar{\nu}_e + \pi^0 \tag{2.2}$$

$$K_L^0 \to e^\mp + \pi^\pm + \nu_e. \tag{2.3}$$

The energetic muon neutrinos, stemming from the decay of pions, are boosted primarily along the direction of the beam axis. The electron neutrinos resultant from the decay of the muons, seen in Equations 2.4 and 2.5, are not as strongly boosted due to the three-body nature of the muon decay.

$$\mu^+ \rightarrow e^+ + \nu_e + \bar{\nu}_\mu \tag{2.4}$$

$$\mu^- \rightarrow e^- + \bar{\nu}_e + \nu_\mu . \tag{2.5}$$

The particles remaining at the end of the decay pipe consist of neutrinos and the hadrons and muons that have not yet decayed. Dolomite rock spans 240 m between the end of the decay pipe and the Near Detector Hall to absorb the undecayed particles and allow the neutrinos to pass through and enter the MINERvA detector.

From March 2009 to June 2012, MINERvA recorded data in the low energy (LE) beam configuration where the graphite target was in the closest proximity to the focusing horns, creating a neutrino beam with a peak neutrino energy flux of around 3 GeV. This thesis uses MINERvA data taken from June 2012 to March 2019 where the NuMI beamline was running in medium energy (ME) beam configuration. In this configuration, the beam had an on-axis flux peak at ~ 6 GeV and was run in both forward horn-current mode (FHC) and reverse horn-current mode (RHC). Figure 2.4 shows the simulated neutrino flux for the FHC beam configuration and Figure 2.5 shows the simulated neutrino flux for the RHC beam configuration

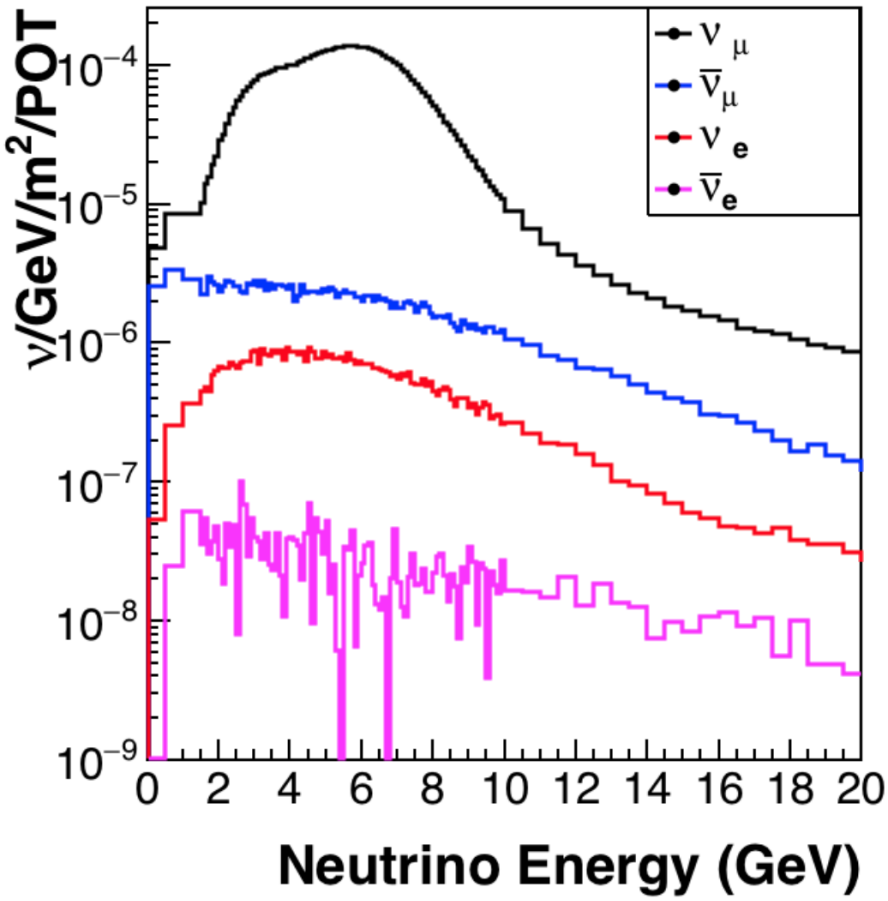


Figure 2.4: Neutrino flux for various species for forward horn current (FHC) for the MINERvA medium energy run. Figure from Ref.[26]

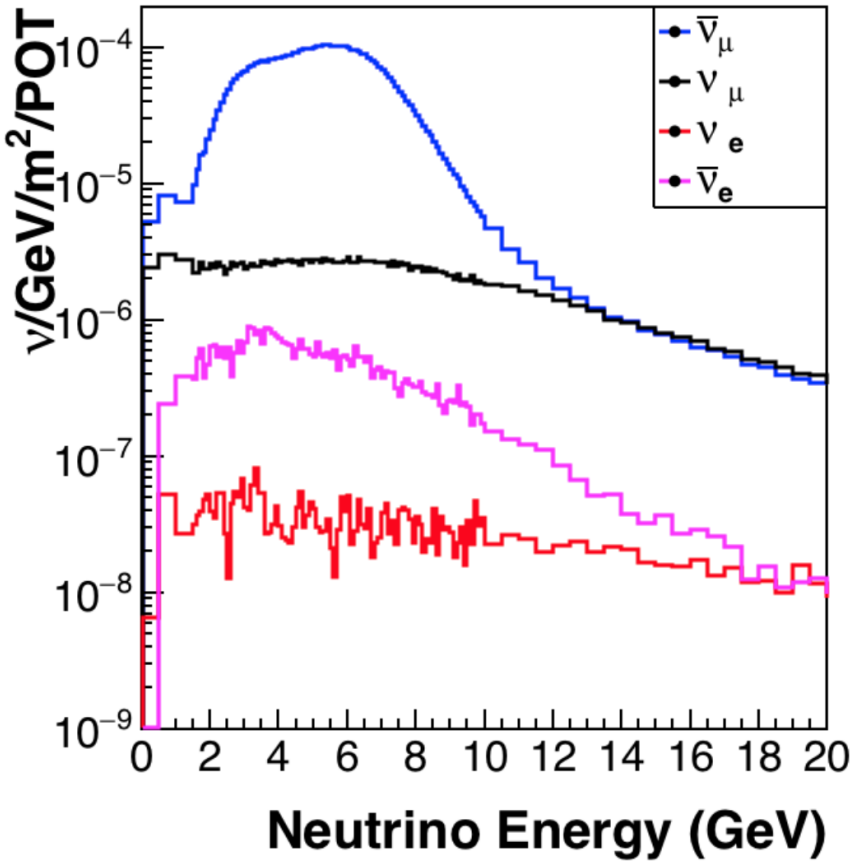


Figure 2.5: Neutrino flux for various species for reverse horn current (RHC) for the MINERvA medium energy run. Figure from Ref. [26]

2.3 Neutrino Flux Simulation

The neutrino-nucleus cross sections reported by MINERvA are normalized to the incident particle flux on the detector from the NuMI beamline. The general formula for cross section calculations shows this relation as

$$(\frac{d\sigma}{dx})_\alpha = \frac{\sum_j (N_{data,j} - N_{bkgd,j})}{A_\alpha \Phi T \Delta x_\alpha} \tag{2.6}$$

for Φ the flux, $U_{j\alpha}$ the migration matrix mapping the jth truth bin to the αth reconstructed bin, and A_α the product of the reconstructed efficiency and detector acceptance for the reconstructed bin α. The cross section details specific to this analysis will be discussed later. However, the formula shows the general dependency of the cross sections on precision measurements of the neutrino flux. Due to the inability to directly measure the neutrino energy spectrum, MINERvA requires a series of simulations to predict the incoming flux accurately. The flux simulations can be broken down into the horn and target geometry configuration, proton interaction, secondary charged particle interactions, and the decay of particles into neutrinos [27]. The horn system and the hadron production cross sections for the NuMI beam are simulated using the GEANT4 simulator with the G4numi software package. Further *in situ* measurements of neutrino-elastic scattering in MINERvA help further constrain the flux prediction.

2.3.1 Hadron Production

Accurate prediction of the neutrino flux is critical to both neutrino oscillation experiments and cross section experiments like MINERvA. The cross sections of π's and K's produced

by high energy proton interactions on carbon (and subsequent decay to neutrinos) are theoretically well understood; however, precise flux predictions are complicated by multiple interactions in the target as well as reinteractions of the materials downstream. MINERvA applies a method, referred to as PPFX (Package to Predict Flux), to constrain the hadronic models using hadron production data reported by the experiments NA49 [28] and MIPP [29]. MIPP's measurement used the NuMI LE target with 120 GeV/c protons. NA49 studied proton-carbon interactions on a thin target with incident proton momentum of 158 GeV/c. The NA49 data is scaled from 150 GeV to the NuMI 120 GeV energy spectrum through the Feynman scaling (χ_F) technique described in Ref. [30]. The ME simulated flux in Figure 2.6 shows the ratio between the corrected and uncorrected *a priori* flux predictions [31].

The flux uncertainties arising from hadronic production and horn focusing are propagated using a "many universe" technique in which a randomly sampled value from the mean in the central value of the given parameter is varied among the pseudo-random throws of the flux simulation. The "many universe" technique is discussed in greater detail in Section 7.2 and the specific flux uncertainties discussed in Section 7.4.1.

2.3.2 Neutrino Flux Constraint

MINERvA applies an additional *in situ* constraint to the flux prediction derived from the measurement of neutrino-electron scattering in the MINERvA detector. The cross section of neutrino scattering on electrons ($\nu + e^- \rightarrow \nu + e^-$), with the Feynman diagram in Figure 2.7, is known to high precision from electroweak theory [32]. Additionally, $\nu - e$ scattering occurs outside the nucleus and does not contain the complication of nuclear

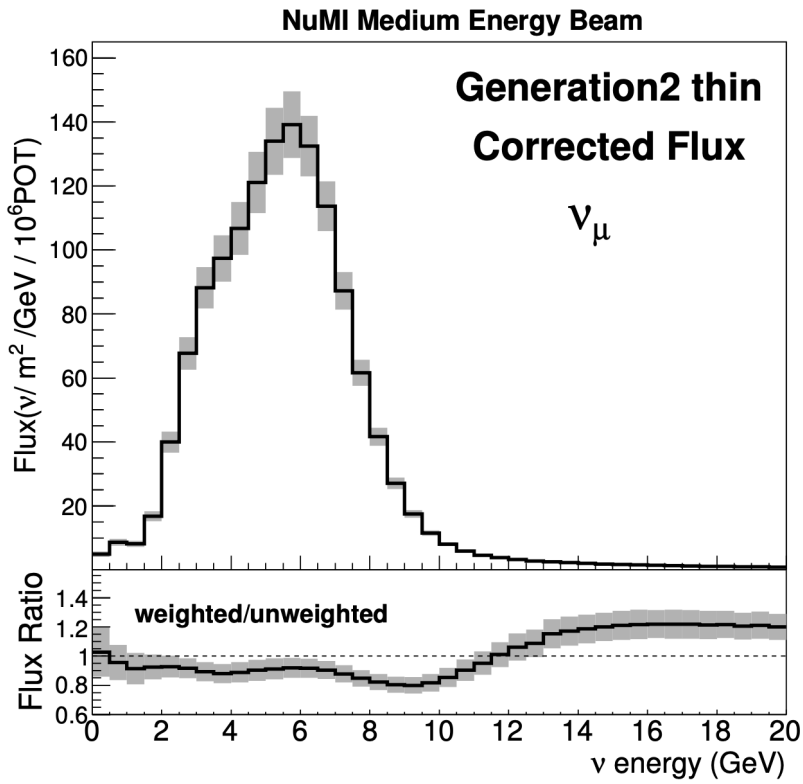


Figure 2.6: Flux distribution with hadron production weight in ME. Figure from Ref. [31].

effects. Measurements of the process took place at MINERνA for the LE [33] and ME [34] beam runs; however, the LE data was hindered by low statistics and is not investigated here. Since the neutrino-electron scattering measurement is used in combination with the *a priori* ME NuMI flux prediction the technique relies on Bayes' theorem.

$$P(theory \mid data) \; \alpha \; P(data \mid theory)\, P(theory) \tag{2.7}$$

where P(theory) is the prior probability of the *a priori* flux model and P(data | theory) is the probability of the measurement for the number of neutrino-electron scattering events

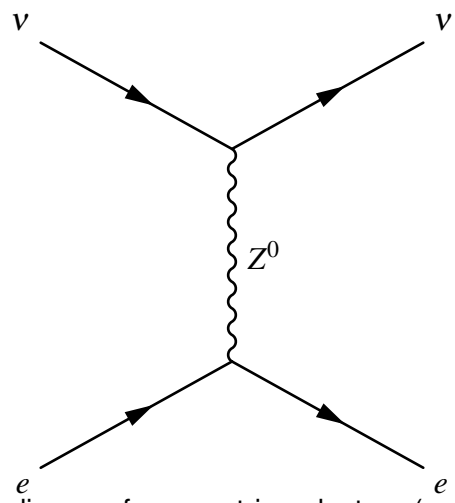


Figure 2.7: Feynman diagram for a neutrino electron ($\nu + e \rightarrow \nu + e$) interaction

given the assumption of the *a priori* flux model. A probability distribution of neutrino flux that accounts for both the *a priori* flux uncertainties and the neutrino-electron scattering measurement is obtained through the likelihood of the neutrino-electron scattering data given the models in that distribution [34].

The neutrino flux predictions are created by varying the uncertainties within the parameters of the underlying flux models and computing the spread of the resulting distributions. Each simulation is weighted by a χ^2-like distribution to construct the probability distribution for the simulated observable constrained by the $\nu - e$ data. The weight is written as

$$W = \frac{1}{(2\pi)^{K/2}} \frac{1}{|\sum_{\mathbf{N}}|^{1/2}} e^{-\frac{1}{2}(\mathbf{N}\text{-}\mathbf{M})^T \Sigma_N^{-1} (\mathbf{N}\text{-}\mathbf{M})} \tag{2.8}$$

for K the number of bins in the measurement, **N** a vector of bin contents in data, **M** a vector of the predicted flux spectrum from the simulation in question and $\sum_N$ the total data covariance matrix describing all uncertainties on **N** [34] [35] . Figure 2.8. shows the comparisons before and after applying the weight calculated using Equation 2.8.

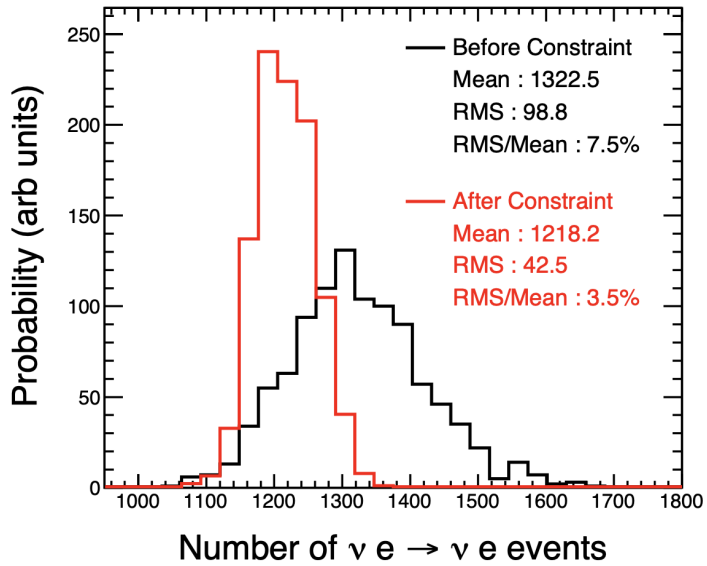


Figure 2.8: Probability distribution for number of $\nu + e$ scatters, before and after constraining the *a priori* flux model. Figure from Ref. [34].

The reduction of neutrino flux uncertainty is, on average, better for values of neutrino energies of $E_\nu < 10$ GeV. However, the highest energy neutrinos are primarily produced from $\pi^\pm$ and $K^\pm$ and are not well constrained using the $\nu - e$ constraint technique. For regions of $E_\nu > 10$ GeV, the flux can be constrained using Inverse Muon Decay (IMD), $\nu_\mu e^- \to \mu^- \nu_e$. Since IMD has a clear signature of a high energy forward-going muon with no additional particles, these specific types of events can be selected with few backgrounds. Like the neutrino electron scattering, the IMD cross section is predicted with small uncertainties and the constraint on the high energy tail of the flux is noticeable. MINERvA carried out an IMD analysis with details of the result found in Ref.[36] and summarized in Figure 2.10.

The last method discussed regarding the flux is termed the "low-ν" method. The "low-ν"

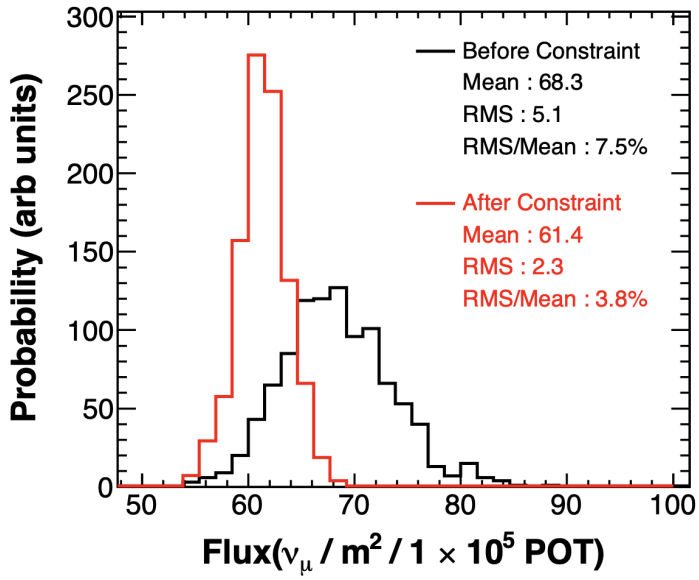


Figure 2.9: Probability distribution of the ν_μ flux between 2 and 20 GeV, before and after constraining the *a priori* flux model. Figure from Ref. [34].

method was first studied in the LE beam configuration [37] and extended to the ME beam configuration, where it was studied in conjunction with a "high-ν" method [38]. The low-ν technique constructs a measurement of the entire flux spectrum using the approximation that for low recoil energy the cross section for neutrino interactions is constant as a function of neutrino energy [38]. The flux fit method provides a way to understand the correlations between the detector and focusing parameters. An initial comparison between the *a priori* and low-ν ME fluxes in MINERvA showed a discrepancy between the two predictions. Significant shape differences in the ratio between the two fluxes indicate a shifted flux peak position. Several parameters were varied in the beamline simulation in an attempt to find agreement between the two fluxes but none were able to explain the shape difference. It was determined that MINERvA must apply a 3.6% shift to the muon momentum reconstruction

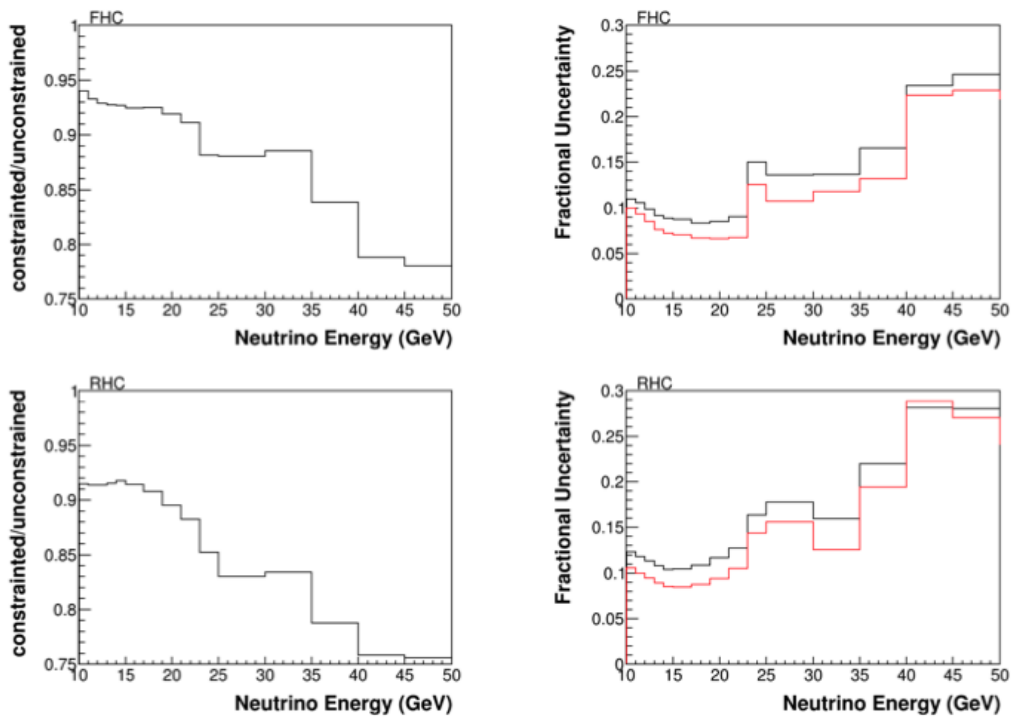


Figure 2.10: The FHC (top) and RHC (bottom) fluxes (left) and uncertainties (right) as a function of E_ν for neutrino energies > 10 MeV. Slight reduction in fractional uncertainty can be seen as a result from IMD flux constraints. Figure from Ref. [36].

in MINOS to reconcile the discrepancy. The effect of the shift can be seen in Figure 2.11 and the details of the study and implementation can be found in Ref. [39].

In conclusion, the neutrino flux used to extract this double differential cross section result must be well understood. Several studies were conducted in an attempt to further constrain the neutrino flux prediction. The final neutrino flux for the ME beam configuration is configured with peak neutrino energy around 6 GeV. The beam composition breakdown can be found in Figure 2.5 and, for this analysis with beam configuration set to RHC, the most dominant contribution is $\bar{\nu}_\mu$ followed by ν_μ, $\bar{\nu}_e$, and ν_e respectively.

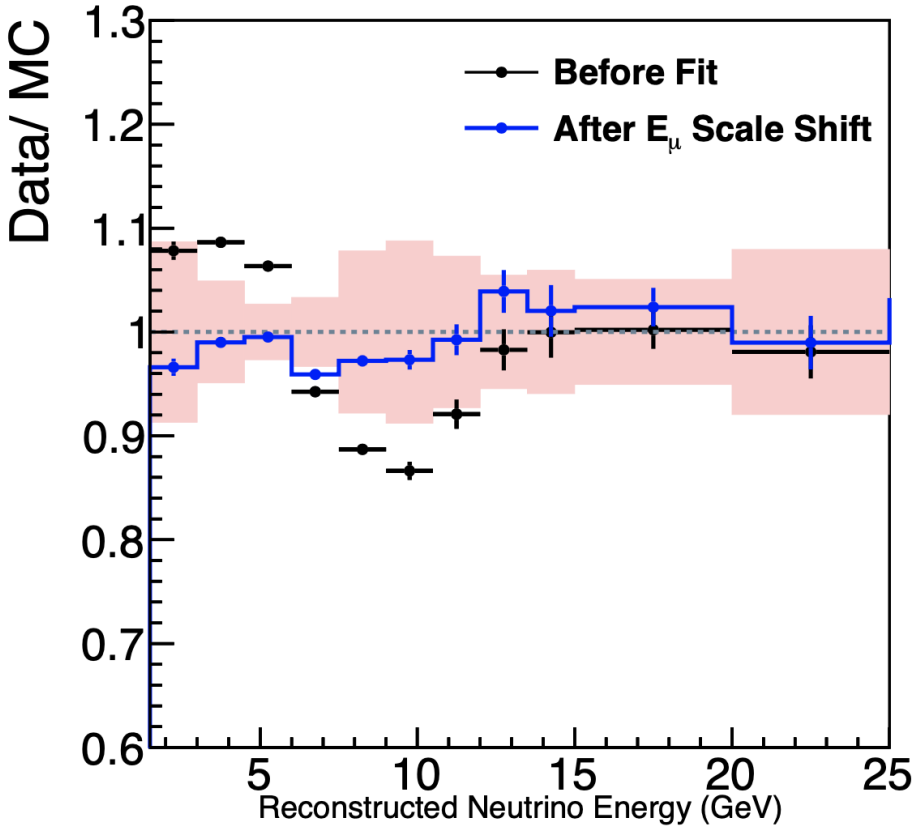


Figure 2.11: Ratio of low - ν events with Data/MC before and after the muon energy scale shift. Figure from Ref. [39].

3. The MINERvA Detector

3.1 The MINERνA Experiment

The MINERvA experiment is designed to take precise measurements of neutrino-nucleus scattering incident at energies of $1 < E_\nu < 20$ GeV. MINERvA studies neutrino scattering on various nuclei: carbon, lead, iron, plastic, and water to help understand the weak interaction in the nuclear environment [40]. MINERvA is located at Fermilab in the NuMI near detector hall located ≈ 1 km downstream of the NuMI graphite target. MINERvA can be broken down into several subdetectors, each of which serves different purposes.

The NuMI beam enters from the left of the schematic seen in Figure 3.1 (defining *downstream* following the direction of the neutrino beam and *upstream* against) and first passes through a veto wall. The purpose of the veto wall is to identify particles entering the MINERvA detector that originated from interactions in the rock upstream of the NuMI near detector hall. These particles, known as rock muons, can partly or entirely traverse the downstream detector and become a background for analyses, especially in the target region. The veto wall is composed of a 5 cm thick steel plate, a 1.9 cm thick plane of scintillator, a 2.5 cm thick steel plate, and an additional 1.9 cm thick scintillator plane. Between the veto

wall and the main detector lies a cubic meter cryogenic vessel filled with liquid helium. The upstream nuclear targets region contains water, liquid helium, graphite, lead and iron targets all separated by scintillator planes allowing for vertex position resolution for interactions originating in the nuclear targets. The central tracking region is surrounded by side and downstream electromagnetic calorimeters (ECAL) and hadronic calorimeters (HCAL). MINERvA sits directly upstream of the MINOS (Main injector neutrino oscillation search) near detector which serves as a muon spectrometer for the MINERvA experiment.

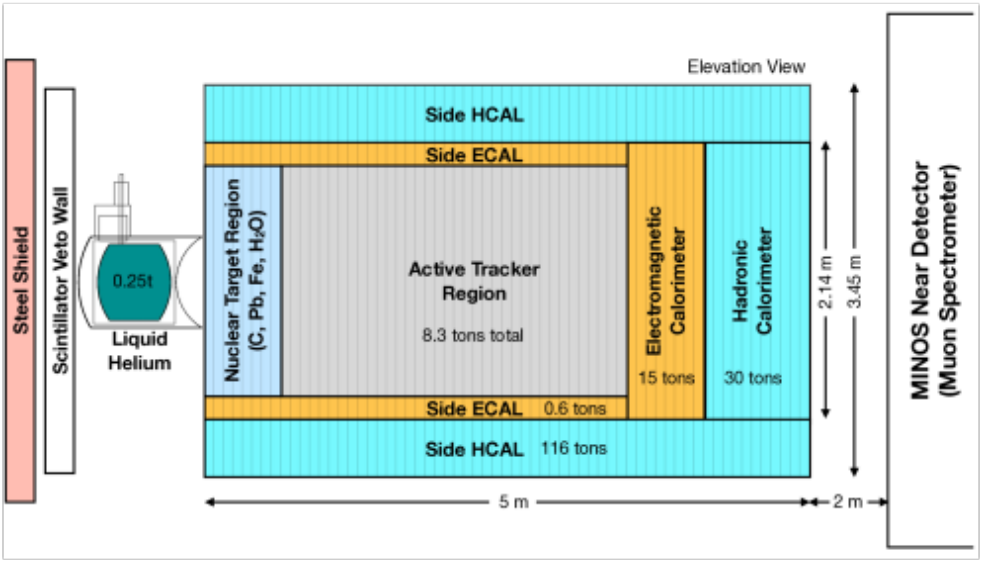


Figure 3.1: Schematic of MINERvA. Figure from Ref. [40]

The sensitive elements of the MINERvA detector are strips of plastic scintillator that emit light when traversed by charged particles. The central tracking region (ID), also known as the tracker, is mainly composed of the plastic scintillating material allowing for

precise measurement of the path and energy deposition of particles. Outside of tracker is the ECAL followed by the outer detector (OD), HCAL. The ECAL contains layers of alternating lead and scintillator, while the HCAL contains layers of iron and scintillator. Their purpose is to contain and report the energy of particles escaping the tracker, primarily those interacting electromagnetically. There is a supporting steel frame for the detector, providing the support structure and alignment for each hexagonal module.

3.1.1 Active plane composition

Each plane is composed of an active, translucent, polystyrene doped with scintillator which absorbs the energy of an interacting particle and emits energy in the form of light. Each plane is composed of 127 strips with triangular cross sections of width 3.3 cm and height 1.7 cm, see Figure 3.2. The strips were extruded with a 0.25 mm coating composed of 85% polystyrene and 15% TiO_2 (by weight) for internal reflection. There is a 3.5 mm hole at the center of each strip where a 1.2 mm diameter wavelength shifting (WLS) fiber is inserted. The WLS fiber spans the length of each strip and is optically coupled to the scintillator through optical epoxy. The scintillator surrounding the WLS fiber is doped with PPO and POPOP fluors and emits blue light when excited by energy depositions. The WLS fiber is doped with Y11 fluor and absorbs blue light emitted by the scintillator and, in turn, emits green light. The green light propagates towards the end of the strip, finding either the electronic readout or a mirrored end coated with aluminum that will reflect the light at 85% efficiency back to the readout end [40]. Using translucent epoxy, 127 strips are glued together to form the hexagonal scintillator planes. A Lexan sheet is adhered to each plane, followed by black PVC electrical tape to shield the planes from ambient light. The

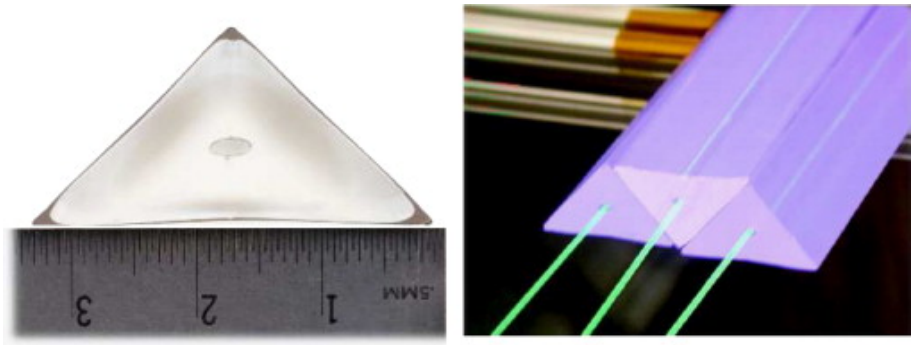


Figure 3.2: Cross section of the MINERvA scintillator strips. Figure from Ref. [40].

chemical composition of the scintillator planes can be found in Table 3.1.

A particle traversing a plane will typically deposit energy in two or more strips because of the overlapping edges from the triangular cross section. With information from the orientations of the planes and the depositions of energy in the strips, MINERvA is able to reconstruct 3D information on the energy and trajectory of a particle in the detector.

Element	C	H	O	Ti	Al	Si	Cl
Mass fraction	87.62%	7.42%	3.18%	0.69%	0.26%	0.27%	0.55%

Table 3.1: Breakdown of the chemical composition of the MINERvA scintillator planes

3.1.2 Coordinate System

Each plane contains a series of stacked triangular scintillating strips, yielding two-dimensional information. To fully reconstruct a particle's trajectory in three dimensions, the planes must be placed in different orientations with respect to one another. In MINERvA, the planes are oriented (termed "views") in an alternating sequence. There are three plane views as defined by gravity: "X" is when the strips are pointing vertically and "U" and "V" are ± 60 deg with respect to X, as seen in Figure 3.3. A module is defined by a group of two

planes with varying views in a repeating pattern. The first module in a sequence is an X plane followed by a U plane. The following module is an X plane followed by a V plane and the pattern is repeated XUXV from upstream to downstream.

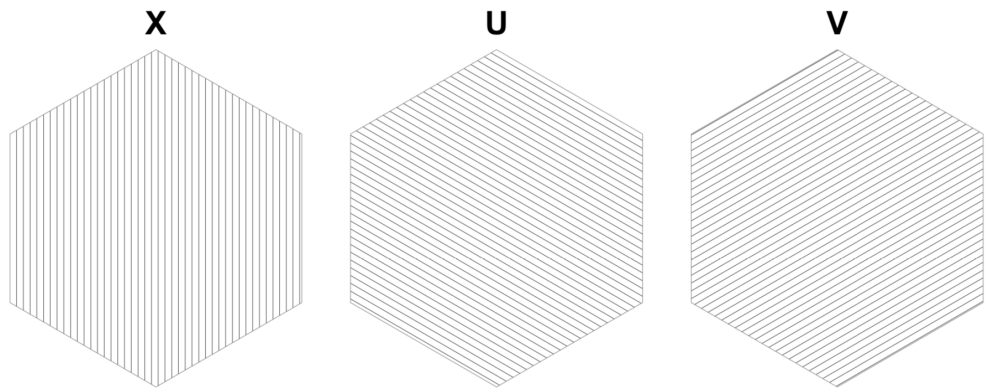


Figure 3.3: Orientation of MINERvA Planes, as defined by the strip orientation. Figure from Ref. [40].

The plane views inform the coordinate system chosen for the detector. As mentioned in Section 2.2, the NuMI beam points slightly downward and forms an angle of about 3° with respect to the $\hat{z}$ direction. The $\hat{z}$ direction is defined as an axis through the center of the scintillating planes, with the $+\hat{z}$ direction from upstream to downstream of the detector. The $+\hat{y}$-direction is the direction “up" or antiparallel to gravity. The $+\hat{x}$-direction points to the left when looking at the detector from upstream to downstream and completes a right-handed coordinate system. The coordinate system of the detector is seen in Figure 3.4. The length of z is 12000 mm from the front of the target region to the front of MINOS and (x,y) = (0,0) corresponds to the center of the plane. The unit vectors corresponding to the strip directions in the U and V planes become $\frac{1}{2}(\hat{x} \pm \sqrt{3}\hat{y})$, respectively.

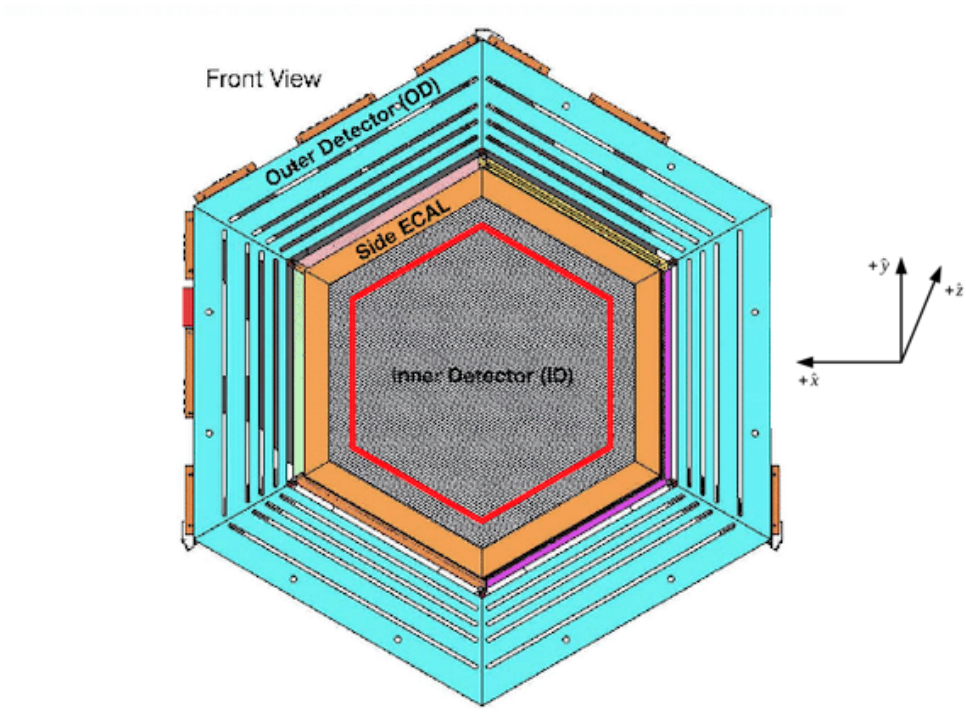


Figure 3.4: MINERvA tracker modules. The X and Y axes of the detector coordinate system are shown where the XY origin is located at the center. Looking downstream, the Z axis points in to the page. Figure from Ref. [40].

3.1.3 Module Assemblies and Nuclear Targets

Each region in the detector (nuclear targets, HCAL, ECAL and tracker) is composed of various modules. At the most upstream end of the detector, the nuclear targets region has nuclear target modules that are single planes of graphite, lead, and/or steel that are edge-welded together in various shapes to form a hexagonal module. The nuclear target regions alternate each of the five nuclear target modules with four hexagonal scintillating modules for a total of 28 modules. This is followed by the tracker region composed of 62 scintillating modules, making up the majority of the detector volume. The remaining modules are located in the ECAL and HCAL with 10 and 20 modules, respectively.

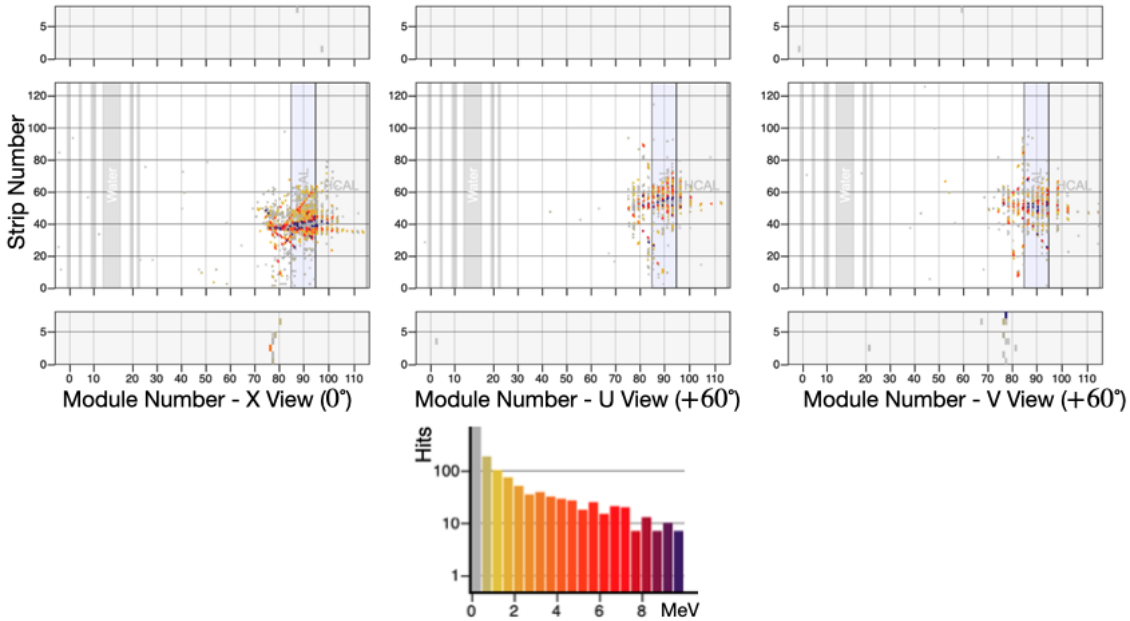


Figure 3.5: Charged current deep inelastic scattering MC event display in the MINERvA detector. Color is the amount of energy deposited in a scintillator strip.

3.1.4 Event Display

MINERvA visualizes the neutrino interaction in event displays like the one in Figure 3.5. The event shows a neutrino entering from the left of the display (+z direction is from left to right) and interacting along the length of the detector producing a proton and a muon. The display portrays a 2D view of the neutrino interaction. The left-most pane is the XZ view, center plane UZ view and right most pane VZ view (for X, U, V being the plane orientations). The color denotes the amount of energy deposited by a particle in a scintillator strip. The horizontal axes are the module numbers and the vertical axes are the strip numbers. Each pane has three subdivisions with the largest center division the view of the planes in the tracker region. As seen in the event displayed, not all energy is contained

in the tracker region and some particles will interact and exit the detector. The smaller top and bottom divisions in the figure represent the HCAL region, with energy deposits from a particle exiting the detector. Event displays are used to help inform analyzers as to why certain events passed or failed some threshold, as well as provide a visual display of data and MC.

3.2 The MINOS Near Detector

Situated about 2.1 m downstream of MINERvA is MINOS, a magnetized iron sampling calorimeter used to inform MINERvA of the charge and momentum of back exiting muon events. MINOS is composed of 282 steel planes, 2.54 cm thick, interleaved with planes of plastic scintillator, 1 cm thick and 4.1 cm wide. The shape and dimensions of the steel can be found in Figure 3.6 and the detector overview in Figure 3.7. Similar to MINERvA, 3D position information is provided from the scintillator planes oriented at alternating +45° ("U") and -45° ("V") with respect to the vertical.

The detector is broken down into two parts: an upstream calorimeter region and a downstream muon spectrometer region. The calorimeter region is composed of 120 steel planes alternating with a single scintillator plane. The muon spectrometer region is composed of 162 steel planes with a scintillator plane for every fifth steel plane. A coil carrying a current of 40 kA, generating a magnetic field of 1.3 T, spans the entire length of MINOS. The direction of the current can be configured to focus or defocus particles with negative or positive charge. Using the scintillator and the magnetic field, the charge and momentum of a muon leaving MINERvA and entering MINOS is measured from its

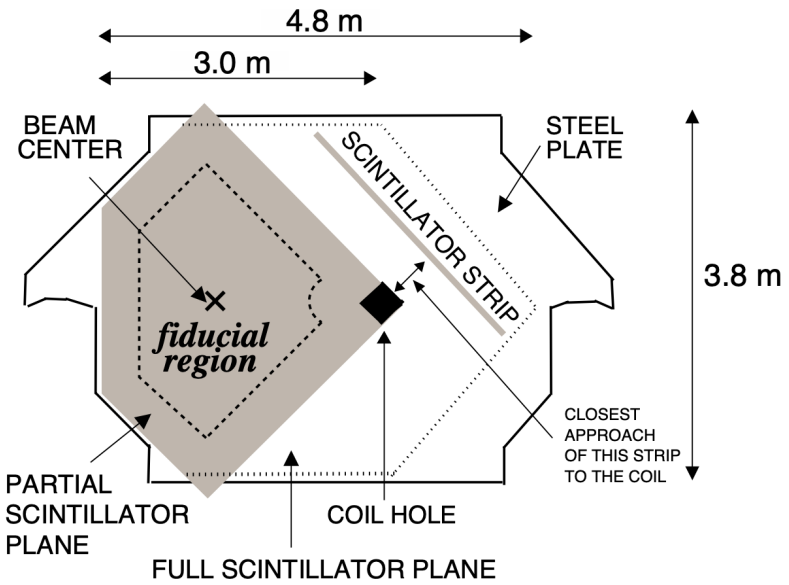


Figure 3.6: Scintillator and steel places on the MINOS detector. Figure from Ref. [41]

curvature in the magnetic field and the distance traveled in the detector.

3.3 Readout Electronics

The light produced by particles in the MINERvA scintillator is converted to a digitized electrical signal. The scintillation light reaches a photomultiplier tube which converts the light into an electrical charge. Front end boards (FEBs) digitize the charge and a data acquisition (DAQ) computer reads out the FEBs with electronics. Each subsystem is explained below, with more details in Ref. [40].

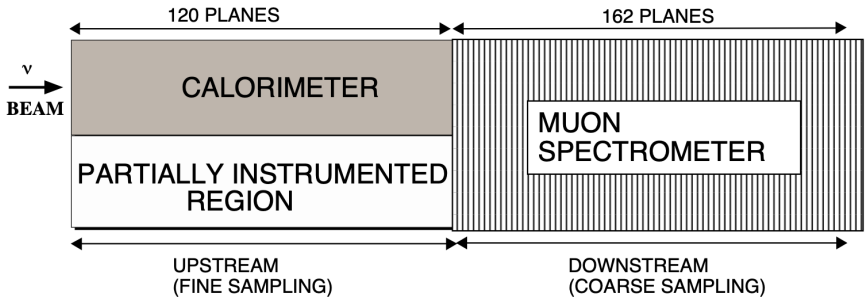


Figure 3.7: Schematic of an overhead view of the MINOS upstream calorimeter and downstream muon spectrometer. Figure from Ref. [41]

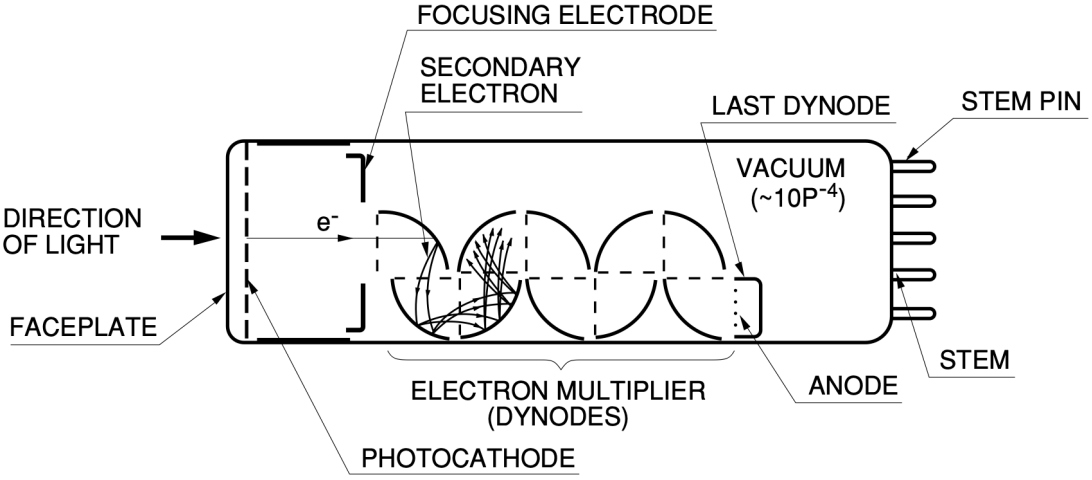


Figure 3.8: Schematic construction of a photomultiplier tube. Figure from Ref. [42].

Photomultiplier tubes

The photomultiplier tube (PMT) is a vacuum phototube containing a photocathode, dynodes and an anode. When the light from the optical fibers reaches the PMT, it hits a photocathode and creates a cascade of electrons through the photoelectric effect, as seen in Figure 3.8. The photoelectric effect occurs when photons are incident upon a photocathode material, thereby ejecting photoelectrons (PE)s. The voltage difference across the PMT is divided between sequential dynodes to maintain a large potential difference between each. As the dynodes are struck by the electron(s), a cascade of electrons occurs resulting in the number of electrons at the end of the PMT to be a factor of 10^5 to 10^6 larger. The PMT is required to be sensitive enough to allow for a timing resolution of better than about 5 ns to distinguish between overlapping events in the detector within a single NuMI spill. For each strip, the readout end of the WLS fiber is coupled into a light-tight capsule containing a Hamamatsu 64-channel multi-anode PMT. The 64 channels are broken down into an 8×8 array on a 2 cm $\times$ 2 cm grid. The 64 channels are interwoven into an arrangement mapping the light from neighboring strips in a single plane to non-neighboring PMT pixels to reduce the amount of ”cross-talk“ between neighboring strips, which occurs when light incident on one PMT channel induces a current in a neighboring channel. The last stage is when the electrons reach the anode and create a current spike easily seen on an electronic readout.

PMT efficiency degrades when exposed to ambient magnetic fields exceeding 5 Gauss; therefore, each PMT is housed in a PMT box with cylindrical steel housing, seen in Figure 3.9. An additional effect of PMTs is afterpulsing, in which latent gas in the PMT ionizes. These ions are positively charged and accelerate the wrong way through the dynodes, eventually colliding with a dynode or photocathode and producing a delayed output signal.

The time delay in this afterpulsing can range from hundreds of nanoseconds to several microseconds.

Additional contamination events include radioactive decays from materials or the passage of cosmic rays through the PMT. Simulated neutrino events are overlayed (referred to as *overlay*) with hits that fall within a dead time period to account for these occurrences.

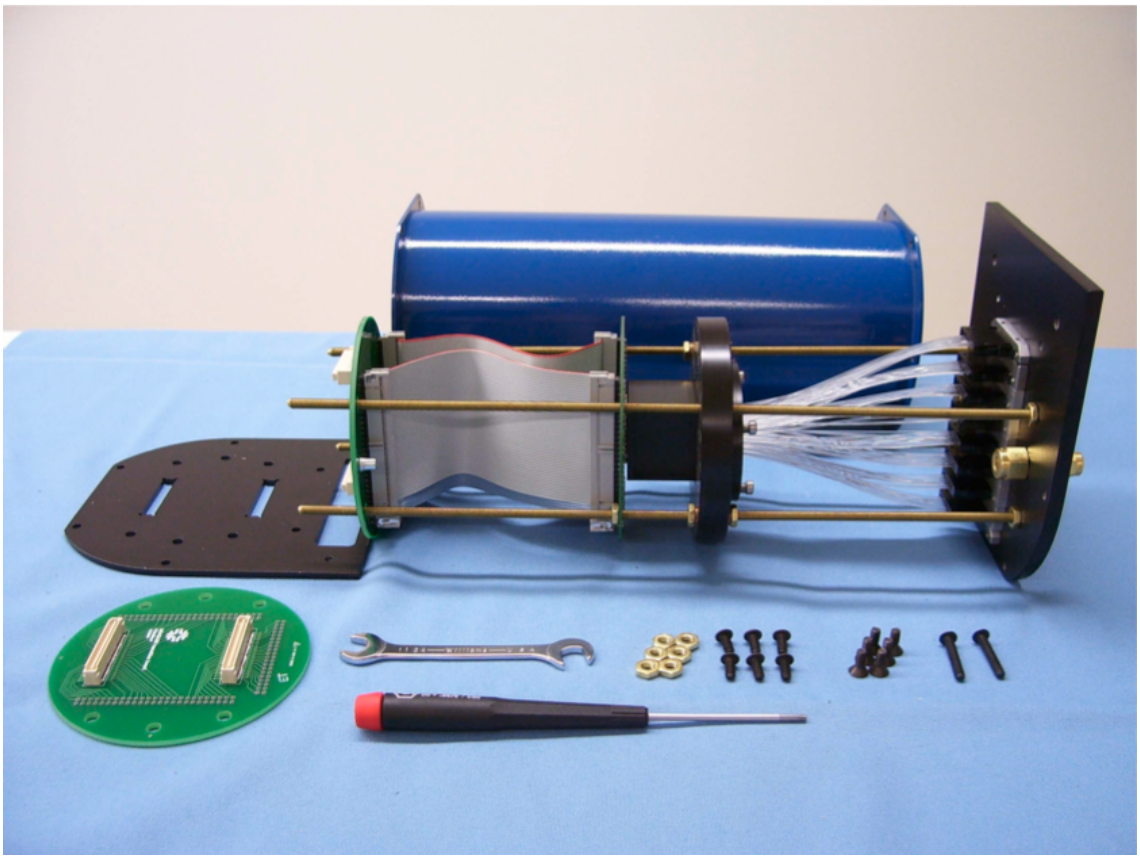


Figure 3.9: Optical PMT box prior to assembly. From left: the end cap, ribbon cable, PMT base board, the PMT (black rectangular box), woven fibers and the PMT circuit board. The blue cylinder is the steel housing. Figure from Ref. [40]

3.3.1 Front-end Electronics

Mounted to the output endplate of the PMT box is a FEB containing 64 channels, one for each anode of the PMT. Each FEB contains a Trigger and Pipeline with timing (TriP-t) Application-Specific Integrated Circuit (ASIC) 32-channel chip that integrates and digitizes charge from each channel into an analog to digital count (ADC). From each FEB channel, the charge from the PMT is divided into TriP-t channels allowing for low, medium and high gain response to the same charge input, increasing the dynamic range of the electronic circuit. The charge collection occurs during a 16 μs window called a "gate" designed to capture the 10 μs NuMI beam spill plus any slower decaying particles. The electronic response is nonlinear; therefore, the gain channels are fitted to a tri-linear function that contains three distinct linear segments [40]. The slopes and starting positions of the linear segments are extracted from each channel and stored in a database allowing for the conversion of the raw ADC count to linearized charge.

If a charge in the channel rises above the saturation level of $\approx$ 2500 ADC counts or one PE at the cathode, the associated gain responses are stamped with a time. The FEBs then integrate charge using a charge integration cycle. If a discriminator fires on one of the 32 channels, one low gain TriP-t chip and two high gain TriP-t chips are charge integrated for 150 ns, immediately followed by a 200 ns reset period in which the integrated charge is stored and the discriminator capacitor is cleared. During this time no incoming charge is recorded, referred to as "dead time." Therefore, if any single channel discriminator fires, then all the channels in the TriP-t group go through the same digitization and reset period, losing any incoming activity in any of those channels. These cases are accounted for by "overlaying" simulation that removes charge when the readout would have occurred during

the deadtime.

3.4 Calibration and Pedestal Subtraction

Once the activity in the detector is recorded in the form of digitized charge called ADC counts, it undergoes a series of calibrations. The digitized charge is not immediately in a form available to analyzers and must be sent to a readout channel and converted to the energy deposited in a strip, as described below and in Ref. [40].

The calibration of ADC counts derives from the independent optical readout channel, c, and the time dependence of the calibration, t. First, the ADC count is converted to PE at the photocathode

$$PE = (ADC - P(c,t)) \times Q(c,ADC) \times \frac{1}{G(c,t)} \tag{3.1}$$

for $P(c,t)$ the pedestal, $Q(c,ADC)$ the ADC to PMT anode charge conversion factor for the FEB channel, and $G(c,t)$ the PMT channel gain. A pedestal is a measure occurring when the ADC counts report a nonzero value even though no current entered the PMT. Removing the pedestal requires measurements to establish where this zero charge point is by measuring events where no detector signal is expected. Non-neutrino beam activity, such as cosmic rays or radioactivity in the detector material may occur during a pedestal measurement. A method called Peirce's Criterion [43] is applied to remove these outlier gates from the distribution. Figure 3.10 shows a pedestal distribution for an individual channel. All outliers from the pedestal distribution are removed, the mean value and standard deviation are calculated, and the values are saved in a database. The tables are

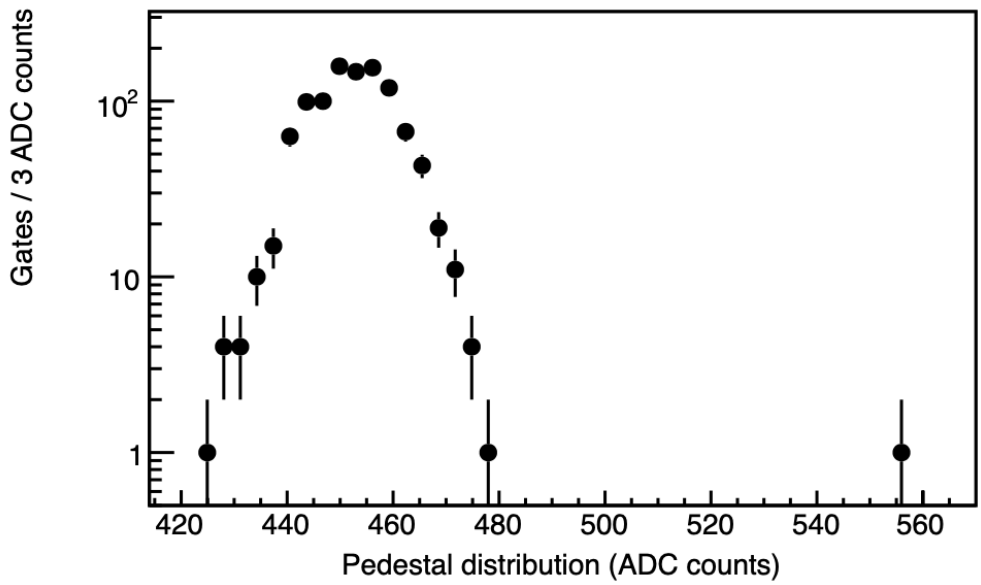


Figure 3.10: A pedestal distribution for an individual channel readout from one 16 μs gate. Figure from Ref. [40]

accessed offline during the calibration stage and hits for a range of ADC values below the pedestal mean or outside of 3 standard deviations of the mean are removed. The mean pedestal rate for each channel is subtracted from the ADC counts in the raw digits beam data. The PMT's energy response drifts over the timescale of weeks and a gain calibration is required to correct for the drift. The gain for each channel in the detector is measured by using a pair of optical fibers to periodically inject low levels of light into the PMT between beam spills. Since the intensity of the injected light is known, the PMT response is measured. The time variation is calibrated by fitting a curve to the light injection distribution and tracking the location of the 1 PE peak [31]. The parameters from the fit are used to find the gain, as seen in Figure 3.11.

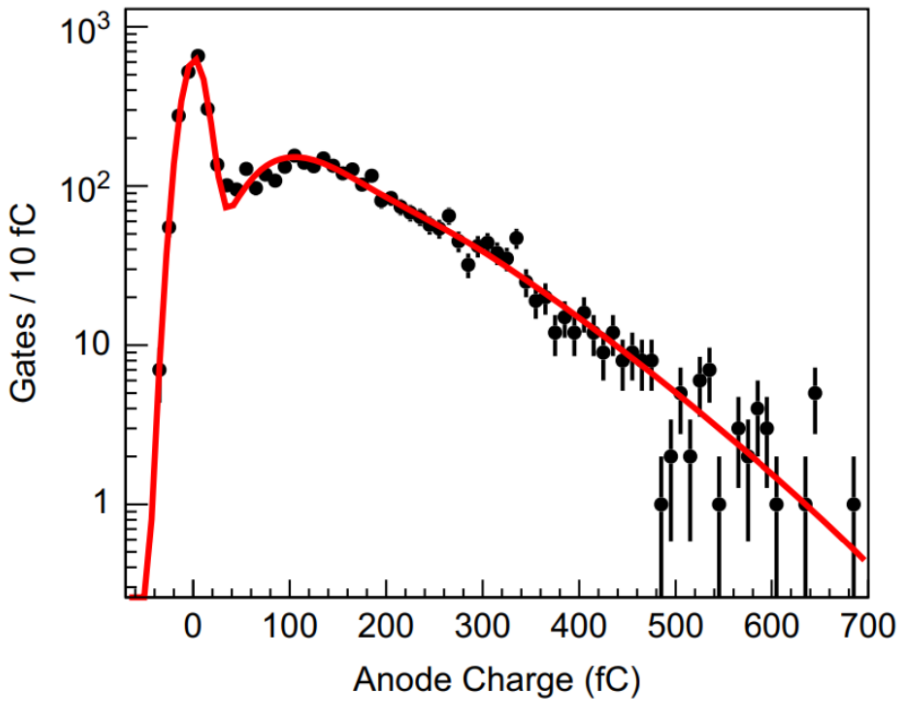


Figure 3.11: Number of hits as a function of charged measured. The first peak on the left is the pedestal (zero). The second peak represents 1 PE. Figure from Ref [40].

Following Equation 3.1, the PEs are then converted to energy depositions by strip

$$E = PE \times A(c)\,S(c,t) \times M(t) \tag{3.2}$$

for $S(c,t)$ the strip response correction, $A(c)$ the correction for the attenuation of light from the WLS and optical fibers, and $M(t)$ the detector energy scale. Data from the FEBs are compiled in memory and written to a binary file on a disk. The raw data contains the timing and pulse height information from the detector in binary format and must undergo many stages of processing and calibration to account for various effects such as light attenuation

along the strip length of the detector and effects from the electronics in the detector.

3.5 Sample Preparation

3.5.1 General Track and Calorimetry Reconstruction

After calibration is completed, the MC simulation and data provide energy information on the neutrino scattering interaction in the MINERvA detector. The next step of processing is to take the position and energy of the energy depositions and reconstruct it to identify particle trajectories. At the time of calibration, the simulation and data are grouped by gate. During that 16 μs window of gate time, $\geq$1 real neutrino interactions are expected in the detector and exactly one neutrino interaction is simulated per gate. Each of the 16 μs gates is divided into "time slices" using an algorithm that clusters peaks of activity within the gate. This serves two purposes: the first is to try and isolate single neutrino events in the case two neutrinos (in data) enter the gate window, leading to pileup (neutrino interactions occurring closely in time) and mis-reconstruction of tracks. The second is for clusters of activity to become their own compiled objects of information sent downstream for use by the analyzer. Further information on the optimization of the time slicer algorithm can be found in Ref. [44]. Time slices are fed to common reconstruction tools where additional algorithms determine what slices should reconstruct to which respective particle tracks along with all kinematic information of the particle and interaction. After consolidating hits to digits and digits to clusters, a calorimetric correction is applied according to the position of the energy deposit within each sub-detector to account for any energy loss that occurred.

Once a gate's data is divided into time slices corresponding to individual particle

interactions, energy clusters are formed using adjacent strips of activity. First, the energy found in a single optical channel readout is recorded as a hit. Hits are grouped together to form digits and digits are formed together to give us a complete picture of how the neutrino interaction occurred. Digits are grouped first in time, then in position and finally into reconstructed objects called clusters.

A set of clusters in the shape of a slightly bending line may form particle trajectory tracks, also known simply as tracks. *Tracks* are three-dimensional objects representing a particles' trajectory through the MINERvA detector. There is standard reconstruction information for muon tracking and energy of the muon systems; however, electron/positron tracking requires additional tools at the level of individual analyses [40] [45] .

3.5.2 Calorimetric and Spline Correction

The hadronic recoil energy is the total hadronic energy deposited inside the detector minus the energy of the lepton (electron). This includes energy deposited into the active regions of the detector, like the plastic scintillator, where the energy can be measured directly. It also includes hadronic recoil energy that goes unaccounted for in the passive regions of the detector, like the lead or epoxy. Therefore, the hadronic reconstruction uses calorimetry measurements to provide precise corrections to account for any missing hadronic energy. Each non-electron energy deposition is corrected based on its energy and location within the detector. The lower fraction of active material, the larger the correction must be to account for unobserved energy. This calorimetric correction is calculated as a constant, C^{sd}, as shown in Equation 3.3 where E_{abs} and E_{scint} the energy lost in a single plane of the absorber and scintillator and f the active fraction of scintillating planes. The corresponding

calorimetric constants for each subdetector are found in Table 3.2.

$$C^{sd} = \frac{E_{abs} + E_{scint}}{f \times E_{scint}} \tag{3.3}$$

Subdetector	MC Constant	Data Constant
Tracker	1.222	1.197
ECAL	2.013	1.998
Side ECAL X	2.805	2.609
Side ECAL UV	4.388	4.339
HCAL	10.314	10.441
OD 123	21.679	24.232
OD 4	42.137	42.660

Table 3.2: Passive material corrections by sub detector

An additional overall scale factor, α, is calculated to account for the energy loss due to leakage outside the detector, neutral particles and binding energy of struck nucleons within the nucleus. This correction needs to be energy dependent since higher energy processes can produce more neutral particles in the final state. The scale factor is calculated by comparing the reconstructed and true energy depositions on a hit-by-hit basis. The energy dependence of α is introduced through the construction of a "polyline", referred to as a spline, essentially a mapping of uncorrected to corrected recoil energy. The spline is constructed through a per-bin energy correction calculated from the fractional energy resolution of the calorimetry energy in bins of true hadronic recoil energy, shown as:

$$\frac{\Delta E}{E_{recoil}} = \frac{(E_{reco} - E_{true})}{E_{true}} \tag{3.4}$$

The distribution of the energy resolution is then fitted to a Gaussian distribution. The scale factor α is then applied to the reconstructed hadronic recoil energy, E_{reco}, to center the

energy resolution at zero. This process is repeated for multiple energy ranges of reconstructed hadronic recoil which extracts scale values for different energy values. Therefore, the calorimetrically corrected hadronic recoil energy is defined as:

$$E_{recoil,reconstructed} \equiv \alpha \times \sum_i C_i^{sd} E_i. \tag{3.5}$$

This weighted sum is executed over the five different sub-detectors, i = tracker, ECAL, SideECAL, HCAL, OD and E_i is the summed energy for the hits in the respective sub detector. The passive material correction, C^{sd}, is mentioned previously in Table 3.2. There are different constants for hits in the nuclear target region in the upstream part of the detector, but this will not be covered because this analysis does not occur in the nuclear target regions.

3.5.3 Kinematic Variable Reconstruction

Many observables are calculated in the general reconstruction phase of the analysis. Among them is the neutrino energy calculated as

$$E_\nu = E_{lepton} + q_0 \tag{3.6}$$

where $q_0 \equiv E_{recoil} =$ calorimetric hadronic energy. The q_0 seen in equation 3.6 is a spline corrected quantity. In this case, the spline correction applied to q_0 corrects for neutrons whose energy typically goes unaccounted for in the reconstruction phase. The splines rely on model predictions to give a total energy estimation for an event; therefore, a quantity like q_0 is deemed *model dependent*. Ideally, kinematics would be derived solely based

on mathematical equations. However, this is not possible due to the reliance on model predictions to describe anything from neutrino flux to vertex reconstruction. This analysis instead chooses to report in a variable called the *hadronic available energy*, E_{avail}, To limit model dependency as described below.

The calorimetric hadronic energy in MINERvA, q_0, is a similar but separate quantity from the visible energy, E_{vis}, and available energy, E_{avail}. E_{vis} is the conversion of photo-electrons to energy using basic calorimetry and conversions like attenuation and pedestals and is not as reliant on model predictions. E_{avail} is the extra energy in the system with only a passive material correction. Since E_{avail} does not receive the spline-derived correction like q_0, it is less model dependent and provides a more direct probe into the RHC $\bar{\nu}_e$ low-recoil system. The hadronic available energy is defined by

$$E_{avail} = \sum T_p + \sum T_{\pi^\pm} + \sum E_{particle} \tag{3.7}$$

for $\sum T_p$ the sum of proton kinetic energy, $\sum T_{\pi^\pm}$ the sum of pion kinetic energy and $\sum E_{particles}$ is the sum of the total energies of everything else, barring neutrons. The last term, $\sum E_{particles}$, is defined by

$$\sum E_{particles} = \sum E_{e^\pm} + \sum E_{\pi^0} + \sum E_\gamma + \sum E_{other} \tag{3.8}$$

where E_{other} is primarily kaons and eta mesons.

The three-momentum transfer of the system is calculated as

$$q_3 \equiv |q| = \sqrt{Q^2 + q_0^2} \tag{3.9}$$

for Q^2 the reconstructed squared four-momentum transferred from the lepton system defined as,

$$Q^2 = 2E_\nu(E_{lepton} - p_{lepton} \times \cos\theta_{lepton}) - (M_{lepton})^2. \tag{3.10}$$

The model dependent variable q_0 is used in part of the calculation to derive the value for q_3, as seen Equation 3.6 and 3.9. The "low-recoil" part of this analysis refers to low-q_3 values to isolate regions in which data are historically discrepant with simulation. The additional kinematic variable important to this analysis is the lepton transverse momentum, p_T. p_T is measured by projecting the electron three-momentum into the plane perpendicular to the incoming neutrino direction, which is angled downwards at 3.3 degrees.

$$p_T = |\vec{p}_T^{lep}| \sin(\theta_{lep}) \tag{3.11}$$

$$\theta_{lep} = \arccos(\frac{\vec{p}^{lep} \cdot \hat{n}_\nu}{|\vec{p}_{lep}|}) \tag{3.12}$$

where $\hat{n}_\nu = \cos(\theta_\nu)\hat{z} - \sin(\theta_\nu)\hat{y}$ is the beam direction in detector coordinate with $\theta_\nu = 3.3$ degrees.

A cross section result in E_{avail} vs p_T is interesting because the reconstruction of p_T is less model dependent than q_3. By definition, the kinematics of the leptonic system can be directly probed without the introduction of any quantities which depend on assumptions or spline corrections, like the dependency of q_3 on q_0. Both q_3 and p_T provide different insights into the low-recoil system and each is reported in this analysis.

3.5.4 Available Energy Correction

No quantity measured in the detector can be reconstructed with exact precision. For example, MINERvA is limited by the energy resolution in the detector, the application of calibration to energy deposits etc. which can lead to some quantities being measured higher or lower than their true value. Therefore a reconstructed quantity can "migrate" in and out of bins leading to an effect referred to as "smearing". Smearing can be characterized by generating a *migration matrix*, in which the relation between true and reconstructed values is determined by a square matrix broken down into equivalent bin sizes for true and reconstructed values. Smearing is discussed in greater detail in Section 6.2. Of particular interest is the migration of events between reconstructed E_{avail} and true E_{avail} since the cross section will be reported in this variable. As seen in Figure 3.12, there is a population of off-diagonal events for higher values of E_{avail} meaning the reconstructed E_{avail} is consistently under-predicting the true value of the E_{avail}. Some kinematics values like q_0 undergo a spline correction to account for some energy loss within the detector. Neutrons are very difficult to identify in the detector and the RHC sample has a non-negligible population. In regards to E_{avail}, by definition, the energy of neutrons is not included nor does it undergo a spline correction to account for undetected neutron activity.

Figure 3.13 shows the effect neutrons have on the calculation of E_{avail}. The distribution of true available energy in bins of reconstructed available energy reveals a large spike of events at a value of true $E_{avail} = 0$. The large spike can be attributed to neutrons because the definition of true E_{avail}, as defined in Equation 3.7, does not include neutron energy. However, the reconstruction algorithm can fail to identify when a neutron passes through the detector and the energy deposited by a neutron may be mis-reconstructed as energy

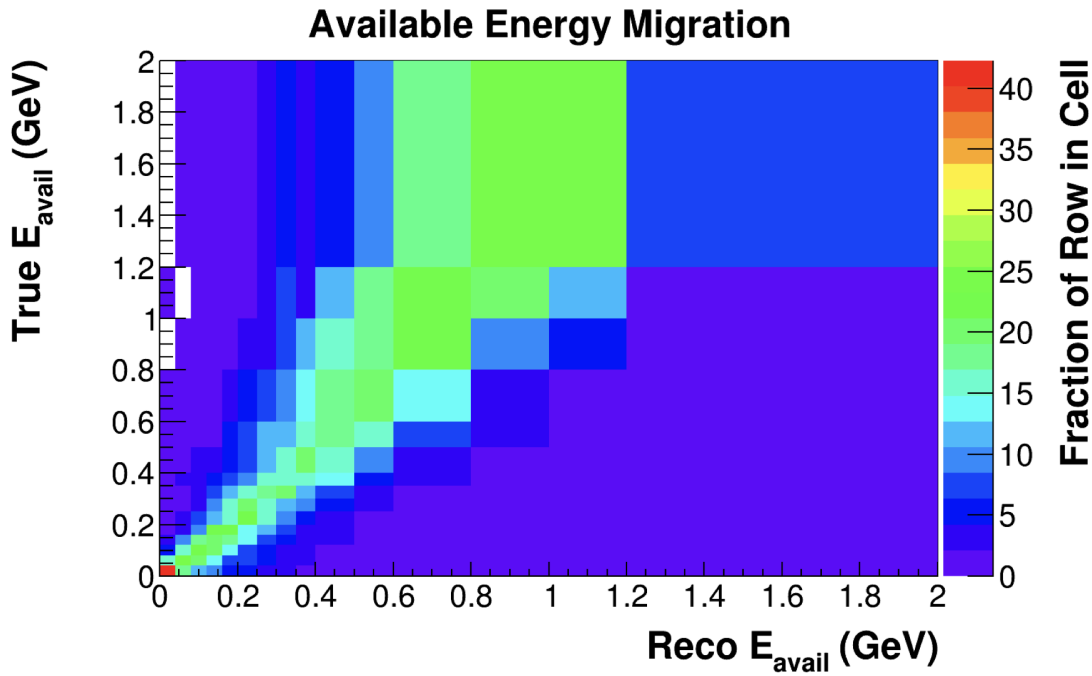


Figure 3.12: E_{avail} migration before additive correction.

for another FSI particle in the event. An additional example is when a π^- is absorbed on a carbon nucleus and produces only neutrons, which would be counted towards the true available energy but not the reconstructed. A correction is formed to get the best estimate of reconstructed E_{avail}. The first step in attempting to correct for the reconstructed E_{avail} is to determine the difference between the true E_{avail} and the reconstructed E_{avail} in bins of reconstructed E_{avail} to identify what type of additive correction can be made.

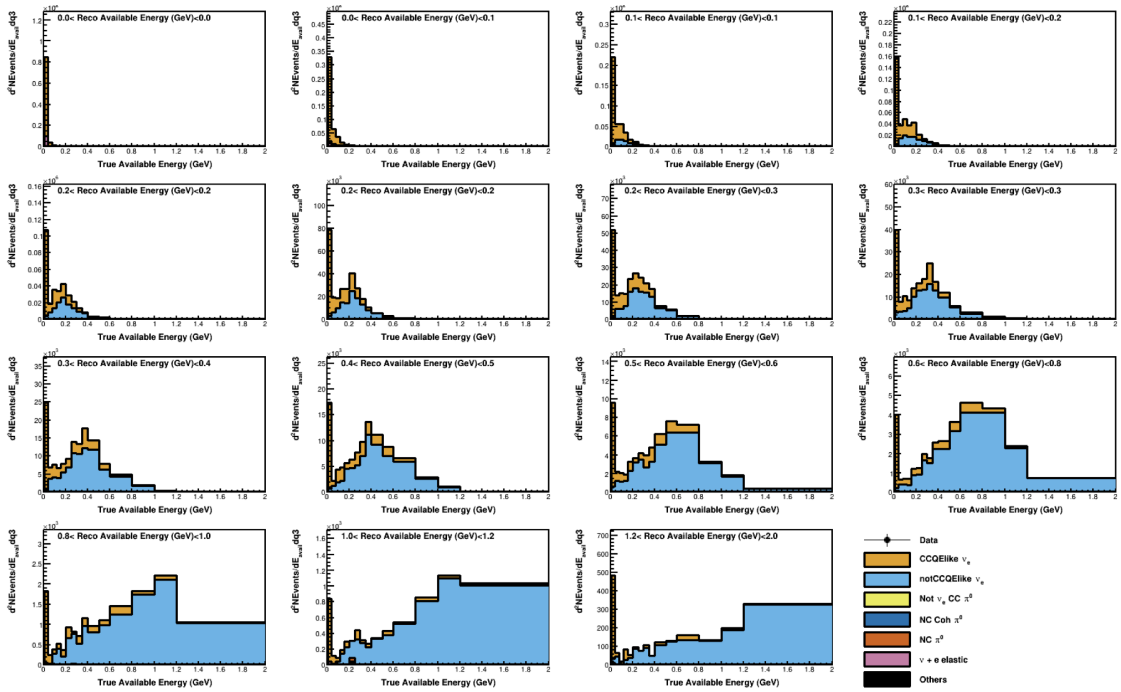


Figure 3.13: True E_{avail} in bins of Reco E_{avail}. The spike population at zero is a neutron population.

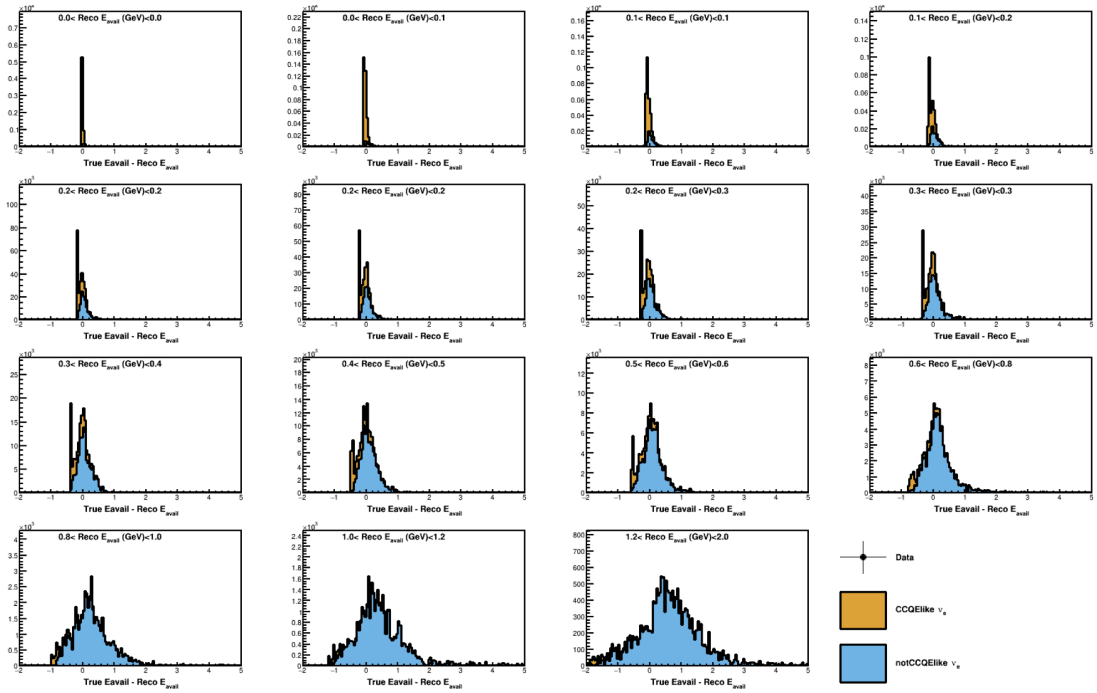


Figure 3.14: E_{avail} Additive correction calculated from the difference of the true available energy and reconstructed available energy. The correction is formed based on the mean of each distribution in bin of reconstructed energy.

Reco E_{avail} bin (GeV)	Additive Correction (GeV)
$[0.0, 0.04)$	0.001
$[0.04, 0.08)$	-0.003
$[0.08, 0.12)$	-0.014
$[0.12, 0.16)$	-0.021
$[0.16, 0.24)$	-0.024
$[0.24, 0.32)$	-0.026
$[0.32, 0.4)$	-0.017
$[0.4, 0.5)$	0.006
$[0.5, 0.6$	0.048
$[0.6, 0.7)$	0.101
$[0.7, 0.8)$	0.141
$[0.8, 0.9)$	0.191
$[0.9, 1.0)$	0.244
$[1.0, 1.1)$	0.30
$[1.1, 2.0)$	0.357

Table 3.3: Additive weights for the reconstructed $E_{available}$ in bins of reconstructed $E_{available}$

The distribution shown in Figure 3.14 is the difference between true and reconstructed E_{avail}. The correction is an average of the distributions in each bin. The corrections found in Table 3.3 are relatively small and near zero with a few bins around low reconstructed E_{avail} being a slightly negative value. Before applying the correction, a smooth function is fitted between all points of the additive correction, found in Figure 3.15. The correction is added to the reconstructed E_{avail} at the event level based on the function. The migration matrix with the correction applied is found in Figure 3.16, with a much more diagonal shape.

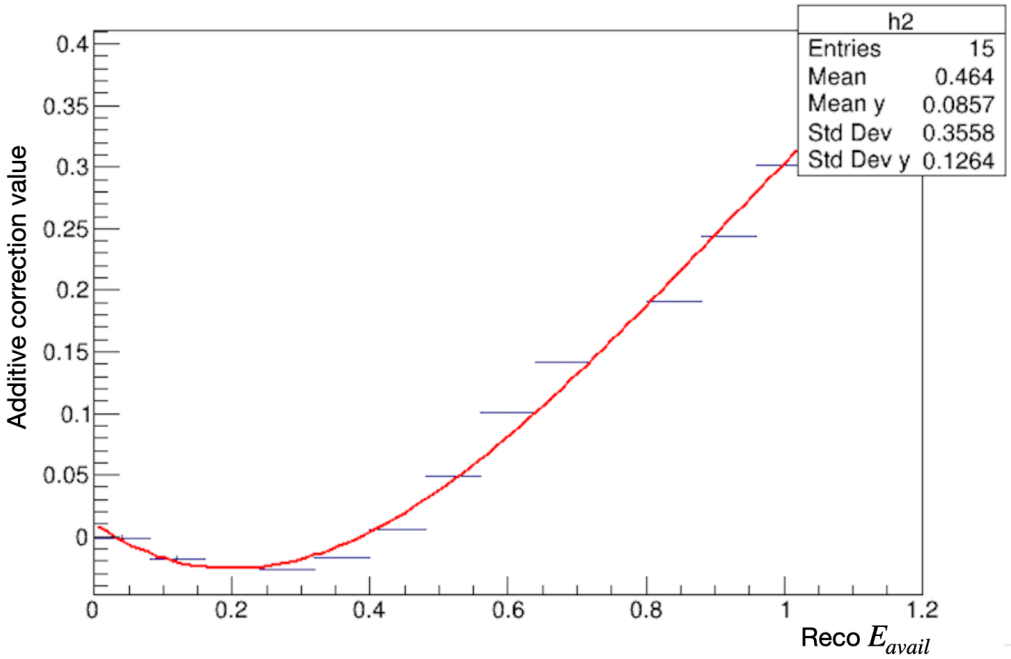


Figure 3.15: Smooth function fit between additive reconstructed E_{avail} values.

3.6 Interaction Models and Simulation

MINERvA relies on multiple simulation packages to predict the neutrino-nucleus interactions occurring inside of the detector. GENIE is a Monte Carlo (MC) event generator providing comprehensive physics models for the simulation of neutrino interactions [46]. It takes an input of the simulated flux of incident neutrinos and generates neutrino interactions inside a specified detector configuration. GENIE's model simulations can be broken down into two major categories: interaction probabilities and final state interaction (FSI) probabilities. FSI is the re-interaction of final state particles inside the nuclear medium before detection. Interaction models control the probability of the initial neutrino-nucleon interactions and FSI models control the probability the particles will undergo additional

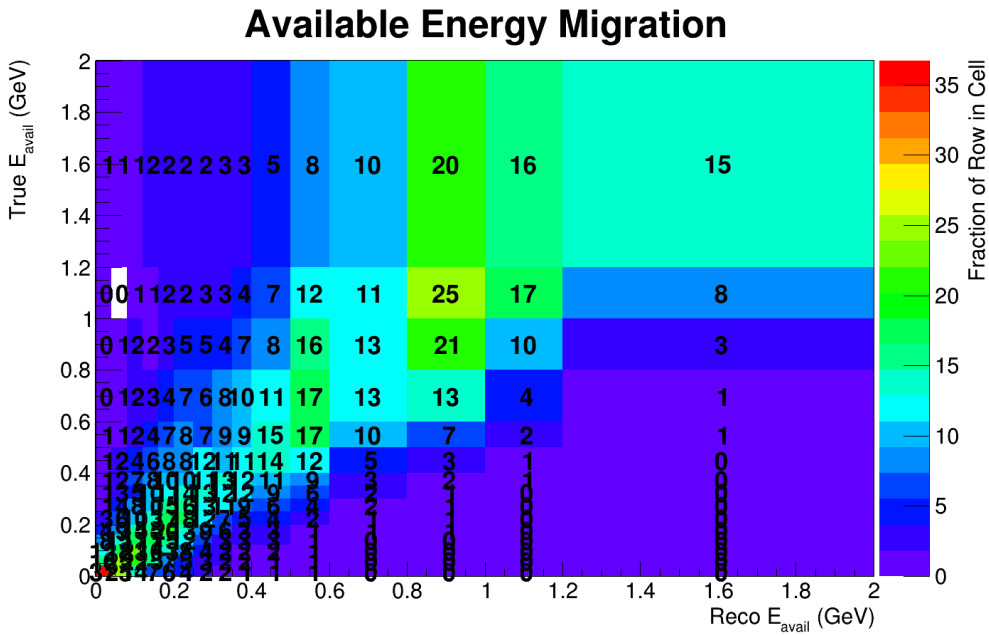


Figure 3.16: E_{avail} migration matrix with additive correction to reconstructed available energy

re-interactions. Additional interactions of hadrons may occur inside of the nucleus itself or on the way out of the nucleus, referred to as the final state. GENIE is a widely used simulation in the neutrino physics community and can be tailored for specific detectors. It handles all neutrino and target information for processes from MeV to PeV scales [46]. This analysis uses the version GENIE 2.12.6. The GENIE simulation uses neutrino kinematics and type to simulate the neutrino-nucleus interaction. The MINERvA detector geometry is used in the simulation to determine where GENIE simulates neutrino-nucleus interactions in the detector volume. GENIE simulates both the neutrino-nucleus interactions and the subsequent processes inside the nucleus. The type of interaction is randomly selected from GENIE's library of allowable interactions for the given flux and material. The probability

of interaction is proportional to the total cross section of the material through which the neutrino passes. The interaction models included are those cross section models, which are the most dominant scattering mechanisms seen by the neutrino physics community.

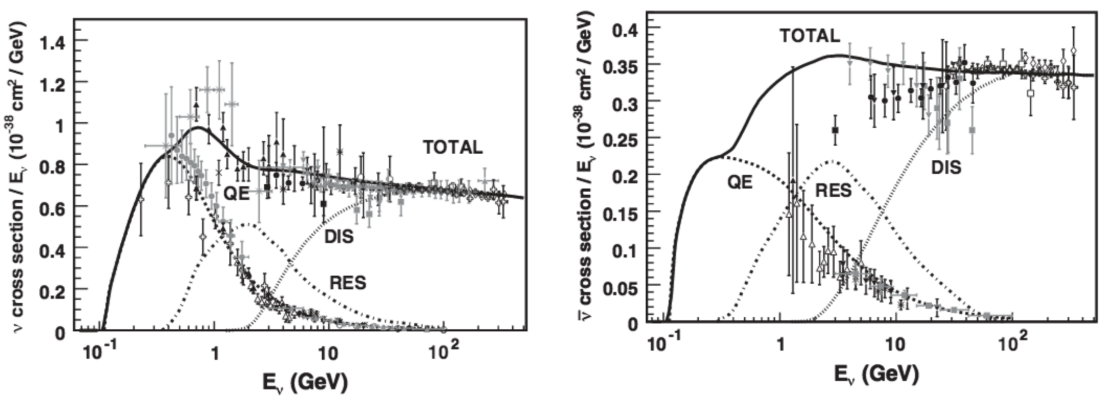


Figure 3.17: Variety of neutrino interaction models contributing to the total charged current neutrino-nucleon cross section, from [47].

3.6.1 Nuclear Model

One of the simplest models used in most neutrino interaction generators is the Relativistic Fermi Gas (RFG) model [48]. Nucleons are confined to a potential well with walls at the nuclear radius. The nucleus contains a bound state of nucleons that must obey the Pauli exclusion principle where two or more identical fermions cannot occupy the same quantum state. The Fermi momentum, $p_F \approx 250$ MeV/c, is derived from the energy of the highest filled quantum state relative to the nucleus rest frame. In the RFG model, nucleons are given an initial flat momentum pulled from the Fermi distribution, 0 - p_F, before the initial scattering occurs. This initial state effect is known as Fermi motion. The Pauli exclusion principle also leads to Pauli blocking, which requires that a nucleon must

have a momentum greater than p_F to be ejected from the nucleus. Those neutrino-nucleus interactions involving less momentum transfer than required to surpass the highest occupied energy level of the nucleus will be suppressed. Therefore, events with low Q^2 values are suppressed.

The RFG model does not account for effects involving interactions between more than one nucleon in the nucleus, also known as multi-nucleon effects. Therefore, it doesn't model those interactions which can lead nucleons to have a momentum higher than p_F. Bodek and Ritchie provide a modification to the RFG model accounting for nucleon-nucleon correlations through the addition of a high-momentum tail to the RFG model [49].

Many models attempt to encompass the distributions of energy and momentum within the nucleus. Some alternative models relax the assumptions that the nucleus has uniform density and that the nucleons have a flat momentum distribution. For example, the Local Fermi Gas (LFG) model averages the initial nucleon momentum over the nuclear interior and weights it by the proton or neutron density [50], where different densities of nucleons produce different Fermi momentum distributions. Another alternate model replaces the step function potential of the RFG model with spectral function models that include the addition of different energy momentum distributions of initial nucleons [51].

A different approach is made for correlations between nucleons through a random-phase approximation (RPA) technique used to incorporate long range nucleon-nucleon correlations. These correlations occur in neutrino-nucleus interactions when certain pairs of nucleons form stronger bound states in the nucleus, allowing one of the struck nuclei to eject the second nuclei from the nucleus. The correlations are compared to the screening effect found from an electric charge in a dielectric. However, in this case the electroweak

coupling is modified by the presence of strongly interacting nucleons inside the nucleus. The long range correlations affect single nucleon knockout, known as one-particle one-hole excitation (1p1h). The use of RPA effects is important to cross sections produced at low energy transfers and is included in the standard MC simulation. Short-range correlations occur between two or more nucleons that are in very close proximity to one another. Neutrino interactions affected by short-range correlations are characterized by a two nucleon knockout, leaving two "holes," also referred to as 2p2h [52]. This process is likely to be a mis-reconstructed of quasi-elastic events, especially when a pair of protons and neutrons is knocked out since most detectors are not efficient at detecting neutrons. Modeling of this process is actively being developed.

The previously mentioned nuclear effects are of great importance when it comes to describing different forms of neutrino-nucleus interactions. For MINERvA, there are five primary models used to characterize the neutrino interactions and the resulting final states: quasi-elastic (QE) scattering, two-particle two-hole (2p2h), resonant (RES) scattering, deep inelastic (DIS) scattering and coherent (COH) pion production. Figure 3.17 shows the leading charged current interactions as a function of neutrino energy. The figure does not include the mentioned 2p2h and coherent pion production since these reactions are not dominant at any neutrino interaction energy.

3.6.2 Quasi-elastic scattering and the 2p2h Model

Neutrino-nucleus interactions that knockout at least one nucleon can be categorized into quasi-elastic (QE) scattering or two-particle two-hole (2p2h) processes. QE scattering can

occur through neutral current (NC)

$$\nu_l + p \rightarrow \nu_l + p$$

$$\nu_l + n \rightarrow \nu_l + n$$

and charged current (CC)

$$\nu_l + n \rightarrow l^- + p$$

$$\bar{\nu}_l + p \rightarrow l^+ + n$$

channels. The charged current process is usually referred to as charged current quasi-elastic (CCQE) scattering while the neutral current is traditionally referred to as elastic scattering. GENIE models quasi-elastic cross sections using Llewellyn Smith formalism, where the differential cross section is grouped in terms of form factors representing the structure of the target nucleon [53]. The CCQE channel is important in neutrino oscillation experiments because the neutrino energy can be reconstructed from only lepton kinematics.

$$E_\nu^{QE} = \frac{m_p^2 - (m_n - E_b)^2 - m_l^2 + 2(m_n - E_b)E_l}{2(m_p - E_b - E_l + p_\mu \cos\theta_l)}$$

$$Q_{QE}^2 = 2m_n(E_\nu^{QE} - E_e)$$

where E_ν^{QE} is the reconstructed energy of the neutrino, m_p is the mass of the initial proton, m_n is the mass of the outgoing neutron, E_b is the binding energy (zero in the case of hydrogen), m_l is the mass of the lepton, and θ_l is the the angle of the muon relative to the direction of the neutrino beam. It is also possible to reconstruct the momentum transfer,

Q^2, between the neutrino and proton.

In the event of neutrino scattering off of a free neutron, the resultant final state includes a negatively charged lepton and a proton. However, in the case of a scatter off of a heavy nucleus, it has been shown that neutrinos can scatter off of a correlated nucleon pair, thereby ejecting both nucleons, referred to as 2p2h [54]. The scattering off of correlated nucleon pairs can be broken down to np, pp and nn pairs. These 2p2h events can be a large background for measurements looking at CCQE interactions on uncorrelated nucleons and overall the 2p2h events are a significant portion of the MINERvA dataset [55]. 2p2h interactions are modeled using the Valencia model [1] [56] [57]. The results of the 2p2h model implementation are discussed further in Section 3.7.1.

Models often predict the probability of neutrino-nucleon scattering but these events are inherently unobservable due to the confusion caused by nuclear effects such as 2p2h. MINERvA often uses the definition of "CCQE-like" to refer to a final state containing any number of reconstructed nucleons (protons or neutrons) and no reconstructed mesons.

3.6.3 DIS and Resonant Scattering

Neutrino interactions creating new particles in the final state include resonant (RES) scattering and deep inelastic (DIS) scattering, seen in Figure 3.18. CC RES scattering occurs when the neutrino strikes a nucleon with enough energy to excite it to an unstable resonant state which quickly decays back into the nucleon and the addition of a pion as shown in Equation 3.13.

$$\nu_l + N \rightarrow l + N^* + \pi \tag{3.13}$$

[1] Typically the 2p2h events are simulated with a three-momentum transfer less than 1.2 GeV. This analysis uses an extended 2p2h sample with a three-momentum transfer up to 2.0 GeV.

There are three different charged current resonant pion processes:

$$\bar{\nu}_\mu + p \rightarrow \mu^+ + p + \pi^-$$
$$\bar{\nu}_\mu + p \rightarrow \mu^+ + n + \pi^0$$
$$\bar{\nu}_\mu + n \rightarrow \mu^+ + n + \pi^+.$$

DIS events occur at high interaction energies where a neutrino scatters off of a quark within the nucleon, allowing the interaction to be identifiable through a hadronic shower

$$\nu_l + N \rightarrow l + N^* + X \tag{3.14}$$

In Equation 3.14 X is the number of hadrons produced in the shower and N refers to a neutron or proton. DIS cross sections are calculated using the Bodek-Yang implementation [58], with the hadronic showers modeled with the AGKY model [59].

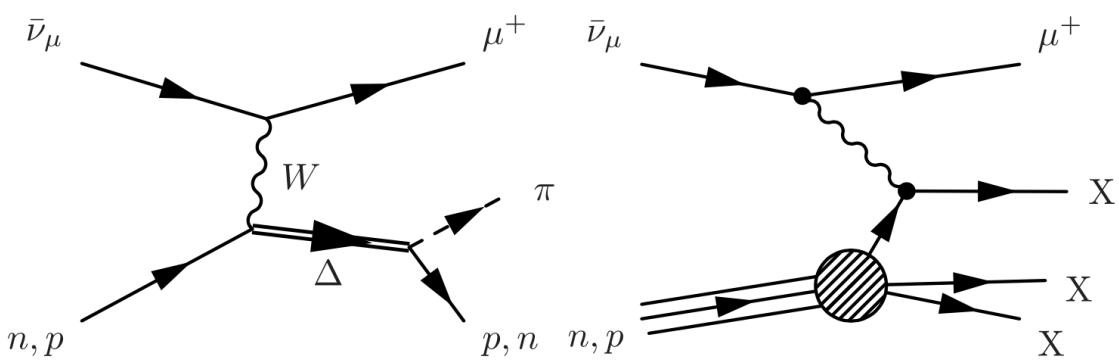


Figure 3.18: RES scattering (left) and DIS scattering (right) Feynman diagrams. Figure from Ref. [45].

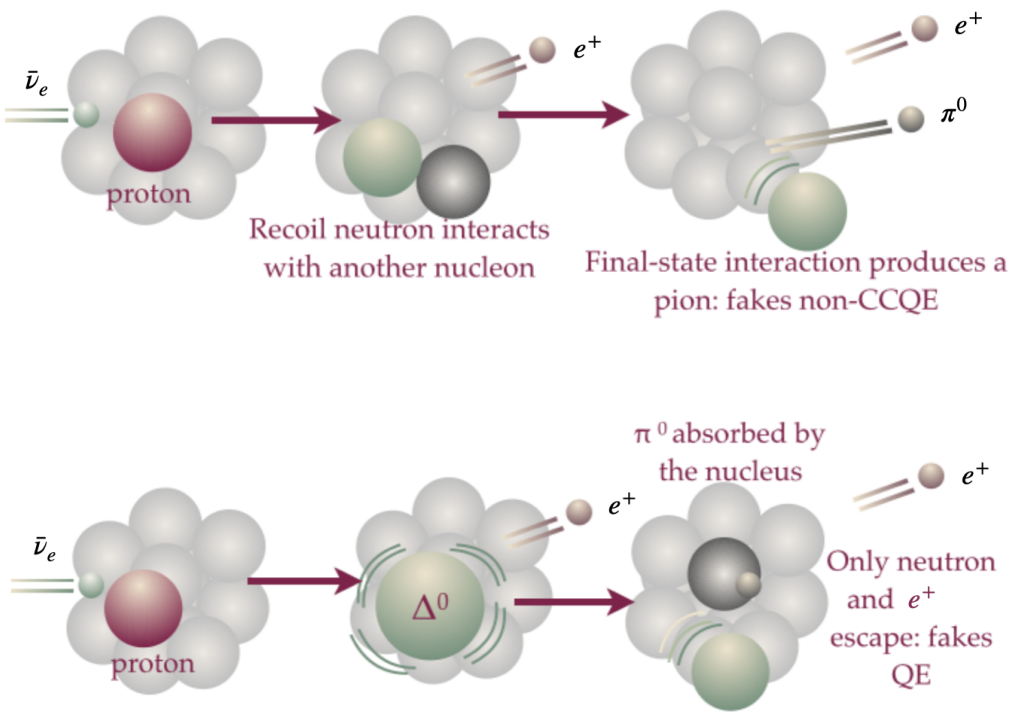


Figure 3.19: Resonant scattering with a not-CCQE-like final state (top) and resonant scattering where pion is absorbed (bottom) creating CCQE-like final state. Figure from Ref. [45].

3.6.4 Final State Interactions

FSI interactions may complicate the interpretation of certain types of interactions in the detector. Hadrons produced by neutrino interactions traverse the remainder of the nucleus and may re-interact or be absorbed within the nucleus before exiting. These occurrences alter the kinematics and multiplicity of the final state hadrons as illustrated in Figure 3.19. GENIE simulates FSI using the INTRANUKE subpackage. After GENIE has produced the neutrino interaction using the neutrino four-momentum and the selected interaction model, INTRANUKE takes those final state particles and steps the hadrons through the nucleus. MINERvA uses a specific model, the "hA" model, where the pions and nucleons are stepped through the nucleus in 0.05 fm increments. The hadrons can undergo re-interaction according to the total hadron-nucleon interaction cross section predetermined from hadron-nucleon scattering data and electron-nucleus elastic scattering data [46]. The simulated type of interaction is determined from hadron-nucleon scattering data or model calculations in the case data is unavailable.

3.7 The MINERvA Tune

While the leading theoretical models governing neutrino-nucleus interactions and particle re-interactions are assembled in GENIE, MINERvA (among other experiments) has found the model predictions in specific regions of phase space to be in tension with data. An analysis from the LE data set was completed by looking at low values of the three-momentum transfer to the nucleus to improve the agreement to both internal and external datasets. MINERvA developed a set of corrections to the base GENIE models to account for discrepancies seen

between the simulated prediction and MINERvA data. The corrections are applied on an event-by-event basis through a set of weighting functions. These corrections, referred to as MnvTunes (MINERvA Tunes) are described Section 3.7.4.

3.7.1 2p2h and RPA Corrections

An LE MINERvA analysis was conducted at low three-momentum transfer (q_3) and parameterized by the available energy E_{avail}. Figure 3.21, shows the selection of events distributed in two q_3 bins as a function of E_{avail}. In the first bin of q_3, there is an excess prediction of MC events in the lowest E_{avail} bins. In the second bin of q_3, there is a deficiency of MC events between the MC delta and MC QE peaks. The MC excess between the QE and RES peaks is believed to be a result of short-range multi-nucleon correlations (2p2h), as mentioned previously. As explained in section 3.7.1, the simulation can be improved by using a calculation that includes the random phase approximation (RPA). The effect of the corrections is seen in Figure 3.20, with a comparison between model predictions using the default GENIE version 2.8.4, RPA model, RPA with the addition of the Valencia model, and only the Valencia model. The RPA model, with the addition of the Valencia model, best agrees with the data. MINERvA modifies the RPA model using a set of reweights as a function of the energy and momentum transferred in the event [60]. An "RPA" weight is applied as part of the MINERvA Tune using the Valencia model.

MINERvA uses the IFIC Valencia model to simulate the 2p2h events in GENIE [54], [61]. Studies done with the MINERvA low energy data showed that the 2p2h events modeled by the IFIC Valencia model better predicted the data than the default empirical model used in GENIE [55]. The same study also showed that while the implementation

of the Valencia model made the MC distribution and data agree better, the 2p2h events predicted by the model had to be enhanced to bring the simulation even closer to the data. The empirical enhancement of the 2p2h events based solely on MINERvA 's measured ν_μ cross section brought the neutrino event distribution in both the FHC and RHC runs into better agreement with data.

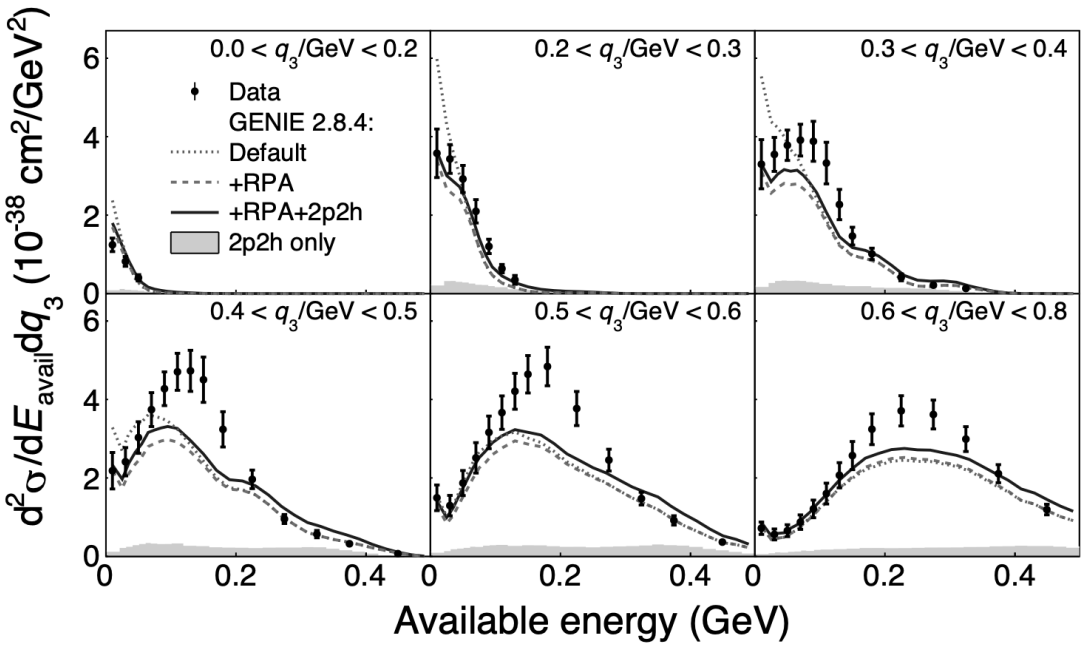


Figure 3.20: The double-differential cross section $d^2\sigma/dE_{avail}dq_3$ compared to GENIE 2.8.4. Adding the RPA effect reduces the over-prediction in the first three q_3 bins. 2p2h modeled by Valencia is shown in solid, with RPA+2p2h corrections the best fit to the data. Discrepancies still remain. Figure from Ref. [55].

3.7.2 Non-resonant Pion Reweight

Assumptions of neutrino-nucleus interactions stem from the studies of neutrino scattering off of a free nucleon at rest. The closest the community has achieved to a free nucleon

scattering analysis has been the deuterium bubble chamber [62]. However, neutrino interaction measurements are inherently more complicated since the nucleons are bound and nuclear effects must be included in the interaction models. As a result, the implementation of non-resonant pion production in GENIE through the Bodek-Yang and AGKY DIS models, has large systematic errors and discrepancies with data. By default, GENIE assigns a 50% uncertainty to the nominal values of the non-resonant parameters. A fit using data from the ANL (Argonne National Lab) and BNL (Brookhaven National Lab) bubble chamber by Wilkinson, Rodrigues *et al* [63], observed that the non-resonant channels of $\nu_\mu + n \rightarrow \mu^- + p + \pi^0$ and $\nu_\mu + n \rightarrow \mu^- + n + \pi^+$ were overpredicted by GENIE. The GENIE weight for non-resonant pion interactions is suppressed by 43% with a $\pm 4\%$ uncertainty to correct for the overprediction.

3.7.3 Low Q^2 Suppression

There exists a disagreement between simulation and data for pion production on scintillator at low Q^2 values, where the model overpredicts the cross section measured in data. This was seen in all four MINERvA LE pion results [64, 65, 66, 67]. MINERvA conducted a tune on four different pion production channels using NUISANCE, a software package allowing for simultaneous comparisons of various neutrino interaction generators [68]. The channels of single charged pion and single neutral pion production through neutrino and anti-neutrino channels ($\nu_\mu CC1\pi^\pm$, $\bar{\nu}_\mu CC1\pi^\pm$, $\nu_\mu CC1\pi^0$, $\bar{\nu}_\mu CC1\pi^0$) were simultaneously fit between data and GENIE pion production parameters as free parameters. The fit between the MINERvA data and the GENIE models, with the addition of ANL/BNL bubble chamber data included as priors on the fit parameters, showed improved agreement between data

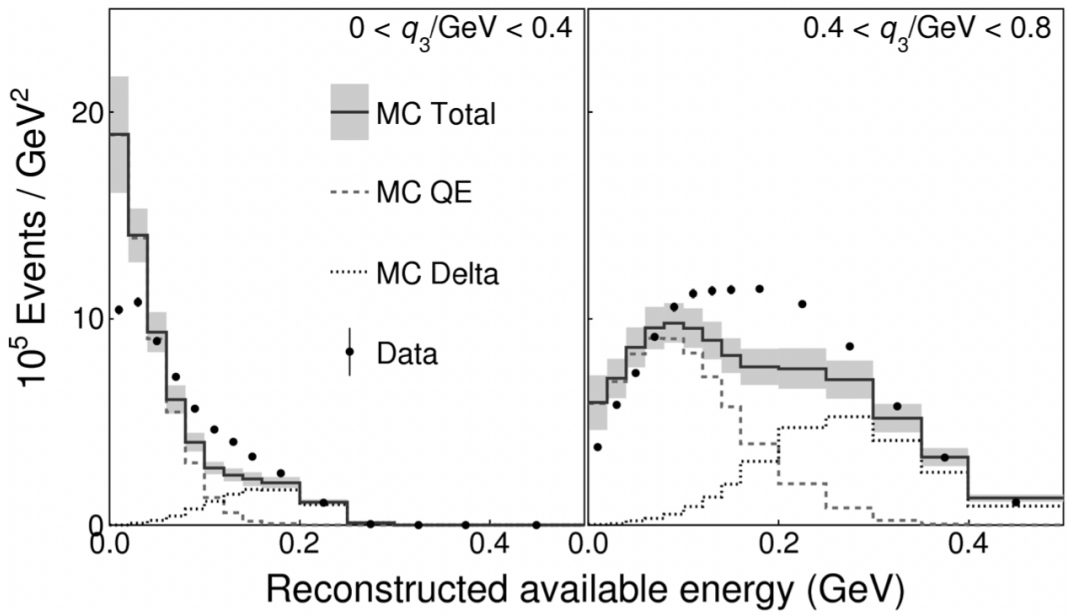


Figure 3.21: The reconstructed available energy in bins of three-momentum transfer. The left panel refers to the MC excess of events and the right panel refers to the MC deficiency seen between the MC QE and MC Delta (RES) peaks. Figure from Ref. [55].

and MC predictions. However, as seen in Figure 3.22, no tune finds perfect agreement in all channels.

3.7.4 MINERvA Tune v1 and MINERvA Tune v2

As a result of the tunes between GENIE model predictions and data, MINERvA extracts a set of reweights. The reweights are directly applied to the GENIE MC simulation, allowing for a better constraint of neutrino-nucleus interactions and the ability to compare reweighted results with other collaborations seeking to tune GENIE predictions further. The modification of GENIE is referred to as the "MINERvA Tune" and this analysis specifically

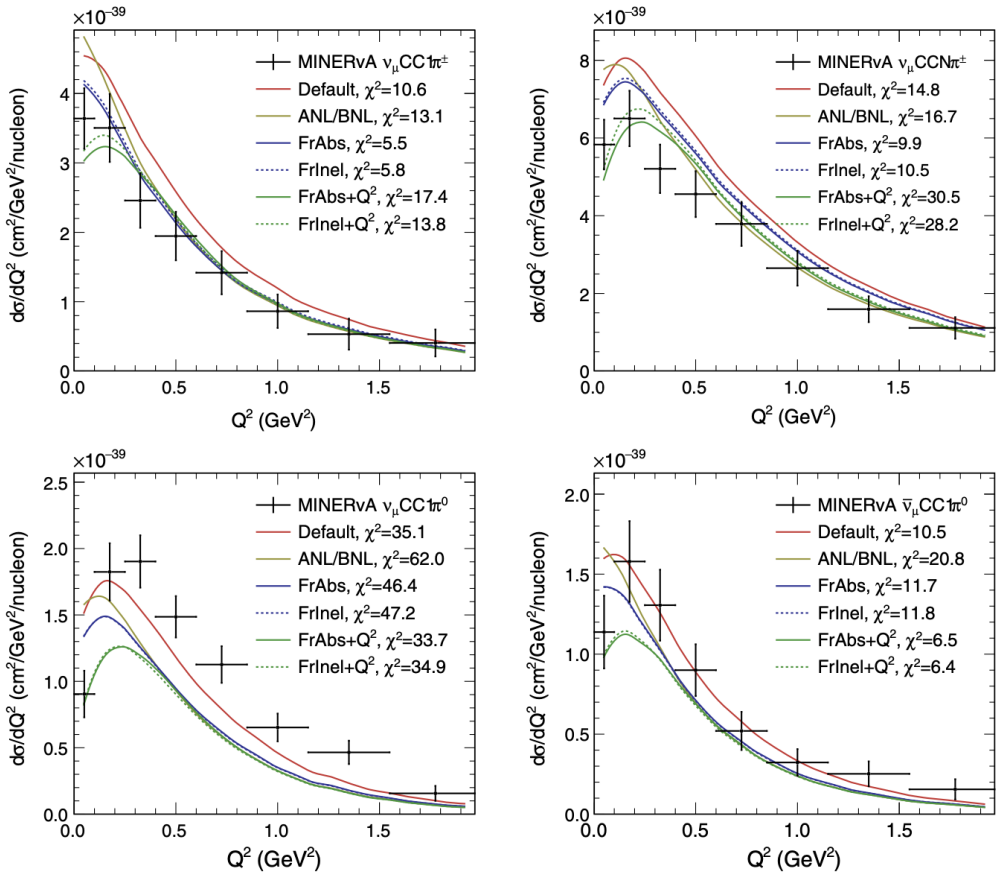


Figure 3.22: Comparisons between the nominal and tuned MINERvA pion production models. Channels of $\nu_\mu CC1\pi^\pm$ (top left), $\bar{\nu}_\mu CC1\pi^\pm$ (top right), $\nu_\mu CC1\pi^0$ (bottom left) and $\bar{\nu}_\mu CC1\pi^0$ (bottom right) distributions are shown as a function of Q^2. the light green model (ANL/BNL) refers to the tuned GENIE model. The FrAbs shows the cross-section prediction with the nominal FrAbs model, and FrAbs+Q^2 with the suppression weight. Figure from Ref. [69].

uses "MnvGENIE-v2." A summary of MnvGENIE-v1 and MnvGENIE-v2 tunes are found in Tables 3.4 and 3.5 respectively.

In conclusion, MINERvA's use of high-intensity neutrino beams to study neutrino-nucleus interactions allows it to provide high-precision cross section measurements as inputs for current and future neutrino oscillation experiments. MINERvA has already produced many cross section results, with a current publication list found in Ref. [70].

MnvGENIE-v1
MC simulation uses GENIE v2.12.6 [46]
Central Value (CV) universes are adjusted with flux corrections from a PPFX package and a correction made to the flux uncertainty [31].
Nonresonant pion reweight based on reanalysis of bubble chamber data [71]
2p2h events simulated using the Valencia 2p2h model and Valencia RPA suppression applied as a weight to q_3 events [72]
2p2h enhancement using an 2D Gaussian empirical fit of the simulation to MINERvA data [61]

Table 3.4: Table summary of Minerva Tune components for the "MnvGENIE-v1" tune.

MnvGENIE-v2
MC simulation uses GENIE v2.12.6 [46]
Central Value (CV) universes are adjusted with flux corrections from a PPFX package and a correction made to the flux uncertainty [31].
Nonresonant pion reweight based on reanalysis of bubble chamber data [71]
2p2h events simulated using the Valencia 2p2h model and Valencia RPA suppression applied as a weight to q_3 events [72]
2p2h enhancement using an 2D Gaussian empirical fit of the simulation to MINERvA data [61]
Suppression of low-Q^2 pion production

Table 3.5: Table summary of Minerva Tune components for the "MnvGENIE-v2" tune.

4. $\bar{\nu}_e$ Low Recoil in Medium Energy

The double differential cross sections reported include a final state of an anti-electron neutrino plus final state interactions. This analysis is an inclusive charged current sample at low three momentum transfer to probe multi-nucleon effects like 2p2h and production of baryon resonances. The analysis is performed in the tracker region of the MINERvA detector and uses the medium energy neutrino flux from the NuMI beamline, which is peaked at around 6 GeV. The cross section is reported in two dimensions, E_{avail} vs q_3 and E_{avail} vs p_T to gain insight into the nuclear effects for the neutrino-nucleus interactions.

This chapter establishes the process by which low recoil $\bar{\nu}_e$ events are selected, referred to from here as Signal events. Due to a large number of backgrounds (not $\bar{\nu}_e$ events) and the nature of electron and positron interactions in the detector, multiple studies were conducted to optimize the selection.

4.1 Characterization of Events in the MINERvA Detector

4.1.1 Minimum Ionizing Particles

The detection of particles traversing the MINERvA detector is delivered through scintillating light as discussed in Section 3.1.1. These nuclear particles transfer energy to the material they encounter through ionization. The Bethe-Bloch formula for mean rate of ionization loss of a charged particle is written as

$$\frac{dE}{dx} = \frac{4\pi N_0 z^2 e^4}{m_e c^2}\frac{Z}{A}[\ln(\frac{2mv^2}{I(1-\beta^2)}) - \beta^2] \tag{4.1}$$

for $\frac{dE}{dx}$ the average energy deposited per length traversed in a material by a charged particle, m_e the electron mass, z the charge, v the velocity, $\beta = \frac{v}{c}$, Z and A the atomic number and mass number of the atoms in the medium, I the mean excitation energy of the material, x is the path length in the medium, and N_0 is Avogadro's number. The total amount of ionization that occurs from an incident high energy particle is dependent upon the $\frac{dE}{dx}$ and the energy required to free an ion pair in the given material. The Bethe-Bloch $\frac{dE}{dx}$ for pions, muons, and protons in different materials can be found in Figure 4.1.

Due to the energies created by the NuMI beamline, muons seen in the MINERvA detector are often referred to as *minimum-ionizing particles* or "MIPs". MIPs are charged particles that primarily interact in the detector through ionization of the atomic electrons, are weakly dependent on momentum and they leave behind deposits of energy near the minimum of their energy loss ($\frac{dE}{dx}$) distribution which is typical for most relativistic particles.

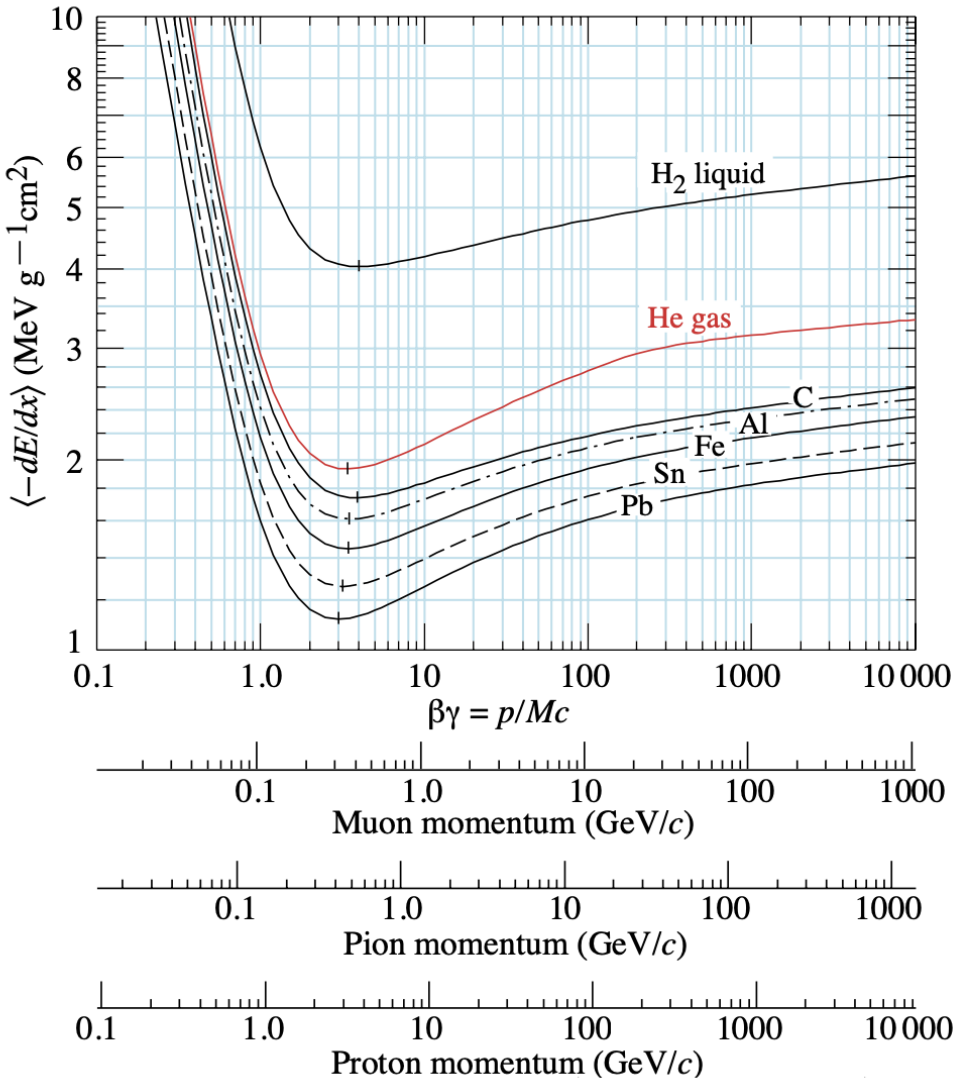


Figure 4.1: Mean energy loss rate for muons, pions and protons in various materials derived from the Bethe-Block equation. Figure from Ref. [73]

.

MIPs in MINERvA, for example muons, tend to produce a clear path of energy deposits in the detector, making their kinematics relatively easy to reconstruct. There are various forms

of the Bethe-Bloch equation; however, the standard equation shown in Equation 4.1 refers to "heavy" charged particles where the mass is much greater than the mass of an electron (μ, π, p, α,... etc). Muons at moderate energy (E $<$ 100 GeV) almost entirely interact through ionization and it is not until high energies that the relationships between the cross sections of different processes come into play. This analysis focuses on $\bar{\nu}_e$ interactions with the production of the light charged particle, e^+.

4.1.2 Electrons and Photons

Electrons (and positrons) lose energy primarily via ionization and bremsstrahlung. Bremsstrahlung, or "breaking radiation", is when the radiative collisions of electrons (or charged particles) occur mainly with the atomic nuclei of the traversed medium and the nuclear electric field of the atomic nuclei decelerates the electron. The change in energy from the decelerated electron comes in the form of radiation, or photons, through conservation of energy. In lead, at low energies ($<$10 MeV) electrons primarily lose energy through ionization and for energies $>$10 MeV bremsstrahlung is the dominant process, as seen in Figure 4.2. In general, bremsstrahlung scales as $\propto Z^2$ while ionization and e^+e^- processes scale as $\propto Z$. Therefore, the point of crossover is dependent on the material through which the particles

Material	Z	E_c [MeV]	X_0 [cm]
C	6	81.7	19.3
Al	13	42.7	8.90
Fe	26	21.7	1.76
Cu	29	19.4	1.44
W	74	8.0	0.35
Pb	82	7.4	0.56

Table 4.1: Critical energy of electrons Ec and radiation length X_0 for different materials.

are traversing, as shown in Table 4.1 [73] [74]. The point of crossover between energy loss through ionization vs bremsstrahlung is referred to as the critical energy, E_c, and occurs around a few tens of MeV for electrons. This analysis requires electrons of energy value greater than 2.5 GeV and therefore the electron energy loss is mainly through the bremsstrahlung process.

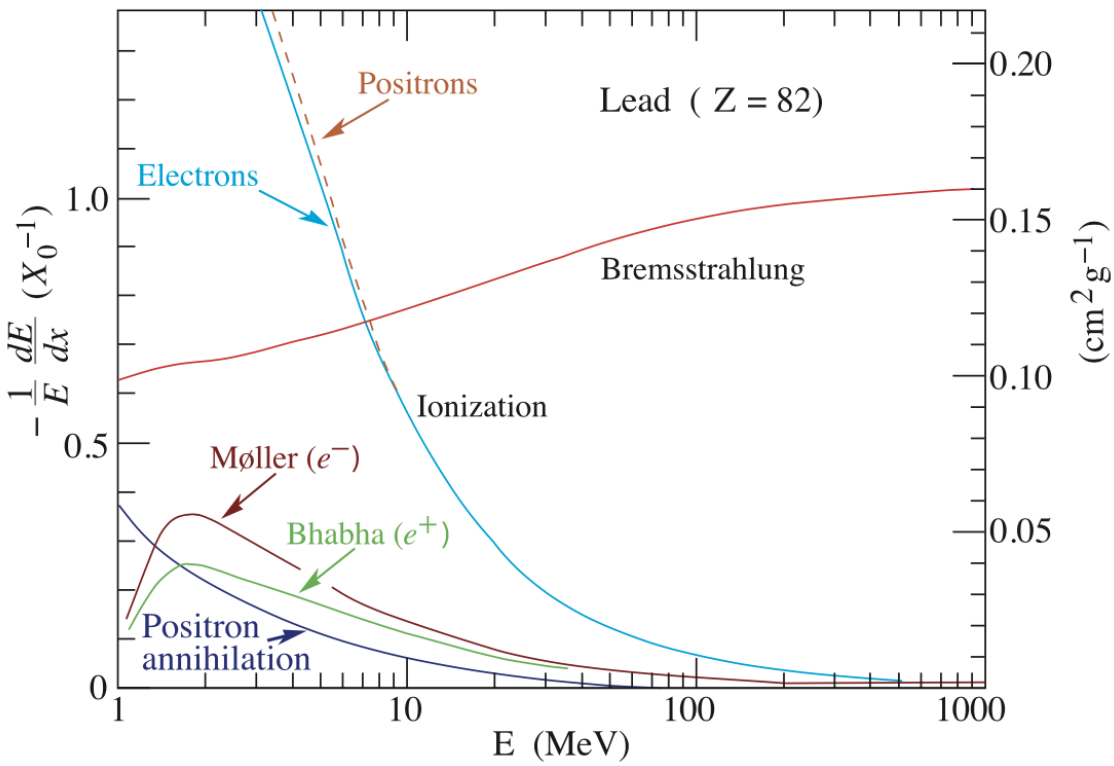


Figure 4.2: Fractional energy loss per radiation length in lead as a function of $e^{\pm}$ energy. The crossover of ionization and bremsstrahlung is referred to as the critical energy, E_c. Figure from Ref. [73]

When traversing a material, photons can interact through the photoelectric effect, Compton scattering (incoherent scattering off of an electron), Rayleigh scattering (coherent

scattering off of an atom), and photonuclear absorption (absorption of a photon and ejection of a nuclear particle). The photoelectric effect occurs when electrons are emitted from a material that has absorbed some amount of electromagnetic radiation. In MINERvA, the PMTs described in Section 3.3 use this photoelectric effect to accelerate electrons through a chain of dynodes, resulting in the collection of electrons at an anode in the detector with a typical gain of about half a million electrons per photoelectron. The photoelectric effect dominates at low photon energies, but as energy increases, a new energy loss process of Compton scattering enters. Compton scattering is the scattering of high energy photons off of free electrons. The photoelectric cross section scales as $\frac{1}{E^3}$ and the Compton cross section scales as $\frac{1}{E}$ for E the photon energy. As seen in Figure 4.3, the point of crossover is material dependent. Pair production becomes dominant in lead at around 10 MeV and in carbon at around 30 MeV.

Pair production is the dominant mode of photon interaction with matter for higher photon energies and defined by the conversion of a photon into an electron-positron pair:

$$\gamma \rightarrow e^- + e^+. \tag{4.2}$$

To conserve energy and momentum, the photon must pair-produce near a nucleus, which receives some recoil. As seen in Figure 4.3, pair production begins to occur around 1.02 MeV because the photon requires the combined rest energy of an electron and positron. However, the dominant reaction for pair production changes according to different media. In the MINERvA detector the threshold for pair production is on the order of 30 MeV, since the Compton scattering is more dominant at lower energies in lower Z material.

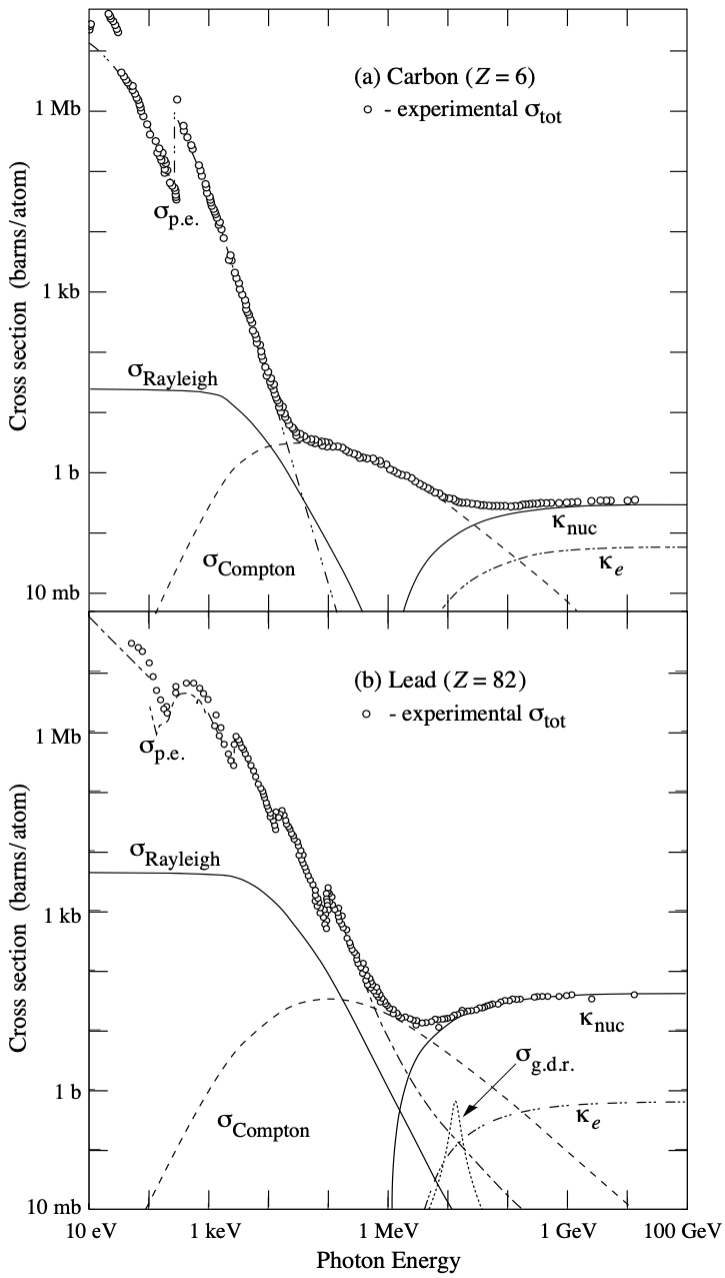


Figure 4.3: Photon total cross section as a function of energy in carbon (top) and lead (bottom). Figure from Ref. [73]. $\sigma_{p.e}$ dominates at low photon energy while κ_{nuc} pair production in a nuclear field dominates at higher energy.

4.1.3 Electromagnetic Cascades

High-energy electrons and positrons mainly lose energy through bremsstrahlung processes, whereas high-energy photon interactions are dominated by electron–positron pair production. Since high-energy electrons and positrons can lose energy in the bremsstrahlung processes, new photons can be produced. In turn, the photons will undergo pair production, producing an electron-positron pair with lower energy. The number of particles increases exponentially with the depth of the medium and continues until the electron falls below the critical energy, where it will lose the rest of its energy through ionization. The result of the two processes is referred to as an *electromagnetic cascade* or *electromagnetic shower*. The shower development typically follows the same pattern regardless of material. There is an initial exponential rise, a broad maximum and then a gradual decline with increasing depth (radiation lengths) in the material [75]. While these electrons, positrons, and photons multiply in number as seen in Figure 4.4, they are also being scattered through matter due to electromagnetic interactions with different nuclei referred to as multiple Coulomb scattering. This Coulomb scattering makes the shower spread transversely, with the spread of the shower defined as a Molière radius R_M written as

$$R_M = 21\frac{X_0}{E_c} \tag{4.3}$$

for X_0 the radiation length and E_c the critical energy. About 90% of the electromagnetic shower is contained within a cylinder of radius 1 R_M [76]. The Molière radius for polystyrene is ≈ 9.4 cm [77].

The electromagnetic cascade begins with a charged particle of starting energy greater

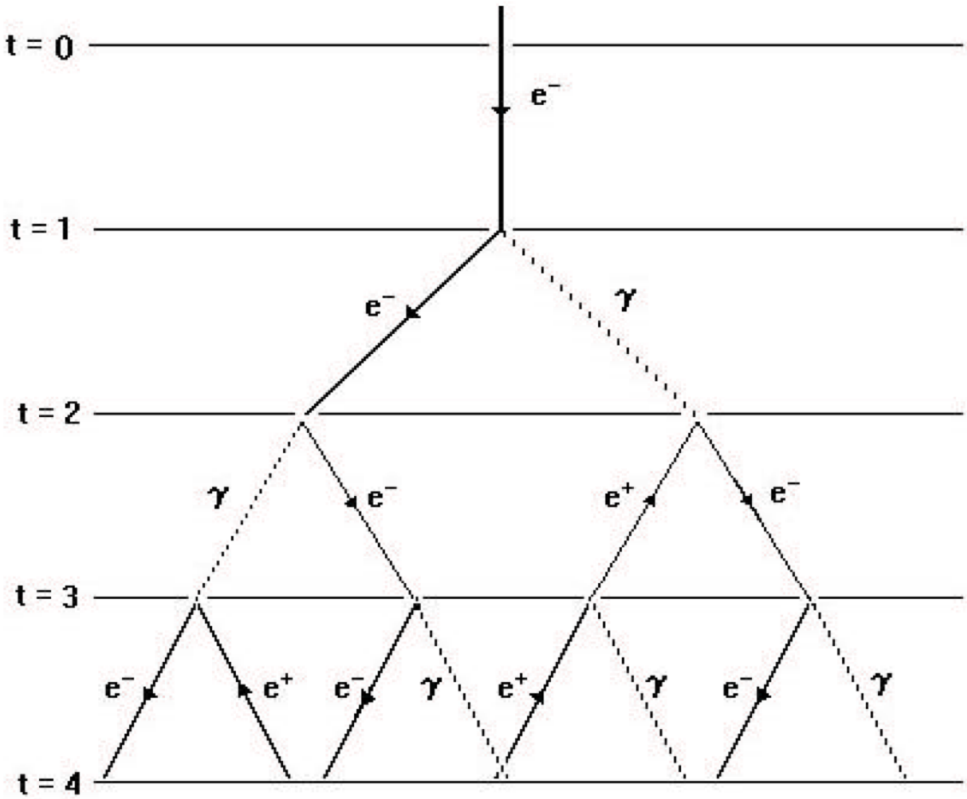


Figure 4.4: Diagram of an electromagnetic shower for t radiation lengths. Figure from Ref. [78].

than the critical energy E_c with pair production and bremsstrahlung dominating the early shower development. Electrons will travel one radiation length and lose about half their energy to a bremsstrahlung photon while $E_0 > E_c$. Each photon produced while $E_0 > E_c$ will travel one radiation length and create an electron-positron pair carrying away the energy of the photon. Electrons will stop radiating when their energy falls below the critical energy and will begin ionizing in the detector becoming the primary mechanism of energy loss. The electromagnetic shower will reach a maximum and then abruptly end at a radiation

length t, written as:

$$t_{max} = \frac{\ln(E_0/E_e)}{\ln 2} \tag{4.4}$$

When traversing the medium, the charged electrons will be deflected by many small angle scatters through Coulomb scattering. Typically, the net displacement of the particles is summarized through multiple individual collisions and can be characterized as Gaussian at small angles.

Electromagnetic cascades are well understood and well modeled. Monte Carlo simulations include the known energy dependent cross sections for the radiation, ionization loss for electrons, the absorption of photons through pair production and other processes [75]. As mentioned previously, MINERvA uses the software package Geant4 which uses Lewis theory to encompass the spatial and angular distributions pertinent to the electromagnetic shower development [27]. A MINERvA MC electromagnetic shower event display is seen in Figure 4.5; note the transverse spread relative to the incoming neutrino direction. The low Z nuclei of the central tracking region allow the dE/dx (energy loss per unit length) near the beginning of the electromagnetic (EM) shower to be a distinguishing metric between the interactions of an electron vs. a photon [40]. This is only true at the beginning of the EM shower. The early shower development in the tracker region is followed by the EM calorimeter (the blue-shaded region in the figure). Due to the higher pair production cross section in the EM calorimeter lead material, the radiation length is shorter (see table 4.1), and the shower develops much faster. Therefore, the EM calorimeter serves as a method to capture the remaining energy of the neutrino interaction within the length limitation of the detector.

In the following sections, many of the methods used to select a ν_e and $\bar{\nu}_e$ signal sample

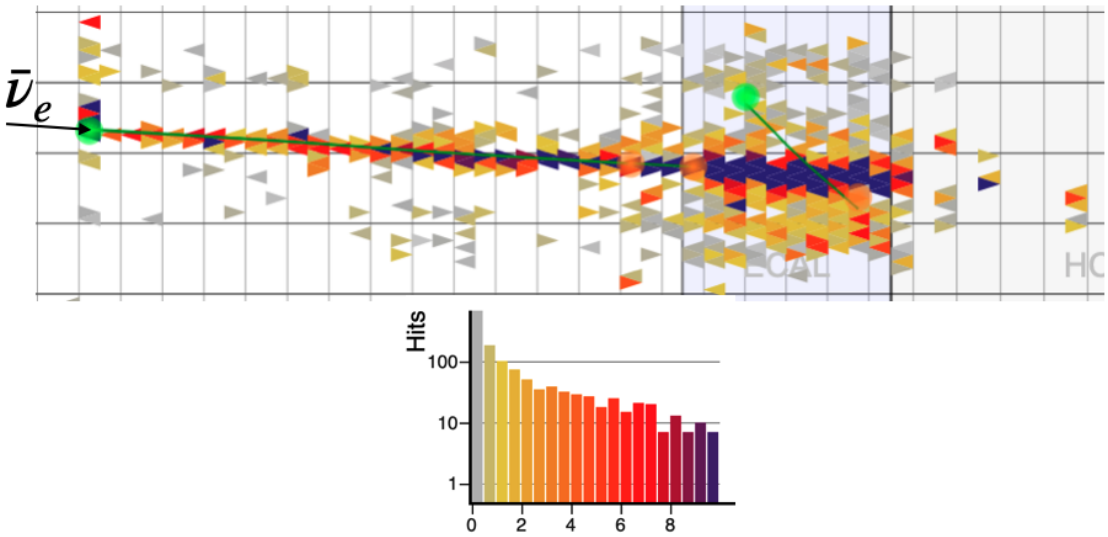


Figure 4.5: A MC event display showing the interaction of a 12.23 GeV $\bar{\nu}_e$. The final state particles are protons, a π^-, and an 11.4 GeV positron which creates an electromagnetic shower in the detector.

were developed in an earlier MINERvA analysis carried out in the low energy era of the experiment with peak flux around 3 GeV. For this analysis, many of the applied techniques and selection requirements were optimized in the LE analysis and can be found in Ref. [79]. Figures borrowed from Ref.[79] will be indicated. Discussed below are the processes required in finding an electron-like interaction in the detector and determining whether that interaction meets the requirements set by the analysis.

4.2 Electron Cone Candidate

4.2.1 Neutrino Type

The goal of the Signal selection is to isolate events containing either an electron or positron. MINERvA is not able to differentiate between the two since it is not magnetized, and all electrons and positrons will shower in the detector before ever reaching MINOS. In practice, selecting events with an electron or positron can be challenging due to the presence of additional particles or particles that mimic the characteristics of an electron or positron. Therefore, there are many layers in the analysis to isolate true electron and positron events.

Muon neutrinos dominate the beam entering MINERvA. When these neutrinos interact in the detector and produce a photon, perhaps from π^0 decay, it can create an electromagnetics shower that mimics the signal for this analysis. Charged current ν_μ interactions that produce a muon are easily identifiable because a large percentage will traverse the entire MINERvA detector and enter the magnetized muon spectrometer, MINOS. The energy deposits are collected from MINOS, reconstructed, and matched with the track exiting MINERvA, allowing for a straightforward way to tag the event as a charged current muon neutrino event. Therefore, in this $\bar{\nu}_e$ analysis it is required that no tracks exit out the MINERvA detector, eliminating approximately 90% of charged current ν_μ and $\bar{\nu}_\mu$ background events. Even after the initial requirements of no particle tracks exiting MINERvA, there is still a significant background for a $\bar{\nu}_e$ analysis. For example, any particles with similar signatures in the detector compared to an electron, like photons from neutral pion decay, require further constraints to remove them from the signal sample.

4.2.2 Electron Cone Construction

After a neutrino-nucleus interaction event has gone through the preselection of containing at least one reconstructed track in the fiducial volume and no tracks exiting MINERvA, the next step is to determine whether the tracks look like an electromagnetic shower. The Arachne event display seen in Figure 4.6 shows the signatures of differently charged particles as they traverse the MINERvA detector. Due to electromagnetic showers, the positron has a visually larger spread compared to the other particles. For this reason, a cone object is constructed around each track to contain the shower.

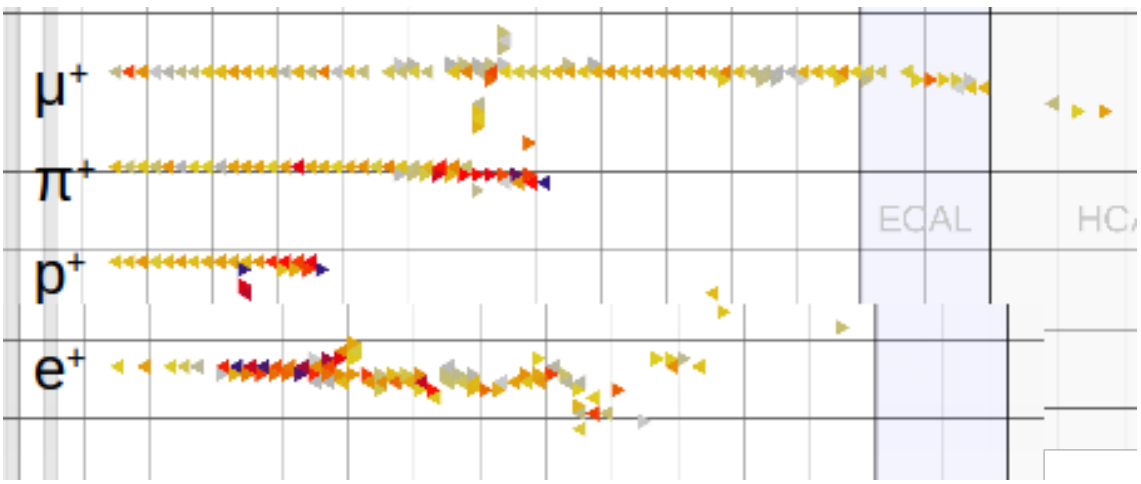


Figure 4.6: Event display showing the signatures of different particles in the MINERvA detector. The triangles are the different energy deposits in the detector.

If there are multiple tracks in an event, the tracks are considered in turn. The vertex of the cone is set on the first node of the reconstructed track and the cone angle opens along the direction of the track. An opening angle of 7.5 degrees was chosen because it reduces the low side tail of the distribution and any greater angle does not contain a more significant fraction of true energy within the cone, as seen in Figure 4.7. Once the cone is

formed around the candidate track, each energy deposit in the event is then classified as either inside or outside of the cone. Once all reconstructed objects are considered, the ones in the cone are consolidated together to undergo further examination.

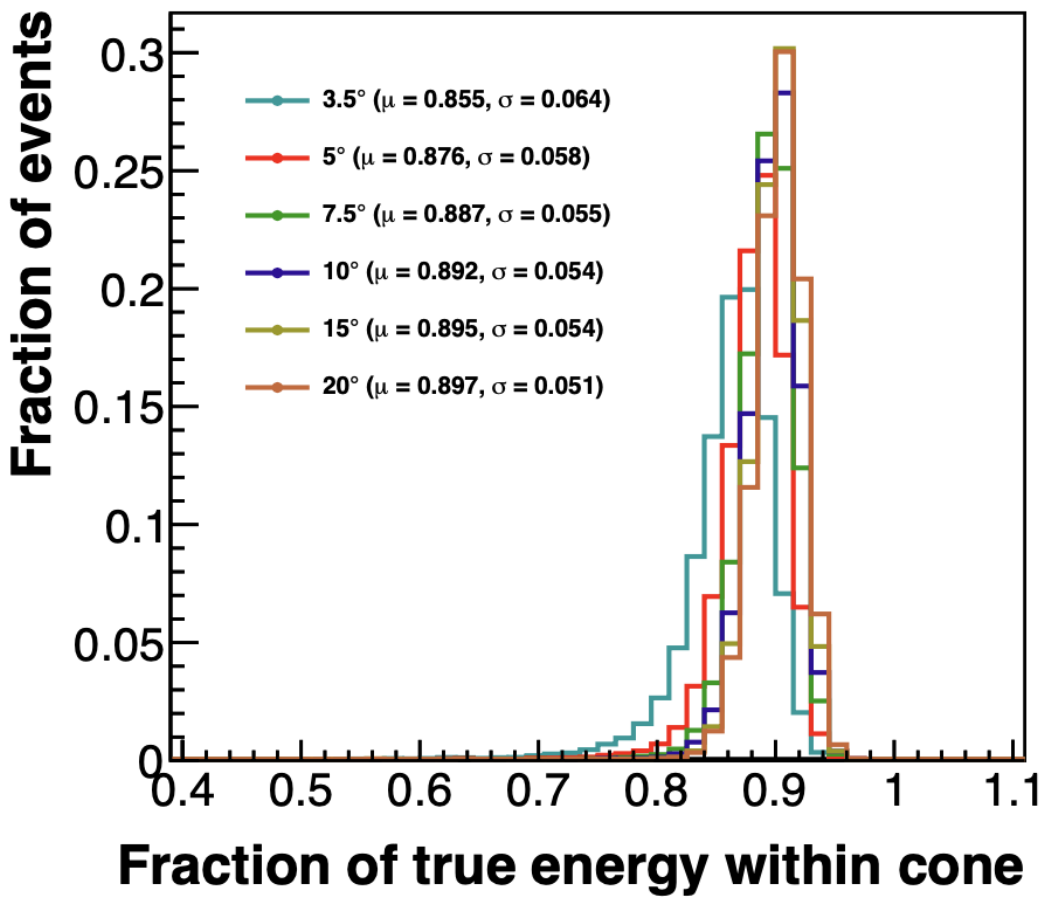


Figure 4.7: Fraction of true energy inside of electron candidate cone with different opening angles as measured in a simulated particle cannon. Figure from Ref. [79].

Electromagnetic cascades will deposit energy at the beginning of their propagation through the detector. Each cluster is projected onto the longitudinal cone axis and sorted

according to distance from the cone vertex to determine whether the clusters in the cone are characteristic of an electromagnetic cascade. Once the cluster's location along the cone axis is determined, the distance between neighboring clusters is calculated. If the cone contained a single photon for example, gaps in the energy profile would be expected. However, due to a large number of particles in an electromagnetic shower, the distance between neighboring clusters along the cone axis is not expected to be significant. If a gap of 3 radiation lengths or more between neighboring clusters exists, all clusters from the gap and beyond are removed from the cone.

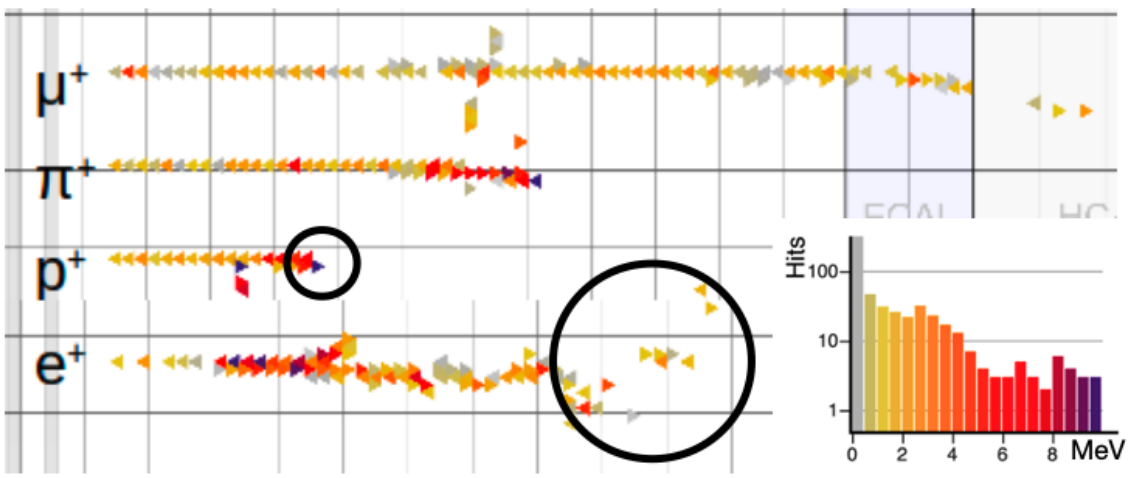


Figure 4.8: An event display of the activity of different particles in the MINERvA detector. The black circles represent the purpose of the endpoint energy fraction, where we can see the different energy deposition profiles at the end of the track. The color represents amount of energy deposited in a scintillator strip.

At this point, the final set of clusters considered part of the cone is determined. With the addition and removal of clusters within the cone object, the original assigned longitudinal cone axis may no longer be the best fit for the electron candidate. The cone axis is reevaluated by creating a new track node using all of the digits within a single plane

inside the cone. A Kalman filter is applied to the clusters in the same manner as general reconstruction and if the filter successfully finds the most optimal track, the cone's vertex and axis are updated [80]. Each cone object is assigned a reconstructed energy with the final cone positioning. The reconstructed energy is calculated by summing the energies of the clusters inside the cone.

4.3 Electromagnetic Shower Classification

Once the electron cone object has been created and the energy inside of the cone is reconstructed, three distributions are used to determine whether the energy profile inside the cone is electromagnetic in origin. The three distributions are the endpoint energy fraction, Mean dE/dx and Median Transverse Width. Each of these distributions is simulated using a particle cannon. A particle cannon is a Monte Carlo simulation that propagates a single particle in the MINERvA detector with custom base requirements such as angle and initial momentum of the particle. In the case of these studies electrons, photons, pions, protons, and muons were all simulated with the particle cannon using initial momenta of 0-10 GeV with angles of less than 45° with respect to the MINERvA defined z-axis, inside of the MINERvA tracker region. By comparing the energy profile of each particle, a requirement is formed in each category to find separation between electrons and other particles. After the requirements are formed, they will be fed into a multivariate analysis classifier (MVA), providing a metric to determine whether the electron cone object qualifies as an electron candidate.

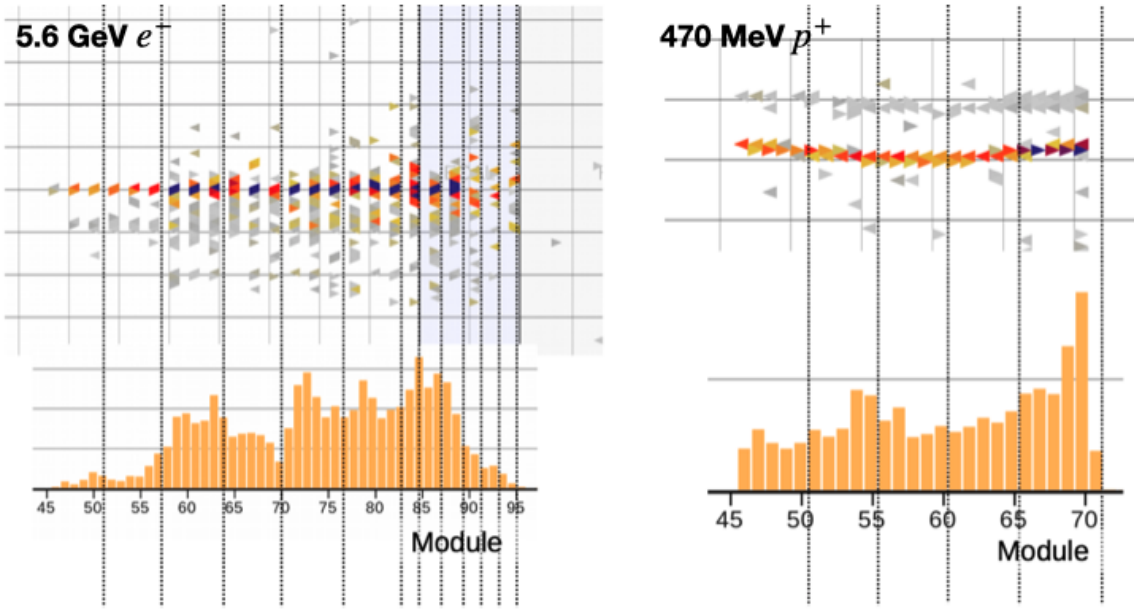


Figure 4.9: A representation of the division of energy deposits into bins of 10 g/cm^2 of areal density. Figure from Ref. [81].

4.3.1 Endpoint Energy Fraction

The endpoint energy fraction is the fraction of the cone's energy deposited at the end of the cone axis. The longitudinal energy profile of an electromagnetic shower rises to some maximum and then more slowly decays as the shower loses energy. This is in contrast to a MIP-like particle in which a large amount of energy is deposited at the end of its track in the detector, as seen in Figure 4.8. The fraction, $f_{endpoint}$, is calculated by projecting the cone's energy onto the cone axis in bins of size 10 g/cm^2, as seen in Figure 4.9.

Any clusters with less than 2 MeV of energy are removed and the median energy is calculated for each bin excluding the last, as represented in Figure 4.10. The endpoint

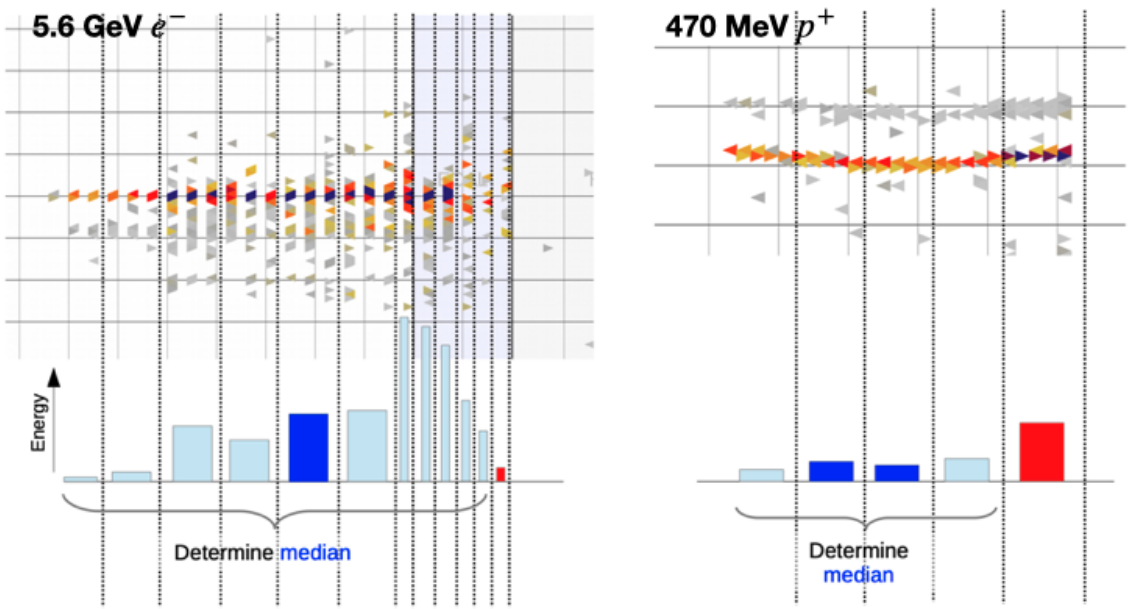


Figure 4.10: Determining the median (blue) value of the energy deposits excluding the last bin (red). Figure from Ref. [81].

energy fraction can be written as

$$f_{endpoint} = \frac{E_N}{E_{median}} \tag{4.5}$$

for E_N the last bin containing energy in the cone. The energy profiles of each respective particle in relation to the endpoint energy fraction are found in Figure 4.11. On average the $\mu^\pm$, p^+ and $\pi^\pm$ have an average $f_{endpoint}$ above 1 whereas the $e^\pm$ and γ have an average $f_{endpoint}$ of $1/2$, though there are large event by event fluctuations.

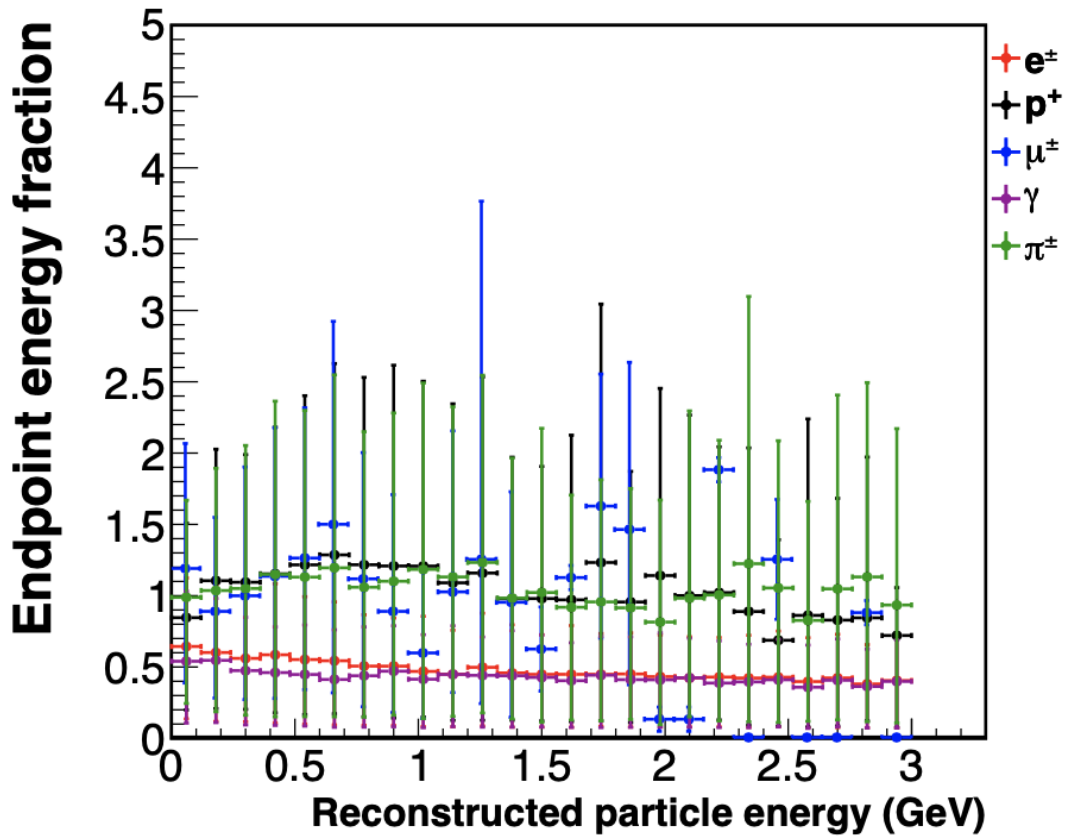


Figure 4.11: Endpoint energy fraction as a function of calorimetrically reconstructed energy. Figure from Ref. [79]. Error bars on each point represent the smallest interval containing 68% of the distribution in that bin.

4.3.2 Mean dE/dx

The longitudinal energy profile inside the cone is determined by considering the amount of energy deposited per unit range dE/dx. The mean dE/dx, in this case, is the total reconstructed energy divided by the total integrated density traversed along the cone axis. Due to the multiplicative showering process, the energy loss of the $e^{\pm}$ and γ occurs at a faster rate than the ionization processes that dominate the $\mu^{\pm}$, p^{+} and $\pi^{\pm}$. Therefore, $e^{\pm}$

and γ deposit a higher amount of energy per unit range and there is again separation of the $e^{\pm}$ and γ profiles as seen in Figure 4.12.

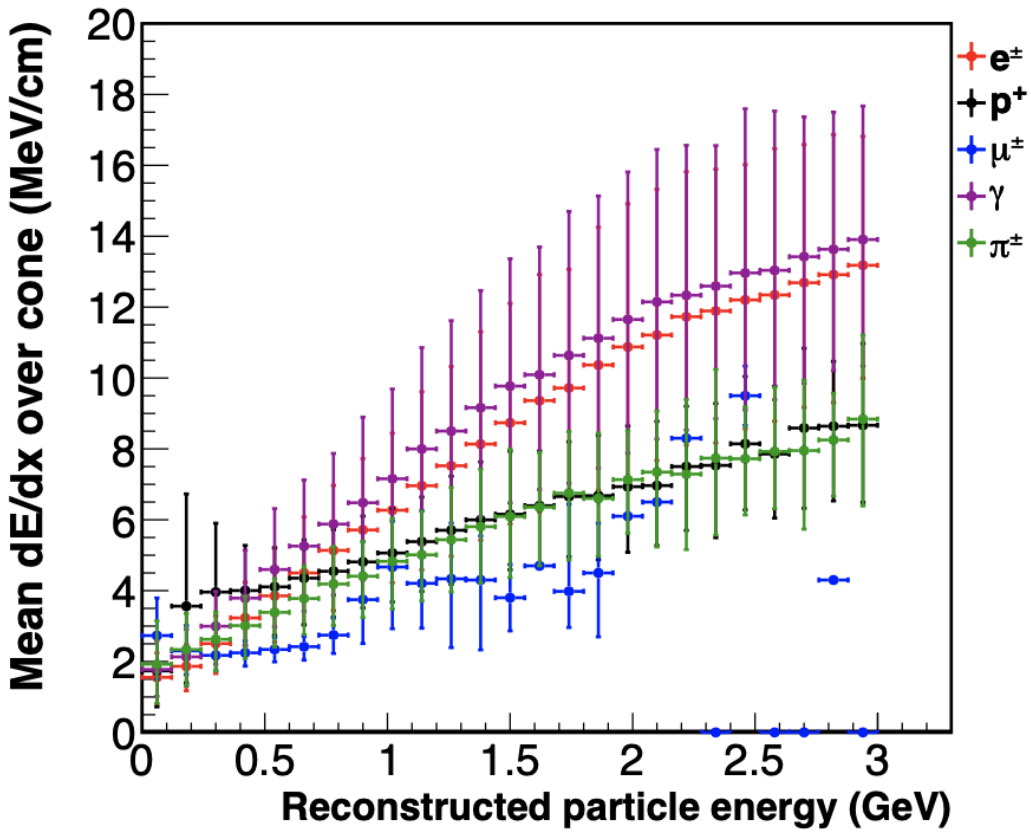


Figure 4.12: Mean dE/dx as a function of calorimetrically reconstructed energy. Error bars on each point represent the smallest interval containing 68% of the distribution in that bin. Figure from Ref. [79].

4.3.3 Median Transverse Width

An attempt is made to characterize the transverse spread of different particles in the detector, expecting electromagnetic shower to have a high spread value. The median transverse width

is calculated by searching for pairs of neighboring Digits in each plane. If the two Digits with the highest energy value within a plane are neighbors, then the Digits are merged into one large Digit. After merging, the weighted standard deviation of the strip numbers are calculated for each plane, using the Digit energy values as the weights. Median transverse width is defined by the median of the standard deviations and is shown in Figure 4.13 in terms of strip number. The large number of particles involved in the multiplicative showering process scatter at different trajectories, creating separation for the $e^{\pm}$ and γ profiles.

4.3.4 MVA Classifier

A k-nearest-neighbor classifier (kNN) provided by the TMVA package from ROOT is used for the training on the three input distributions [82]. The kNN discriminator trains on N variables simultaneously by dividing the samples into the training signal ($e^{\pm}$) and background category and populating an N-dimensional space. It is determined what fraction of the distance between the points in N-dimensional space are true signal events. Weights are then constructed to classify whether an event is deemed a candidate electron event. The classification output of the kNN algorithm is referred to as ξ_{kNN}. The three variables mentioned in Sections 4.3.1, 4.3.2 and 4.3.3 are the inputs to the kNN discriminator.

The minimum value of ξ_{kNN} is determined using $e^{\pm}$ events from the particle cannon simulation to isolate electron candidate events. To determine where the cut value lies, the product of efficiency ε and purity π serves as a metric to determine what threshold keeps the most Signal events with the least amount of background contamination. The distributions of the EM-like and non-EM-like events in the ξ_{kNN} output and the corresponding $\varepsilon \times \pi$

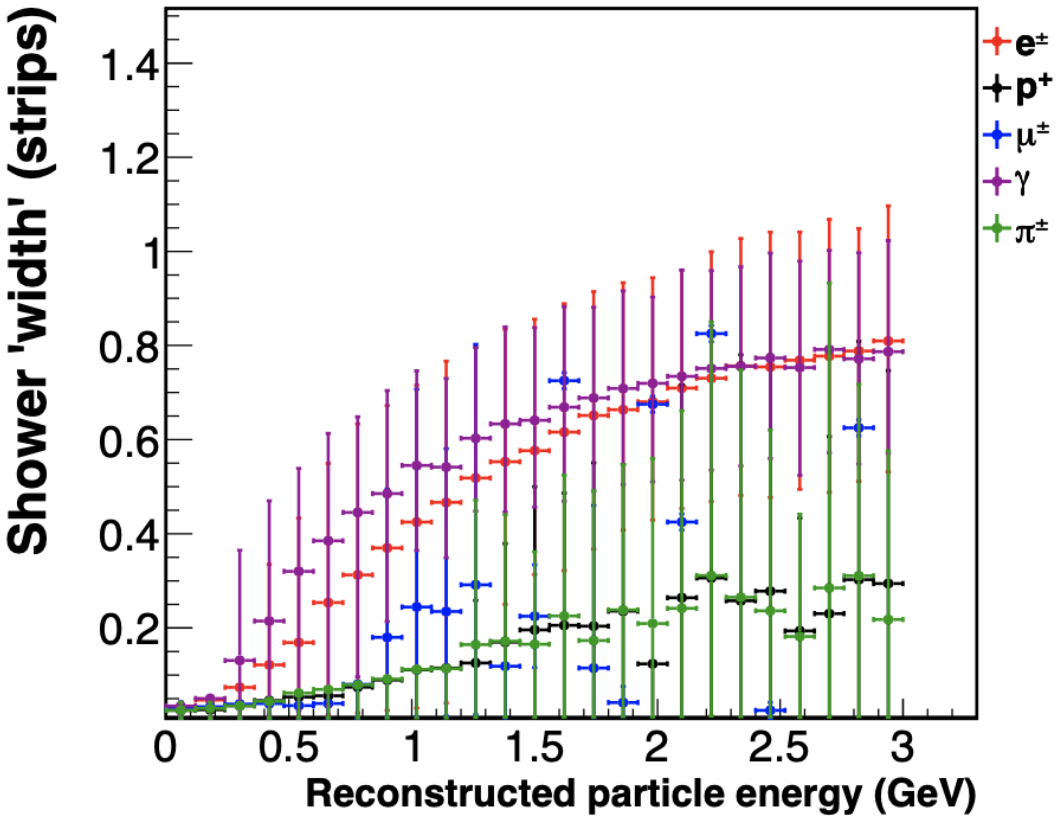


Figure 4.13: Median transverse width as a function of calorimetrically reconstructed energy. Figure from Ref. [79].

values corresponding to various cuts of ξ_{kNN} are shown in Figure 4.14.

If the cone has a value of $\xi_{kNN} \geq 0.7$ then it is considered EM-like. Any cone objects not meeting this threshold will have all clusters released and fed back into the reconstruction. The $\xi_{kNN} \geq 0.7$ will be referred to as the particle identification (PID) requirement in future sections since it distinguishes electrons and photons from other particles.

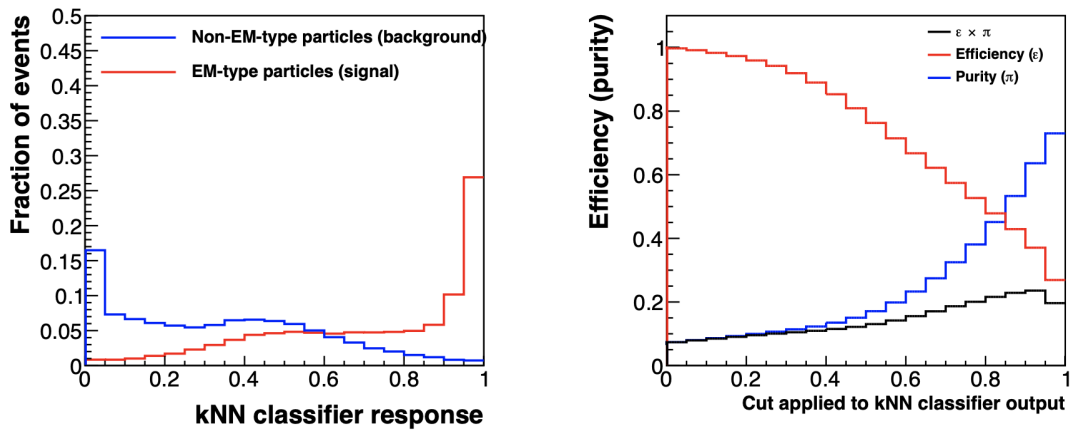


Figure 4.14: The ξ_{kNN} values (left) from the kNN classifier and the corresponding efficiency-purity curve (right). Figure from Ref. [79].

4.3.5 Fraction of Energy in non-MIP-like-clusters

A metric called "non-MIP-cluster" energy fraction is formed to reduce background contamination from low energy electromagnetic showers. Low energy showers created from low energy electrons tend to burn themselves out before having a large transverse spread. However, information from the longitudinal energy profile can prove to be insightful. In the case of a MIP it is expected that the energy deposits left by the particle would be considered "Trackable" using the Cluster classification system from the reconstruction stage. An electromagnetic shower would show fluctuations in the longitudinal energy profile, suggesting multiple particles are overlapping within an event. These types of clusters are considered "non-Trackable". A metric is formed by dividing the total energy of the non-Trackable clusters within the electron candidate cone by the total energy in the electron candidate cone. As seen in Figure 4.15, a fraction value of $\frac{E_{reco,non-MIP}}{E_{reco,total}} > 0.4$ is chosen in an (unsuccessful) attempt to separate out any non-MIP like events.

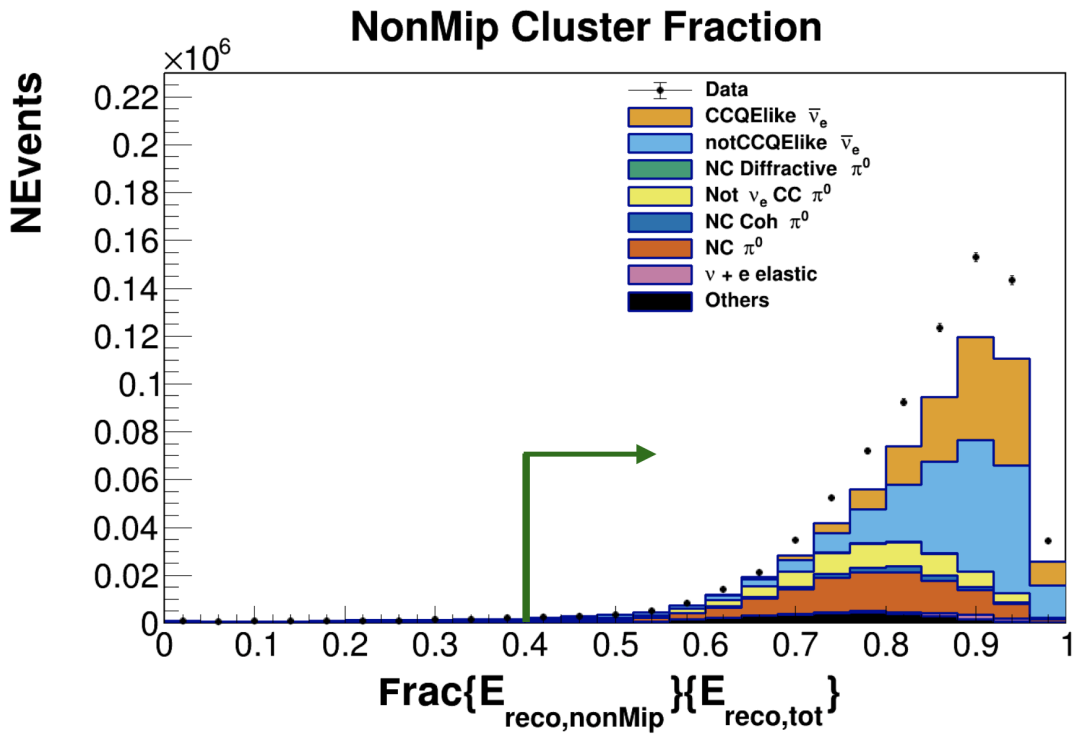


Figure 4.15: Distribution of Non-MIP-cluster fraction.

4.4 Event-level Cleanup

Once the cone contains a signal candidate event as determined by the PID value, the cone becomes an electron candidate. Events that meet this criterion are then subjected to further cleaning to aid in the better reconstruction of the event kinematics. Any events outside of a [-20,+35] ns window with respect to the electron candidate's time are eliminated. Events are eliminated if non-electron candidate tracks exit the inner detector or penetrate the HCAL. The cone object vertex is required to be within the fiducial volume. Additionally, PMT afterpulsing is removed using a metric called *first-fire fraction*. Afterpulsing in the detector occurs from positively charged ions in the PMT traveling in the wrong direction causing a

delayed output signal. Afterpulsing is not in the MINERvA simulation chain and must be removed from a first-fire fraction determined by examining the distribution of the fraction of signal events created by afterpulsing from hand scanned event displays [79]. The fraction is determined from the Digits belonging in the cone and how many stem from the same channel within the same gate. To remove fake signal events that arise from the afterpulsing the optimized cut requirement becomes $f_{first-fire} < 0.25$.

There are also containment requirements to ensure that the electron candidate is contained within the detector. The ratio of electron candidate energy in the downstream HCAL to the downstream ECAL must be less than 0.2. Also, the ratio of electron candidate energy in the OD to Side ECal is required to be less than 0.1. It is required that no more than 6 reconstructed tracks originate from the electron candidate track starting point. This is because the granularity of MINERvA limits the capability of multiple tracking, typically resulting in an algorithm failure to set a high multiplicity vertex. As a result, the reconstructed tracks are unreliable and these events are eliminated. This concludes the evaluation required to construct the final cone object around EM-like particles.

4.5 Leakage Study

Once the cone object is optimized to capture EM-like particles, some corrections must be made for energy deposits that have not made it into the cone object. For CCQE-like events, where the dominant reaction is $\bar{\nu}_e p \rightarrow e^+ n$, we expect very little visible energy not to be contained inside the cone; however, energy leakage (energy clusters that are deposited as a result of a true electron but find themselves outside of the cone) does occur and tends to

increase with increasing shower energy. A correction for the energy leakage outside the cone is measured using an electron particle cannon simulation. In the case of this particle cannon study, there were 10^6 events with a flat initial momentum of 0-10 GeV in the central tracking region of MINERvA with angles up to 30°. By definition, all energy from an electron particle cannon should be found inside the cone object. The correction to the energy leakage can be calculated from the energy outside of the cone divided by the true energy of the electron, as seen in Equation 4.6. The leakage energy correction is applied to correct the electron energy and the hadronic recoil energy.

$$E^{corr} = E_{outside}/E_{inside} \tag{4.6}$$

$$E^{corr}_{leak} = E_e * E^{corr} \tag{4.7}$$

$$E'_e = E_e + E^{corr}_{leak} \tag{4.8}$$

$$E'_{recoil} = E_{recoil} - E^{corr}_{leak} \tag{4.9}$$

where $E_{outside}$ refers to energy outside of the electron cone. Figure 4.16 shows the energy correction (E^{corr}) calculated from the particle cannon is relatively small, only 0.8% on average. Figure 4.17 shows the difference of the leakage energy $(E_{outside})$ and the corrected leakage energy value (E^{corr}_{leak}) to test how well the leakage energy correction represents the leakage energy. The mean of the distribution after the correction is nearly centered at 0, indicating the correction is properly calculated.

A similar study was also done in bins of electron energy and theta to see if the correction should be applied on a bin-by-bin basis. Figure 4.18 and 4.19 show the output of these studies with the conclusion that there is not much variance when breaking down the

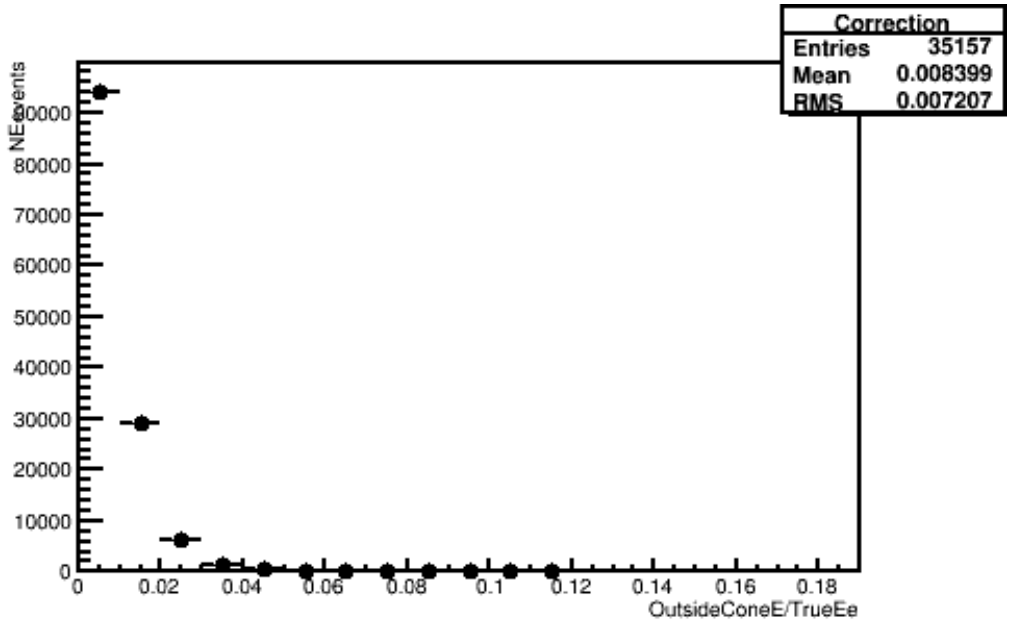


Figure 4.16: Calculation of leakage correction from electron particle cannon

correction into bins. A study was also done in bins of transverse momentum, p_T, where the leakage varies with some correlated combination of energy and angle. Figure 4.20 shows the results of the p_T study with the same conclusion. These studies show that no refinement of the correction is needed as a function of electron energy or direction.

4.6 Separation of Events

The following section describes distributions that are optimized to further purify the electron candidate sample at low reconstructed electron candidate energies. The primary motivation behind the sample requirements is to isolate the signature of an electron or positron (both $e^{\pm}$ are included because they are indistinguishable by the MINERvA detector), indicating an $\bar{\nu}_e$ (ν_e) interaction. A final set of requirements is carried out to purify the Signal sample

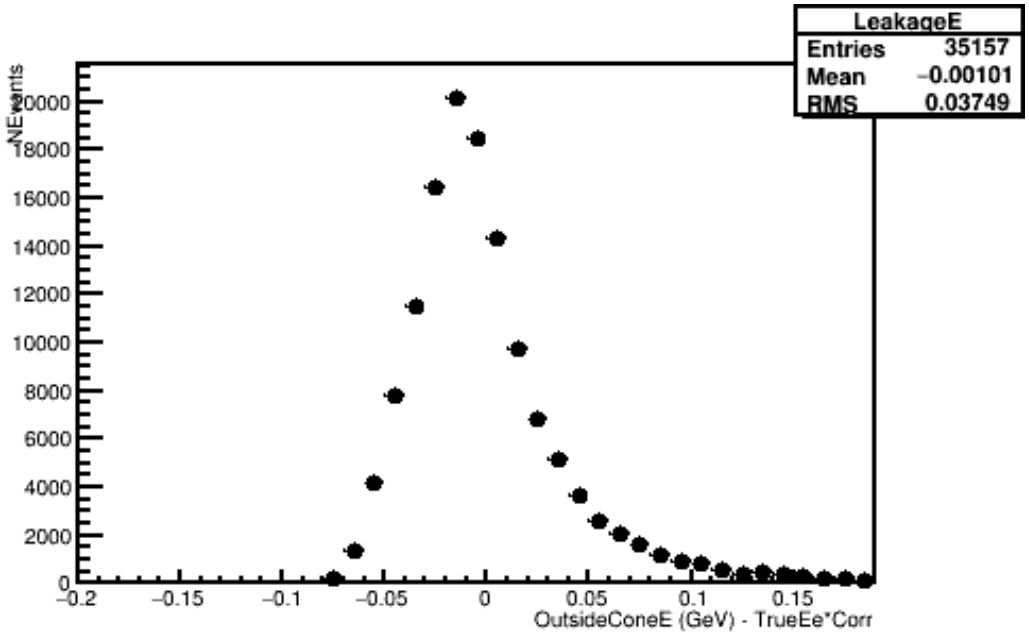
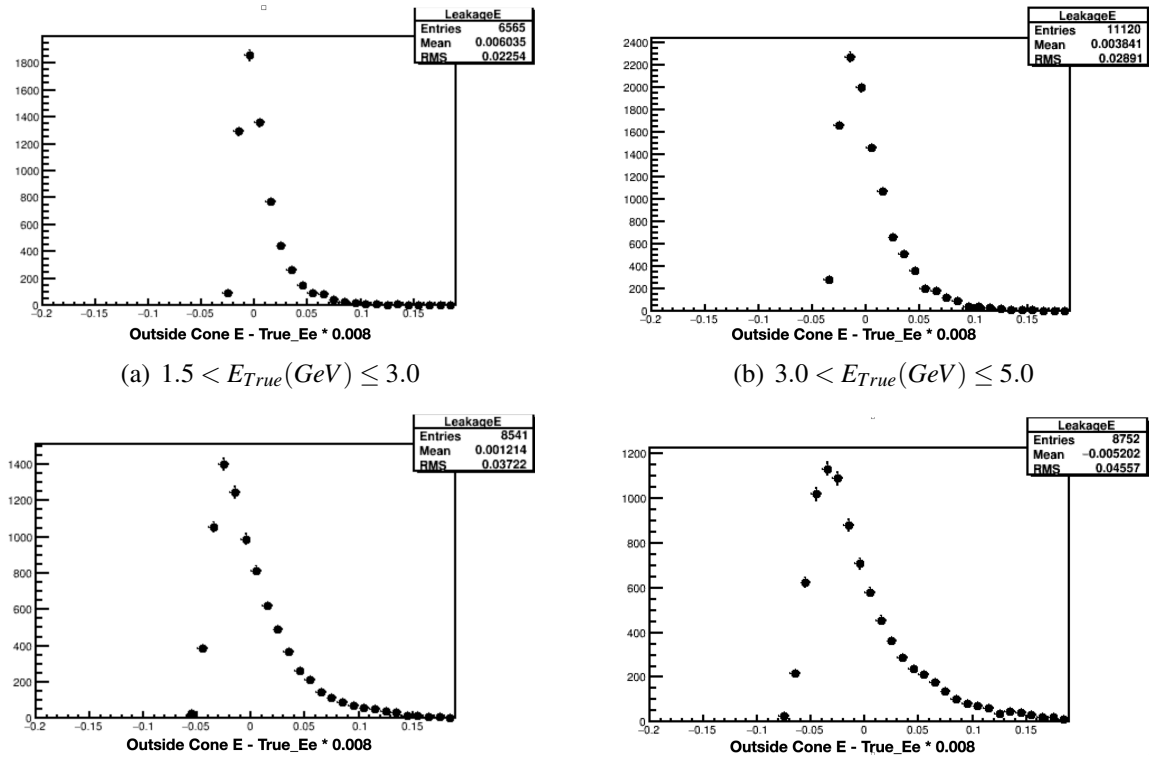


Figure 4.17: Difference of leakage energy and corrected leakage energy from electron particle cannon

from photons that mimic an electron signal and from particles that make it into the electron candidate cone that are not in fact electrons.

4.6.1 Electron-Photon Discrimination: dE/dx

Photons are a large source of background for this analysis. As mentioned previously, photons will interact via pair production in the detector. As the photon interacts with the strong electric field around a nucleus, it is transformed into an electron and positron pair. The electron and positron pair are boosted in the forward direction, causing them to appear as if they overlap in the detector. The electron-positron pair will then reinteract in the detector and an electromagnetic shower begins. The issue with pair production is that



(a) $1.5 < E_{True}(GeV) \leq 3.0$

(b) $3.0 < E_{True}(GeV) \leq 5.0$

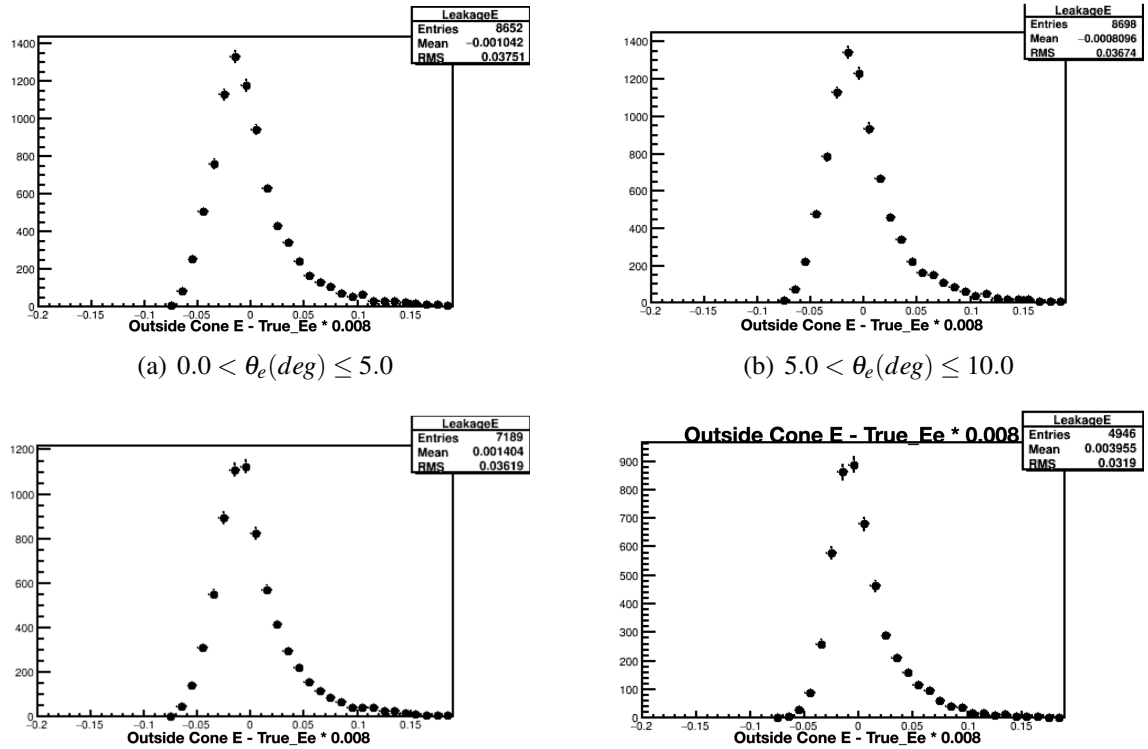
(c) $5.0 < E_{True}(GeV) \leq 7.0$

(d) $7.0 < E_{True}(GeV) \leq 10.0$

Figure 4.18: Difference of cone leakage and corrected leakage in bins of true electron energy

since the electron-positron pair typically appear to be overlapping, it becomes challenging to distinguish between a single electron and an overlapping electron and positron in a reconstructed event. In both cases, the algorithms will reconstruct the event to look like a single track that ends in an electromagnetic cascade, akin to the Signal definition. The best metric to separate the two is to look at the energy loss per unit path length or dE/dx.

The energy profile of the electron-positron pair will have about twice the amount of energy that a single particle would. However, a separation between an overlapping electron positron pair must be made from the event of a true electron with the addition of some

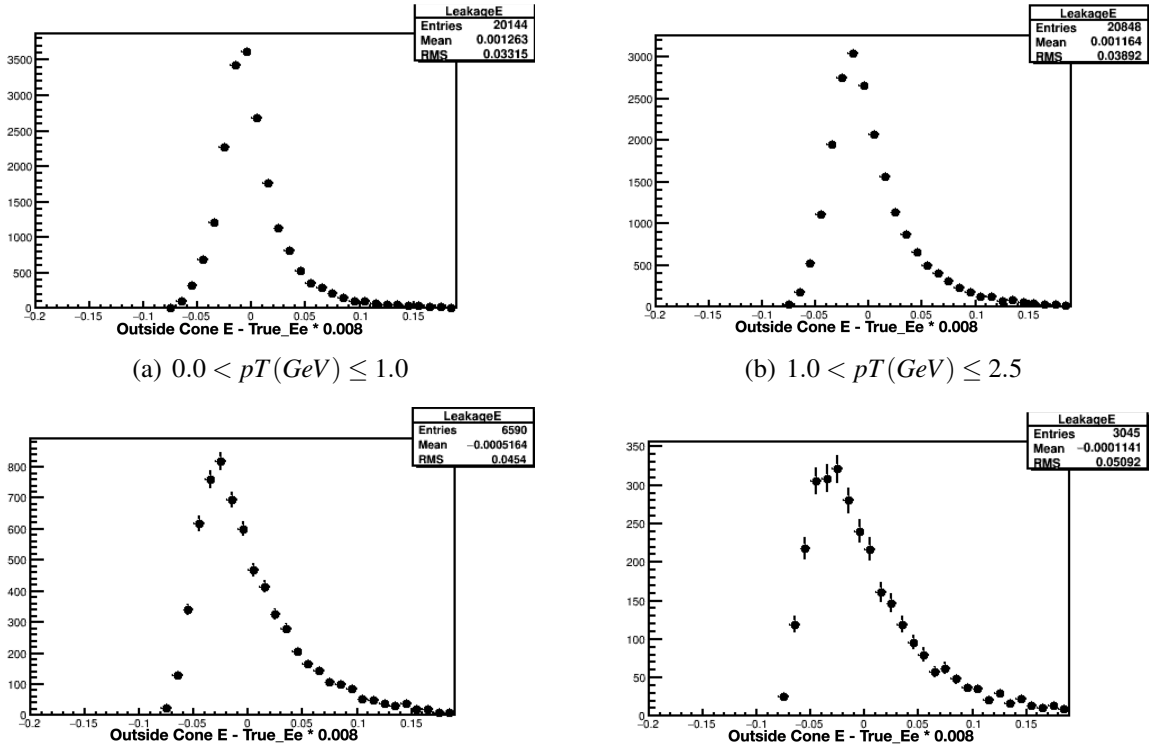


(a) $0.0 < \theta_e(deg) \leq 5.0$ (b) $5.0 < \theta_e(deg) \leq 10.0$

(c) $10.0 < \theta_e(deg) \leq 15.0$ (d) $15.0 < \theta_e(deg) \leq 20.0$

Figure 4.19: Difference of cone leakage and corrected leakage in bins of true electron theta

other particle coming out of the nucleus and overlapping with the electron. In an attempt to distinguish between these cases, the dE/dx is measured in a sliding region, searching for a minimum that should have the least disruption from other particles near the vertex or from the multiplying shower far from the vertex. First, the clusters inside of the electron cone candidate are projected onto the cone axis. Next, the dE/dx is computed in a sliding window of 100 mm blocks which are moved in 25 mm increments from the vertex position down the cone axis until 500 mm (the radiation length in the Tracker material is about 400 mm) is reached or until the end of the Tracker region, as shown in Figure 4.21. The energy



(a) $0.0 < pT(GeV) \leq 1.0$

(b) $1.0 < pT(GeV) \leq 2.5$

(c) $2.5 < pT(GeV) \leq 3.5$

(d) $3.5 < pT(GeV) \leq 5.0$

Figure 4.20: Difference of cone leakage and corrected leakage in bins of transverse momentum

inside the 100 mm block is saved for each 25 mm increment down the electron candidate axis, and the minimum dE/dx value among the 100 mm blocks is chosen as the classifier. Choosing the minimum dE/dx allows for a way to overcome any nuclear activity near the cone vertex while still examining the tail end of the showers towards the end of the cone. Minimum dE/dx is used to separate events with an electron Signal from those of photons. A particle cannon simulation is used once again to determine the optimal requirement. An electron and photon particle cannon with a flat angle up to 30 degrees and initial flat energy between 0-10 GeV is simulated to determine an optimal value for dE/dx. As seen in Figure

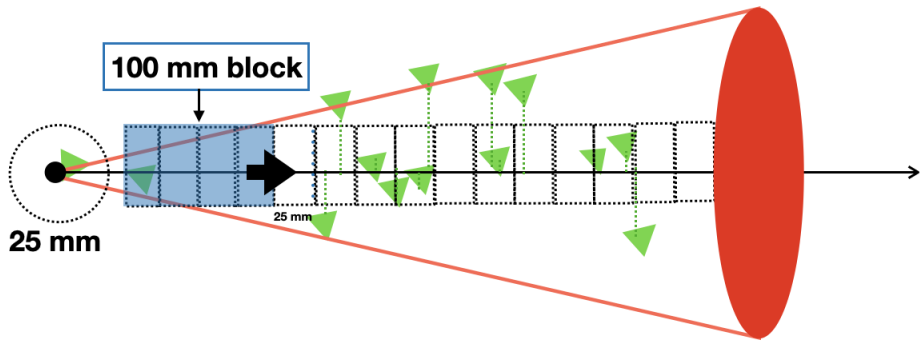


Figure 4.21: Diagram of the dE/dx calculation. The clusters inside of the electron cone candidate are projected onto the cone axis. The dE/dx is calculated inside of a 100 mm block sliding in 25 mm increments. There is a 25 mm exclusion radius around the vertex.

4.22, the optimal requirement for an electron candidate Signal becomes a dE/dx $<$ 2.4 MeV/cm, while (as discussed later) the sideband region of the analysis is constrained by dE/dx $>$ 2.4 MeV/cm.

In Figure 4.23 it is apparent that there is significant disagreement between simulation and data, which peaks at around 3 MeV, indicating that the simulation is not predicting enough of a photon-like background. There is also an additional discrepancy at low dE/dx values, which is assumed to be attributed to signal mis-modeling, though, even if a correction is made to account for the discrepancy at low dE/dx there would still be a large data and MC disagreement at high dE/dx. Additionally, the processes that make up the region of dE/dx $>$ 2.4 MeV/cm also contribute to the signal region. Therefore, it is concluded that this excess of data events must be well understood. The data and MC disagreement is discussed in depth in Chapter 5. A summary of the selection cuts can be found in Table 4.2.

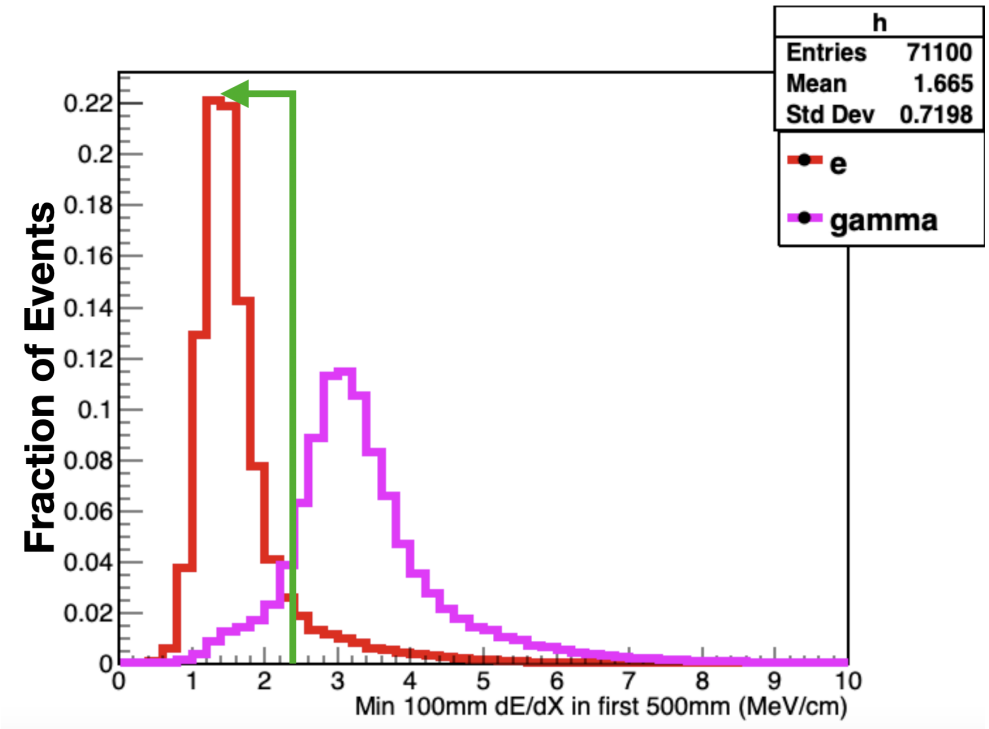


Figure 4.22: Optimization of the minimum dE/dx (MeV/cm) requirement used to isolate signal electron events. e is the electron particle cannon, gamma is for the photon particle cannon

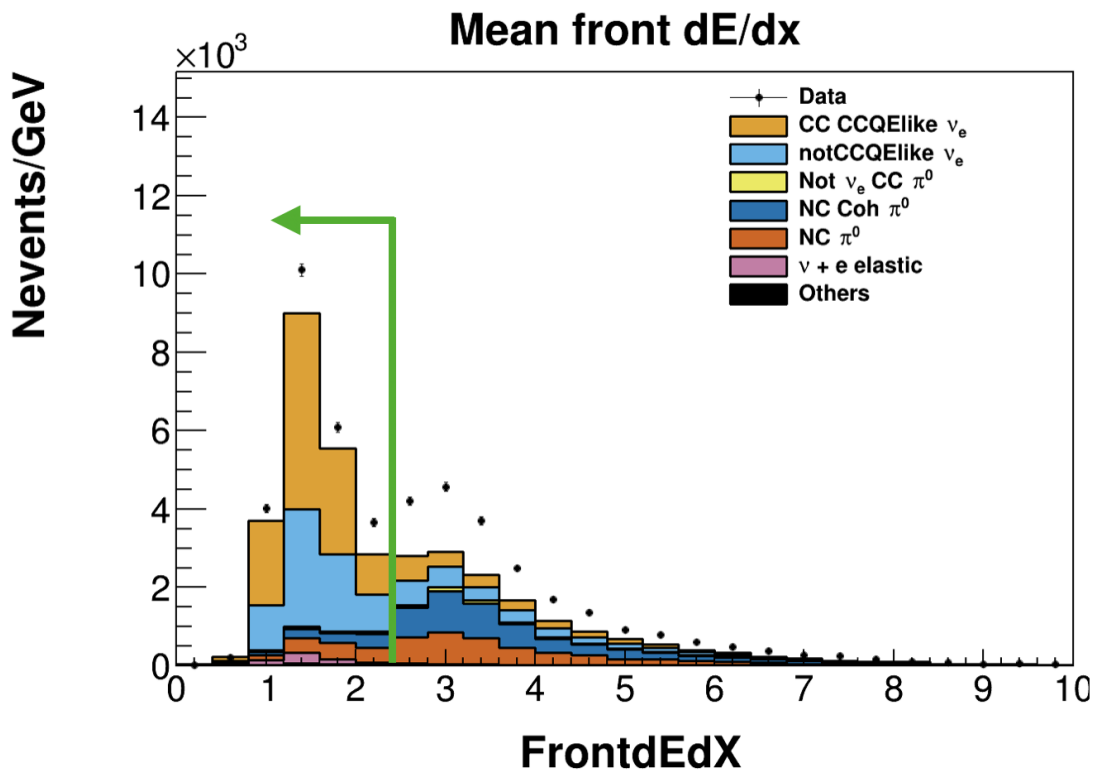


Figure 4.23: Minimum dE/dx distribution (MeV/cm), showing the signal optimization

Vertex Requirements	Vertex and track in fiducial region $<$ 6 reconstructed tracks originating from electron track start point
Muon Elimination	No MIP-like tracks exiting the detector
Visible Energy in Calorimeters	$HCALE_{vis}/ECALE_{vis} \leq 0.2$ $ODE_{vis}/SideECALE_{vis} \leq 0.1$
Electron Object Selection	KNN PID $>$ 0.7
Electron Quality Selection	First-fire fractions $\geq$ 0.25 Energy fraction of non-MIP cluster to total $>$ 0.4
Rejection of photons	Minimum 100mm dE/dX over first 500mm $<$ 2.4 MeV/cm
Kinematic Selection	$E_e > 2.5$ GeV

Table 4.2: Cuts and requirements used to form base sample used in the analysis

5. Charged Current Low Recoil $\bar{\nu}_e$

This chapter explains the methods to remove the remaining background events from the signal sample. The categorization of the event sample used in this section is found in Table 5.1, with the $\bar{\nu}_e$ CCQE-like and $\bar{\nu}_e$ not CCQE-like categories equating to the desired signal definition.

5.1 Diffractive and Coherent Pion Production

Up to this point, the remaining background processes that require discussion are coherent and diffractive pion production. Coherent pion production from neutrino-nucleus interactions is an inelastic process where a lepton and a pion are produced in the forward direction while leaving the struck nucleus in the ground state. The signature of this process in MINERvA would appear as two forward tracks originating from a single vertex with no additional energy deposition near the vertex. The charged current (CC) coherent processes are written as

$$\nu_l + A \rightarrow l^- + \pi^+ + A \qquad \bar{\nu}_l + A \rightarrow l^+ + \pi^- + A \tag{5.1}$$

while the neutral current (NC) processes are written as

$$\nu_l + A \rightarrow \nu_l + \pi^0 + A \qquad\qquad \bar{\nu}_l + A \rightarrow \bar{\nu}_l + \pi^0 + A \tag{5.2}$$

The square of the four-momentum transfer, $|t|$, must be small to preserve the initial state of the nucleus. In the instance of a ν_μ CC coherent π^0 interaction, a muon will leave a minimum ionizing track in MINERvA as it makes its way from the interaction vertex out the back of the detector and into MINOS. The pion will also leave an ionizing track and will eventually stop or interact hadronically within MINERvA. For an ν_μ NC Coherent π^0 interaction, the only signature of the interaction is $\pi^0 \rightarrow \gamma\gamma$, since the nucleus does not receive much energy and the final state neutrino typically leaves without reinteracting. Therefore, a coherent pion interaction defined as a forward-going lepton and pion with an intact nucleus would be distinguishable with the tell-tale signature of the lepton, pion, and a small $|t|$ with the absence of additional particles.

For CC coherent interactions the calculation of $|t|$ comes from the measured four-momenta of the charged lepton and pion. Assuming the nucleus is at rest, zero energy is transferred to the nucleus and, by conservation of four-momentum, $|t|$ is defined as

$$|t| = |(p_{\nu_l} - p_l - p_\pi)^2| \approx (\sum_{i=\mu,\pi} E_i - \vec{p}_{i,L})^2 + |\sum_{i=\mu,\pi} \vec{p}_{i,T}|^2, \tag{5.3}$$

for p_{ν_l}, p_l and p_π the four-momenta of the neutrino, charged lepton, and pion respectively and $\vec{p_T}$ and $\vec{p_L}$ are the transverse and longitudinal momenta with respect to the incoming neutrino direction. By assuming the incoming neutrino direction as parallel to the direction of the neutrino beam, the neutrino four momentum can be calculated from the measured

charged lepton and pion energies,

$$E_\nu = E_l + E_\pi \tag{5.4}$$

Since MINERvA's tracker material is a CH-based hydrocarbon, there is a possibility a neutrino interaction will occur on a free proton, referred to as diffractive pion production. Diffractive pion production, also known as diffractive scattering, produces a forward going pion and lepton while leaving the proton in the ground state. CC diffractive processes are defined by

$$\nu_l + p \rightarrow l^- + p + \pi^+ \qquad \bar{\nu}_l + p \rightarrow l^+ + p + \pi^- \tag{5.5}$$

and NC diffractive processes defined by

$$\nu_l + p \rightarrow \nu_l + \pi^0 + p \qquad \bar{\nu}_l + p \rightarrow \bar{\nu}_l + \pi^0 + p \tag{5.6}$$

The process is similar to coherent pion production with the production of a lepton and pion in the forward direction. While coherent pion production occurs on a nucleus, diffractive pion production occurs on a single proton, leaving the proton intact. However, since the proton contains much less mass than a carbon nucleus, the proton recoils from the energy imparted to the neutrino. The energy of the recoil is large enough to be detected in the MINERvA detector. By conservation of momentum, the recoiling proton deposits its energy upstream of the interaction vertex.

In the MINERvA detector, diffractive pion production is indistinguishable from coherent

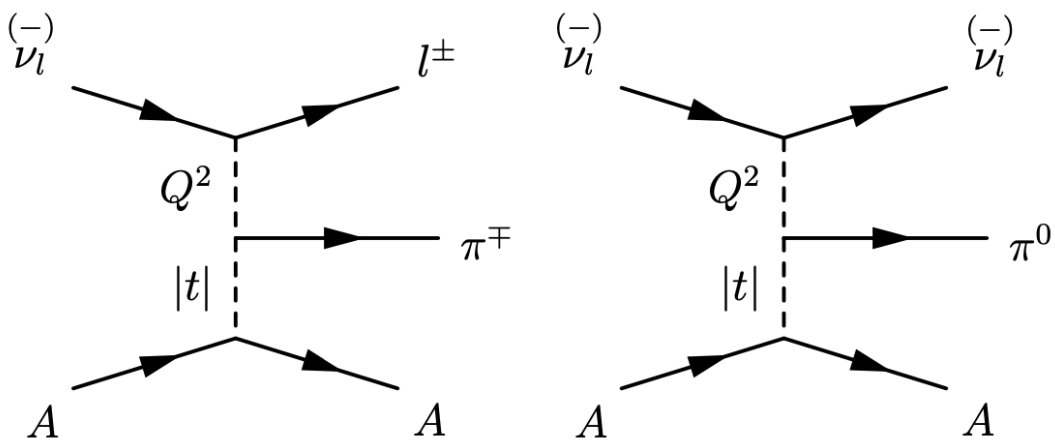


Figure 5.1: Feynman diagrams for CC coherent pion production (left) and NC coherent pion production (right)

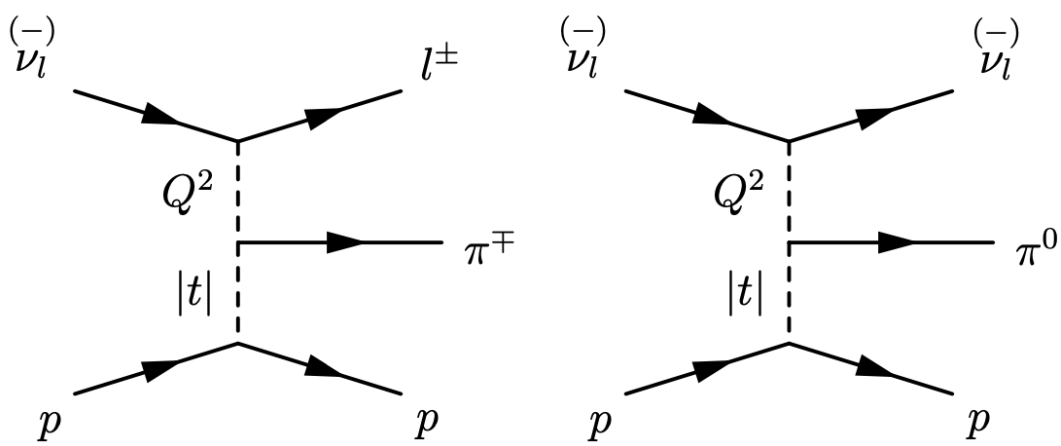


Figure 5.2: Feynman diagrams for CC diffractive pion production (left) and NC diffractive pion production (right)

pion production when the recoiling proton is below the detection threshold. A possible way to find a separation between coherent and diffractive scattering is from the $|t|$ dependence of the cross section. The diffractive cross section is extrapolated from the coherent cross section and the coherent cross section is modeled using the Rein-Sehgal Partially Conserved Axial-Vector Current (PCAC) Coherent model [83]. As an overview, the PCAC coherent model uses the $|t|$-dependence derived from the pion-proton/nucleus elastic scattering cross section which is predicted to fall exponentially with respect to $|t|$ as $\exp(-b|t|))$ where b is defined as

$$b = \frac{1}{3} R_0^2 A^{\frac{2}{3}} \tag{5.7}$$

for A the number of nucleons in the target and $R_0 \approx 1$fm the nuclear length scale. The predicted slope for coherent pion production on carbon (A=12) is calculated to be ≈ 40 $(\frac{GeV}{c})^{-2}$ whereas the predicted slope for diffractive scattering (A=1) is $\approx 8(\frac{GeV}{c})^{-2}$. By comparison, the diffractive cross section is expected to fall more slowly with $|t|$ compared to the coherent cross section. The relationship between the kinetic energy of the recoiling proton and the squared four momentum transfer with the target proton is written as

$$|t|_{diff} = |(p_{\nu_l} - p_l - p_\pi)^2| = |(\vec{p}_{p,f} - \vec{p}_{p,i})^2| = 2m_p T_p, \tag{5.8}$$

for $\vec{p}_{p,i}$ the initial state proton four momentum vector and $\vec{p}_{p,f}$ the final state proton four momentum vector.

In summary, diffractive processes in MINERvA can occur because the experiment's tracker material is composed of a CH-based hydrocarbon. Due to the reduced mass of the

proton compared to a nucleus, the proton is able to recoil from the neutrino interaction with enough kinetic energy to be seen in the MINERvA detector. This allows for an important distinction between diffractive and coherent NC pion processes, whereas in the case of coherent production, the nucleus is much more massive than a proton and, therefore, absorbs a large amount of the momentum imparted from the neutrino.

There exists a model for an NC Diffractive scattering process in GENIE, from the work of Rein [62], though it has not been thoroughly vetted against data to become a default model for GENIE. However, a special MC sample can be created using the Rein diffractive model which is combined with the standard set of Monte Carlo simulations allowing for a single compilation of all MC models. These special diffractive samples can then be fed into the reconstruction and analysis chain.

5.1.1 Construction of Sidebands

Category label	**Definition**
$\bar{\nu}_e$ **CCQE-like**	Charged current anti-neutrino events with any number of nucleons but no other hadrons and no photons
$\bar{\nu}_e$ **not CCQE-like**	A charged current anti-electron neutrino event with any number of nucleons and any number of other hadrons and photons
NC diffractive π^0	Neutral current diffractive pion production with a π^0 in the final state
Not $\bar{\nu}_e$ **CC** π^0	Charged current event with a π^0 in the final state that is not a ν_e or $\bar{\nu}_e$ event
NC Coh π^0	Neutral current coherent interactions with a π^0 in the final state
NC π^0	Neutral current interactions that are not coherent, nor diffractive, and which produce a π^0 in the final state
$\nu + e$ **elastic scattering**	Elastic scattering of neutrinos from atomic electron
Other	Remaining events in none of the categories above

Table 5.1: Categorical breakdown of selection.

The selection is broken into four components: the Signal region and the NC Diffractive π^0, NC Coherent π^0, and π^0 sidebands. Shown in Figure 5.3 is Front dE/dx distribution with the signal region defined by Front dE/dx < 2.4 MeV/cm to isolate the $\bar{\nu}_e$ CCQE-like and $\bar{\nu}_e$ not CCQE-like selection. After cuts, the remaining signal sample still contains background

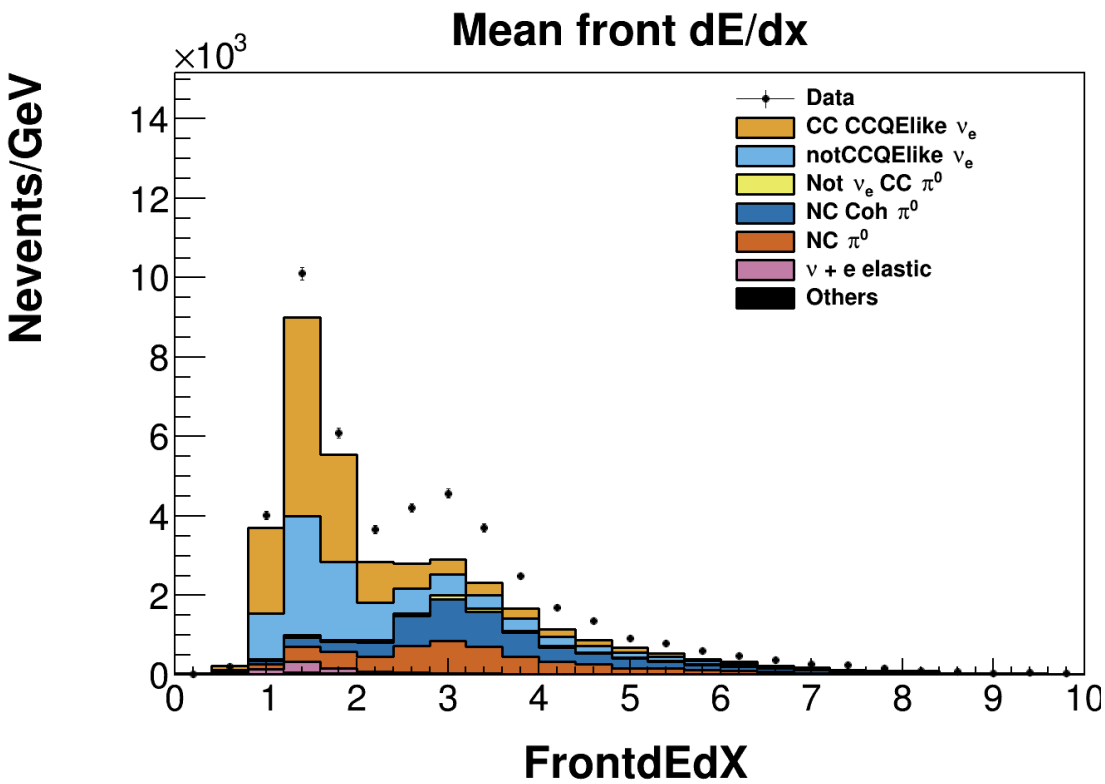


Figure 5.3: Mean dE/dx distribution used to determine sideband definitions

contamination which must be subtracted. The background is subtracted with the help of simulated data and to limit the dependency on the simulation models, the background is tuned. This is done by finding criteria that provide pure samples of background processes. These regions are optimized to provide the best prediction for backgrounds in the signal region and are referred to as sidebands. Once a prediction is formed, the background process

for each respective sideband is extrapolated into the signal region and subtracted from data to measure the desired cross section. The background model prediction's uncertainty is minimized by tuning the MC estimate of the background rates to data in the sideband region as discussed in Section 5.2.2.

To derive a cross section result, background processes must be subtracted out of the signal region before any analysis can be carried out. Sidebands form a constraint on the background processes that leak into the signal region. The sidebands are chosen based on an optimization between purity (the ratio of the number of signal events that pass signal cuts to the total number of events that pass signal cuts) and efficiency (the ratio of the number of signal events that pass signal cuts to the total number of signal events). The intent is to keep as many signal events as possible in the kinematic regions where the cross section is being reported but not at the cost of allowing an abundance of background events to enter. The sidebands are chosen to obtain the optimal division between different background processes in an attempt to get a handle on each individual process to be ultimately subtracted from signal events.

The background processes can be divided into two major categories. The first is when a π^0 is the only particle produced in the neutrino interaction (NC coherent pion production and NC diffractive pion production). The second is when photons are produced from a π^0 and additional particles are also produced simultaneously (NC incoherent pion production and non-electron neutrino CC pion production). It is expected that a π^0 inside of the electron candidate cone will have a larger amount of energy that ends up leaking outside the cone compared to an actual electron. In case one (or two) photons are inside the cone, a discriminator is needed to separate photons from electrons.

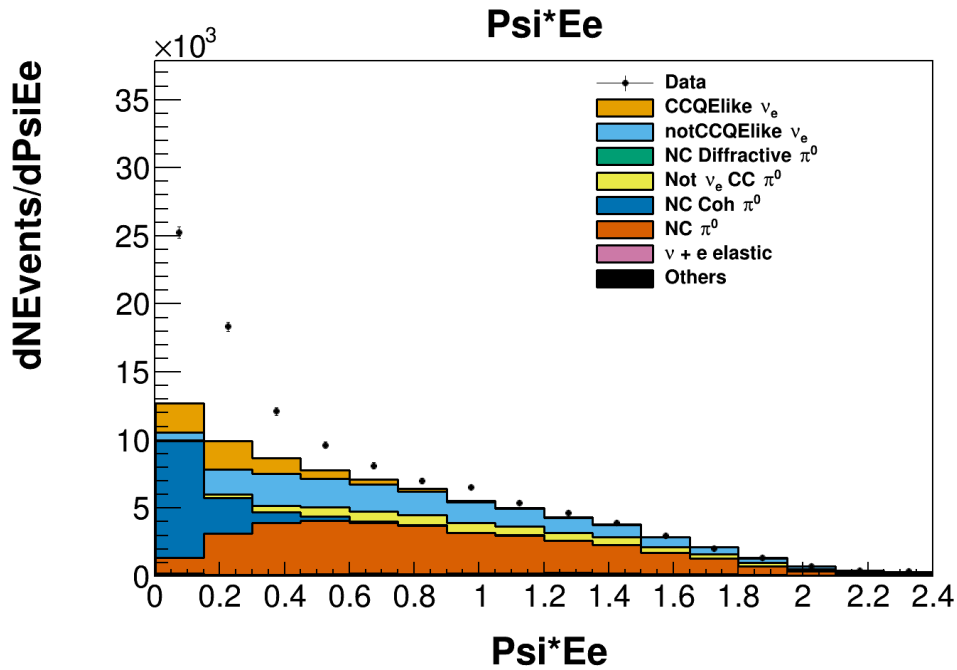


Figure 5.4: $\psi * E_e$ for events of dE/dx $>$ 2.4 MeV/cm

For this analysis, the sidebands are determined by looking at two key distributions: Mean Front dE/dx and $\psi * E_e$. Mean Front dE/dx is the best discriminator to separate photons from electrons. $\psi * E_e \equiv E_{other}$ is defined as any energy not associated with the electron. This variable is labeled $\psi * E_e$ and not as E_{other} for historical reasons beginning with the adaptation of definitions from the low energy CCQE-like electron neutrino analysis [79]. All background processes are first constrained by requiring a cut of Front dE/dx $>$ 2.4 MeV/cm. As seen in Figure 5.3, a cut of Front dE/dx $>$ 2.4 MeV/cm isolates a large portion of background processes and simultaneously isolates a region in which there is a large discrepancy between data and MC. Termed the *Excess* region, due to the large excess of data events compared to MC events, further studies must be conducted to determine how to constrain this unknown population. By looking at the $\psi * E_e$ distribution for values of

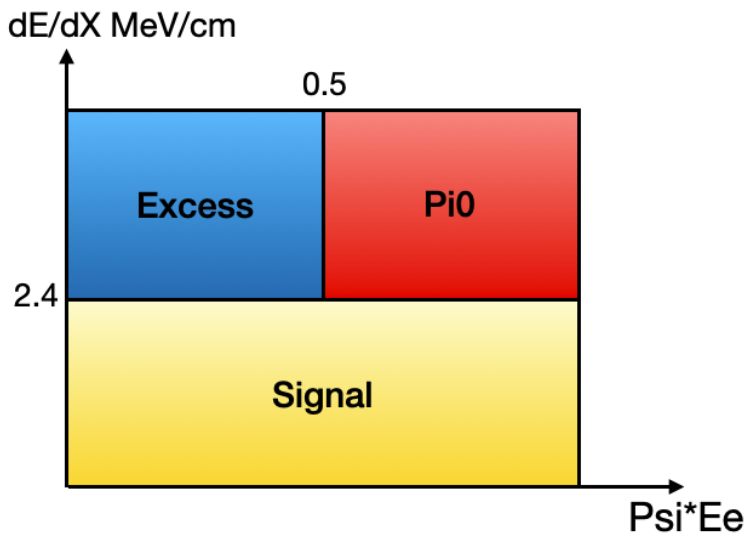


Figure 5.5: Representation of signal and sideband regions with respective cuts.

Front dE/dx $>$ 2.4 MeV/cm as seen in Figure 5.4, it becomes apparent that a cut of $\psi * E_e <$ 0.5 GeV is required to isolate the excess region.

The regions discussed so far can be summarized as the signal region (dE/dx $<$ 2.4 MeV/cm), the excess sideband (dE/dx $>$ 2.4 MeV/cm, $\psi * E_e <$ 0.5 GeV), and the π^0 sideband (dE/dx $>$ 2.4 MeV/cm, $\psi * E_e >$ 0.5 GeV). Figure 5.5 shows a visual representation of these regions. The remainder of this chapter will provide further details about each sideband and the method used to determine how the excess of data events is best described. Section 5.2.2 describes how the excess and pion regions are constrained through a fitting technique in preparation for their extrapolation into the signal region.

π^0 Sideband

The π^0 sideband is defined by dE/dx $>$ 2.4 MeV/cm and $\psi * E_e >$ 0.5 GeV. Figures 5.7 and 5.8 show the π^0 sideband in the E_{avail} vs p_T and E_{avail} vs q_3 regions of phase space. It

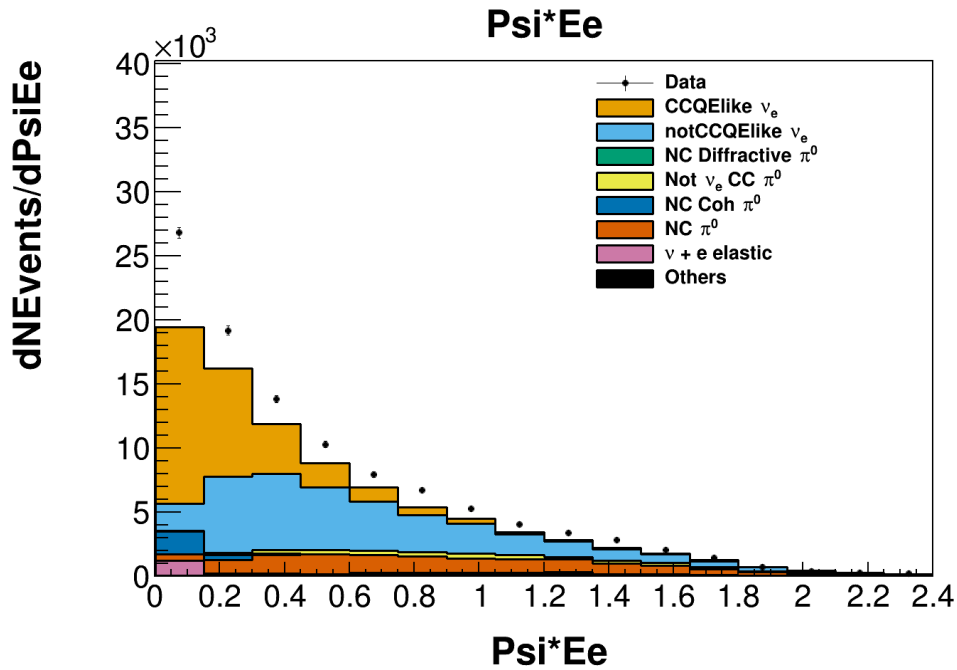


Figure 5.6: $\psi * E_e$ distribution in the Signal region of Front dE/dx $<$ 2.4 MeV/cm

is important to note that while there is a large population of NC incoherent π^0 and CC π^0 events in this region by construction, there are a non-negligible amount of signal events as well. Additionally, in the lowest q_3 bins in Figure 5.8, there is a population of NC coherent π^0 events. This indicates that during the background fitting stage, a global fit will be required to predict signal and background events in each of the sideband regions.

5.1.2 Excess Sideband

As seen in Figure 5.3, in the region of high dE/dx, dE/dx $>$ 2.4 MeV/cm and $\psi * E_e < 0.5$ GeV, there is noticeable tension between data and MC. The excess was first noted in the MINERvA low energy era [79] and many of the studies used in characterizing the excess of

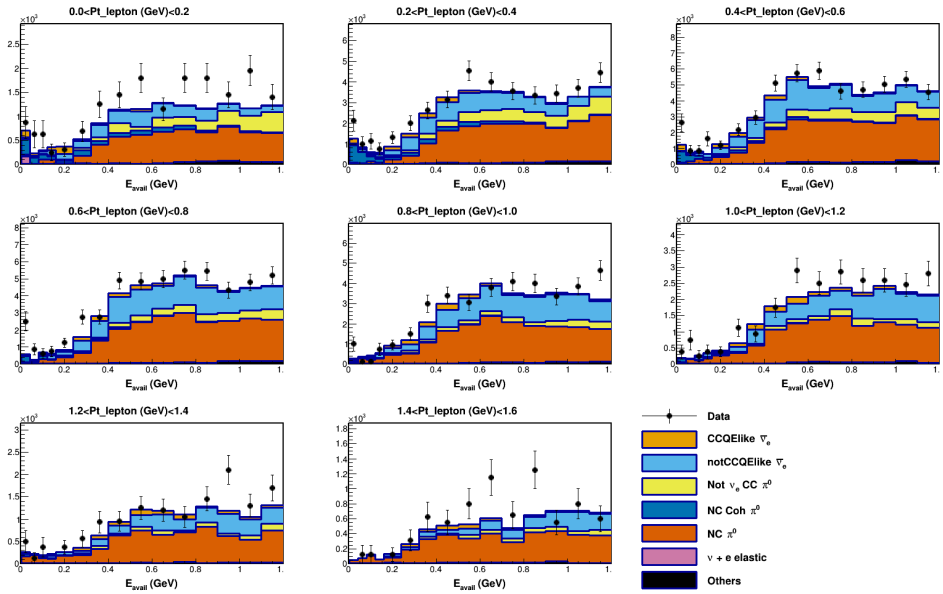


Figure 5.7: E_{avail} vs p_T distribution for events in the π^0 sideband defined by dE/dx $>$ 2.4 MeV/cm and $\psi * E_e > 0.5$ GeV.

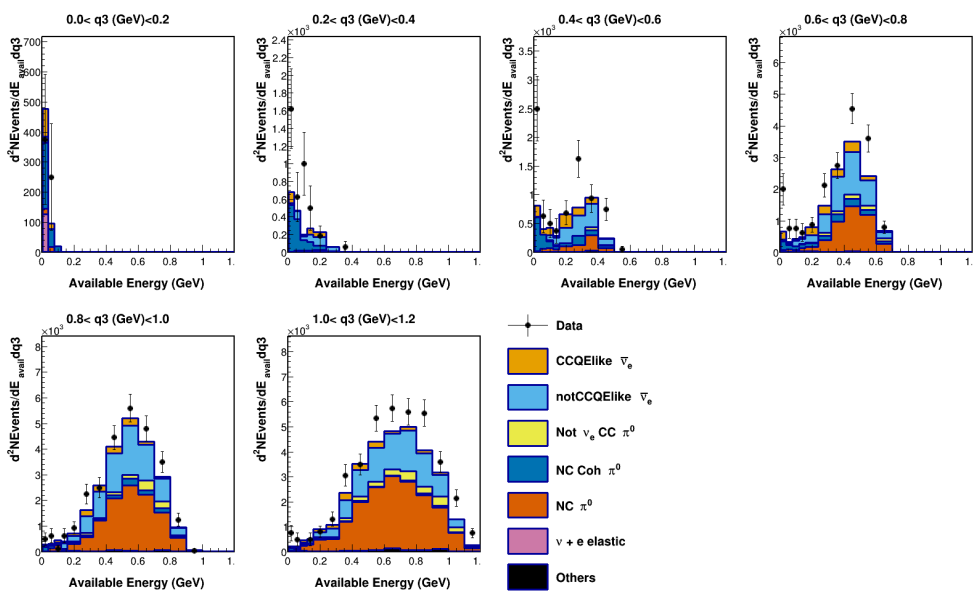


Figure 5.8: E_{avail} vs q_3 distribution for events in the π^0 sideband defined by dE/dx $>$ 2.4 MeV/cm and $\psi * E_e > 0.5$ GeV.

data are also used in this analysis. It must be determined whether the tension between data and MC stems from some under-prediction of π^0 events or whether there exists an entire process GENIE does not simulate. The first step is to evaluate the shape of the excess to try and gain insight into which process could explain the excess. The "shape" of the excess refers to the subtraction of MC from data, allowing for a direct comparison against *ad hoc* models later discussed in Section 5.1.2.

Characterization of the Excess

First, the reconstructed vertex of the electron candidate cone and the shower direction are evaluated to rule out any algorithm failures or any external particles unknowingly entering the fiducial volume. If a neutrino interacted outside of the defined fiducial volume, then the z position of the vertex should be reconstructed further upstream of the detector, either near the front end of the fiducial volume or near the outer edges of the detector. The reconstructed vertex value of R^2 is defined by

$$R^2 = x^2 + y^2 \tag{5.9}$$

for the x and y positions of the vertex. Both R^2 and z are two quantities in which event vertices should be uniformly distributed, up to acceptance effects. It can be seen in Figures 5.9 and 5.10 that the distributions do not indicate the excess favoring any specific vertex reconstruction position.

Additional metrics of inline upstream energy and extra energy are formed to probe the location of the energy depositions surrounding the electromagnetic shower profile described in Section 4.1.3. Inline upstream energy is defined as the energy inside of a reversed

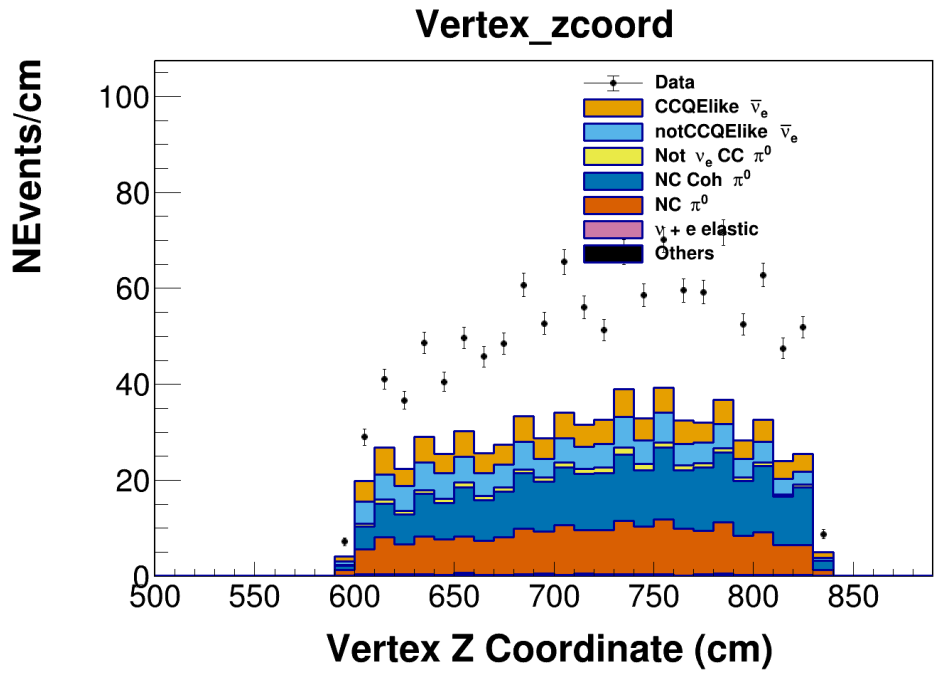


Figure 5.9: Vertex z position for dE/dx $>$ 2.4 MeV/cm and $\psi * E_e < 0.5$ GeV

electron candidate cone as seen in Figure 5.11. The distribution of inline upstream energy is found in Figure 5.12 and it will become a metric of great importance as discussed in Section 5.1.2. In inline upstream energy, about $\frac{2}{3}$ of the excess is seen in the first bin of < 1 MeV. The additional $\frac{1}{3}$ of excess events appear in the remaining bins > 1 MeV. The energy outside the electron candidate cone and the vertex region is called the extra energy, seen in Figure 5.13, where most events with an excess populate the first few bins. From these distributions, it is apparent that the typical excess event has little non-shower activity and the profile of the non-excess energy is consistent with the shower being the only energy [79]. All distributions indicate that a neutral current process is needed to explain the excess events.

To probe the shape of the excess itself, several different final states were simulated using

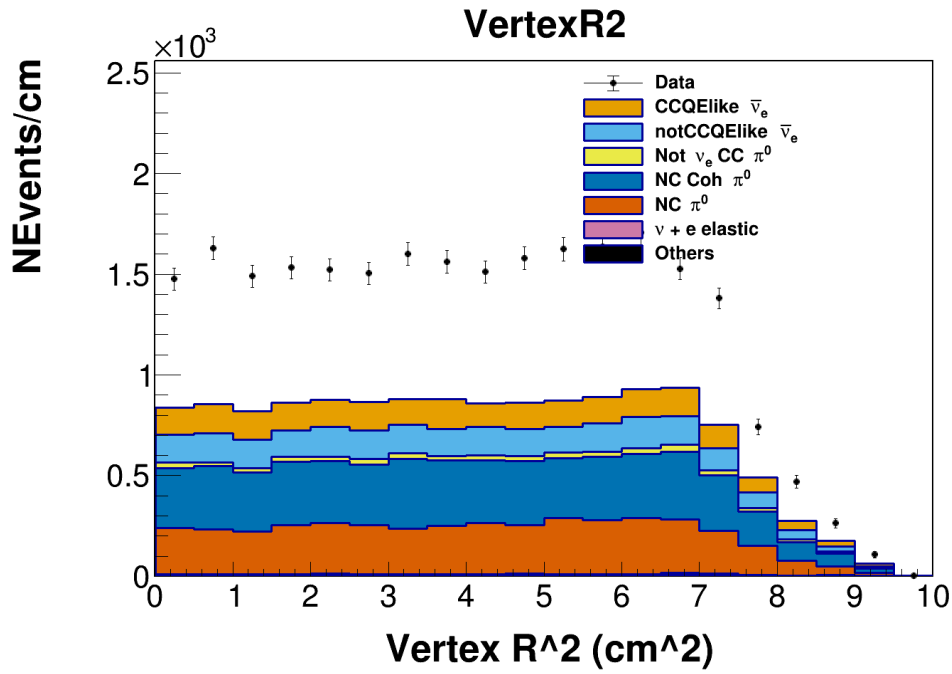


Figure 5.10: Vertex $R^2 = x^2 + y^2$ for dE/dx $>$ 2.4 MeV/cm and $\psi * E_e <$ 0.5 GeV

particle cannons with the approach to treating them as *ad hoc* models created specifically to match the final state kinematics. Since no particular distribution investigated up to this point allowed for differentiation between the characteristics of the excess, the *ad hoc* models chosen encompass both single and multiple photon showers by constructing three-particle cannon simulations. The particle cannon samples simulate 10^6 events of γ, $e^{\pm}$, and π^0 events with a flat distribution of 0-10 GeV and a maximum angle of 30°. These *ad hoc* models are compared against each other to determine whether single-photon or multi-photon showers best simulate the shape of the data in the excess sideband. Additionally, the three models were each weighted in bins of reconstructed electron energy such that the weighted energy matched the energy distribution observed in the data. The *ad hoc* models were then compared in the following two distributions: transverse shower asymmetry and

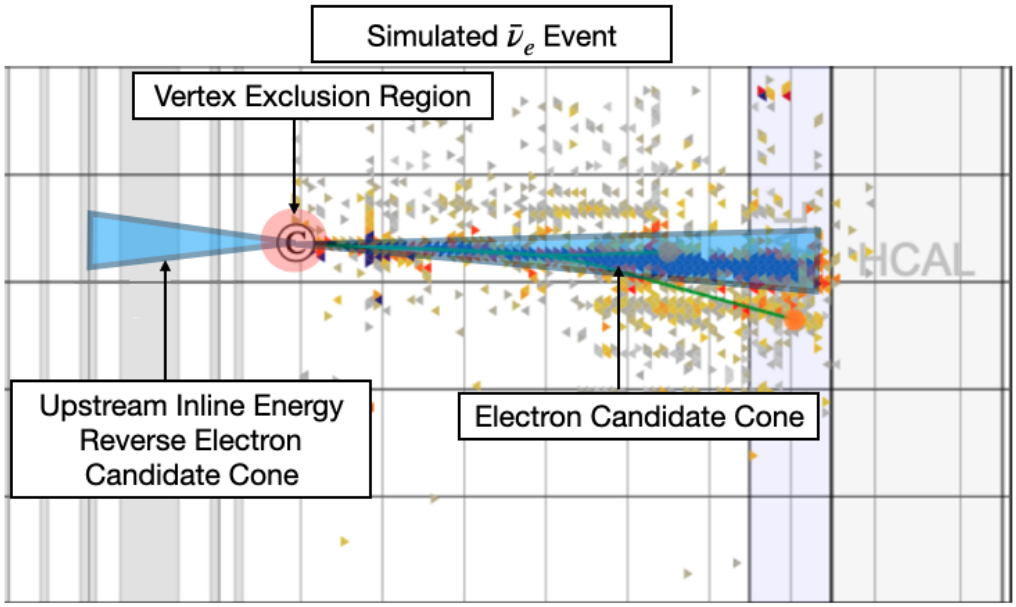


Figure 5.11: Inline Upstream Energy diagram

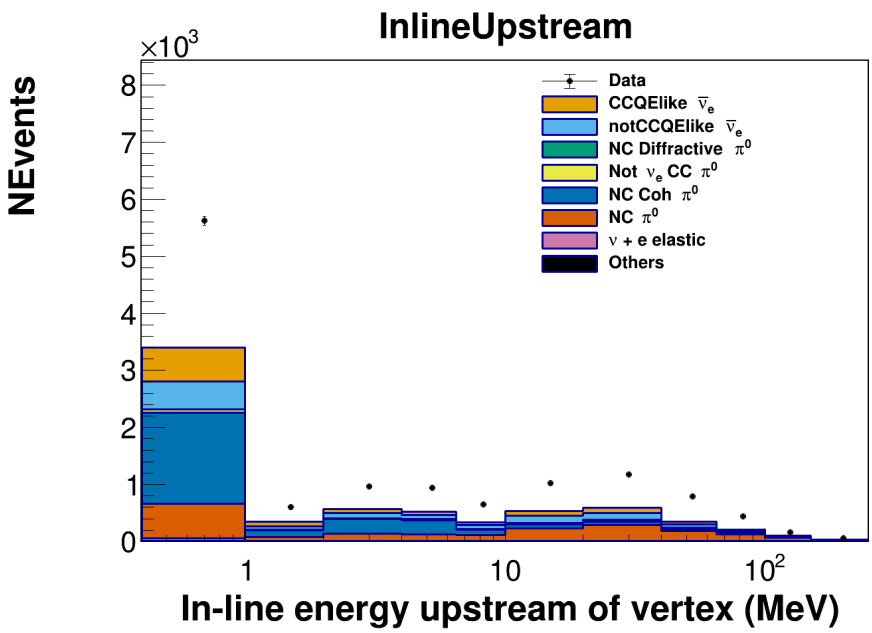


Figure 5.12: Energy inside a 7.5 degree cone pointing upstream.

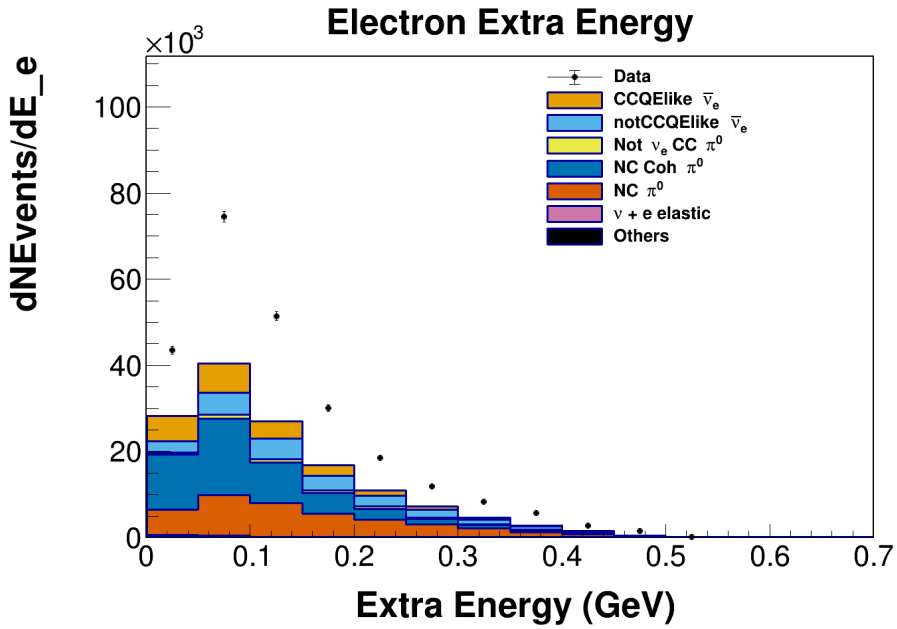


Figure 5.13: Energy outside of electron candidate cone and vertex region

median shower width.

Excess Shape Comparisons

Since electromagnetic cascades spread out transversely to the direction of propagation, there is a range of energy where single-photon showers can be distinguished from multi-photon showers based on transverse size. Median shower width, or median transverse width, provides the extent to which an electromagnetic cascade spreads transversely to its direction of propagation. In Figure 5.14, the excess shape seems to prefer the π^0 *ad hoc* model over the single photon.

The transverse shower asymmetry is formed to probe the transverse energy distribution of single vs. multi-photon showers. Due to the stochastic nature of photon conversion

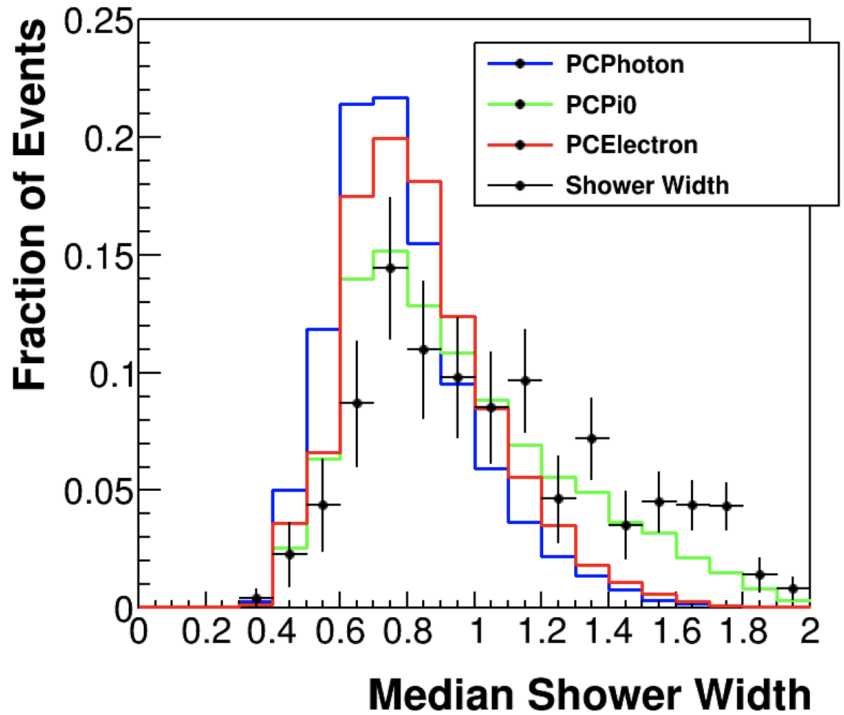


Figure 5.14: Shapes of particle cannon samples in median shower width

into electron-positron pairs for a multi-photon shower, there is some longitudinal distance between each photon shower. The track reconstruction algorithm will effectively pick the first photon that converted, forming a cone around that track. If the angle between the two converted photons is more than a few degrees, then the energy within the cone will be distributed asymmetrically around the cone axis. The transverse shower asymmetry is formed as a reconstructed quantity where the cone axis is projected into each plane for which a scintillator strip is illuminated. The point of intersection between each strip and the projected axis is recorded. The energy weighted mean of the difference between the illuminated strips, x, and the projected intersection point of the axis, x_p, gives an asymmetry

score for the view for i over the strips

$$A_{X,U,V} = \frac{1}{\sum_i E_i} \sum_i (x_i - x_p) E_i \tag{5.10}$$

Summing across the detector $\hat{z}$ coordinate leaves two degrees of freedom which can be represented as an x and y coordinate asymmetry.

$$(\hat{x}, \hat{v}) = \frac{1}{2}(\hat{x} \pm \sqrt{3}\hat{y}) \tag{5.11}$$

$$\hat{x} = \hat{u} + \hat{v} \tag{5.12}$$

$$\hat{y} = \frac{1}{\sqrt{3}}(\hat{u} - \hat{v}) \tag{5.13}$$

However, X planes enter the coordinate twice as frequently as the U+V combination, and so the error weighted mean becomes

$$A_{\hat{x}} = \frac{1}{(\frac{1}{\sqrt{2}})^{-2} + 1^{-2}} (A_X (\frac{1}{\sqrt{2}})^{-2} + (A_U + A_V)(1)^{-2}) = \frac{1}{3}(2A_X + (A_U + A_V)) \tag{5.14}$$

$$A_{\hat{y}} = \frac{1}{\sqrt{3}}(A_U - A_V) \tag{5.15}$$

The x and y coordinate systems are then constructed as

$$A_{\hat{x}} = \frac{1}{3\sum_{allstrips} E_i} (2 \sum_{Xstrips} E_i(x_i - x_p) + (\sum_{Ustrips} E_i(x_i - x_p) + \sum_{Vstrips} E_i(x_i - x_p))) \tag{5.16}$$

$$A_{\hat{y}} = \frac{1}{3\sum_{U,Vstrips} E_i} (\sum_{Ustrips} E_i(x_i - x_p) - \sum_{Vstrips} E_i(x_i - x_p)) \tag{5.17}$$

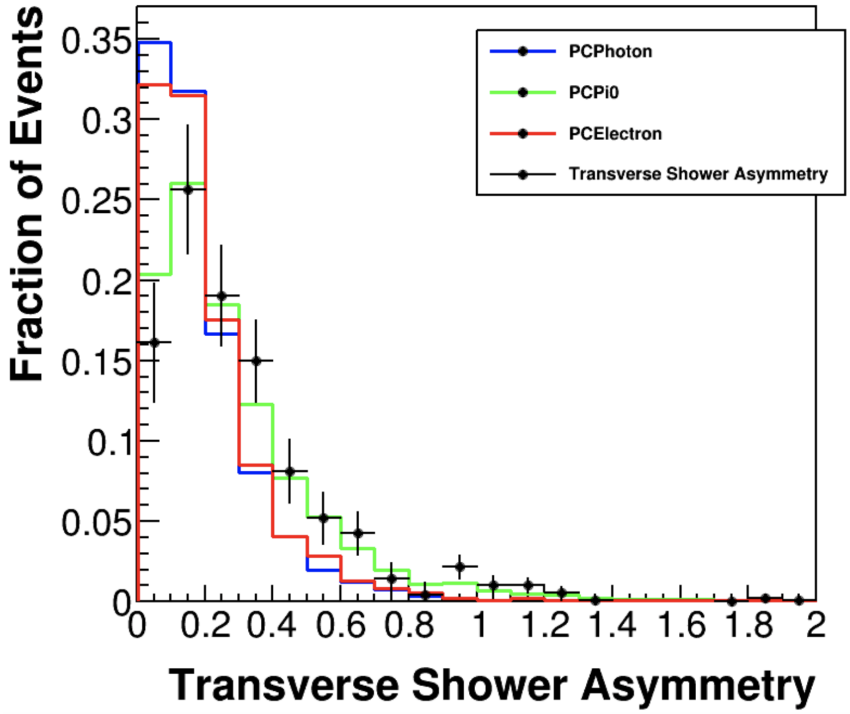


Figure 5.15: Transverse shower asymmetry shape distributions of ad hoc sample compared to the data excess

forming a single metric, referred to as Transverse Shower Asymmetry defined by

$$A = \sqrt{(A_{\hat{x}})^2 + (A_{\hat{y}})^2} \tag{5.18}$$

The transverse shower asymmetry shown in Figure 5.15 shows the shape of the excess as slightly biased towards the π^0 final state compared to the electron and single photon *ad hoc* models.

The distributions in median shower width and transverse shower asymmetry appear to rule out the single photon *ad hoc* hypothesis. This section explores the π^0 hypothesis as a candidate explanation for the excess between data and MC. The following section will

discuss the π^0 *ad hoc* model implementation in the general sample.

Neutral Current π^0 ad hoc model

After evaluating the *ad hoc* models as candidate explanations for the disagreement between data and MC in the region of the excess sideband, it appears that the best candidate to explain the excess is an NC π^0 process. This is a surprising result, as it is expected that GENIE models NC π^0 events to better precision than the (approximately) factor of two suggested by this excess. To determine which NC π^0 process best describes the excess, several different final states are considered: those in which there are no extra particles in the final state besides a π^0, NC coherent pion production, and reactions with multiple particles in the final state characteristic of incoherent pion production. Several metrics are formed to simultaneously find regions of separation between the possible NC π^0 final states as well as an agreement with the shape of the excess. The first metric under consideration is ψ which describes how much extra energy is found outside of the electron candidate cone and is defined by

$$\psi = \frac{E_{other}}{E_{inside}} \tag{5.19}$$

for E_{other} the energy outside of the electron cone and E_{inside} the energy inside the electron cone. As seen in Figure 5.16, it is apparent that the excess is best described by the shape of the NC Coherent π^0 events.

To learn further information about the characteristics of the extra energy, the positions of the extra energy clusters are individually analyzed. Given the vertex location and cluster location with respect to the detector center as well as the θ and ϕ of the vector axis, a

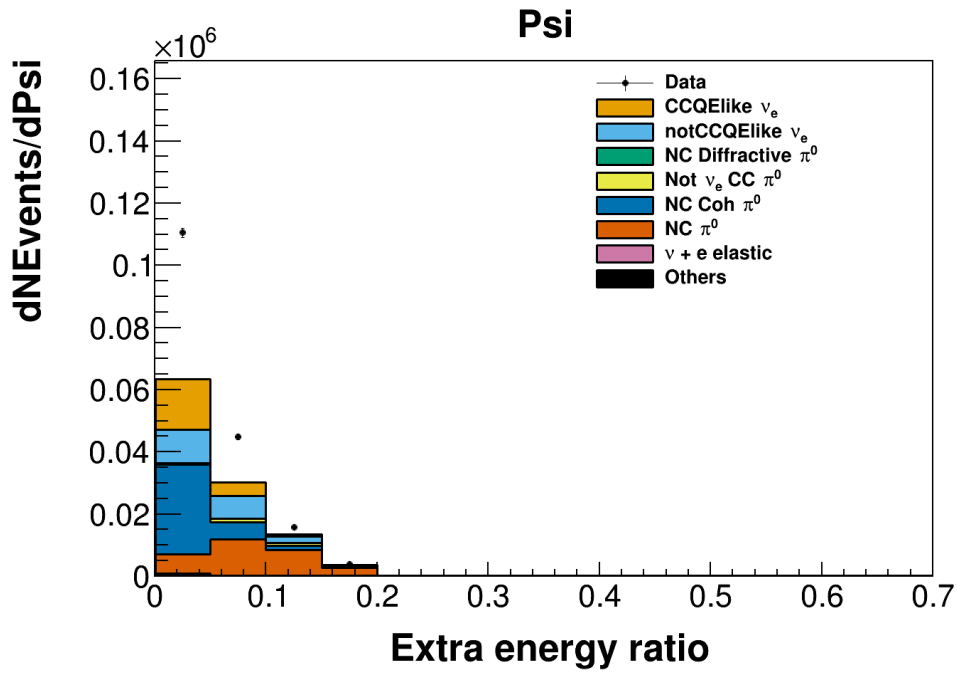


Figure 5.16: Extra energy ratio, ψ, where the excess is best described by the shape of the coherent π^0 event

metric is formed for the longitudinal and transverse distance of the extra energy clusters with respect to the electron candidate axis written as

$$\vec{v}_1 = \vec{v}_{Cluster} - \vec{v}_{Vertex} \tag{5.20}$$

$$\vec{v}_2 = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta) \tag{5.21}$$

$$d_{long} = v_1.Dot(v_2) \tag{5.22}$$

$$d_{trans} = \sqrt{v_1 Dot(v_1) - (v_1 Dot(v_2))^2} \tag{5.23}$$

for d_{long} and d_{trans} the longitudinal and transverse distances. A schematic representing the metric is seen in Figure 5.17, with the MC distribution seen in Figures 5.18 and 5.19. Once

again the excess shape favors the NC Coherent π^0 events.

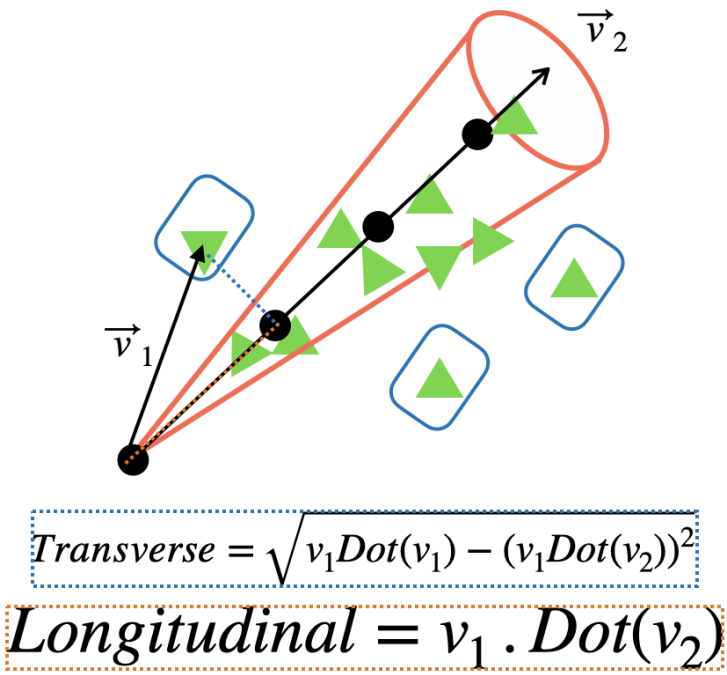


Figure 5.17: A schematic showing the calculation of the transverse and longitudinal distance of clusters. The green triangles represent the clusters of energy resultant from a neutrino interaction. The red cone represents the electron candidate cone formed around the optimized axis. The clusters outside of the cone are assigned a longitudinal (red dotted line) and transverse (blue dotted line) distance with respect to the cone axis.

The first theoretical description of NC coherent pion production was published by Lackner in 1979 [84] and, while studying a sample of single π^0s produced in CERN's neutrino and anti-neutrino beams, the first observation of coherent pion production was reported in 1983 by the Aachen-Padova spark-chamber experiment [85]. After the Aachen-Padova discovery, Rein and Sehgal built on Lackner's work and developed a theoretical model for coherent pion production [86]. Several additional experiments observed and measured NC and CC coherent pion production in a range of energies with neutrino and

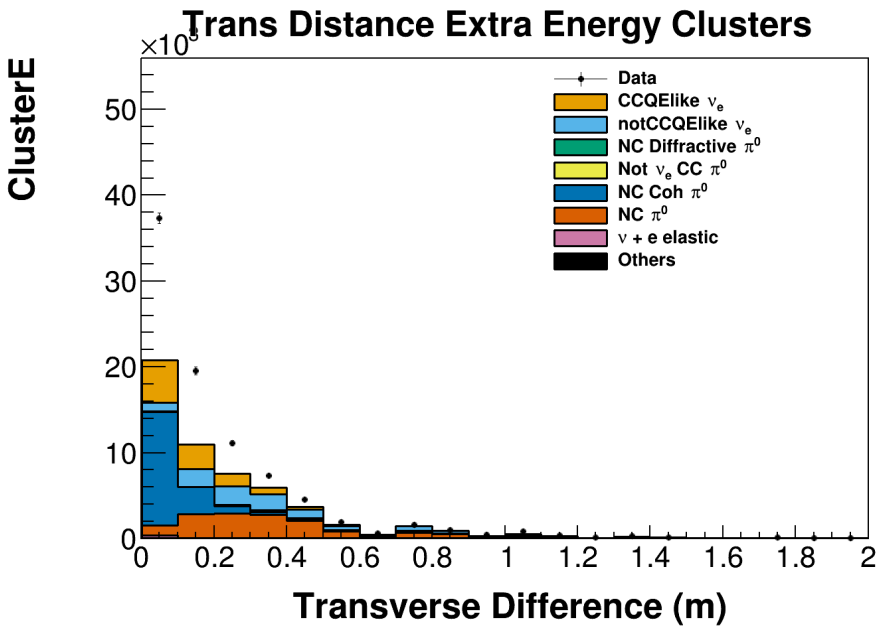


Figure 5.18: Transverse distance of extra energy clusters with respect to cone axis

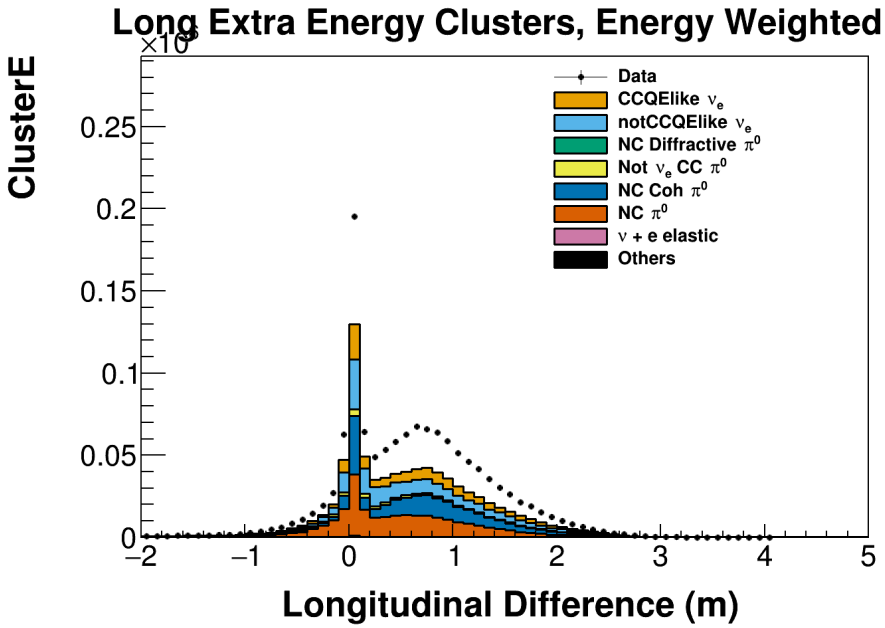


Figure 5.19: Longitudinal distance of extra energy clusters with respect to cone axis

anti-neutrino beams, for all results consistent with the predictions of the coherent Rein-Sehgal model. These experiments which observe the NC coherent production typically work in the variable $E_{\pi^0}(1-\cos\theta_{\pi^0}) \approx E_{\pi^0}\theta^2_{\pi^0}$ ([87, 88, 89, 90, 91, 92]). It is anticipated that a coherent process would require a small momentum transfer such that its value in $E_{\pi^0}\theta^2_{\pi^0}$ would be relatively small compared to other processes. As seen in Figure 5.20, the $E_{\pi^0}\theta^2_{\pi^0}$ excess is peaked near zero with the overall shape preferring that of the NC coherent process.

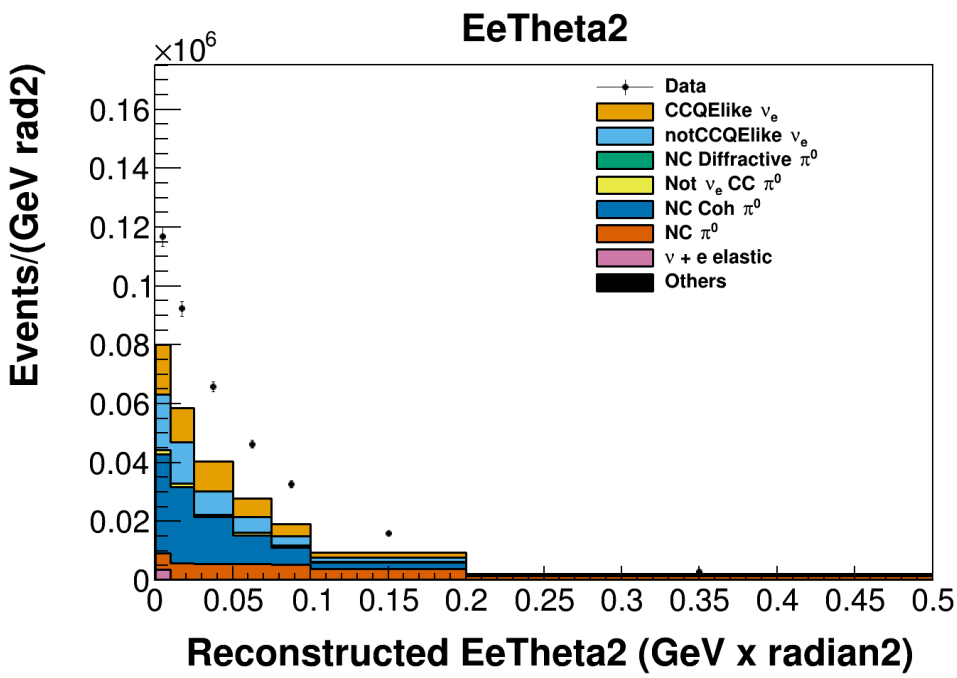


Figure 5.20: $E_e\theta^2(GeV \text{ x } radian^2)$

After ruling out vertex mis-reconstruction, γ interactions, and NC non-coherent π^0 processes, all of the evidence indicates that the data excess seen in the region of dE/dx > 2.4 MeV/ is best explained by a neutral current coherent process. However, there are a few issues with this conclusion. The first is that the same model used to predict neutral-

current coherent pion production also predicts charged-current coherent pion production. MINERvA has made precise measurements of the CC Coherent π^0 process and finds that the data and MC predictions are in very good agreement for pion energies above 0.5 GeV [93]. Additionally, from looking at the inline upstream energy distribution in Figure 5.12, there is an excess of events at both small and large values of inline upstream energy which indicates that a tune to only the NC coherent process would not be sufficient. Therefore, it is concluded that NC coherent π^0 production is not the best candidate to explain the disagreement between data and MC events.

Excess Sideband Division

With the implementation of the NC diffractive π^0 events, the excess sideband needs to be revisited. Recall that the purpose of sidebands is to divide and constrain background processes. Therefore the excess sideband requires an additional cut to separate diffractive and coherent events in preparation for their extrapolation into the signal region.

The inline upstream energy is the best discriminator available between coherent and diffractive events. This metric, a reverse electron candidate cone pointing upstream of the candidate vertex as illustrated in Figure 5.11, allows for the capture of the recoiling proton energy in the case of a diffractive event. Assuming the recoiling proton is above the detection threshold of the detector, it is expected that NC diffractive π^0 events will have more inline upstream energy than NC coherent π^0 events where the nucleus receives no recoil. With the addition of NC diffractive π^0 events to the MC simulation and applying an overall scale factor of 12 [1], the contributions of both processes take form in inline

[1] An overall scale factor of 12 was applied in original studies. In this case, it is included for comparison purposes only and discussion of the final scale factor is discussed later in the chapter.

upstream energy as seen in Figure 5.21. It becomes apparent that a scale factor constant

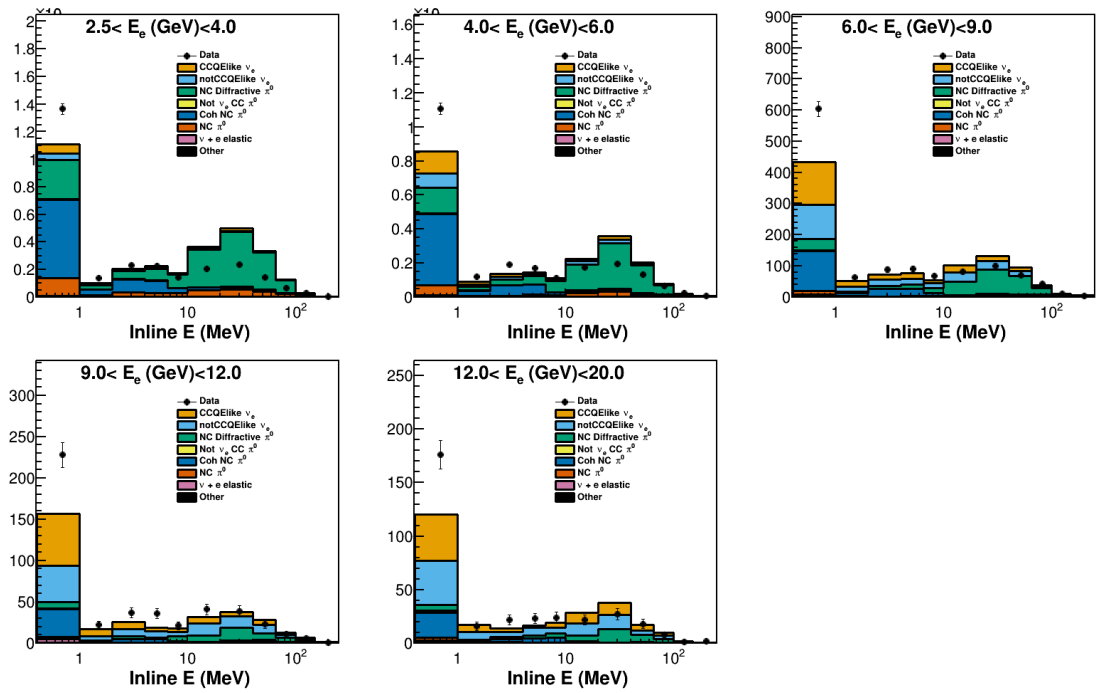


Figure 5.21: Inline Upstream Energy in bins of E_e in the excess sideband, NC Diffractive has a scale factor of 12.

in E_e applied to the diffractive pion model is insufficient in describing the excess's shape. This will be addressed in the background tuning, where the NC diffractive π^0's will be tuned as a function of electron energy. Additionally, the NC diffractive π^0 events have a larger contribution at higher inline energies of > 10 MeV and NC coherent π^0 events are largely dominant at lower inline energies < 10 MeV. Since sidebands are formed to constrain individual background processes, the excess sideband is divided at an inline upstream energy value of 10 MeV. Figure 5.22 shows the sideband re-definitions with the Excess sideband broken into two regions with the inline energies < 10 MeV termed the NC coherent π^0 sideband and inline energies of > 10 MeV the NC Diffractive π^0 sideband.

Even though the NC diffractive π^0 and NC coherent π^0 processes are most dominant in the excess sideband, there still exists a population of additional background processes like NC π^0 production. This suggests that a comprehensive tune, simultaneously fitting all regions, is required. The final section of this chapter, 5.2.2, will show the approach taken in applying a global tune to the signal and sideband regions. Before applying the final background tuning to the signal and sideband regions, the remaining background, ν_e events, must be addressed.

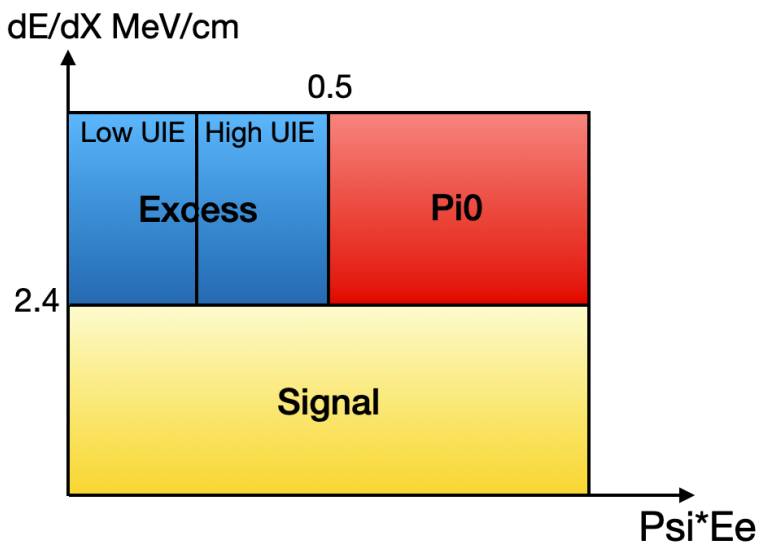


Figure 5.22: Representation of signal and sideband regions with respective cuts. The excess sideband has been subdivided into low and high upstream inline energy.

5.2 Final Analysis Corrections

In a typical muon analysis for MINERvA, an analyzer can determine what charge of muon occurred in an event since a muon will typically traverse the MINERvA detector and enter

MINOS. MINOS, being a magnetized detector, will curl the muon in a direction depending on the charge of the muon and muon neutrino analyzers can separate the differently charged muons. As mentioned previously, there is an inability to distinguish between $\bar{\nu}_e$ and ν_e interactions in the MINERvA detector. This section addresses the correction made for the ν_e contamination found in this $\bar{\nu}_e$ analysis.

5.2.1 FHC Correction

While $\bar{\nu}_e$ and ν_e events are indistinguishable in data, the MC simulation allows for the separation of these events. Since the FHC sample has a much larger population of ν_e events due to the horn configuration, the goal is to form an estimator based on FHC data and MC that gives a prediction of the ν_e background found in the RHC sample. This is accomplished by taking the FHC ν_e interaction sample and normalizing it to the event rate in the RHC sample, giving a ν_e background prediction for the RHC sample. First a ratio of RHC/FHC ν_e events distributed in true neutrino energy is formed, seen in Figure 5.23. Each bin of true neutrino energy serves as a scale factor that can be applied to the FHC sample to normalize it to the RHC ν_e event rate. The scale factors will be applied to the FHC sample on an event-by-event basis from the RHC/FHC ν_e ratio in true neutrino energy. However, while true neutrino energy is known in MC, only reconstructed information is available in the data. Therefore to apply the proper scale factor per neutrino energy bin, a neutrino energy estimator is developed out of the reconstructed available energy and the reconstructed electron energy,

$$E_{est} = E_e + E_{avail} \tag{5.24}$$

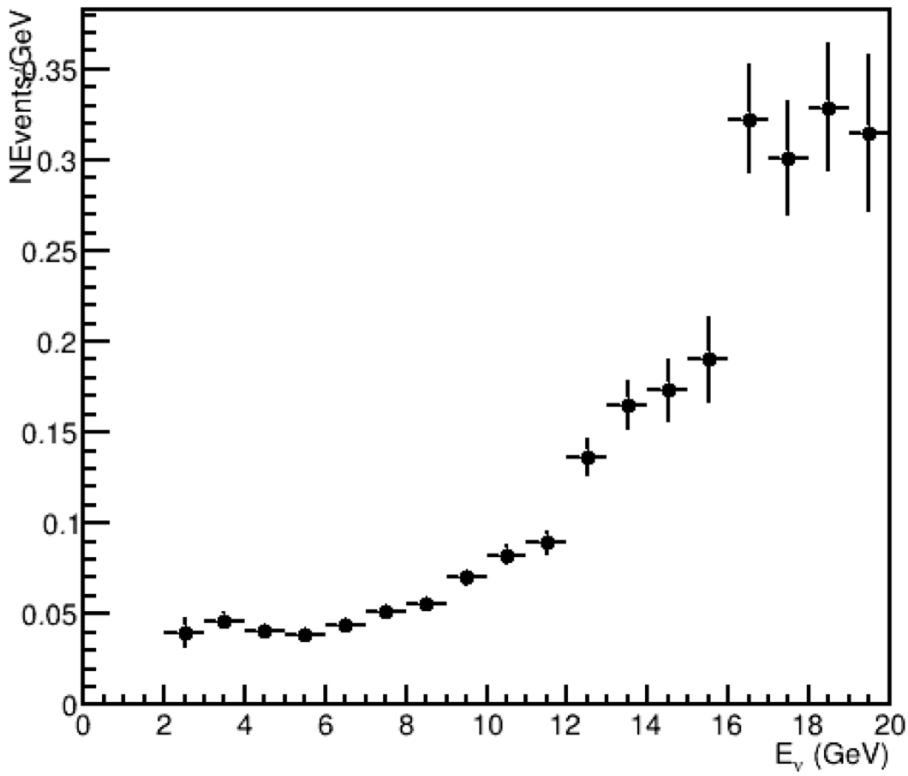


Figure 5.23: Ratio of RHC/FHC ν_e in True Neutrino Energy

The accuracy of the formed energy estimator was vetted against the true neutrino energy as seen in Figure 5.24. A perfect estimator would be represented as a single diagonal line for each bin of the energy estimator matching each bin of the true neutrino energy. It can be seen that there are some off-diagonal populations; however, the estimator is deemed sufficient.

For each FHC event, a neutrino energy estimator is formed from the sum of reconstructed E_{avail} and q_3. The energy estimator value is then used to find the respective scale factor based on the RHC/FHC ν_e ratio in true neutrino energy. The scale factor is applied to the

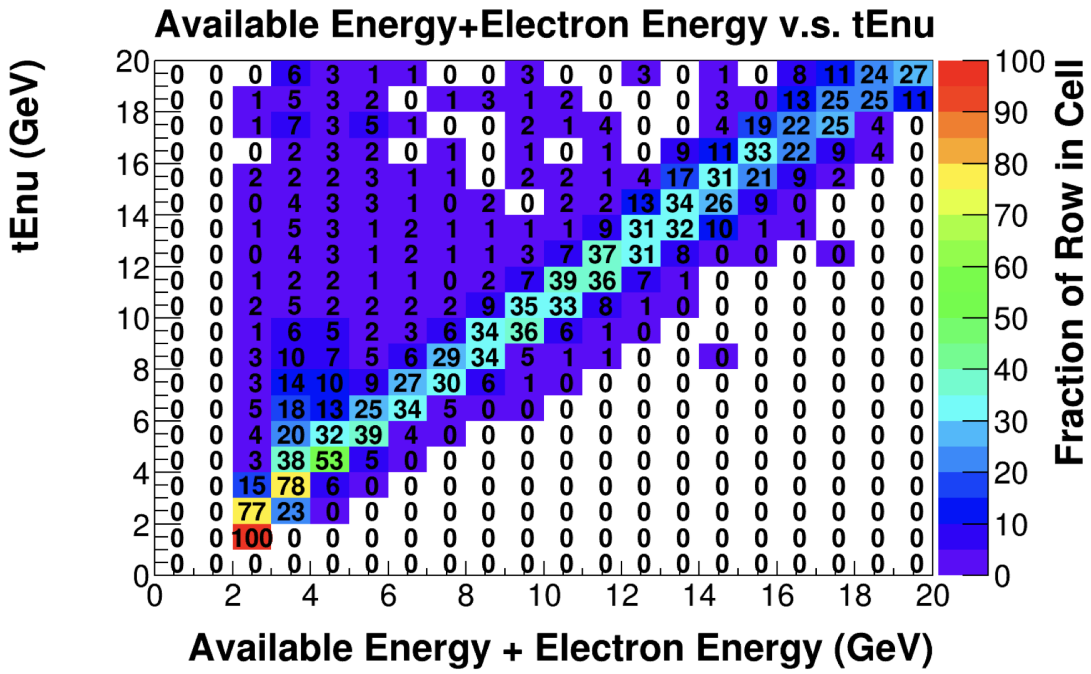


Figure 5.24: Neutrino energy estimator vs true neutrino energy

rate of FHC events and, as a result, a background ν_e prediction is formed for the MC RHC sample by subtracting all FHC $\bar{\nu}_e$ and background scaled events from the FHC scaled data, seen in Equation 5.25.

$$FHC\ \nu_e\ MC\ Prediction = FHC\ Data\ - FHC\ \bar{\nu}_e - \ FHC\ Background \tag{5.25}$$

The scaled FHC ν_e sample was compared against the RHC ν_e in bins of E_{avail} vs q_3 and E_{avail} vs p_T to ensure it was a reasonable prediction based on shape and percentage of total events. Figure 5.25 shows the scaled FHC sample broken into electron neutrino and anti-electron neutrino contributions in bins of E_{avail} vs q_3 and Figure 5.26 shows the RHC sample. Figure 5.27 shows a comparison between the FHC electron neutrino prediction for the RHC background compared to the original RHC electron neutrino background for a randomly chosen bin in q_3. This comparison was performed for q_3 bins and it was determined that the FHC electron neutrino background prediction is reasonable. The FHC background prediction is created for all distributions needed to carry the analysis through to a cross section result.

Before adding the FHC ν_e prediction to the RHC sample, the original ν_e contamination of the RHC sample is removed from the MC. Then the scaled FHC ν_e distributions are added to the respective RHC distributions under the category of “Other”, allowing for the treatment of ν_e events as a background going forward into background fitting and subtraction. When adding in the FHC sample to the RHC sample, it must be properly

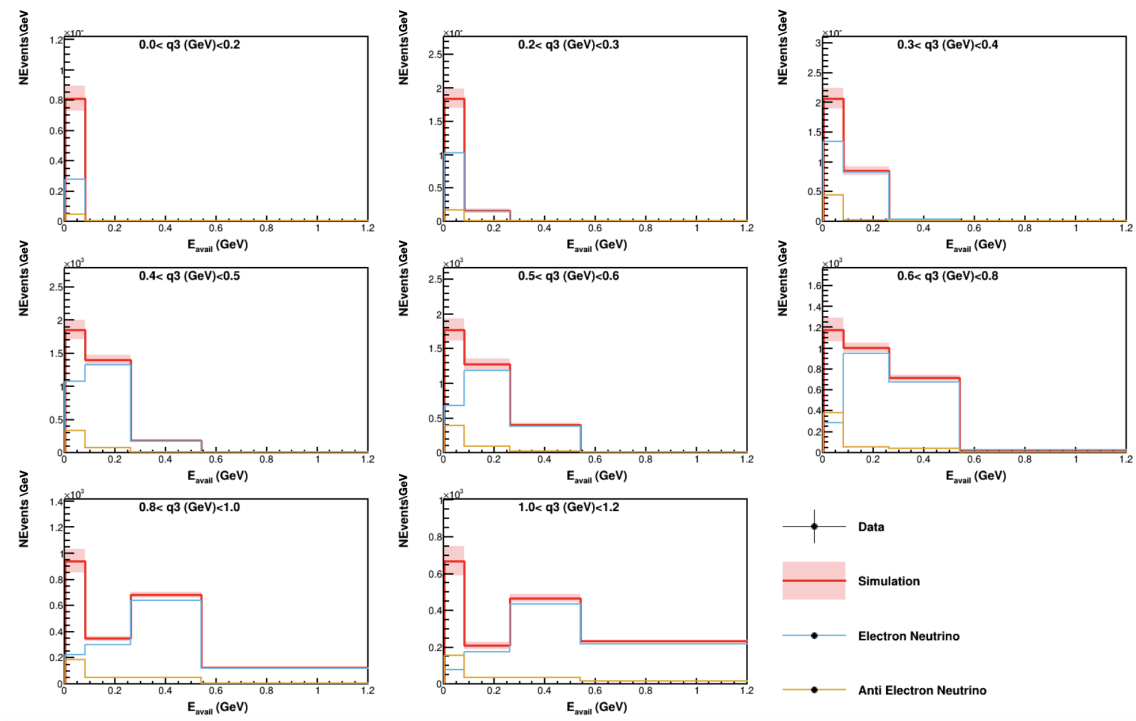


Figure 5.25: The scaled FHC sample providing an electron neutrino (blue) prediction for the RHC electron neutrino background.

normalized as

$$\frac{RHC_data_pot}{FHC_data_pot} * \frac{FHC_data_pot}{FHC_MC_pot} \rightarrow \frac{RHC_data_pot}{FHC_MC_pot} \tag{5.26}$$

With all backgrounds properly accounted for, the next step is to apply a fitting technique.

5.2.2 Background Fitting

The $\bar{\nu}_e$ candidate sample contains backgrounds that must be subtracted from data to precisely calculate the cross section. The MC estimates the rate of background events expected in

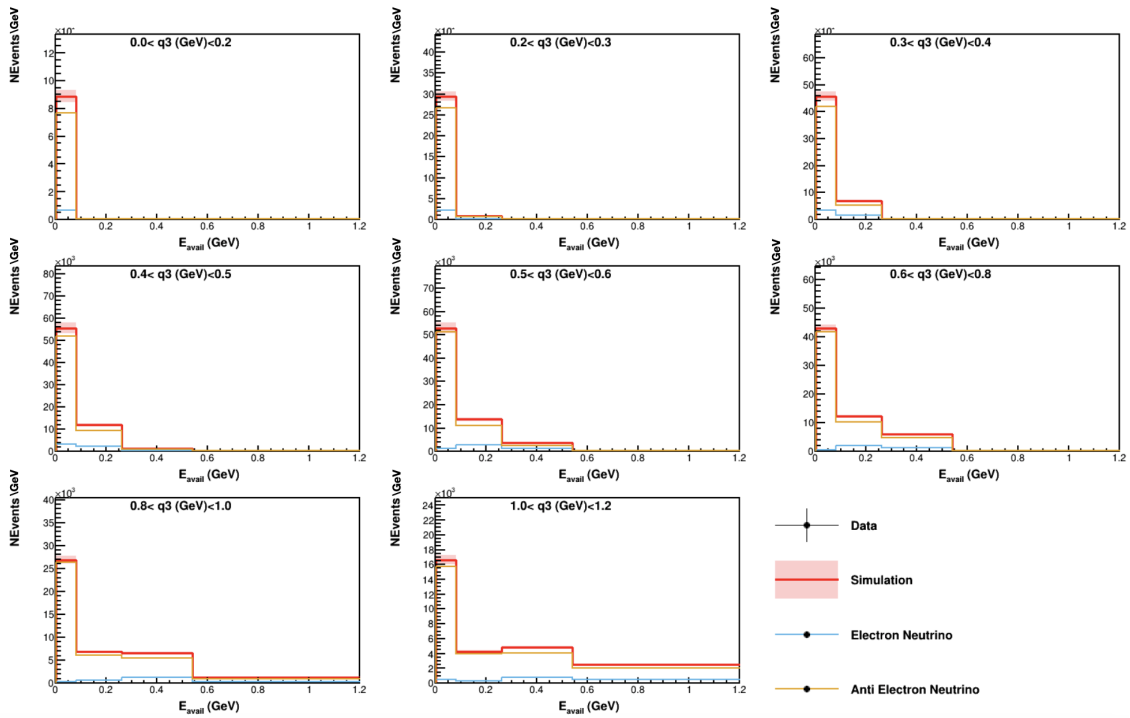


Figure 5.26: A breakdown of the RHC sample into electron neutrino and anti-electron neutrino contributions.

the signal sample, which can have large uncertainties due to uncertainties in the underlying background models and flux predictions. These uncertainties are minimized by tuning the MC estimator of the background rates in the sidebands. Background populations in the signal region mean that these events have passed the reconstruction cuts even though they are not $\bar{\nu}_e$ events. MC allows for the identification of which reconstructed events are signal and which are background by looking at the truth information provided by GENIE.[2] The data background is predicted using the assumption that the fraction of MC background

[2]Truth information refers to generator-level information about all simulated interactions stored in a separate Truth ntuple.

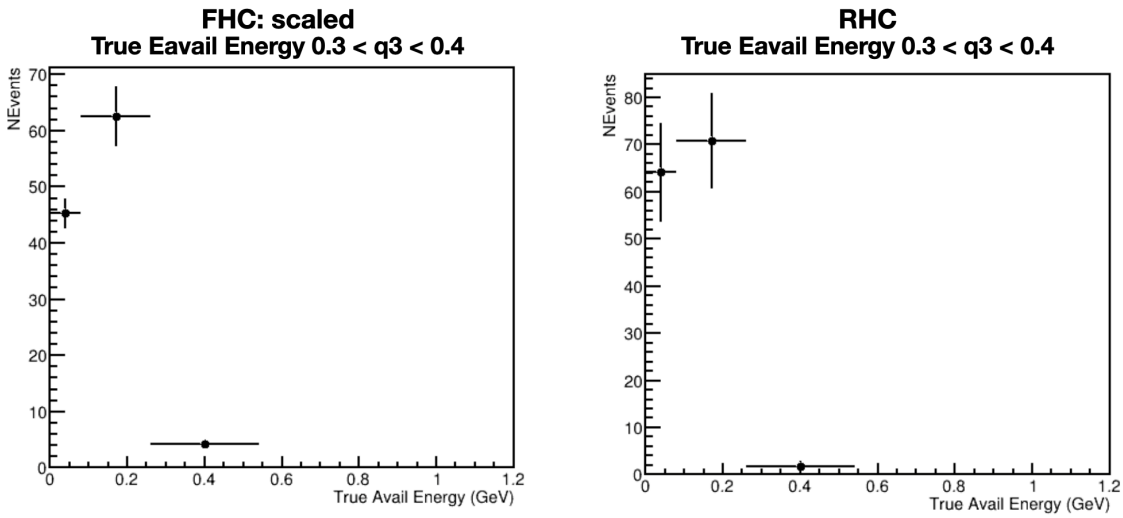


Figure 5.27: A comparison of the predicted electron neutrino contribution from the scaled FHC sample (left) to the RHC electron neutrino background prediction (right) for a randomly chosen bin of $0.3 < q_3 < 0.4$. The number of events (y - axis) is comparable between the two samples.

events to MC signal events is the same as it would be for data. The fraction is formed as

$$N^{bkgd}_{data,ij} = \frac{N^{bkgd}_{MC,ij}}{N_{MC,ij}} \times N_{data,ij} \tag{5.27}$$

for i, j the respective reconstructed bins. However, in using this approach to formulate a data background prediction, there is an inherent reliance on the MC's ability to predict both the background and signal processes accurately. Therefore, the relative fraction of signal and background processes found in data are determined through a data-driven fitting procedure in which the signal and sideband regions are varied to best optimize the shape of the data [45].

The data-driven fitting procedure, referred to as background fitting, is used to extract a

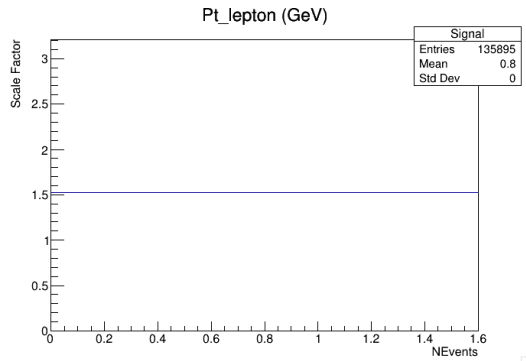
correction to the normalization of the respective backgrounds. The corrections, or scale factors, are determined by simultaneously varying a fit between data and MC. In the case of this analysis, the major backgrounds are the NC Diffractive π^0, NC Coherent π^0, CC and NC incoherent π^0 processes. Therefore, the background fit is carried out in four regions: signal (dE/dx $<$ 2.4 MeV/cm), Pi0 (dE/dx$>$2.4 MeV/cm and $\psi * E_e > 0.5$ GeV), excess low inline upstream energy (dE/dx$>$2.4 MeV/cm, $\psi * E_e < 0.5$ GeV, and inline upstream energy $<$ 10 MeV), and excess high inline upstream energy (dE/dx$>$2.4 MeV/cm, $\psi * E_e <$ 0.5 GeV, and inline upstream energy $<$ 10 MeV).

The regions are fit simultaneously in the distributions. The first global fit is done in $\psi * E_e$ vs p_T in bins of E_e which optimizes the NC coherent and diffractive processes and the second global fit is done in E_{avail} vs p_T in bins of p_T which optimizes the NC and CC π^0 processes. The fit is performed using ROOT's Minuit2 fitter, a numerical minimization library. The library provides minimization algorithms that search for parameter values which minimizes a user defined function by finding the minimized χ^2 value between the input histogram's data and MC values, universe by universe [3]. The χ^2 minimized in fitting the background scale factors is calculated as

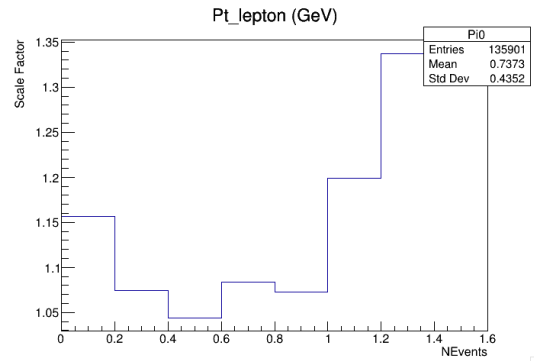
$$\chi^2 = \sum_i \sum_j \frac{(N_{ij}^{data} - \sum_k \alpha_{ik} N_{ijk}^{MC})^2}{\sum_k \alpha_k N_{ijk}^{MC}} \tag{5.28}$$

for i the p_T bins and j the corresponding E_{avail} for the respective input distribution of E_{avail} vs p_T. The k is the MC event category, α_{ik} is the background scale factor with respect to the background category k for bin i of p_T, N_{ijk}^{MC} is the number of MC events for background category k given for ij bins with respect to E_{avail} vs p_T, and N_{ij}^{data} is the number of data

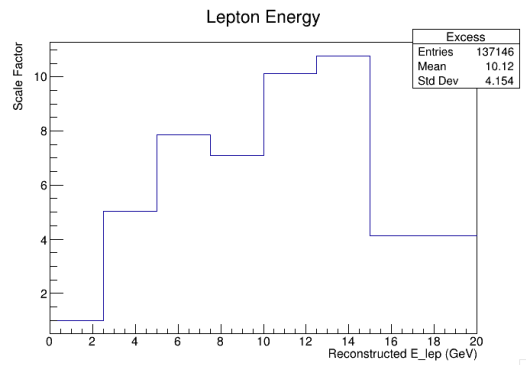
[3] MINERvA uses a "many universe" statistical technique described in detail in Section 7.2



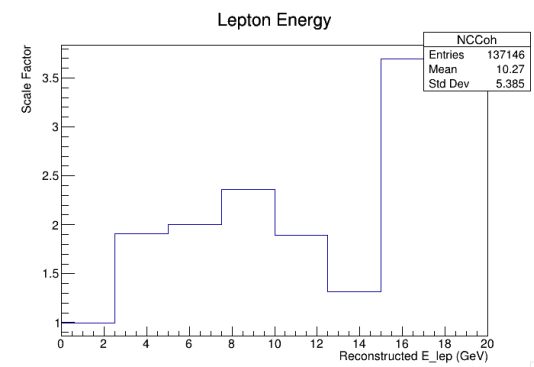
(a) Signal



(b) π^0



(c) NC Diffractive π^0



(d) NC Coherent π^0

Figure 5.28: Scale factor values for signal (top left) and π^0 backgrounds (top right) as a function of p_T, NC Diffractive π^0 background (bottom left) and NC Coherent π^0 backgrounds (bottom right) as a function of E_e as calculated by a global fit

events in bins i and j. For the background tuning, the signal region and the π^0 background are optimized for a single bin spanning all values of p_T and the excess backgrounds are fit in bins of p_T.

Figure 5.28 shows the resulting scale factors for each fit. As a note, the scale factor applied to the signal region is only applied to constrain the background and is removed during the background subtraction process.

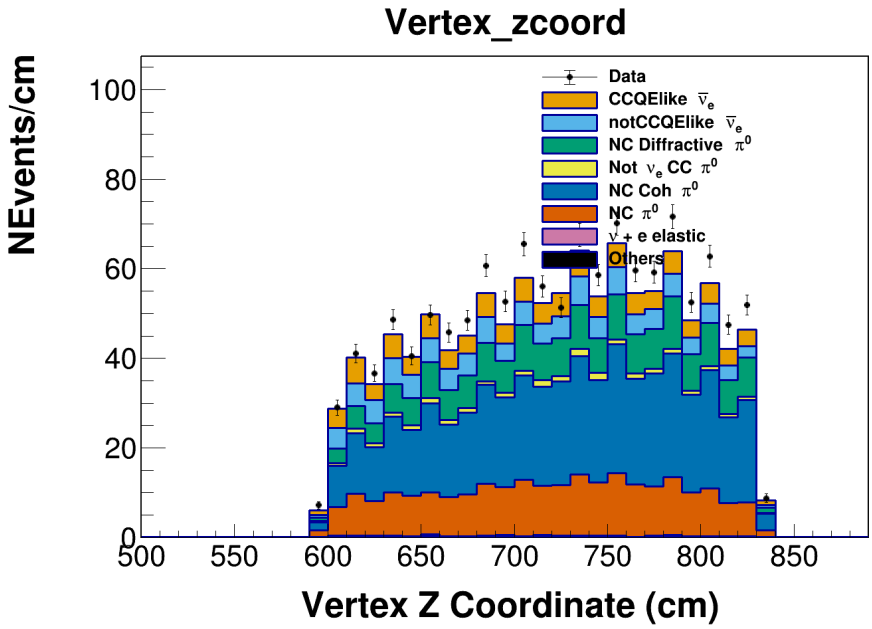


Figure 5.29: Vertex z position in the Excess sideband.

5.2.3 Post Background Tune

After properly applying the final scale factors for the background tunes, the question remains whether the NC diffractive π^0 process is a good estimate for the shape of the excess region. Revisiting the distributions originally used to isolate a candidate process for the excess of data over MC shows how the diffractive process does indeed provide a good description. Figures 5.29, 5.30, 5.31, 5.32, 5.33, 5.34, 5.35, 5.36 show these distributions for dE/dx $>$ 2.4 MeV/cm and $\psi * E_e < 0.5$ GeV with the scale factors applied from the background fit.

The scale factors are computed for the given signal and background event processes and applied on an event-by-event basis. Once these scale factors have been applied, this allows for their subtraction from the data. Figures 5.37 through 5.44 show the Prefit and Postfit for

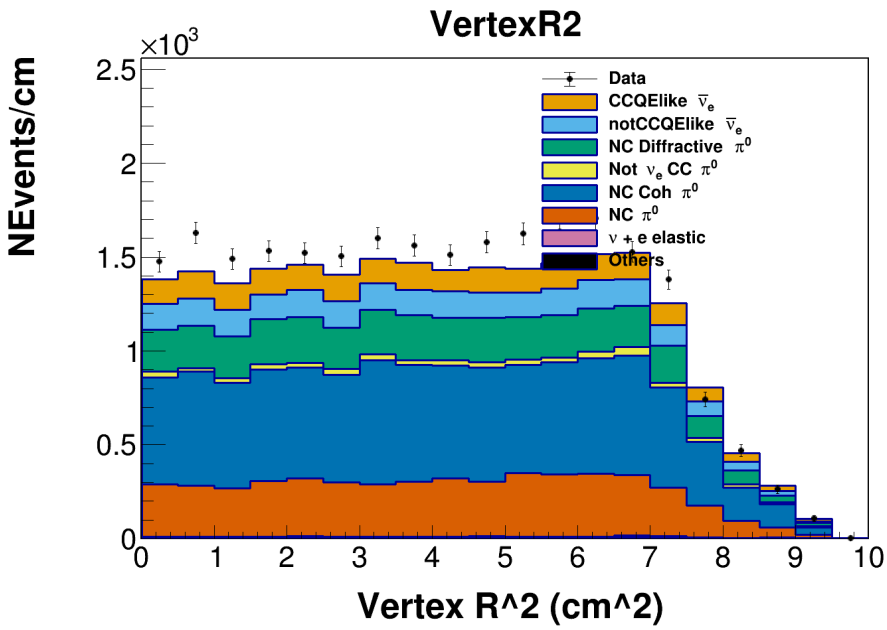


Figure 5.30: Vertex $R^2 = x^2 + y^2$ in the Excess sideband.

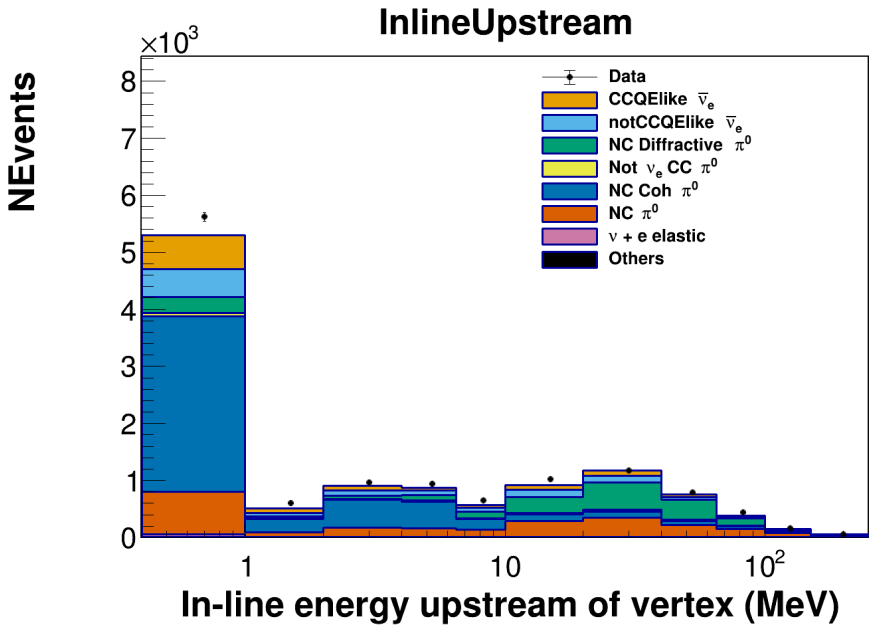


Figure 5.31: Energy inside a 7.5 deg cone pointing upstream in the Excess sideband.

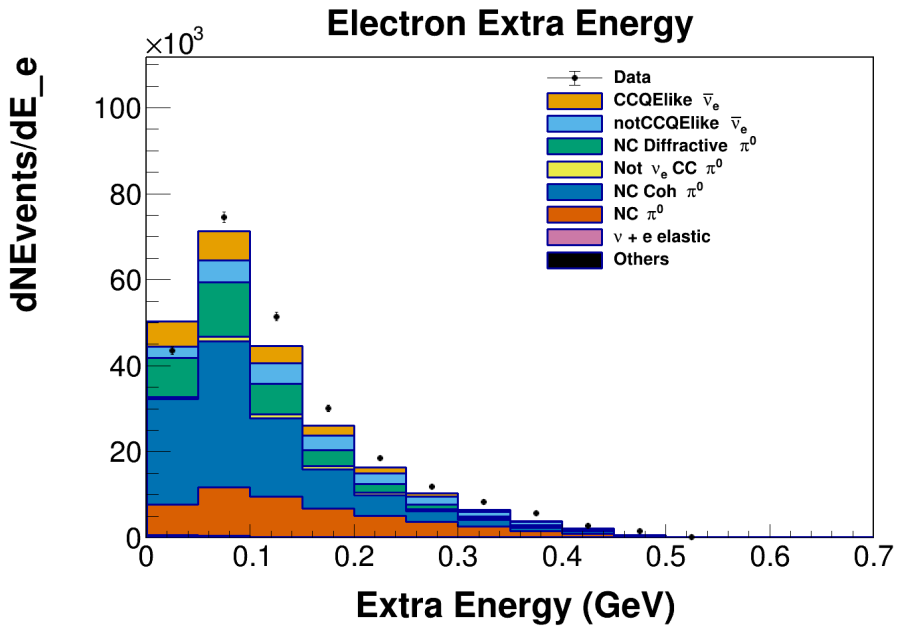


Figure 5.32: Energy outside of electron candidate cone and vertex region in the Excess sideband.

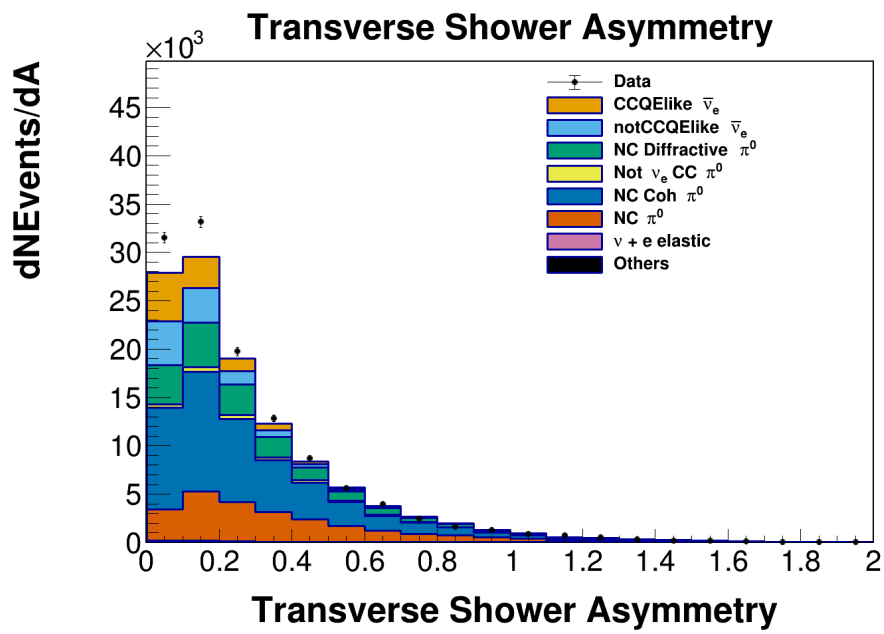


Figure 5.33: Transverse shower asymmetry in the Excess sideband.

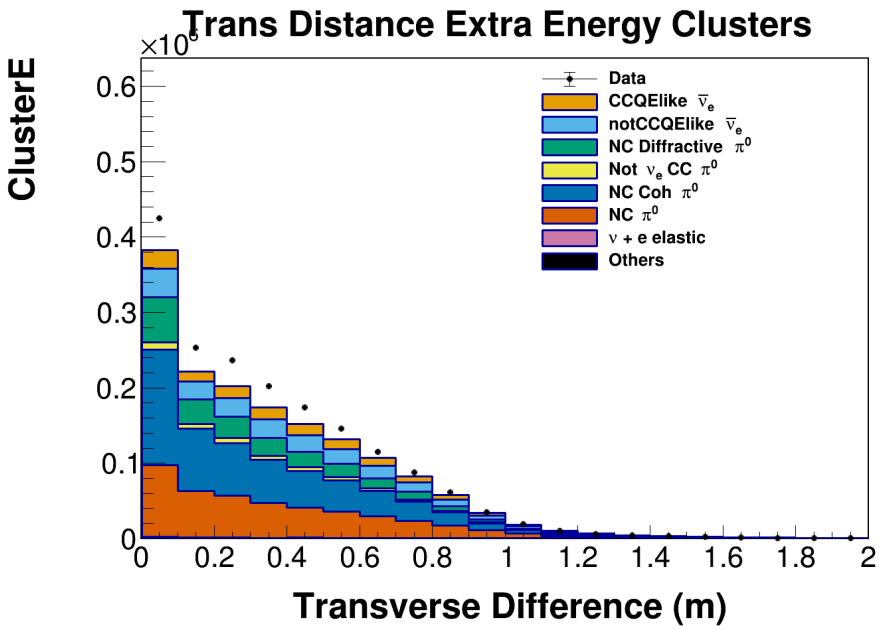


Figure 5.34: Transverse distance of extra energy clusters with respect to cone axis

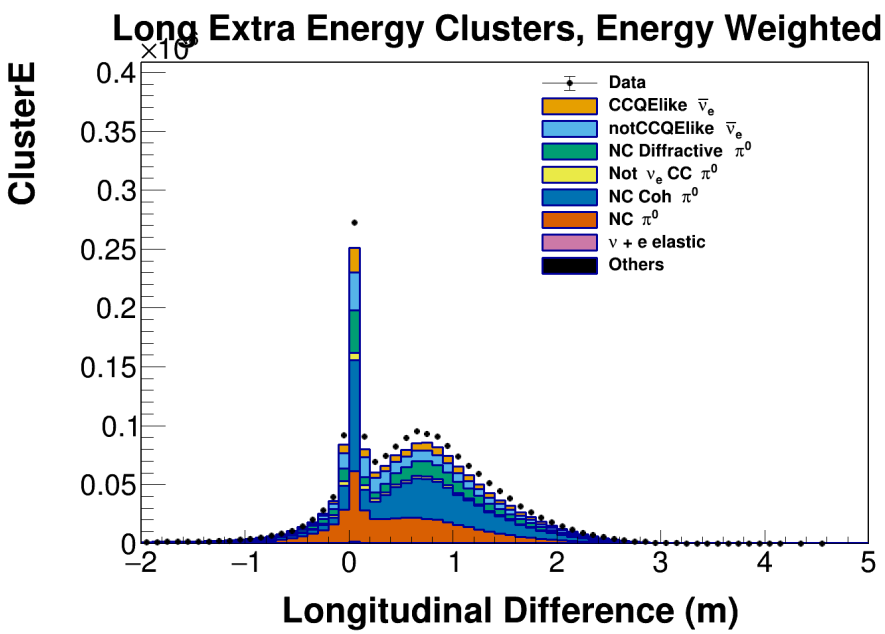


Figure 5.35: Longitudinal Difference (m)

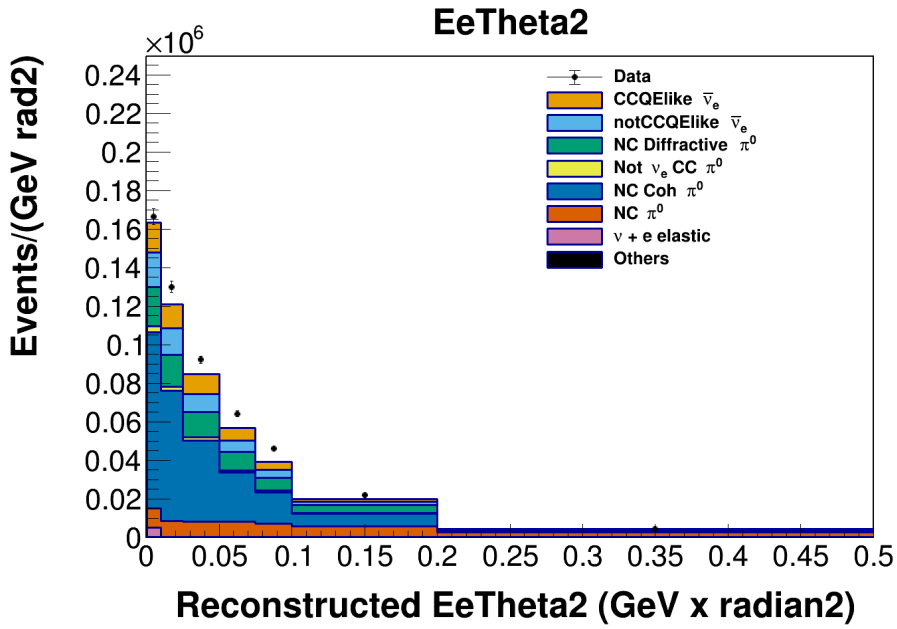
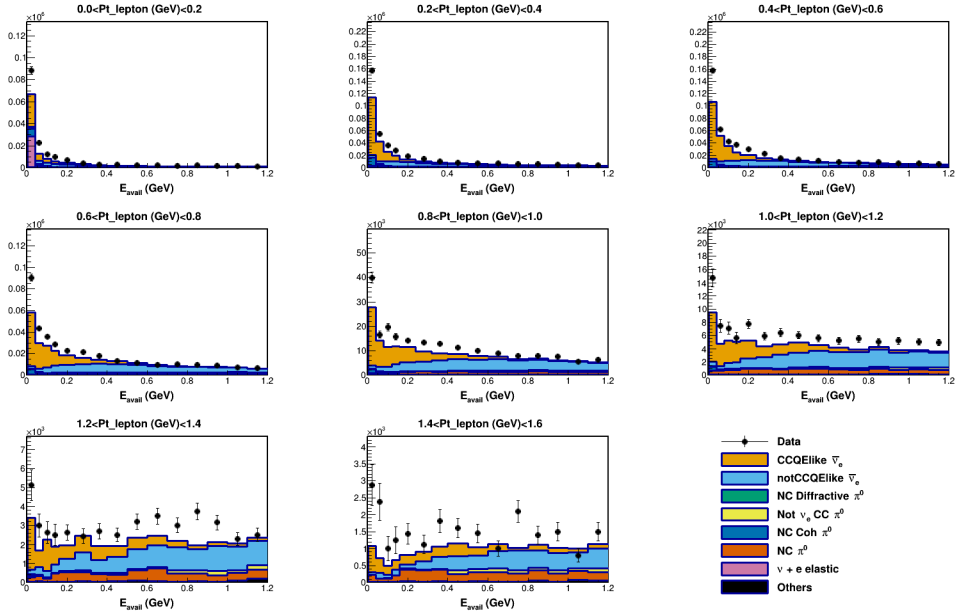


Figure 5.36: $E_e\theta^2$

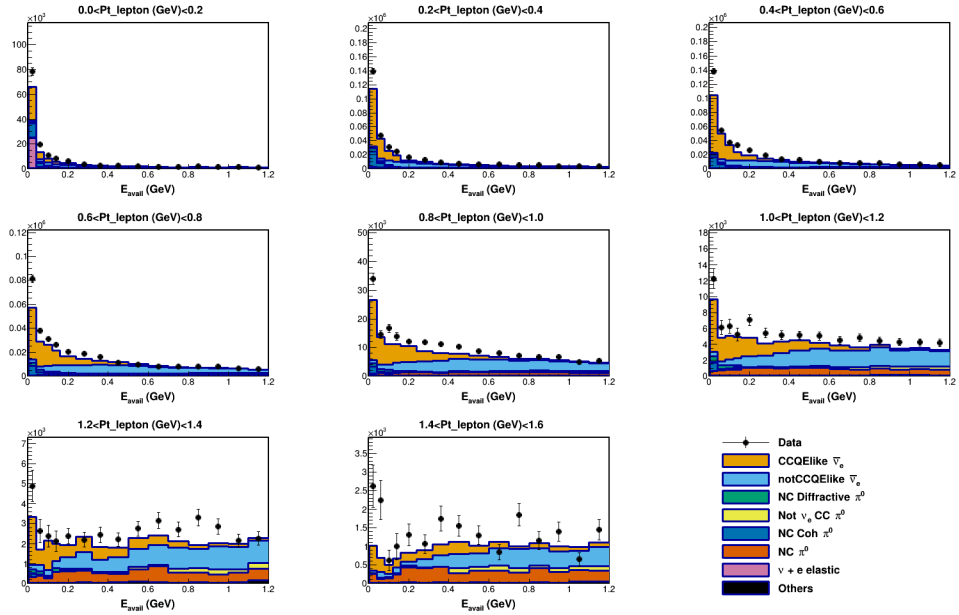
the distributions of importance E_{avail} vs p_T and E_{avail} vs q_3. As seen in Figures 5.38 and 5.42 in the region of excess low inline upstream energy (dE/dx>2.4 MeV/cm, $\psi * E_e <$ 0.5 GeV, and inline upstream energy $<$ 10 MeV), there is good agreement between data and MC except for the first bins of E_{avail} in higher lepton p_T bins of $>$ 0.8 GeV/c (and higher q_3 for values of $>$ 0.8 GeV/c as well). For Figures 5.39 and 5.42 in the region of excess high inline upstream energy (dE/dx>2.4 MeV/cm, $\psi * E_e <$ 0.5 GeV, and inline upstream energy $<$ 10 MeV) there is good agreement up until the last three bins of lepton p_T between 1.0 GeV/c and 1.6 GeV/c. This is mostly due to limited statistics in those bins. (good agreement is seen in all bins of q_3.) Finally, for the π^0 region (dE/dx>2.4 MeV/cm and $\psi * E_e >$ 0.5 GeV) in Figures 5.40 and 5.44 there is good agreement between data and MC except for the first bin of both p_T and q_3 where there is a large MC over-prediction

in the first E_{avail} bin as well as the last two bins in $p_T > 1.2$ GeV/c and the last bin in q_3. However, there is a noticeable amount of other background processes, like NC coherent π^0 events or signal events in these regions.

In summary, the large excess of data over MC in the region of Front dE/dx > 2.4 GeV/c was first seen in MINERvA's low energy CCQE ν_e analysis [79]. The first step was to determine what type of process best fit the shape of the excess through a series of distributions. The conclusion was that a neutral current π^0 process was the best *ad hoc* candidate to explain the shape of the excess. Since the NC coherent π^0 process is well modeled and understood, a similar process known as NC diffractive π^0 production was compared to the excess shape. Through a series of distributions, and the essential inline upstream energy distribution, it is determined that the NC diffractive π^0 process does a good job at predicting the shape of the excess. However, significant scale factors must be applied in order to fill the magnitude of the excess. As a result, three sideband regions are used in this analysis: the π^0 sideband and the excess sideband divided into the two regions of high and low inline upstream energy. A scale factor is also determined for the signal process but only used as a prediction in the sideband regions. With good data and MC agreement in the sideband regions from the completion of background tuning, the resultant tuned estimate for the number of background events $N_{\alpha\beta}^{bkgd}$ is subtracted from the number of signal selected events $N_{\alpha\beta}^{data}$

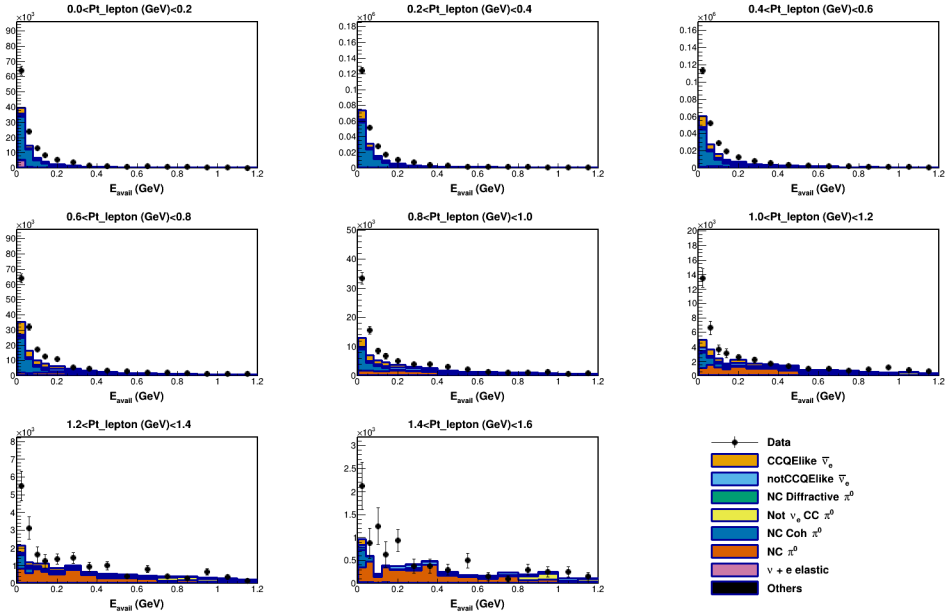


(a) Signal Pretune E_{avail} vs p_T

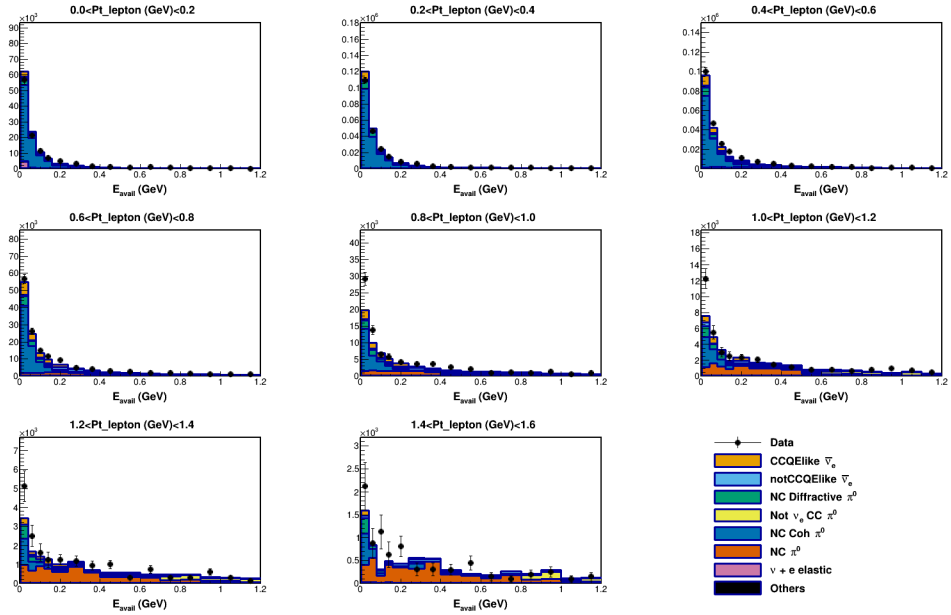


(b) Signal Posttune E_{avail} vs p_T

Figure 5.37: Signal distributions (dE/dx $<$ 2.4 MeV/cm) Pre and Post background tuning. These figures do not have the Signal scaled factor applied.

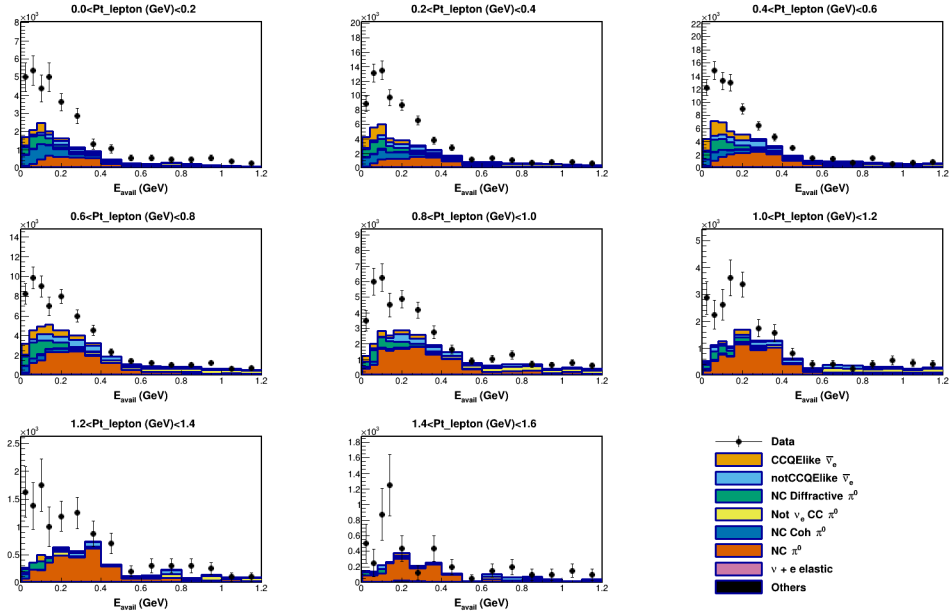


(a) Low Inline Pretune E_{avail} vs p_T

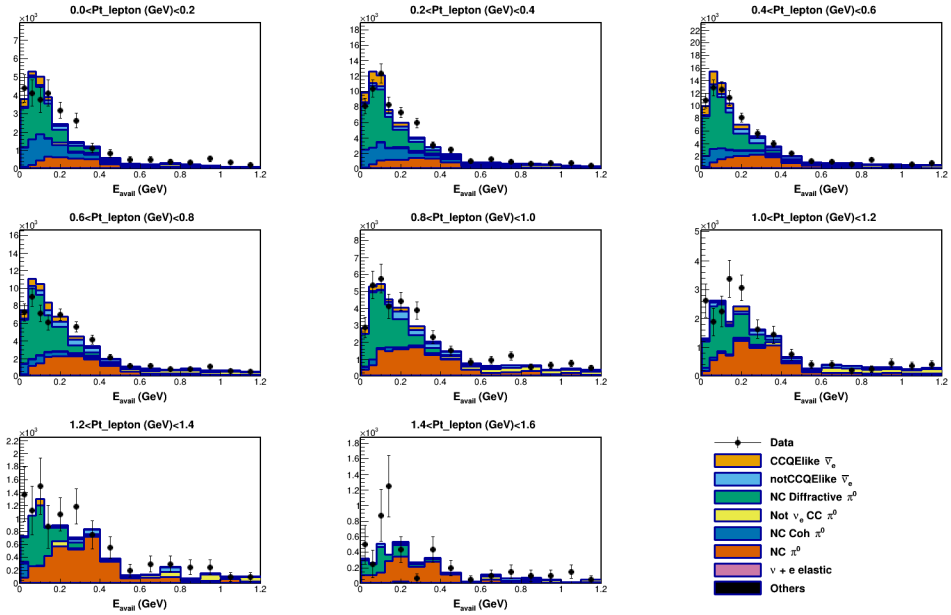


(b) Low Inline Posttune E_{avail} vs p_T

Figure 5.38: Pre and Post background tuning in the NC Coherent π^0 sideband (dE/dx $>$ 2.4 MeV/cm, $\psi * E_e < 0.5$ GeV, Inline upstream energy $<$ 10 MeV).

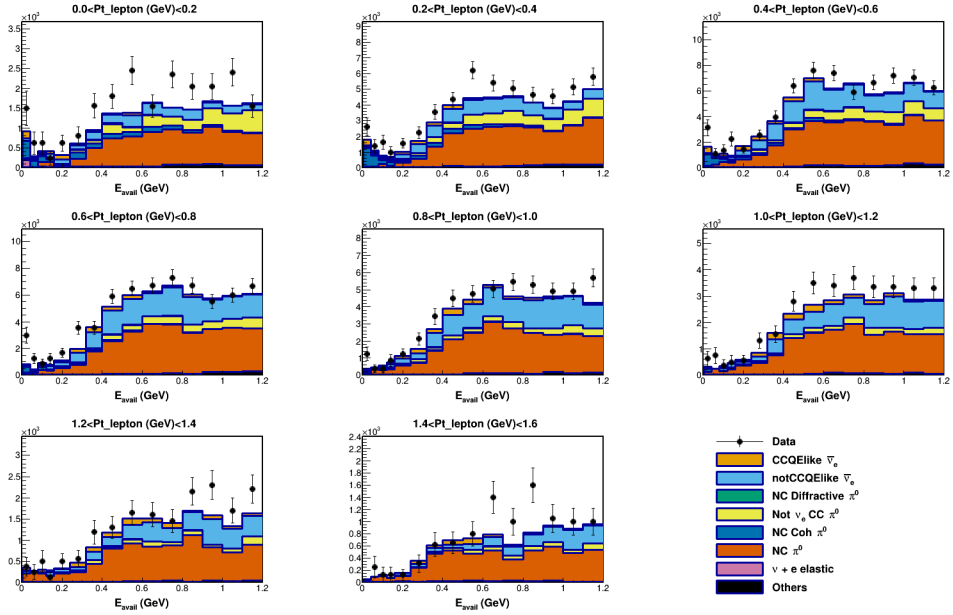


(a) High Inline Pretune E_{avail} vs p_T

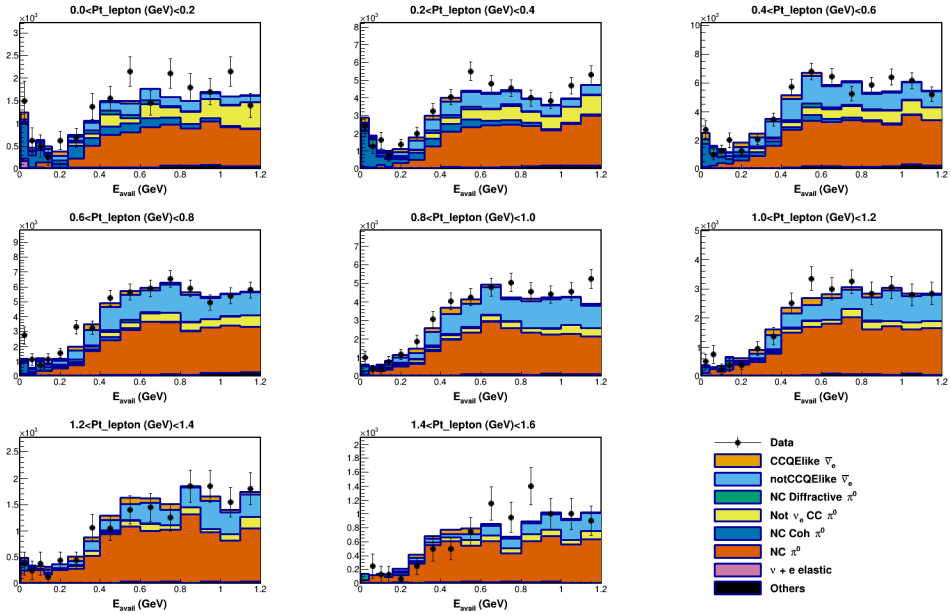


(b) High Inline Posttune E_{avail} vs p_T

Figure 5.39: Pre and Post background tuning in the NC diffractive π^0 sideband (dE/dx $>$ 2.4 MeV/cm, $\psi * E_e < 0.5$ GeV, Inline upstream energy $>$ 10 MeV).

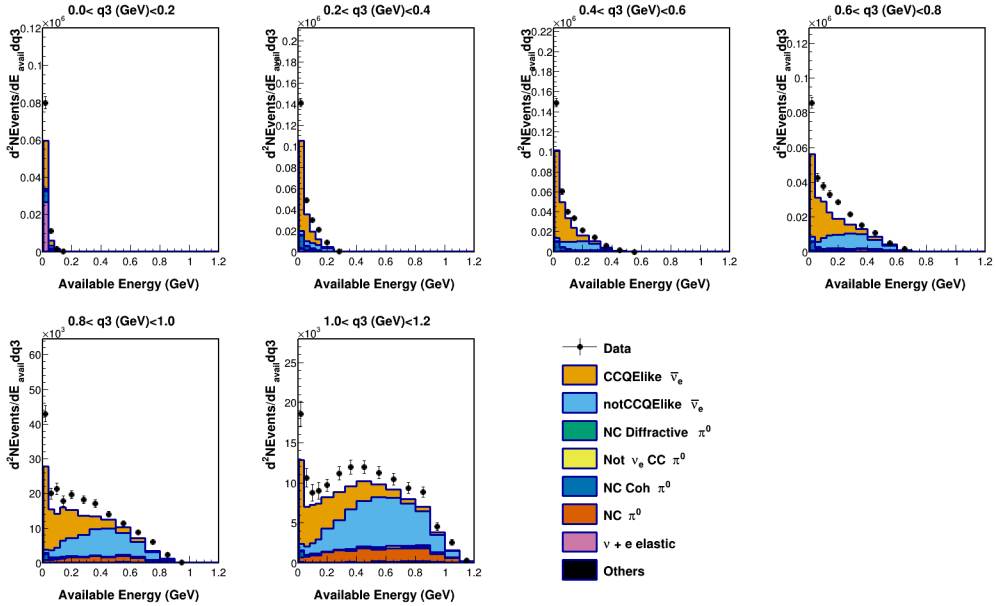


(a) High Inline Pretune E_{avail} vs p_T

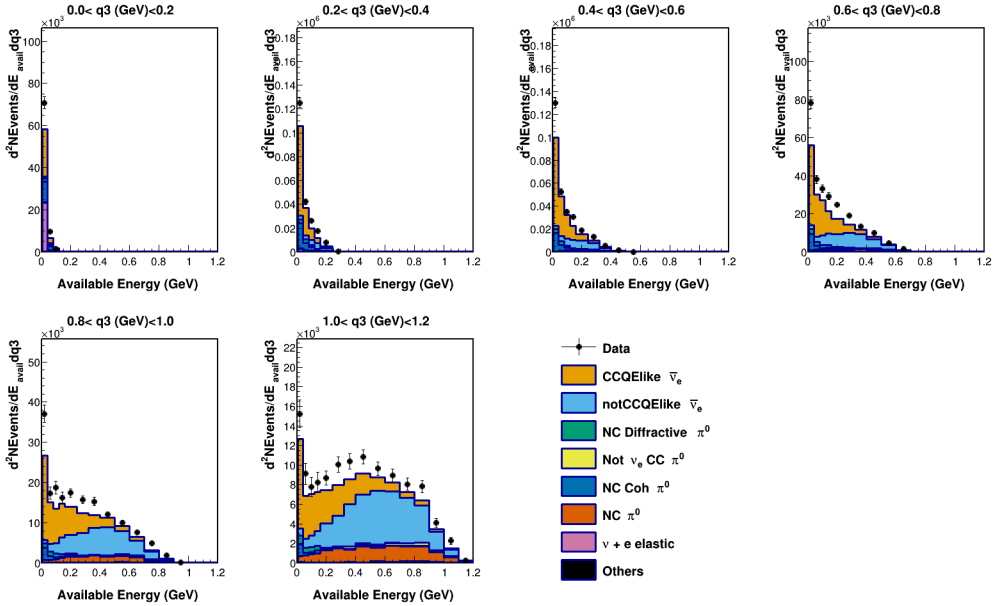


(b) High Inline Posttune E_{avail} vs p_T

Figure 5.40: Pre and Post background tuning in the π^0 sideband (dE/dx $>$ 2.4 MeV/cm, $\psi * E_e > 0.5$ GeV).

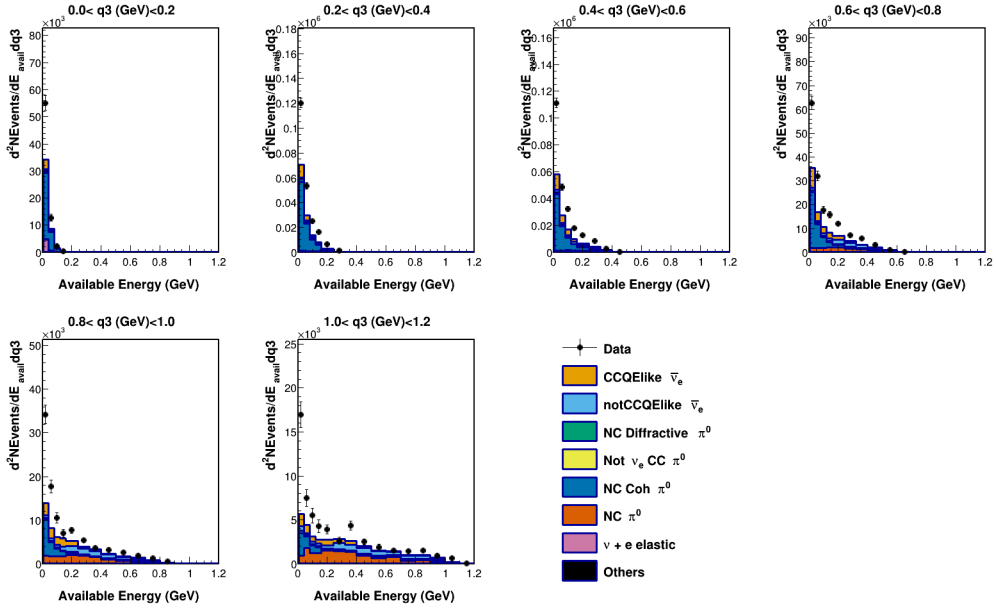


(a) Signal Pretune E_{avail} vs q_3

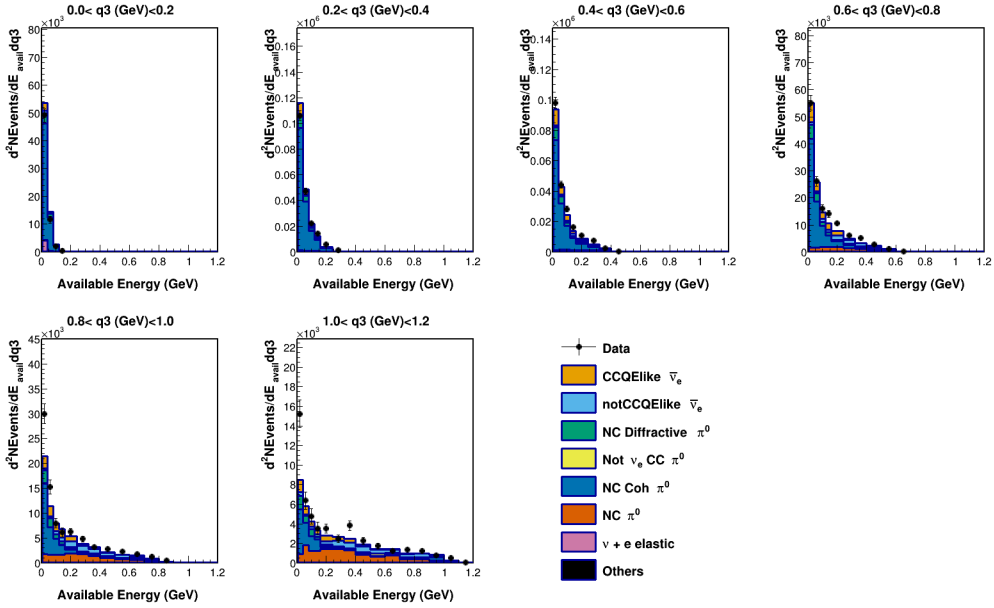


(b) Signal Posttune E_{avail} vs q_3

Figure 5.41: Signal distributions (dE/dx $<$ 2.4 MeV/cm) Pre and Post background tuning. These figures do not have the Signal scaled factor applied.

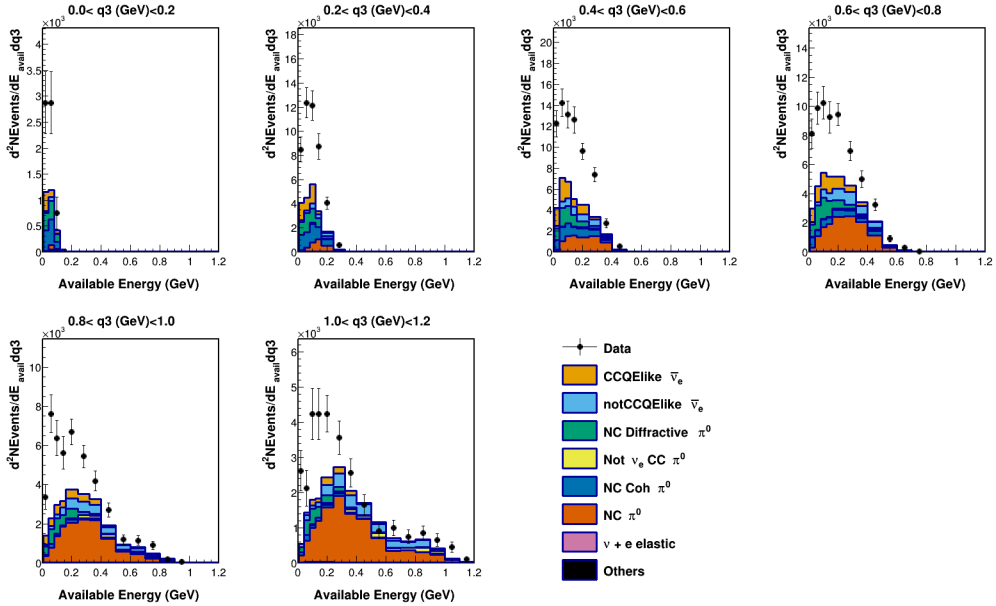


(a) Low Inline Pretune E_{avail} vs q_3

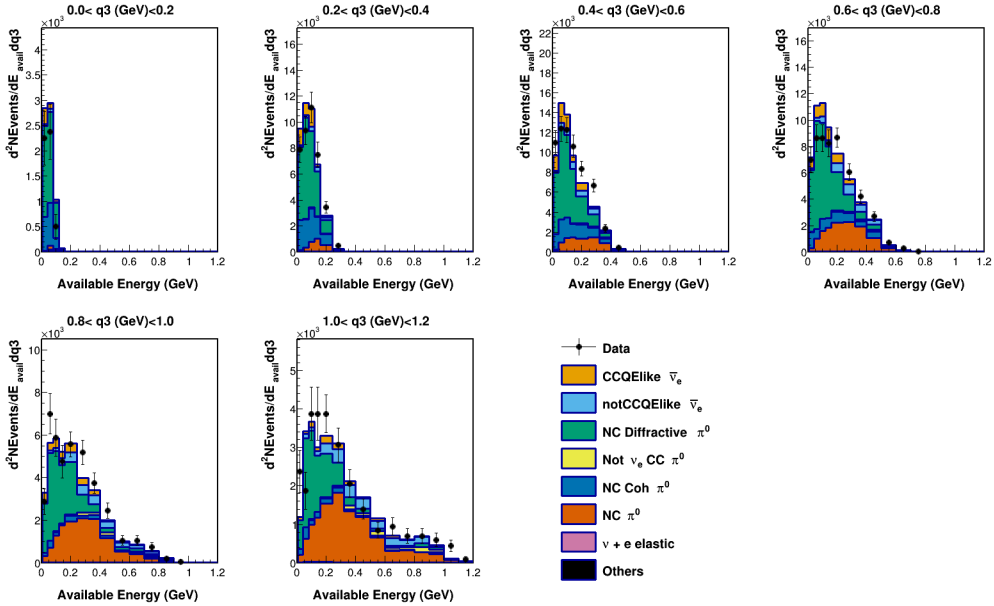


(b) LowInline Posttune E_{avail} vs q_3

Figure 5.42: Pre and Post background tuning in the NC Coherent π^0 sideband (dE/dx $>$ 2.4 MeV/cm, $\psi * E_e < 0.5$ GeV, Inline upstream energy $<$ 10 MeV).

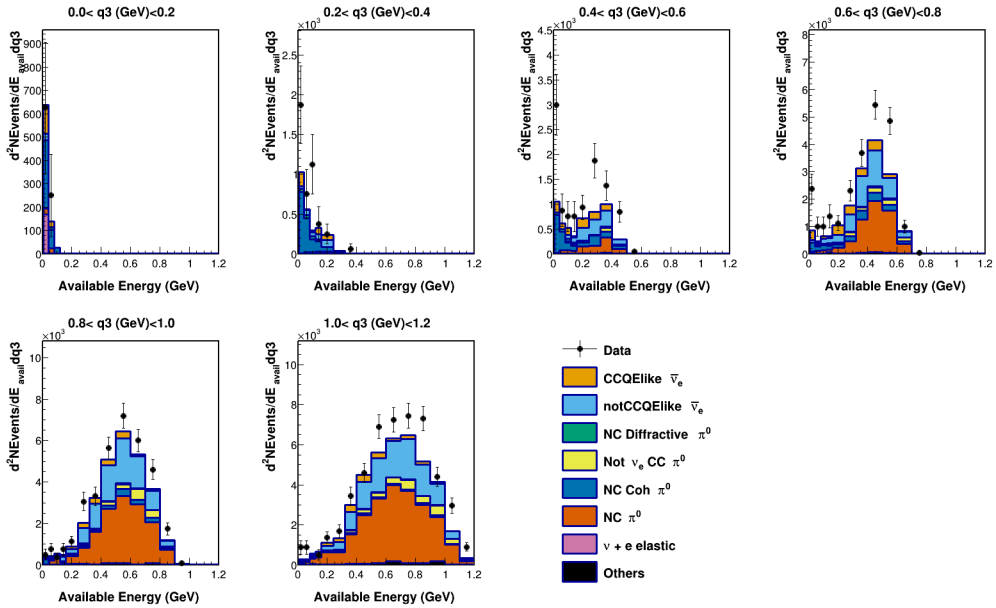


(a) High Inline Pretune E_{avail} vs q_3

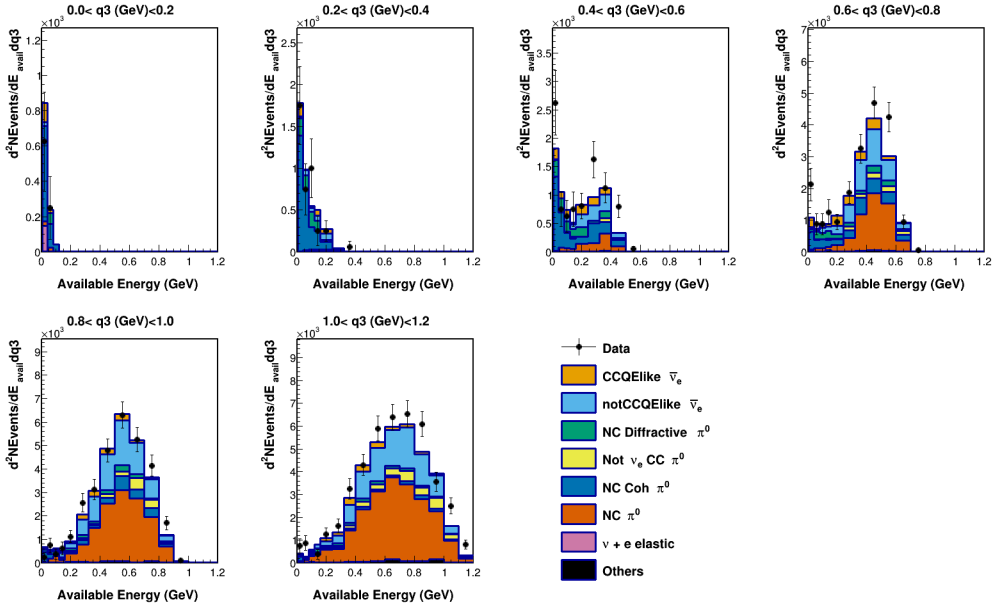


(b) High Inline Posttune E_{avail} vs q_3

Figure 5.43: Pre and Post background tuning in the NC diffractive π^0 sideband (dE/dx $>$ 2.4 MeV/cm, $\psi * E_e < 0.5$ GeV, Inline upstream energy $>$ 10 MeV).



(a) High Inline Pretune E_{avail} vs q_3



(b) High Inline Posttune E_{avail} vs q_3

Figure 5.44: Pre and Post background tuning in the π^0 sideband (dE/dx $>$ 2.4 MeV/cm, $\psi * E_e >$ 0.5 GeV).

6. Double Differential Cross Section

The analysis is performed in the medium energy (ME) beam configuration with total data P.O.T of 1.22e+21 and total MC P.O.T of 4.99e+21. Data and MC are subdivided in to playlists to alleviate the processing load. The playlist POTs are found in Table 6.1. For the simulation RPA, 2p2h and Non-resonant pion weighting are applied (MnvGenie-v2).

Playlist	Data	MC Generic	MC NC Diffractive	Extended 2p2h	High stat ν_e
ME5A	5.45e+19	2.14e+20	1.92e+21	0.87e+21	2.16e+21
ME6A	1.63e+20	6.38e+20	4.68e+21	2.95e+21	5.01e+21
ME6B	1.03e+20	4.16e+20	2.96e+21	1.98e+21	4.08e+21
ME6C	1.10e+20	4.99e+20	3.76e+21	1.98e+21	4.92e+21
ME6D	1.16e+20	4.51e+20	3.60e+21	1.97e+21	4.59e+21
ME6E	8.68e+19	4.04e+20	3.12e+21	1.99e+21	4.12e+21
ME6F	1.40e+20	6.19e+20	2.40e+21	2.99e+21	6.22e+21
ME6G	7.38e+19	2.09e+20	3.20e+21	2.00e+21	4.11e+21
ME6H	1.08e+20	5.09e+20	x	x	4.97e+21
ME6I	7.86e+19	4.09e+20	x	x	3.98e+21
ME6J	9.07e+19	4.08e+20	x	x	4.08e+21

Table 6.1: Protons on target for respective data and MC playlist production. Generic is all interactions (except NC and CC Diffractive π^0). The remainder of the playlists are separate MC productions that are integrated into the analysis: NC diffractive π^0 events, Valencia 2p2h modeled up to 2.0 GeV and a high statistical production of ν_e events created for efficiency corrections.

6.1 Data and MC Sample

In this section we will present two double-differential cross sections of E_{avail} versus q_3 and E_{avail} versus p_T. The formula for calculating a differential cross section is

$$(\frac{d^2\sigma}{dxdy})_{ij} = \frac{\sum_{\alpha\beta} U_{\alpha\beta ij}(N_{data,\alpha\beta} - N^{bkgd}_{data,\alpha\beta})}{\varepsilon_{ij}(\Phi T)(\Delta x_i \Delta y_j)} \tag{6.1}$$

Variable	Definition
x	True bins ij in variable x
y	True bins ij in variable y
$N_{data,\alpha\beta}$	The number of data events in bins $\alpha\beta$
$N^{bkgd}_{data,\alpha\beta}$	Estimated number of background events in reconstructed bins (α,β)
$U_{\alpha\beta ij}$	Simulated migration matrix that maps reconstructed bin (α,β) to true bin (i,j)
ε_{ij}	Product of reconstructed efficiency and detector acceptance for events in true bin (i,j)
Φ	Integrated neutrino flux
T	Number of targets within the fiducial volume
Δx_i	Width of bin i
Δy_j	Width of bin j

Table 6.2: Definitions of the variables measured in the analysis.

6.2 Unfolding

Due to the limitations in detector resolution and inaccuracy of reconstruction techniques, reconstructed events differ from their true kinematic values. As a result, instead of a sharply peaked kinematic distribution at the true value, reconstructed values tend to be distributed with higher variance with respect to the true value. Deviation from the true kinematic value is referred to as “smearing”.

The number of events for a reconstructed bin N_j^{reco} for a given set of true events N_i^{true} can be calculated using a smearing matrix U, also known as a migration matrix calculated as

$$N_j^{reco} = \sum_i U_{ji} N_i^{true}$$

where U is constructed using the distribution of events given their true and reconstructed kinematics. To undo the effects of smearing, events are migrated from reconstructed bin j back to truth bin i in a process called unfolding. In theory, the true distribution can be reconstructed by taking the inverse of U

$$N_j^{true} = \sum_i U_{ji}^{-1} N_i^{reco}$$

However, the inversion of the migration matrix inflates the statistical uncertainty of the unfolded distribution. MINERvA analyses solve this problem by using a numerical iterative process referred to as D'Agostini iterative regularization method [94] [95]. The process of de-smearing the migration matrix will be referred to as "unfolding" and regularization is implemented by truncating the number of numerical iterations.

The method uses a process that estimates the contribution from true bins to reconstructed bins over multiple iterations using a technique similar to how Bayes theorem updates predicted probabilities [96]. In this approach with m number of iterations

$$n^{m+1}(C_i) \approx \sum_j M_{ij} * n_i^m \tag{6.2}$$

$$M_{ij} = P(C_i|E_j) * P_0(C_i) \tag{6.3}$$

for $P(C_i|E_j)$ the probability of reconstructing in true bin i given reconstructed bin j. With each unfolding iteration, the modeling bias is reduced but the statistical uncertainty increases. An infinite number of iterations is mathematically equivalent to taking the inverse of the migration matrix. A divergence in the χ^2 value with increasing iterations means either that the distribution being modeled is too dissimilar from the migration matrix or there isn't enough precision in low populated bins. The number of unfolding iterations serves as a regulation parameter, controlling a trade-off between bias and variance. With each additional iteration, the bias is reduced but the variance is increased.

For a 2D analysis such as this, the unfolding process requires a 4D unfolding matrix, $U_{\alpha\beta ij}$. $U_{\alpha\beta ij}$ is represented as a histogram that maps the x-axis corresponding to the bins $\alpha\beta$ for which events are reconstructed to the y axis corresponding to the bins ij for which events are generated. A single bin within the unfolding matrix represents a relation between a bin of reconstructed E_{avail} and p_T (q_3) and a bin of true E_{avail} and p_T (q_3).

The migration matrix for E_{avail} vs p_T is shown in Figure 6.1 and the migration matrix for E_{avail} vs q_3 is shown in Figure 6.2. The x-axis shows the bins for the reconstructed values and the y axis shows bins for true values. The subplots occur on the x-axis where each bin of E_{avail} is repeated for each bin of p_T (q_3). The-y axis follows the same binning scheme for truth bins. Events populate the diagonal of the matrix when they are reconstructed into the correct truth bin for both E_{avail} and p_T (q_3). If an event is reconstructed into the wrong bin of E_{avail} but the corrected bin of p_T (q_3), then that event will be in a different bin within the same subplot. In the case that an event is reconstructed into the correct bin of E_{avail} but the wrong bin of p_T (q_3), then the event will be in the same relative position in a different subplot [45]. The migration matrix is row normalized such that the sum of each row is 100.

Therefore, each cell within the matrix represents the probability of an event being produced for a given E_{avail} and p_T (q_3) to be reconstructed into each respective E_{avail} and p_T (q_3) bin.

It is simpler to understand the 2D projections of the unfolding matrix. A value reconstructed to its true value is represented on the diagonal line in a migration matrix. The farther away from the diagonal, the larger the smearing of the reconstruction variable. The 2D projection is found by summing over all values of E_{avail} for a given q_3 (p_T) bin.

The q_3 migration matrix found in Figure 6.3 is relatively diagonal, though it has some populations of events immediately above and below the matrix diagonal. This is expected because, as discussed in Section 3.5.2, the kinematic variable q_3 is reliant upon the calorimetrically corrected variable q_0, making it model dependent. The p_T migration matrix found in Figure 6.4 is more diagonal, indicating little smearing between bins. The E_{avail} migration matrix found in Figure 6.5 has a large bias at high true E_{avail} and reconstructed E_{avail} for reasons discussed in Section 3.5.3. As shown, the migration matrix in E_{avail} has poor resolution compared to q_3 or p_T. Therefore, it is anticipated that unfolding the 4D migration matrix of E_{avail} vs q_3 and E_{avail} vs p_T might require a higher number of iterations to unfold the kinematics.

Suppose the migration matrix is a highly accurate representation of the smearing between true and reconstructed kinematics and the signal model is correct. In that case, the D'Agostini unfolding method will unfold the simulated data in a single iteration. In reality, signal models don't precisely predict signal data and more iterations are required to unfold data distributions correctly. The same simulation used to create the migration matrix cannot also be used to determine the number of unfolding iterations. Therefore, the

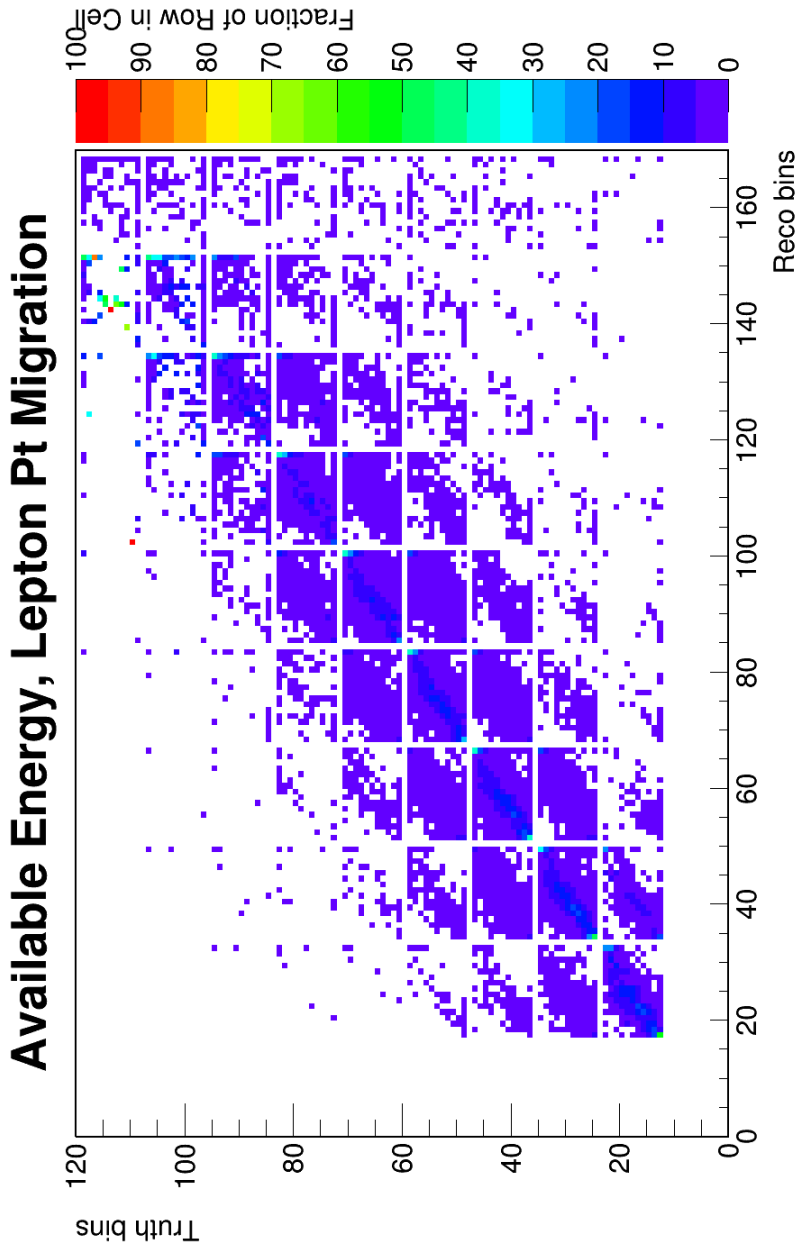


Figure 6.1: Migration matrix for E_{avail} vs p_T distribution. The x axis is reconstructed bins and the y is for true bin.

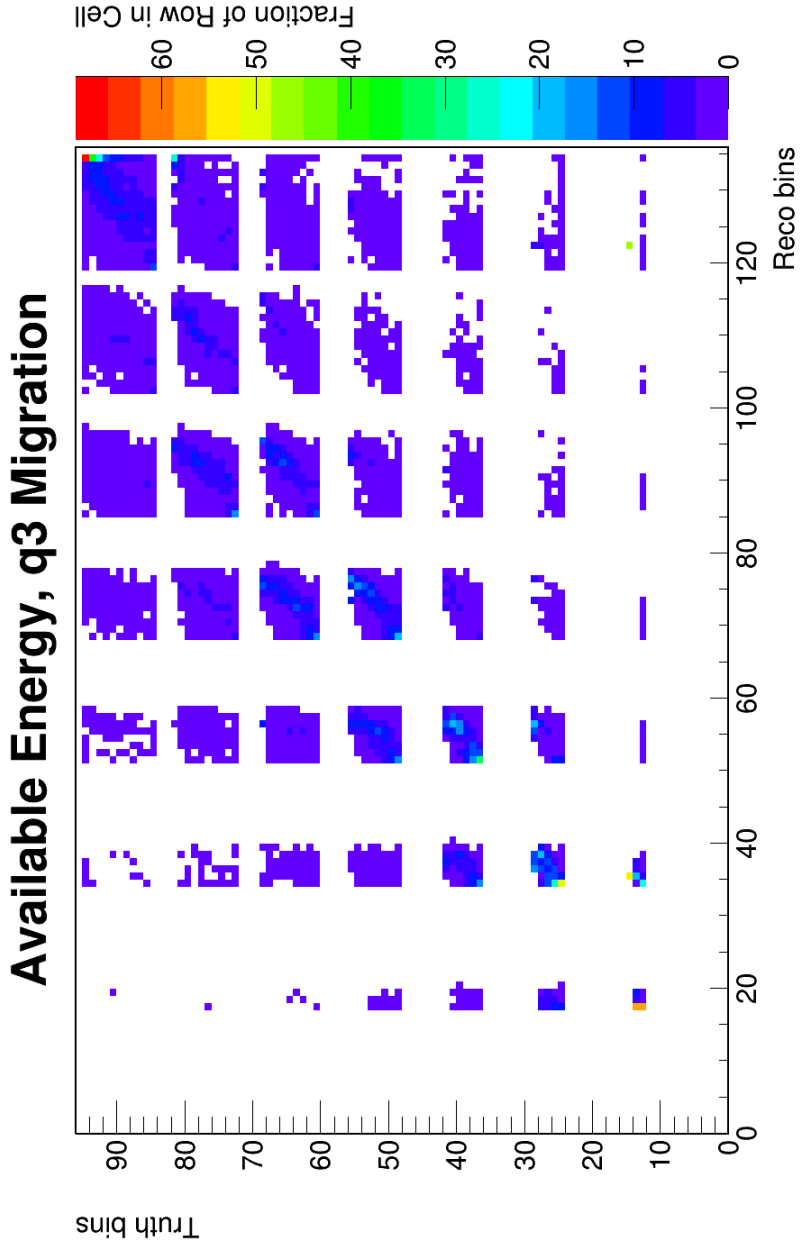


Figure 6.2: Migration matrix for E_{avail} vs q_3 distribution. The x axis is reconstructed bins and the y axis is for true bins.

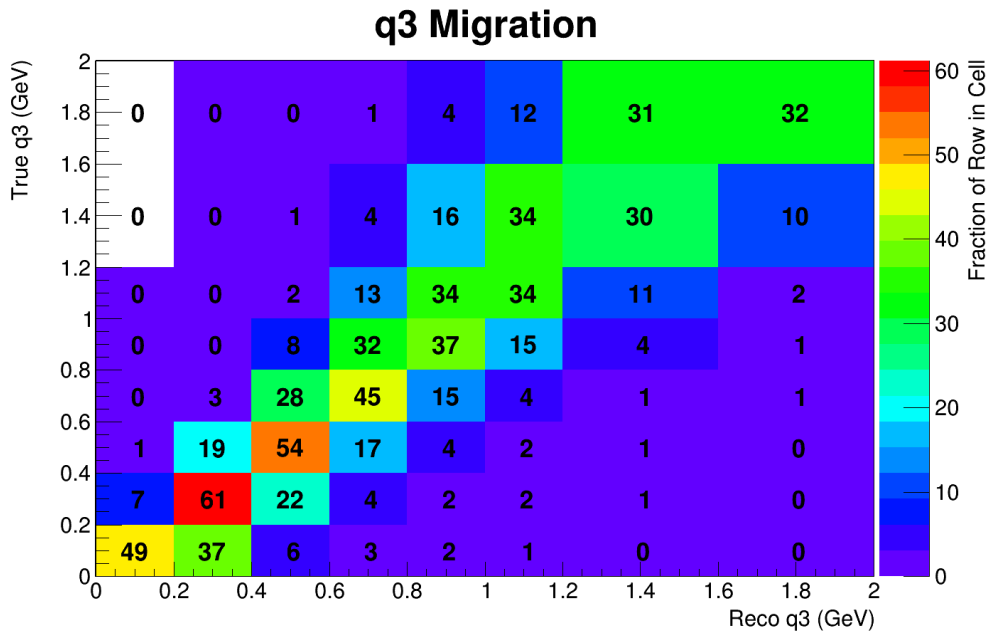


Figure 6.3: Migration matrix for E_{avail} vs q_3 distribution. The x axis is reconstructed bins and they axis is for true bins.

unfolding procedure uses a variety of samples from alternative models and the number of unfolding iterations is chosen based on stability in the difference between the unfolded and true distributions for each model.

6.2.1 Warping Studies

The determination of the number of unfolding iterations is done through what's called a "warping" study. The warping test is a fake data test that evaluates the bias and variance by unfolding an ensemble of measured histograms resulting from a true distribution that is known but different from a prior distribution. The MC simulation is reweighted to produce alternative models where the newly "warped" reconstructed histograms are then compared

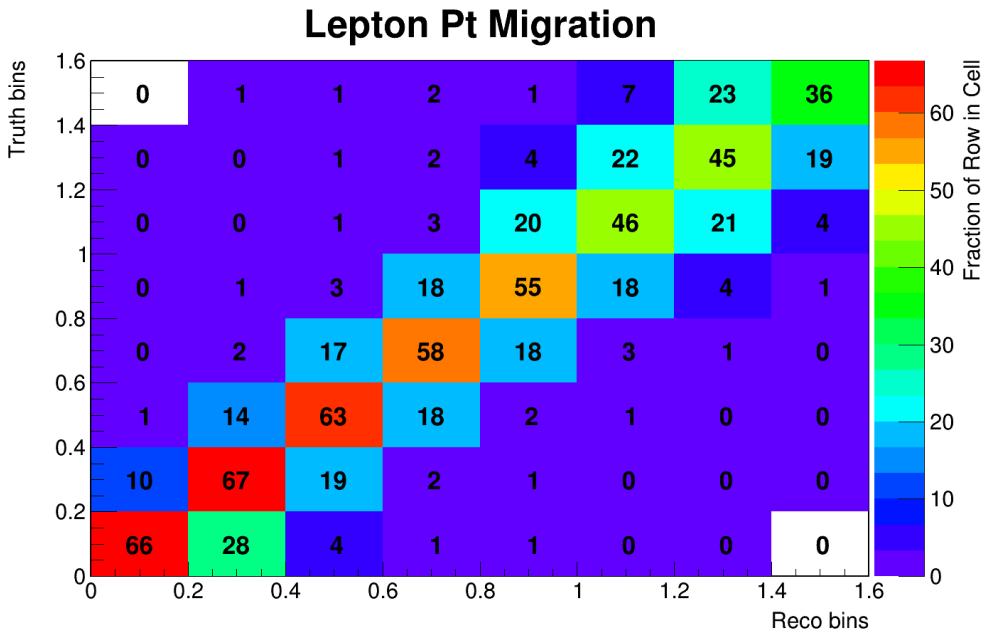


Figure 6.4: Migration matrix for E_{avail} vs q_3 distribution. The x axis is reconstructed bins and the y axis is for true bins.

to true histograms.

A variety of reweightings are applied to the alternative models. Certain reweightings aim to produce the distributions of the alternative models. Other reweightings aim to make the models more reflective of the data in certain variables. Additionally, models are built by exaggerating some features of the model to test the sensitivity of unfolding that feature.

A summary of the reweights included in this analysis is summarized in Table 6.3. The warped reconstructed distributions are unfolded using the original migration matrix and are compared to warped true values using a χ^2 calculation as a function of unfolding iteration. The warped distribution is unfolded in many fake data sets with independent statistical fluctuations. The statistical fluctuations are simulated by randomly setting the bin contents

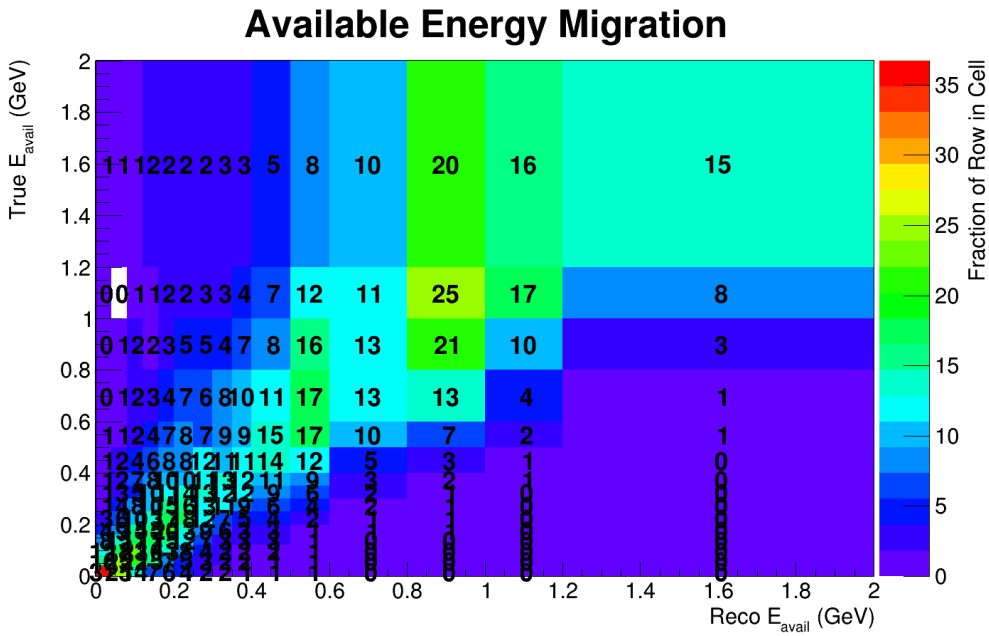
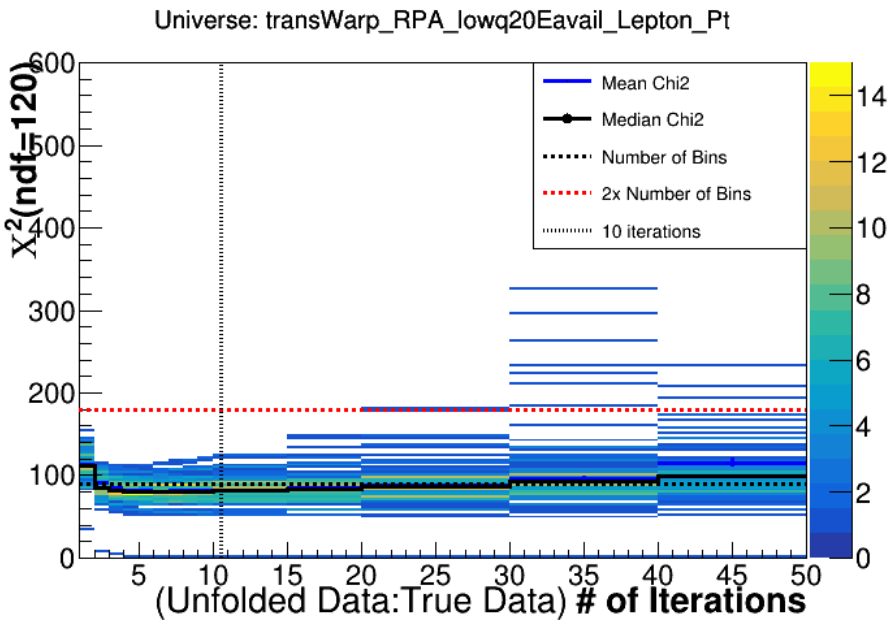


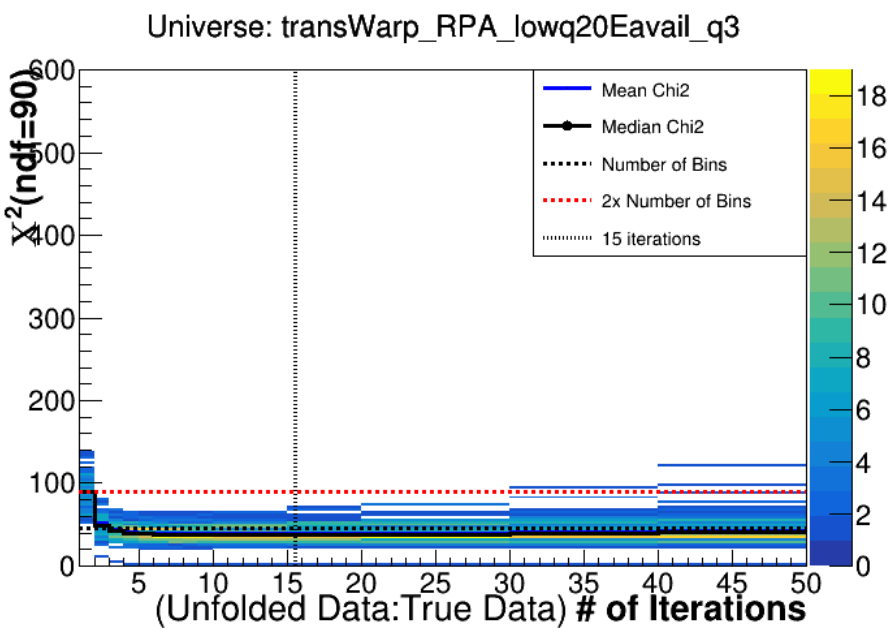
Figure 6.5: Migration matrix for E_{avail} vs q_3 distribution. The x axis is reconstructed bins and the y axis is for true bins.

using a Poisson distribution whose mean is the original content of the bin. The chosen number of iterations used to unfold the unwarped distribution for the cross section is the iteration with the lowest median χ^2 value across all model warpings.

The warping study results are shown in Figures 6.6-6.9. The warping results show the distribution of χ^2 among pseudo experiments, as well as the median and mean of the distributions, as a function of number of iterations. The dashed lines denote the expected effect of the statistical uncertainties in the case of perfect unfolding. The red dashed line represents a threshold where the unfolded model is acceptably close to the true model. The black dashed line is the expected mean χ^2 value when the unfolded model has converged to the true model. The mean (or median) value of χ^2 represents the bias since it is a measure

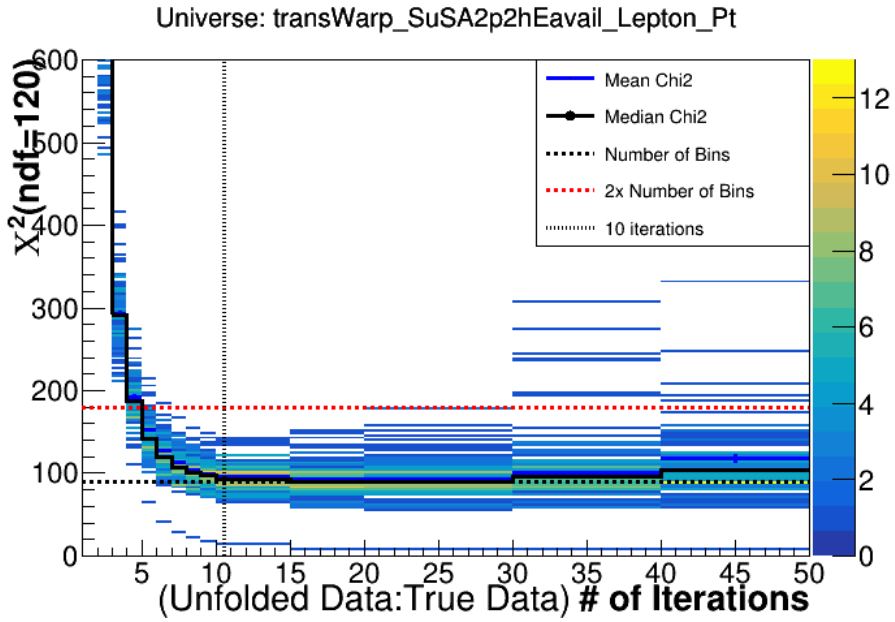


(a) E_{avail} vs p_T

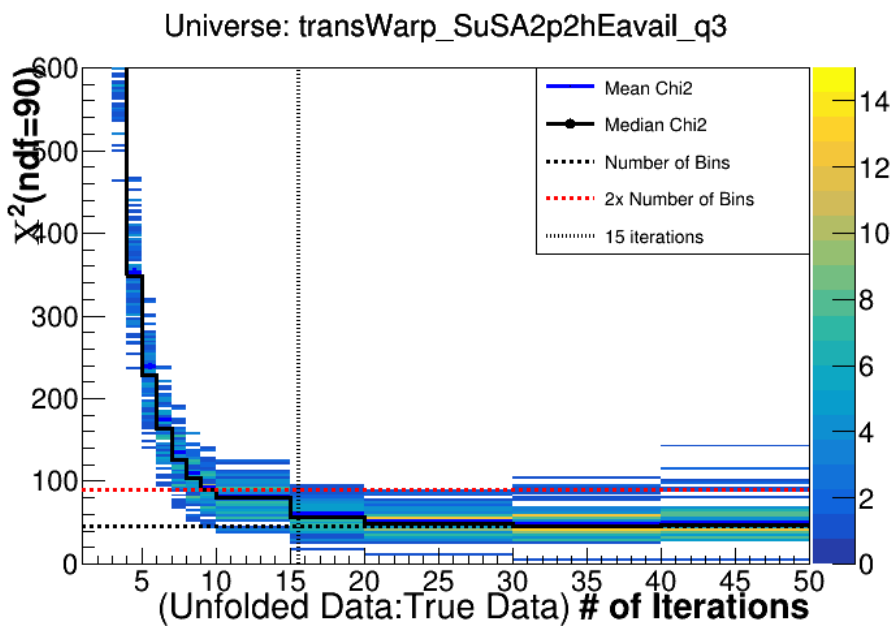


(b) E_{avail} vs q_3

Figure 6.6: Warping study using the RPA model.

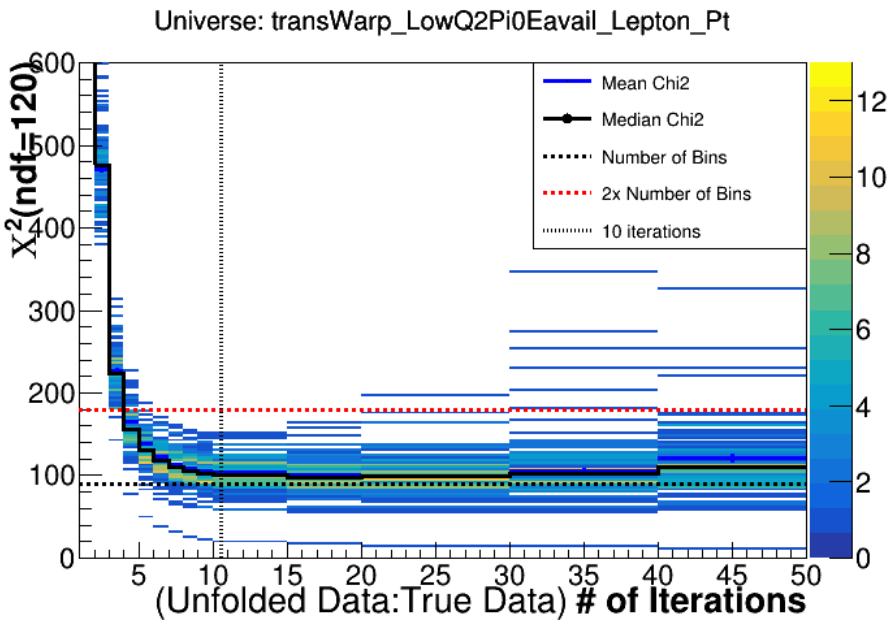


(a) E_{avail} vs p_T

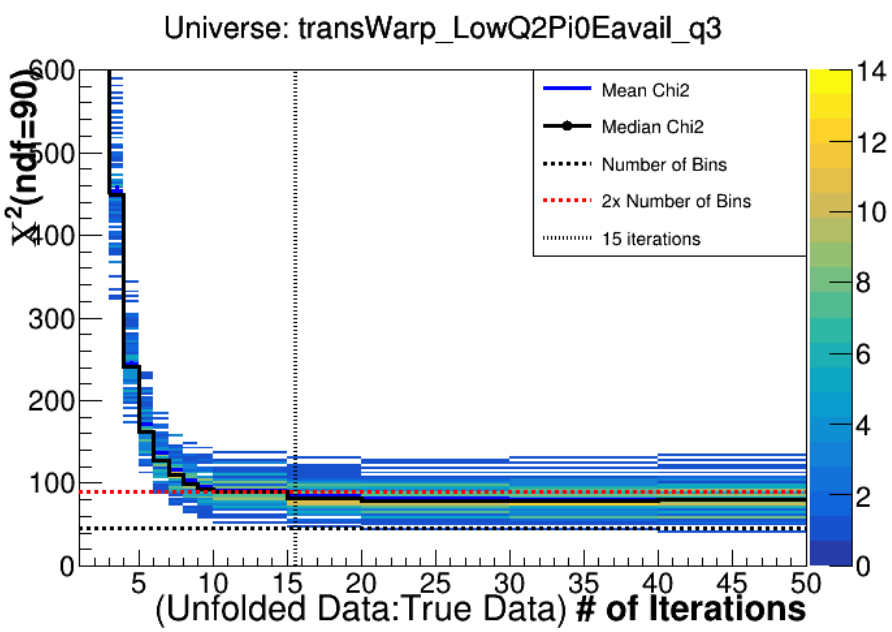


(b) E_{avail} vs q_3

Figure 6.7: Warping study using the SuSA model.

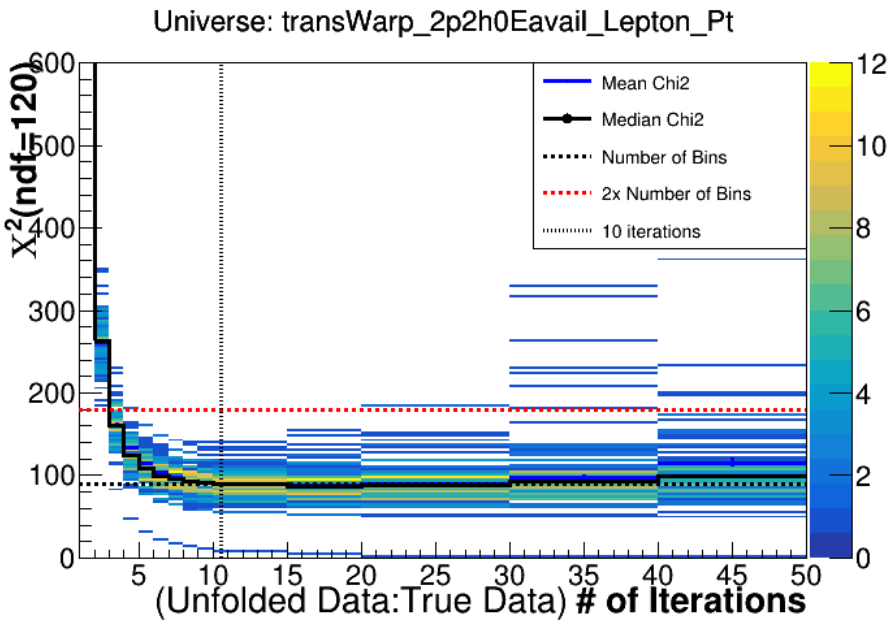


(a) E_{avail} vs p_T

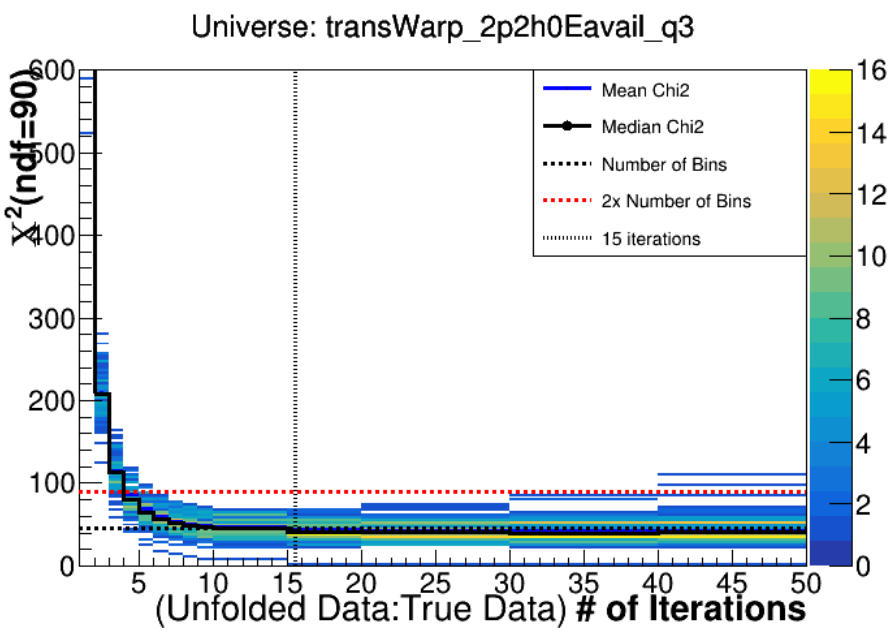


(b) E_{avail} vs q_3

Figure 6.8: Warping study using the Low Q2 model.



(a) E_{avail} vs p_T



(b) E_{avail} vs q_3

Figure 6.9: MINERvA 2p2h Tune Variant: nn pair.

Model	Description
RPA Warp	The Valencia model includes a random phase approximation correction for QE interactions, which is not available in the GENIE model at the time of simulation. MINERvA reweights QE events as function of 4-momentum transfer to approximate Valencia model prediction as mentioned in Section 3.7.1.
2p2h Warp	A MINERvA low-recoil analysis found a data excess in the dip region as mentioned in Section 3.7.1. MINERvA reweights 2p2h events by a 2D Gaussian function of q_0, q_3 to fill in the gap, resulting in roughly 50% increase of simulated 2p2h event rate.
Non-Resonant Pion Warp	The nonresonant pion production weight determined in the deuterium bubble chamber fits, described in Section 3.7.2.
Resonant Pion Warp	MINERvA implemented an *ad hoc* Q2 dependent tuning of resonant pion production to improve data and MC predictions in pion production cross sections [69].

Table 6.3: List of warp models used in the unfolding study.

of the difference between the unfolded model and the measured model. The chosen number of unfolding iterations should be large enough to reduce the mean χ^2 to be statistically consistent with the correct model while small enough to avoid uncertainty blowing up. However, this may not always be possible.

The low recoil analysis is sensitive to changes in the 2p2h model and gives a large χ^2 value for relatively mild model variations. Several previous distributions show an increase in the χ^2 for higher numbers of unfolding distributions. This is often caused by the Poisson throw placing the content of a bin in the distribution near zero, which affects the unfolding process. The large value of median χ^2 at a large number of iterations is more noticeable in unfolded E_{avail} vs q_3 distributions because of the low number of events for high E_{avail}

vs q_3 bins. Therefore, the result of the warping study indicates that a different number of unfolding iterations are required for the two different distributions based on the different minimum χ^2 values. It is decided that E_{avail} vs p_T will be unfolded with 10 iterations and E_{avail} vs q_3 unfolded with 15 iterations.

6.3 Efficiency Correction

Once the distributions have been unfolded to the best estimate of their true values, we have to apply a correction to account for the true signal events that were either not reconstructed or did not pass the selection cuts. The efficiency correction is determined by dividing the number of observed events by the selection efficiency. For bins i, j the selection efficiency is given by

$$\varepsilon_{ij} = \frac{N_{ij}^{\alpha}}{N_{ij}^{\beta}} \tag{6.4}$$

N_{ij}^{α} is referred to as the efficiency numerator and is found by selecting all events in the reconstructed MC that pass signal cuts. N_{ij}^{β} is referred to as the efficiency denominator and is populated with generated signal events before they go through the reconstruction process. Therefore, the efficiency denominator is independent of event reconstruction.

In this analysis, a few bins on the edge of the phase space contain low statistics, leading to large statistical uncertainty in the efficiency estimation. To account for this, the efficiency in the low statistic bins are estimated by the average of adjacent bins. Figure 6.10 shows calculated efficiency as a function of E_{avail} in bins of q_3. Figure 6.11 shows the calculated efficiency as a function of E_{avail} in bins of p_T. In both cases, the efficiency decreases at higher E_{avail} values. This is most likely because it is more difficult for the tracking

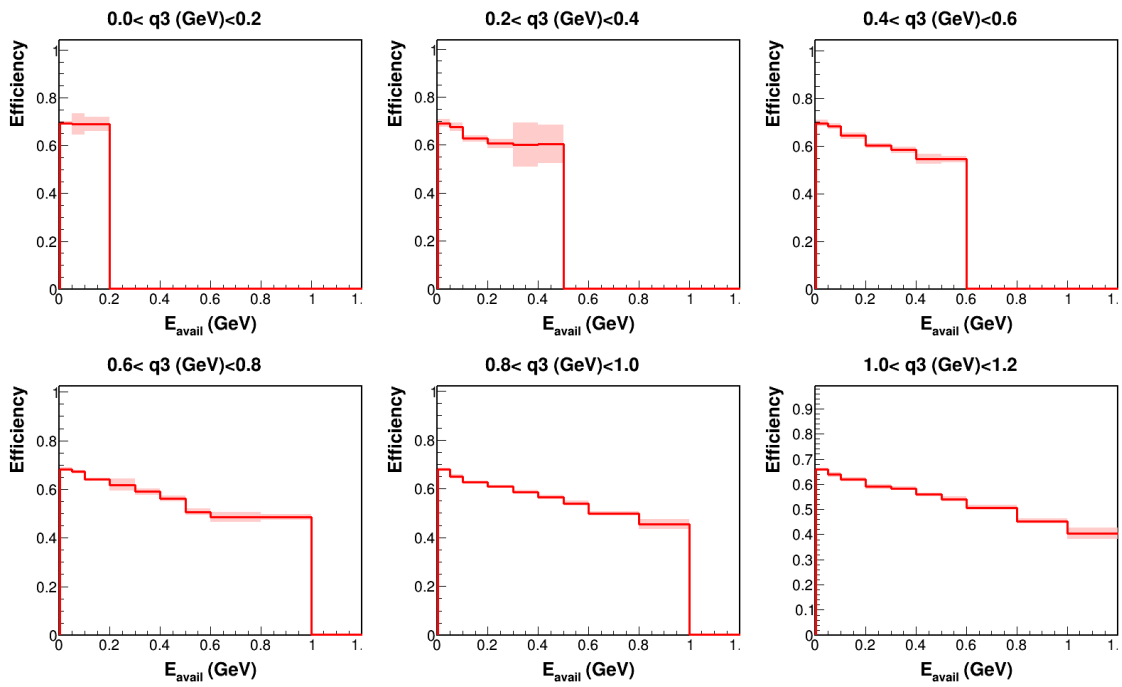


Figure 6.10: Efficiency for E_{avail} vs q_3 distribution. The x axis is reconstructed bins and the y axis is for true bins.

algorithm to reconstruct a proper electron candidate track at higher E_{avail} due to the greater amount of hadronic activity overlapping with EM showers.

6.4 Cross Section Normalization

The final steps in the cross section extraction are to divide the unfolded and efficiency corrected signal data events by the integrated anti-electron neutrino flux and the number of targets in the fiducial volume.

The targets for neutrino interactions are the nucleons within the fiducial volume. Using the detector geometry and material modules, the number of targets is estimated to be

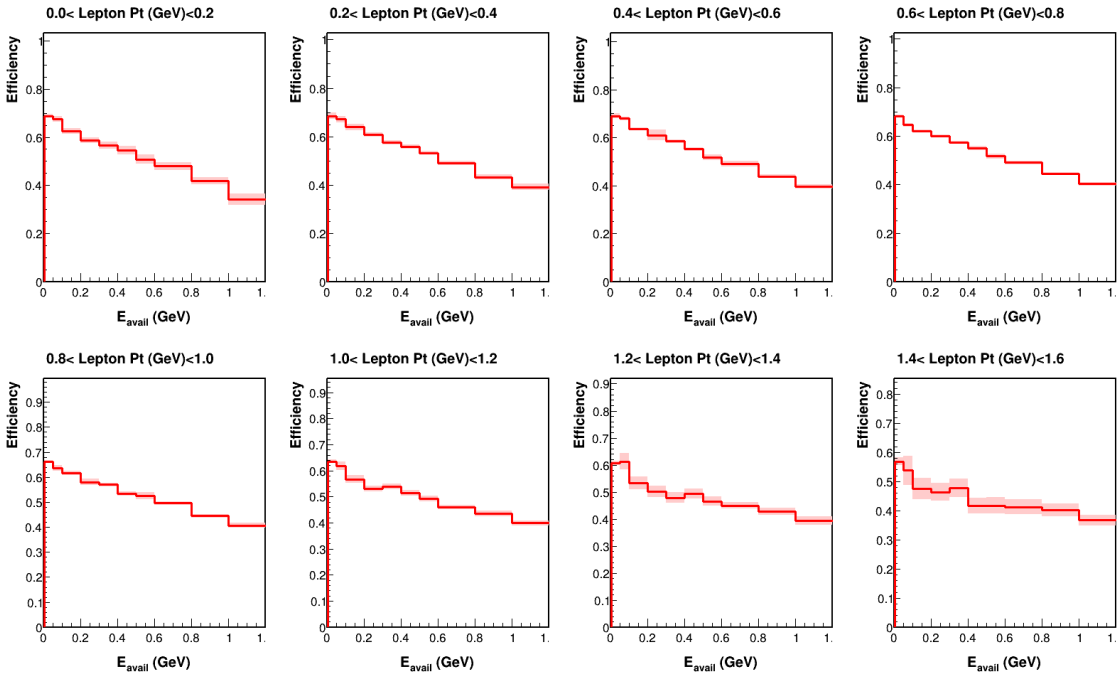


Figure 6.11: Efficiency for E_{avail} vs q_3 distribution.

3.234×10^{30} nucleons [40]. The uncertainties on the number of nucleons within the fiducial volume is propagated to the differential cross section through the target normalization, as discussed in Section 7.4.4.

Target	Num of Nucleons	data POT	MC POT	Integrated Flux ($\bar{\nu}_e/cm^2/POT$)
Tracker	3.234×10^{30}	1.22×10^{21}	4.99×10^{21}	4.68×10^{-10}

Table 6.4: Cross section normalization values for the tracker region

The integrated flux is estimated by simulation and corrected by the flux constraints described in Section 2.3.2. The flux normalization takes the integral of the predicted flux between the values of E_ν from 0 to 100 GeV and divides the efficiency corrected distribution by the integrated flux prediction, scaled by the number of protons on target in the data. The total integrated flux value is 2.34×10^{12} cm^{-2} giving an integrated flux per POT value of 4.68×10^{-10} $\bar{\nu}_e/cm^2/POT$. The uncertainty on the integrated flux will be further discussed in Section 7.4.1.

6.5 Double Differential Cross Section

The final cross section results are presented in this section. Figures 6.12 and 6.13 show the cross section in bins of E_{avail} vs q_3 and Figures 6.15 and 6.16 show the cross section in bins of E_{avail} vs p_T. Cross sections are shown on a linear and log y scale. The ratio of data to the reference cross-section prediction used in the simulation for the cross section results is shown in Figures 6.14 and 6.17. Figures 6.18 and 6.19 show the simulated contribution to the cross section from various channels as defined by Table 6.5. Figures 6.23 and 6.20 show the fractional contribution of signal processes to the cross section.

The error bars are statistical only on the cross-section prediction from the model, and

these are small because there is far more simulated events used to predict this cross section than there are events in the data. The measured cross sections from the data contain both statistical and systematic uncertainties. The systematic uncertainties on the measured cross section are discussed in Chapter 7 with the interpretation of the results discussed in Section 8.1.

Interaction Type	Definition
CC ν_e - QE	Charged current quasielastic scattering
CC ν_e - Res	Charged current resonant pion production from σ decay
CC ν_e - DIS	Charged current deep inelastic scattering
CC ν_e - 2p2h	Charged current two particle two hole
CC ν_e - Other	Other mainly encompasses charged current coherent pion production

Table 6.5: Definitions for the signal component breakdown.

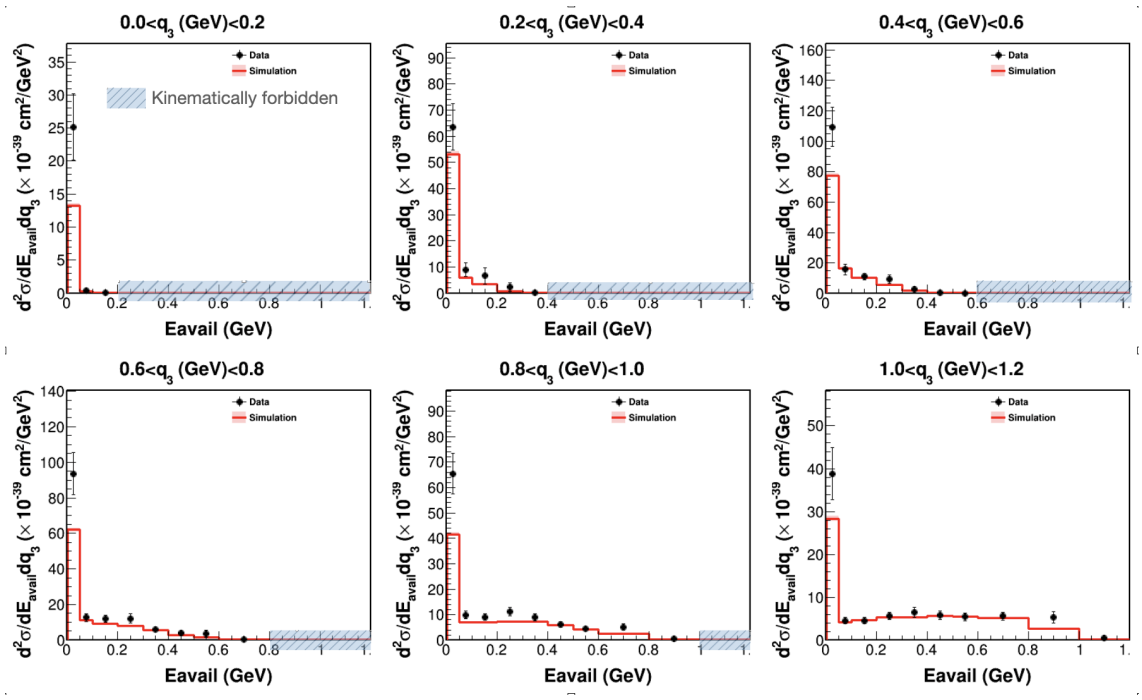


Figure 6.12: Cross section result in E_{avail} vs q_3. The blue regions in each q_3 bin are kinematically forbidden.

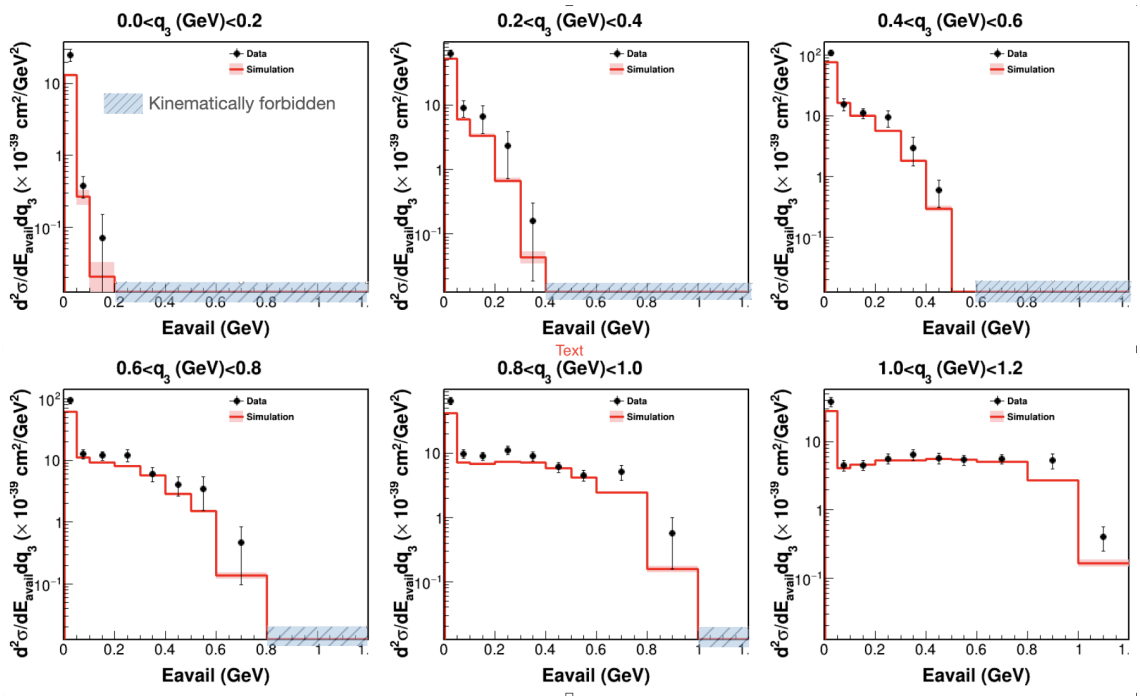


Figure 6.13: Cross section result in E_{avail} vs q_3. The y axis is on a log scale truncated at $10^{-}2$ to enable a better view of the tail end of the cross section.

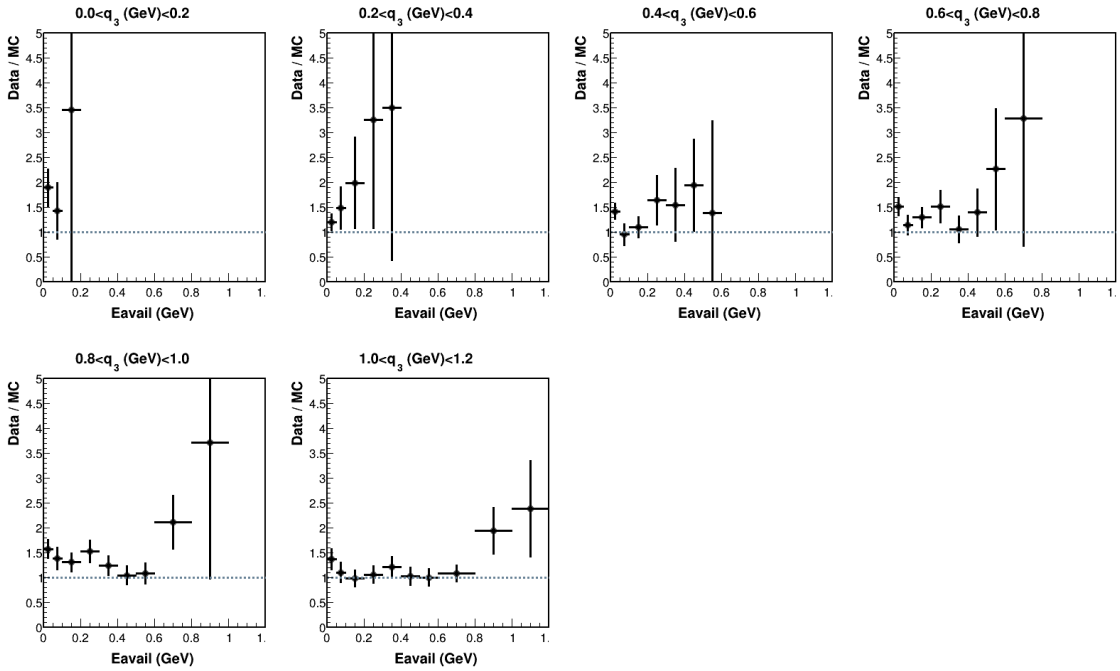


Figure 6.14: Data/MC cross section ratio for E_{avail} vs q_3 distribution.

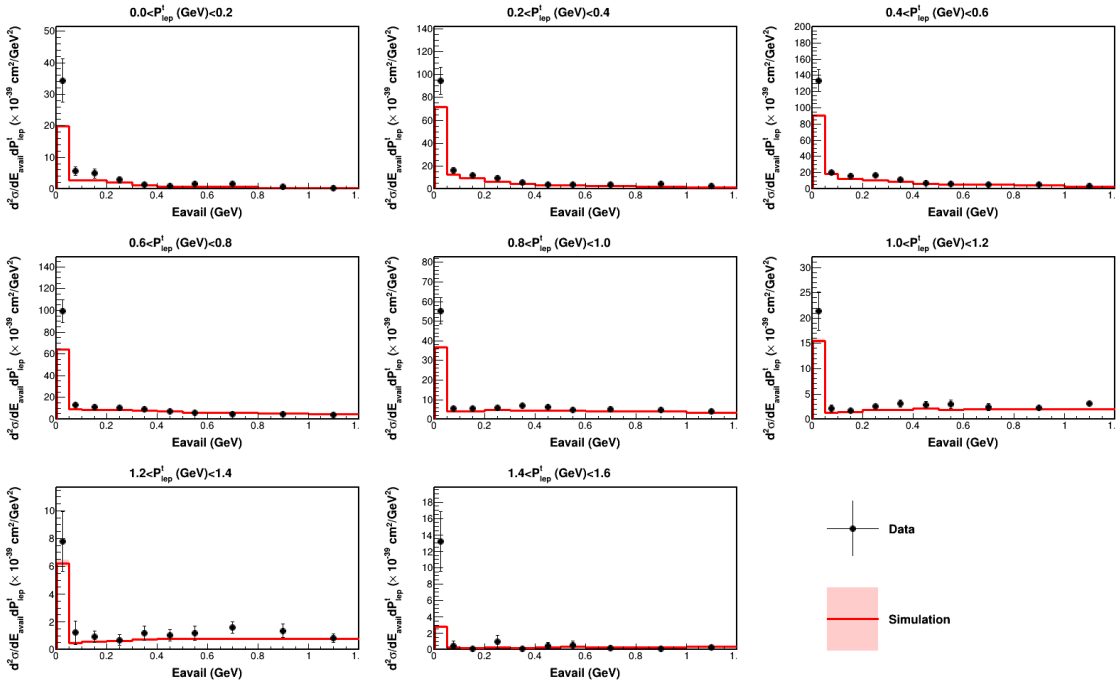


Figure 6.15: Cross section result for E_{avail} vs p_T distribution.

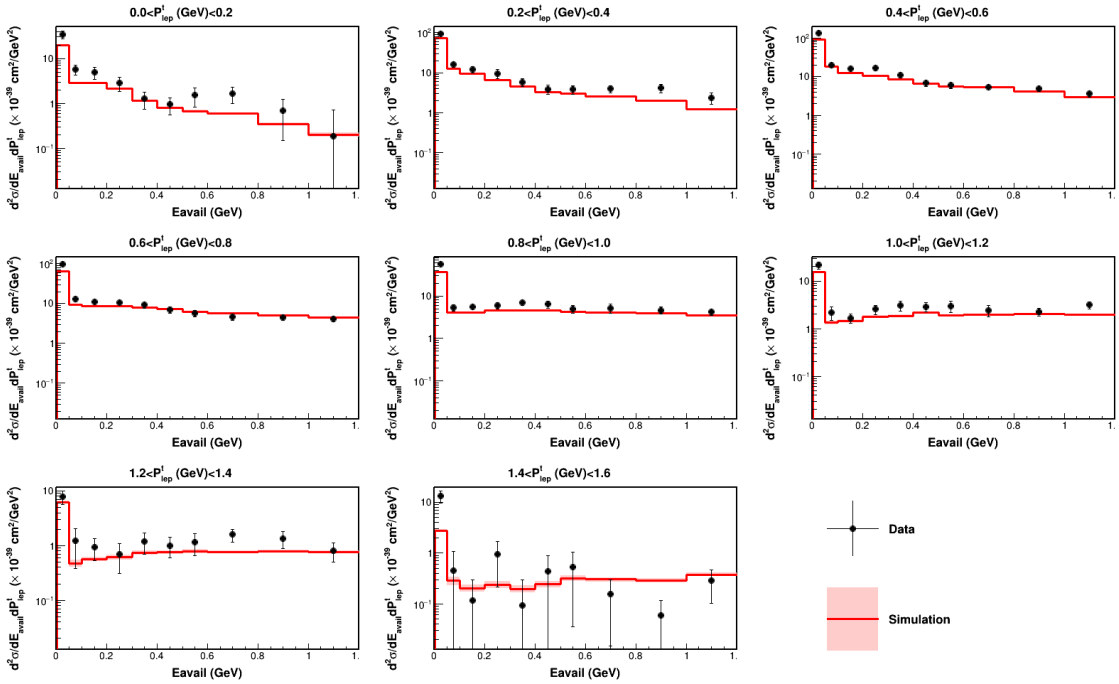


Figure 6.16: Cross section result for E_{avail} vs p_T distribution shown on a log y scale.

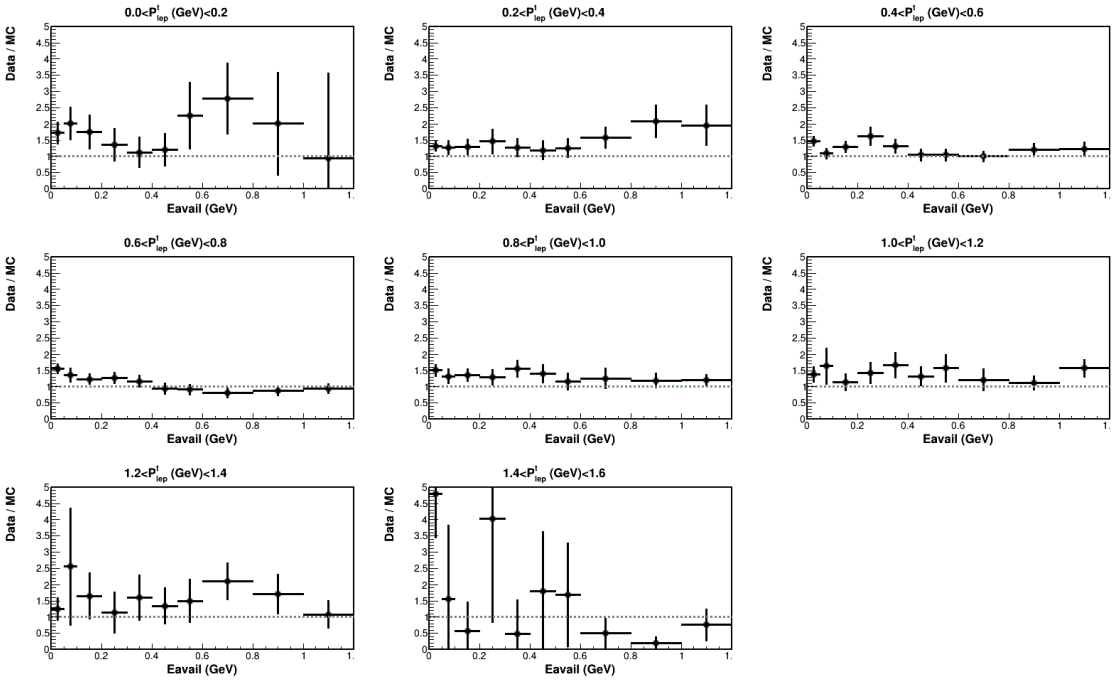


Figure 6.17: Data/MC cross section ratio for E_{avail} vs p_T distribution.

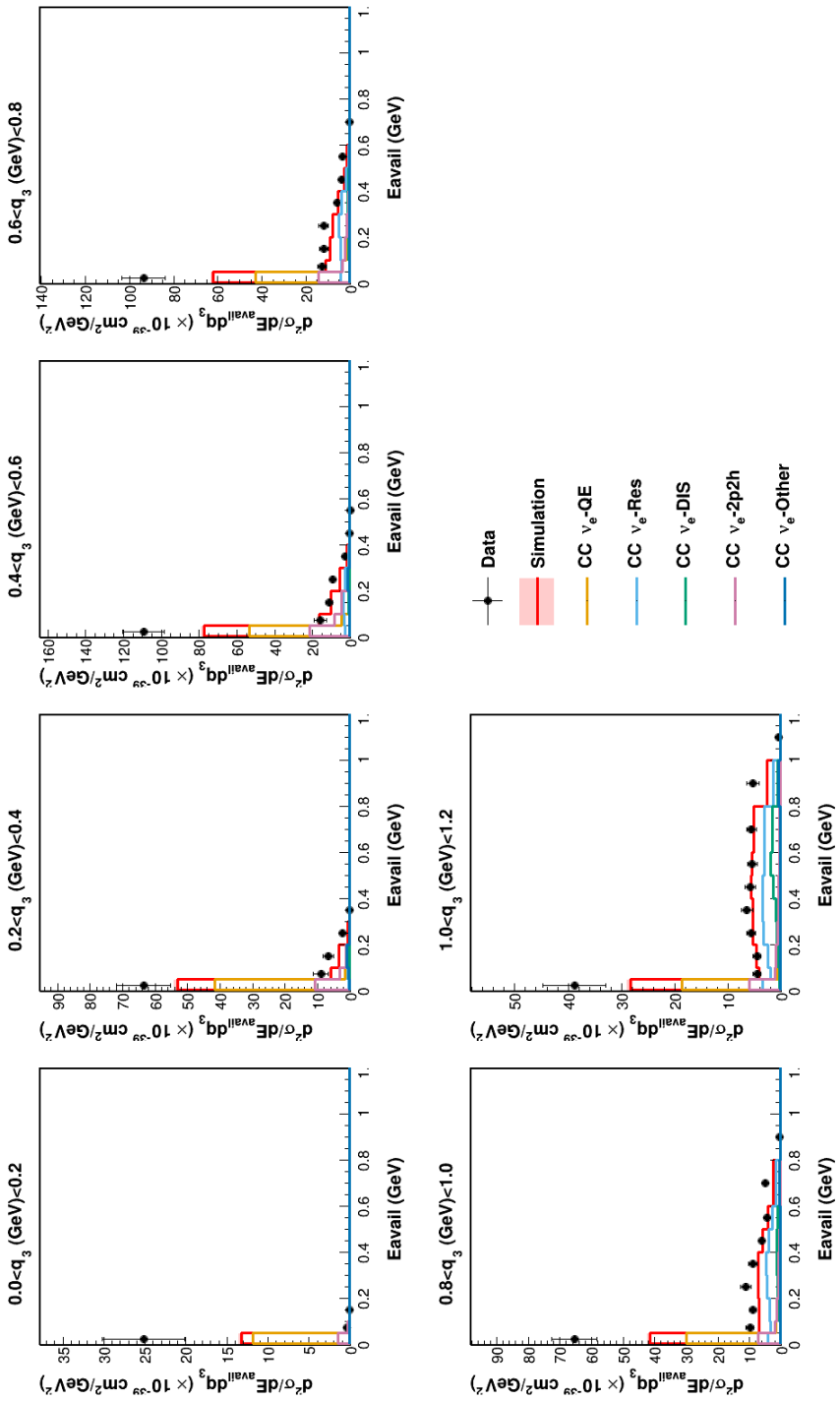


Figure 6.18: A decomposition of the cross section result into contributing interaction types in E_{avail} vs q_3.

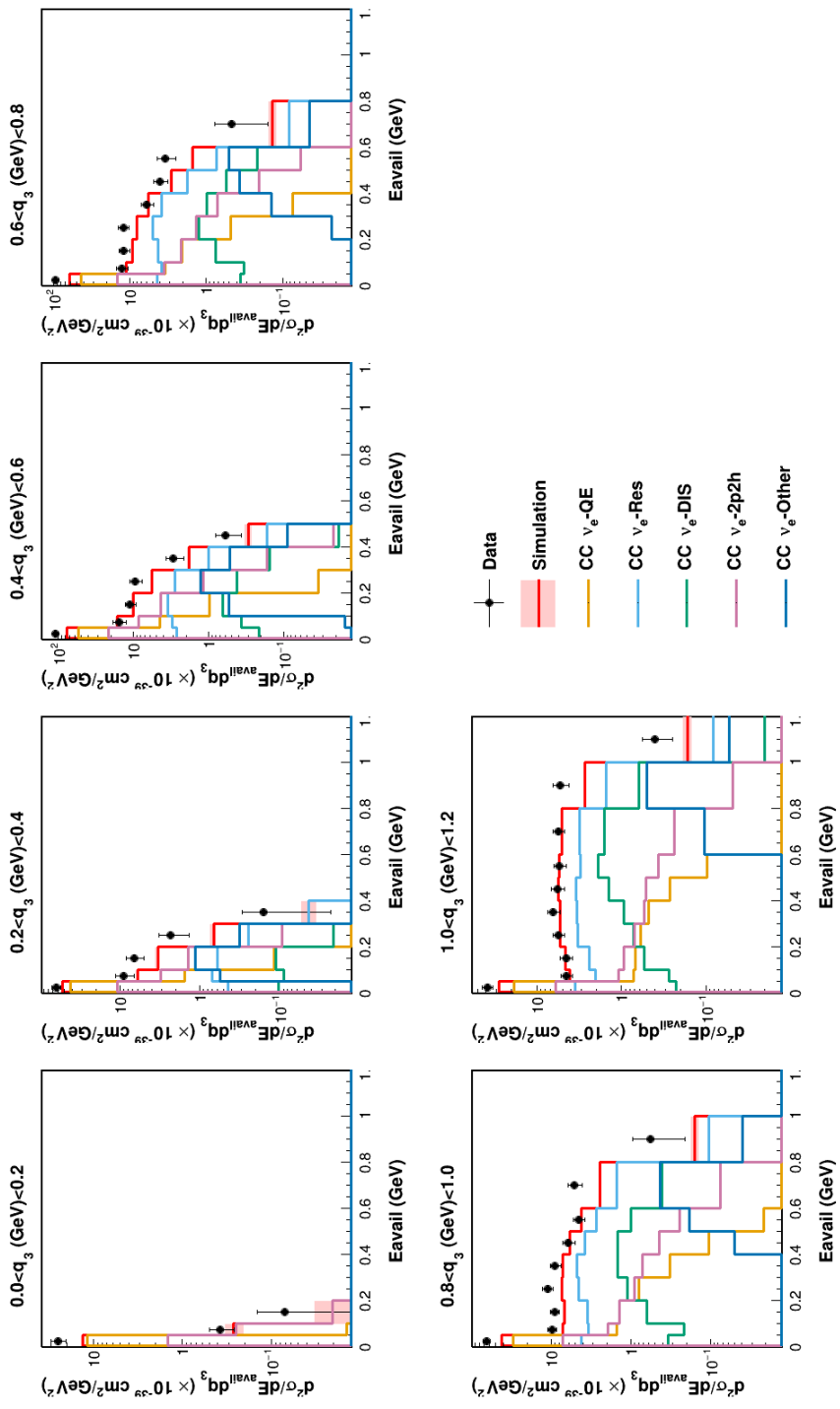


Figure 6.19: A decomposition of the cross section result into contributing interaction types in E_{avail} vs q_3 on a y log scale.

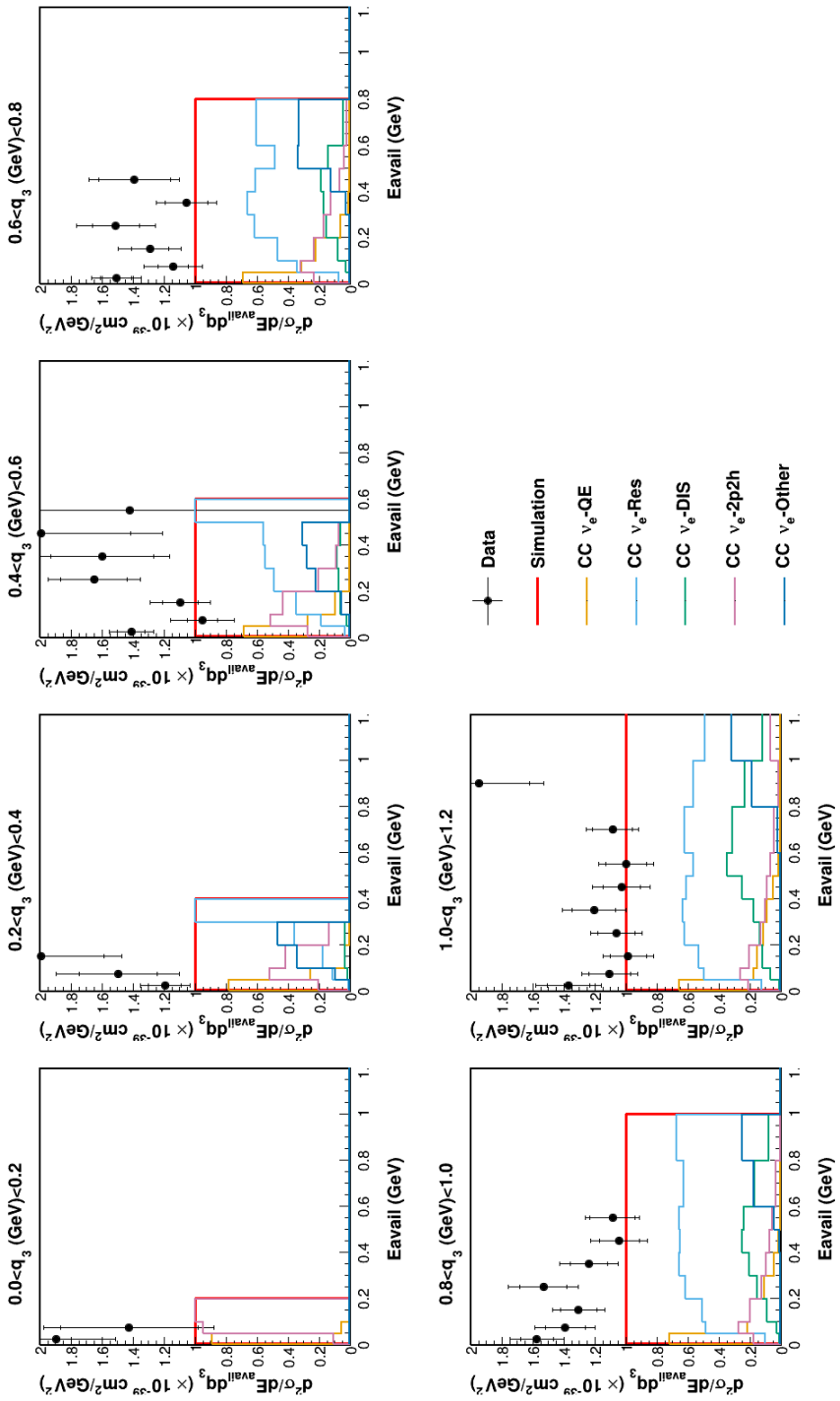


Figure 6.20: Data/MC ratio including fractional contribution of signal processes shown in E_{avail} vs q_3.

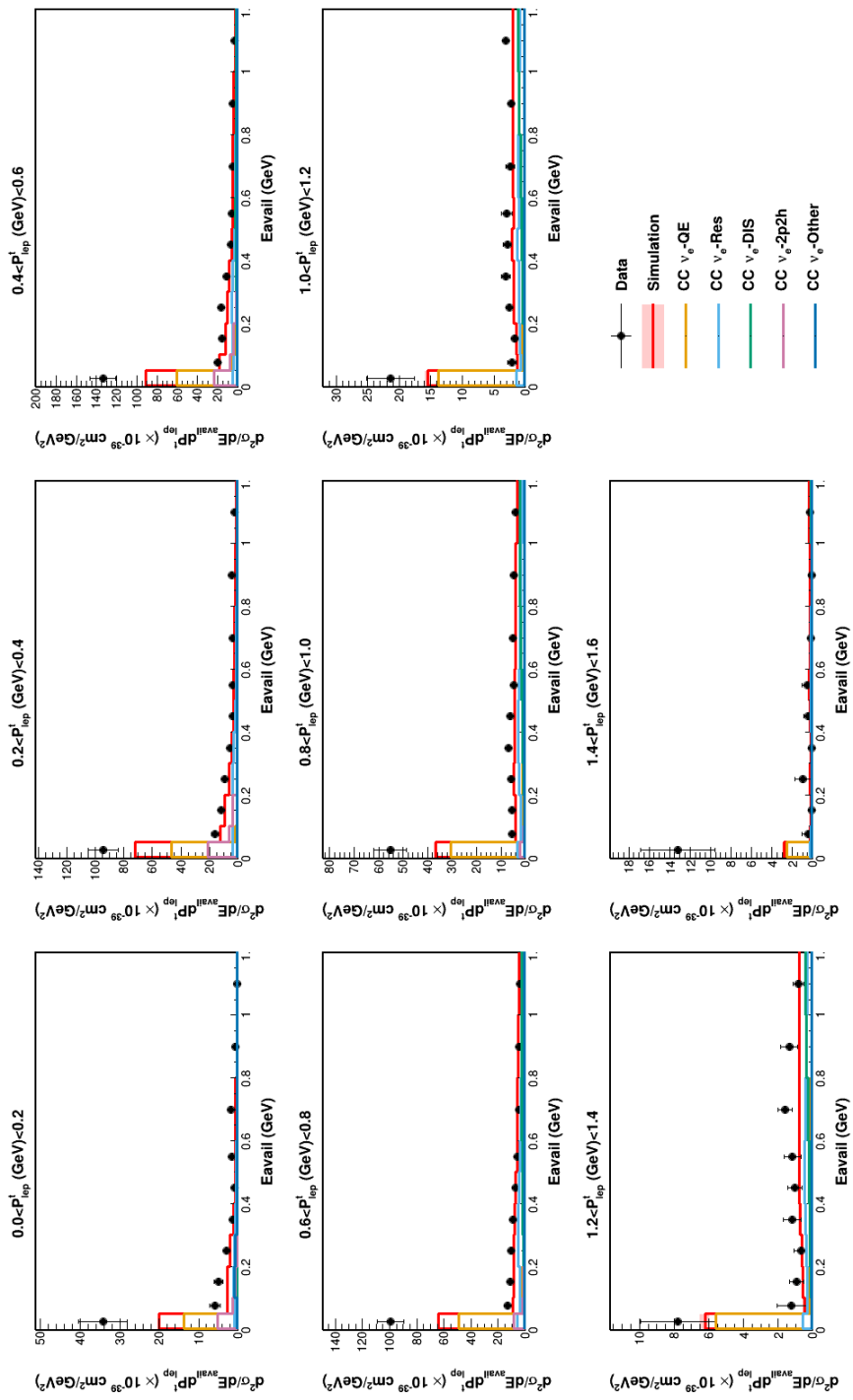


Figure 6.21: A decomposition of the cross section result into contributing interaction types in E_{avail} vs p_T.

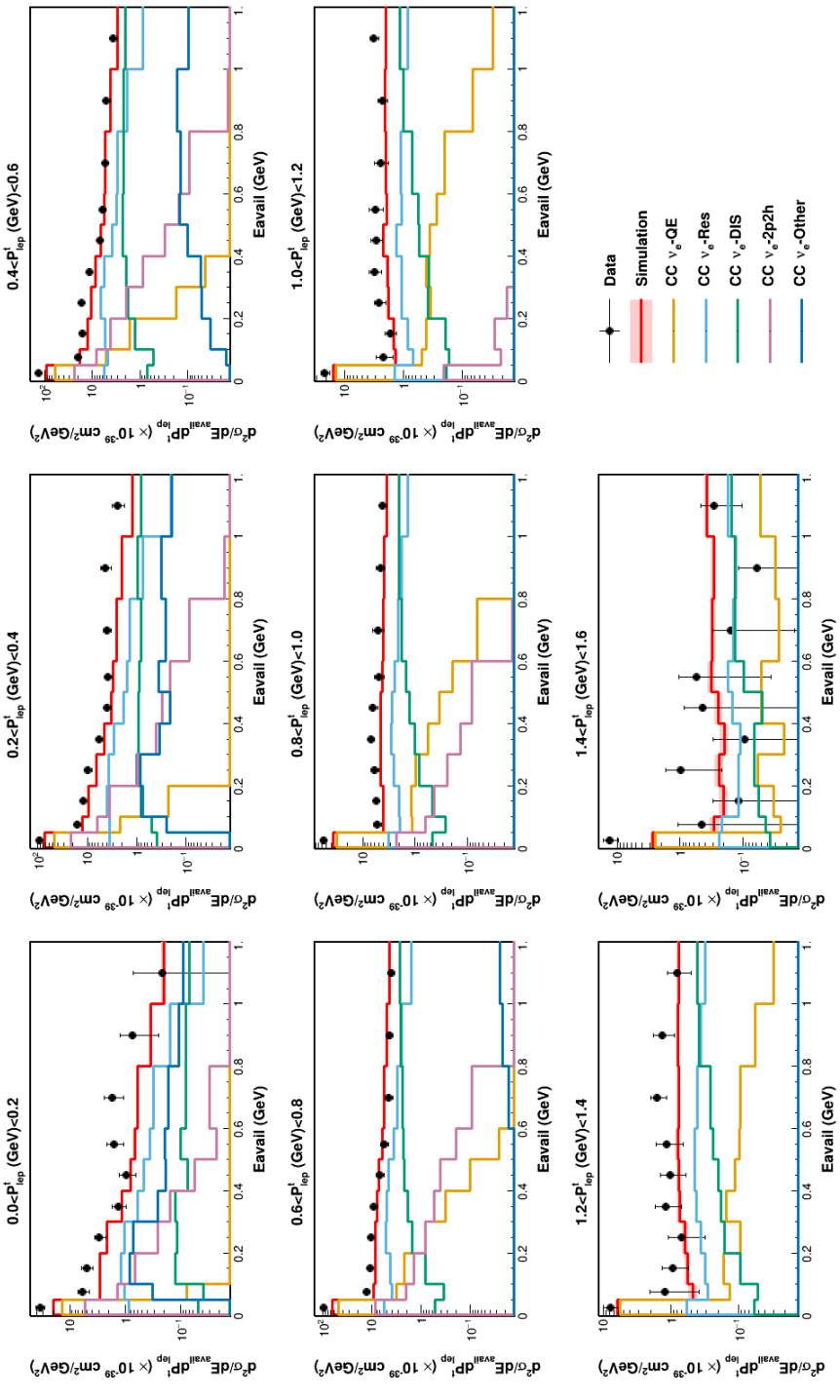


Figure 6.22: A decomposition of the cross section result into contributing interaction types in E_{avail} vs p_T on a y log scale.

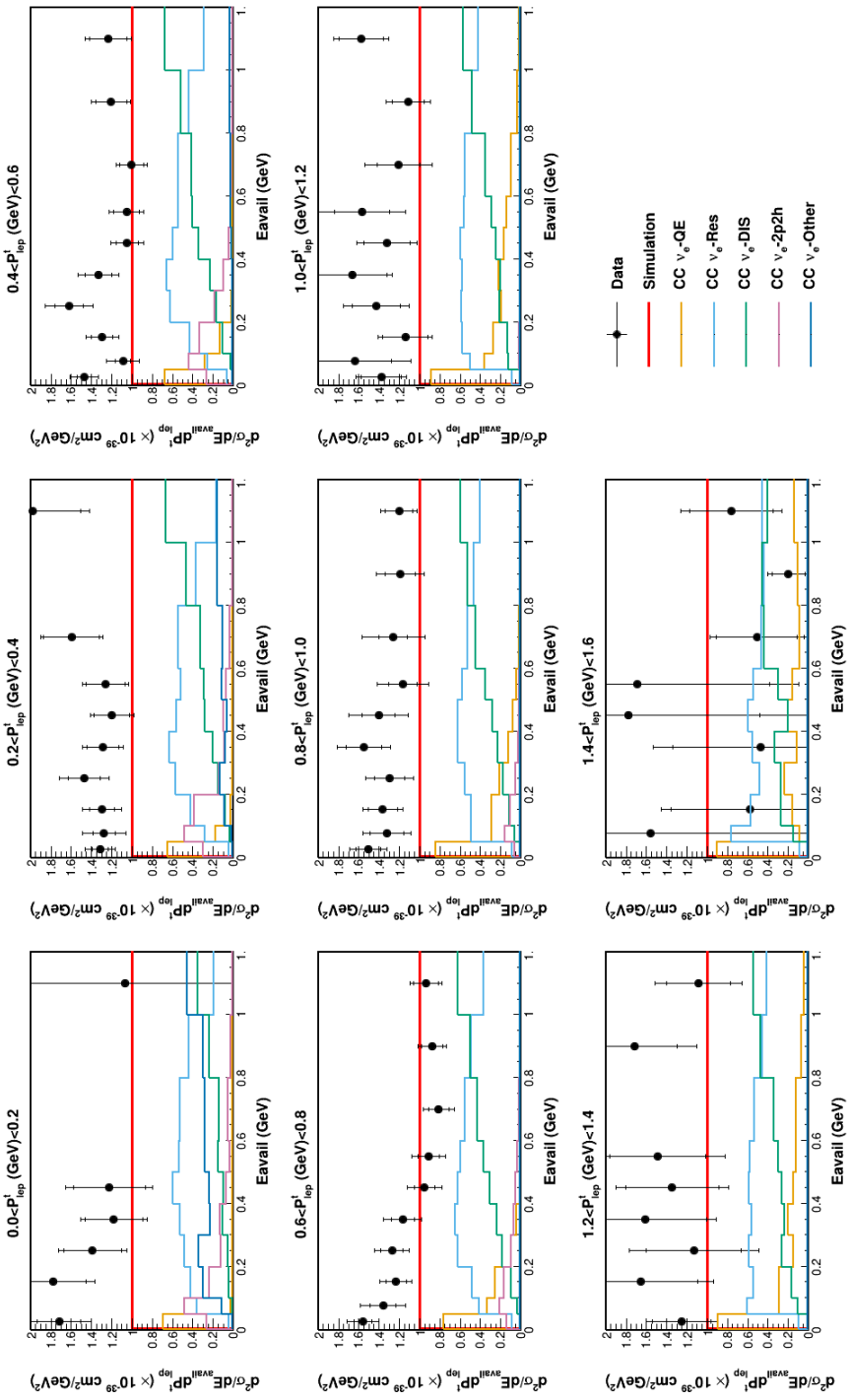


Figure 6.23: Data MC ratio including fractional contribution of signal processes shown in E_{avail} vs p_T.

	All	nn	np	QE
N	10.5798	8.58724	17.0344	5.38719
μ_0 (GeV)	0.254032	0.23626	0.289916	0.213611
μ_3 (GeV)	0.50834	0.502603	0.532062	0.396522
σ_0 (GeV)	0.0571035	0.072291	0.074685	0.0496312
σ_3 (GeV)	0.129051	0.154832	0.137321	0.125062
C	0.875287	0.789796	0.836689	0.806659

Table 6.6: Parameters for the 2p2h reweighting function. The column "All" is used in the reference model while QE is used as the alternative model comparison. Table from Ref. [99].

6.5.1 Model Comparison

The plots in this section compare the measured cross sections to the cross section of alternative models.

The Super-Scale Approach (SuSA) model is a phenomenological model which provides, by construction, a good description of inclusive electron scattering data in the quasi-elastic region [97]. The SuSA 2p2h alternative model implements 2p2h predictions into SuSA with an enhancement of the reference model Valencia 2p2h in the low recoil region of q_3 [98]. The 2p2h QE tune is MINERvA's reweighting of 2p2h events between the peak of QE and resonance interactions in q_0, q_3 phase space, referred to as the dip region [55]. Previous results have shown the 2p2h component of these interactions are underestimated, as seen in Figure 6.24. The reweight results in about a 50% increase of simulated 2p2h events. The weight function is given by

$$w(q_0,q_3) = N \times exp[-\frac{1}{2(1-C^2)}(\frac{(q_0-\mu_0)^2}{\sigma_0^2} + \frac{(q_3-\mu_3)^2}{\sigma_3^3} - 2C\frac{(q_0-\mu_0)(q_3-\mu_3)}{\sigma_0\sigma_3})]$$

with the parameter values listed in Table 6.6. The MK model is a single pion production

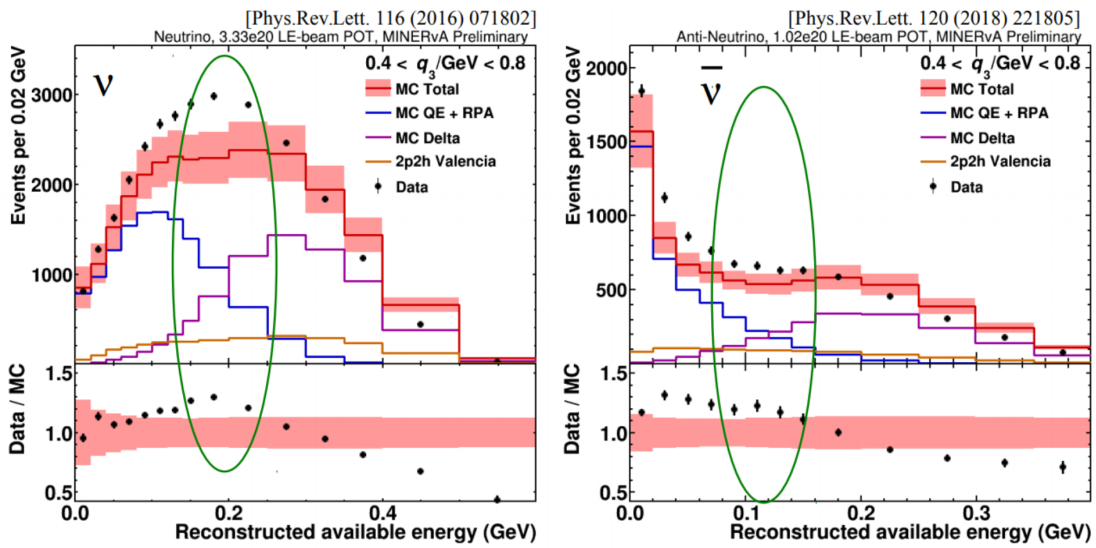


Figure 6.24: Reconstructed available energy for neutrino events (left). Reconstructed available energy for anti-neutrino events (right). The contribution of 2p2h events is in yellow. The green circle represents the regions where simulations underestimate data. Figure from Ref. [102].

model that includes lepton mass effects, as well as resonant and non-resonant background pion production [100]. The main difference compared to the default resonant pion production is in the incorporation of non-resonant pion production and interference terms in the model. During the LE era, it was found that the GENIE FSI of final states from pion absorption was being over simulated. Specifically, GENIE was producing too many np final states from π^+ absorption. Additionally it was found that QE FSI events with highly transverse protons were under-simulated as well as a population of events with very narrow angular distribution [101]. A "bug fix" was implemented to account for both of these findings. The difference between the MnvTune v1 alternative model and the base model of MnvTune v2 was discussed in Section 3.7.4.

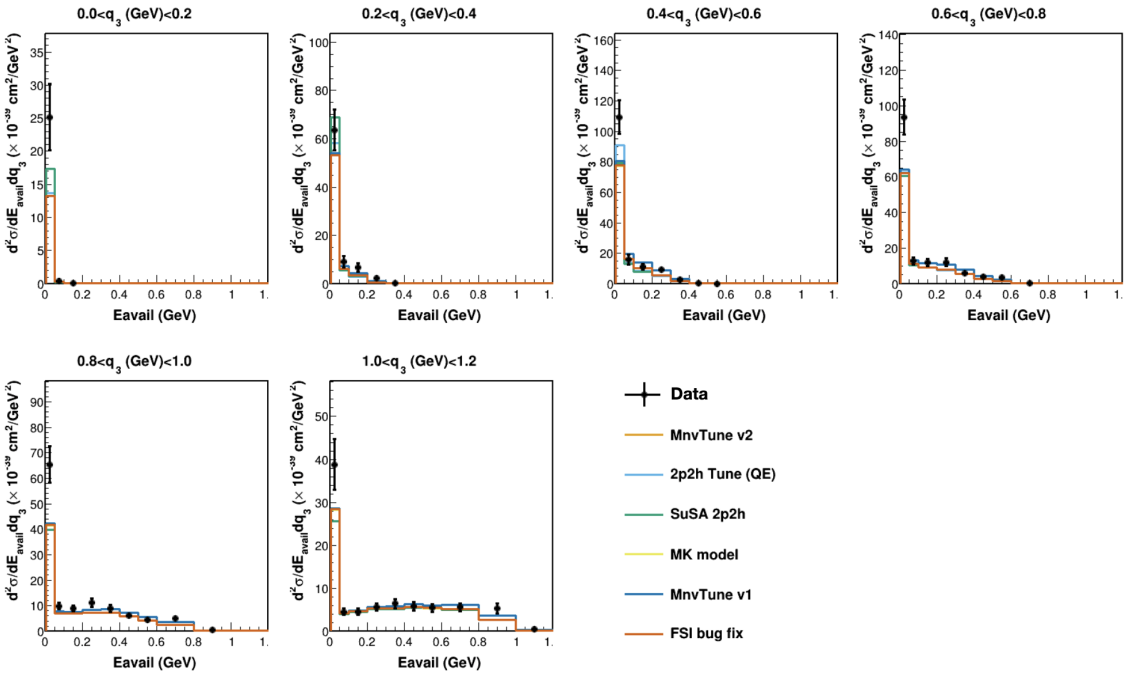


Figure 6.25: Comparison of cross section result to alternative models in E_{avail} vs q_3.

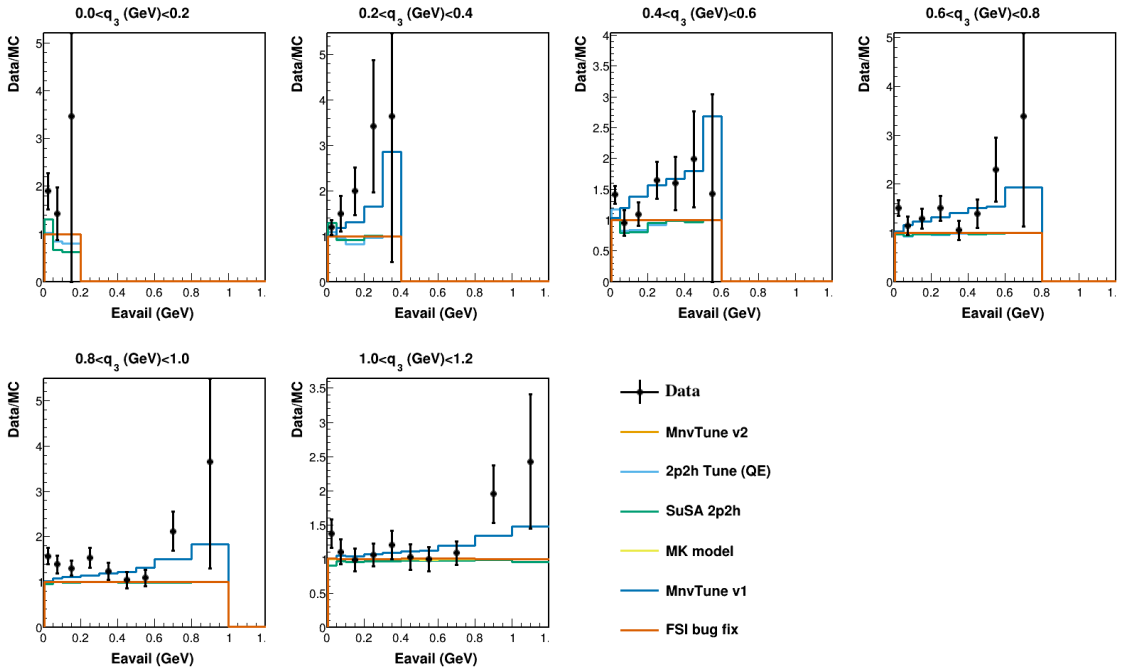


Figure 6.26: Ratio of cross section and alternative model MC to data in E_{avail} vs q_3.

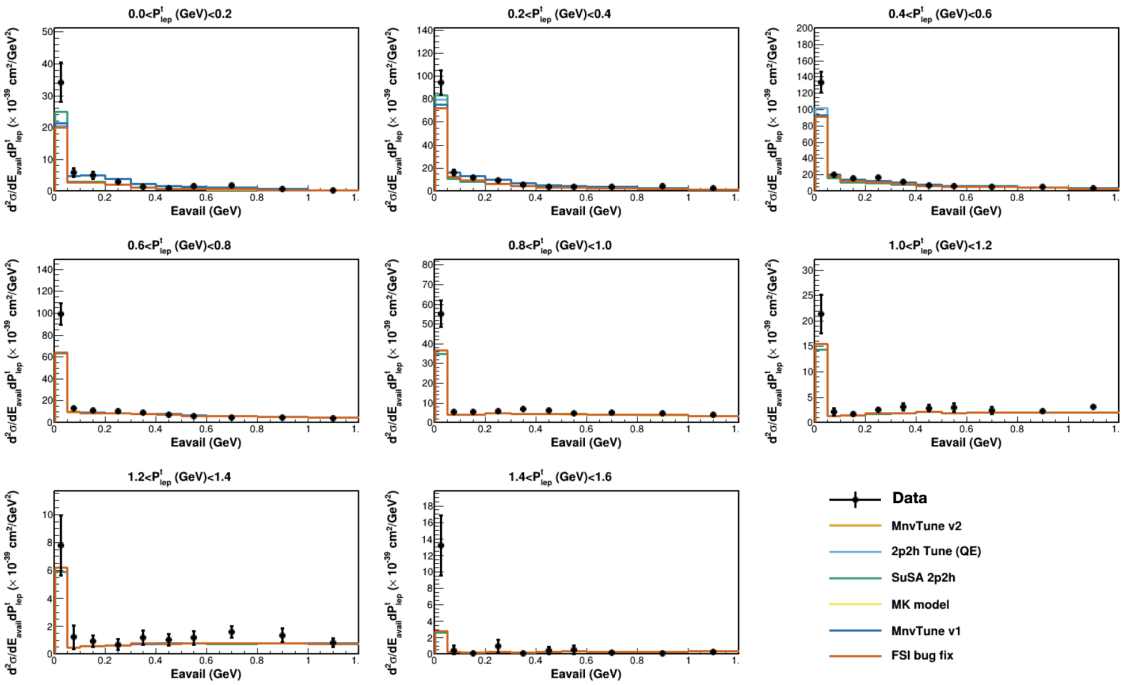


Figure 6.27: Comparison of cross section result to alternative models in E_{avail} vs q_3.

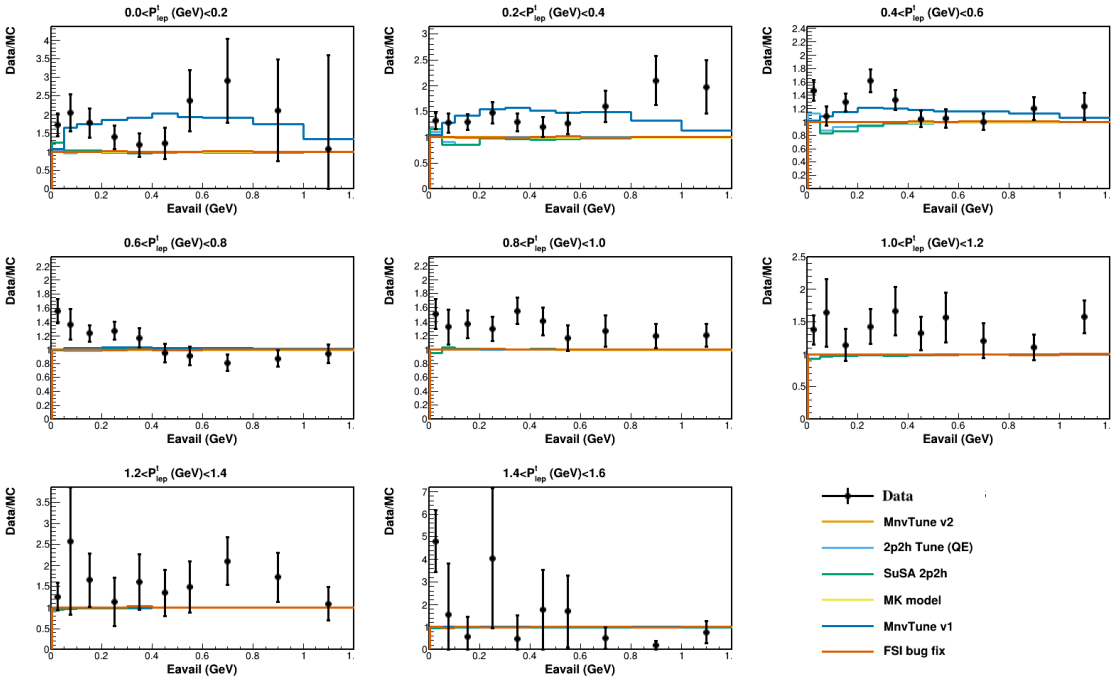


Figure 6.28: Ratio of cross section and alternative model MC to data in E_{avail} vs p_t.

7. Systematic Uncertainties

7.1 Uncertainties

The derived cross section measurement relies on MC simulation in many aspects ranging from background estimates, signal selection efficiency, kinematic parameter reconstruction, detector response, and flux predictions among other things. The uncertainties in the predictions of the Monte Carlo simulations are carried through to uncertainties in the cross section results. Therefore, accurate reporting of the uncertainties and the work done to reduce their magnitude is critical for their use in improving neutrino interaction models. The uncertainties come in two forms: statistical and systematic. Since MINERvA is essentially a counting experiment, the event counts in different energy or momentum bins are based on random processes that lead to uncertainty in the precision of measurements made in MINERvA. Statistical uncertainties apply to MC as well as data and are typically the limiting factor in the total uncertainty and are calculated by adding the data and MC uncertainties in quadrature. Uncertainties represent the precision to which a quantity is measured given an underlying theoretical model. This chapter will cover the general approach taken regarding systematic uncertainties and those of direct impact on this specific

analysis.

7.2 The Many Universe Method

Systematic uncertainties are evaluated for the underlying model predictions used in the simulations and the ability to accurately identify and reconstruct kinematic quantities from an event. Some examples for systematic uncertainties may be the degree to which a track is reconstructed, how much material is in the detector, and the uncertainty in the rate of final state interactions. Due to the maturity of the MINERvA detector, many of these uncertainties have already been calculated from either previous analyses or the MINERvA test beam program [103]. Therefore, once each uncertainty has been calculated, the effects of the uncertainties on the cross section must be evaluated. This evaluation is done through a "many universe" method that is applied when running the Monte Carlo simulation. Both a central value event distribution and altered event distributions are produced, as explained below.

For each MC simulation of a physics interaction in the MINERvA detector, there is a "Central Value" (CV) universe defined as a compilation of the best predictions and models given the physics parameters of the measurements. This can encompass anything from the best estimate of the mass of the detector to the models governing the rates of specific final state interactions. The CV serves as the base prediction used to compare results to data as well as the basis used to evaluate systematic uncertainties. The many universe method works by taking the CV universe and varying a specific underlying assumption, whether it be a physics model or a parameter, and measuring the impact of the variation. The strategy

is to simulate several universes, each of which is some variation of a few parameters p_α determined by their values from Gaussian distributions g(μ_α,σ_α) where the mean μ_α and standard deviation σ_α are already determined. In the MINERvA analysis framework, there is a custom package specifically built to hold all the shifted measurements in histogram form corresponding to the different systematic uncertainties in every universe. This streamlines the operations that must be performed on every histogram, such as background subtraction, unfolding, and efficiency correction that are required to be carried out in every universe to extract the uncertainties on a cross section result.

Typically, a single systematic uncertainty corresponds to a pair of systematic universes where the underlying parameters are varied by $\pm\sigma$. However, some uncertainties, like the flux, require special treatment because simply shifting the flux universes up and down by 1σ does not accurately represent the range of values regarding the flux spectrum. Therefore, the flux is simulated in 500 universes where each key parameter gets evaluated and compared to the CV. The flux uncertainty is then computed based on the spread of all varied distributions.

There are two different types of shifts when it comes to systematic uncertainties. The first is a lateral shift or lateral error band that affects reconstructed values like vertex reconstruction or track angles. The lateral shift of the parameter might cause the event to migrate between bins and move in or out of defined event selections. For example, if the parameter is electron energy, E_e, then shifting the electron energy by some percentage might cause it to fail the selection cut of $E_e > 1.5$ MeV in one of the simulated universes. The second is a vertical shift that changes the probability of expectation for a particular event type to occur. In this case, the shifted distributions are generated by assigning a multiplier to each event which corresponds to its new probability. For example, if the final

state interaction probability for pions is increased, this would give an increased weight, or probability, to all events with a pion in the final state while having no effect on events with no pion. When multiple universes are generated to study systematic effects, the magnitude of the systematic uncertainty in each bin corresponds to the average difference between each shifted parameter and the central value according to an individual bin.

7.3 Covariance Matrix

The systematic uncertainty is evaluated by forming a covariance matrix. The matrix tells us about the value of a particular systematic uncertainty for a given bin, as well as correlations that may be occurring between bins. If there are universes where the event count in bin i increases and leads to an increase in event count in bin j, then these two bins might be positively correlated.

The covariance matrix is an N $\times$ N matrix, for N the total number of bins in the analysis' two-dimensional distribution. In the case of E_{avail} vs q_3, N is the total number of E_{avail} bins multiplied by the total number of q_3 bins. The covariance matrix can be represented as

$$M_{ij} = M_{ji} = \frac{\sum_{k=1}^{n} (x_{ik} - \bar{x}_i)(x_{jk} - \bar{x}_j) w_k}{\sum_{k=1}^{n} w_k} \tag{7.1}$$

for M_{ij} the matrix element corresponding to the covariance between bins i and j, n the total number of universes, x_{ik} the event count in bin i in universe k, w_k the weight applied in the case of a weighted universe k and $\bar{x}_i$ the mean event count in bin i averaged across all

universes written as

$$\bar{x}_i = \frac{\sum_{k=1}^{n} x_{ik} w_k}{\sum_{k=1}^{n} w_k} \tag{7.2}$$

The diagonal elements of the covariance matrix provide information about the magnitude of uncertainty in each bin. The magnitude of uncertainty in bin i is given by the square root of the matrix element $[i,i]$

$$\sigma_i = \sqrt{\frac{\sum_{k=1}^{n} (x_{ik} - \bar{x}_i)^2 w_k}{\sum_{k=1}^{n} w_k}} \tag{7.3}$$

and for an unweighted universe by

$$\sigma_i = \sqrt{\frac{\sum_{k=1}^{n} (x_{ik} - \bar{x}_i)^2}{\sqrt{n}}} \tag{7.4}$$

In MINERvA these uncertainties are shown by using error bands that are vertically above and below the data point, representing the diagonal entries of the covariance matrix. This is done to provide a clear visual of the uncertainty in relation to the CV; however, in turn the information pertaining to the correlations between all bins is neglected. Therefore, a secondary approach, called a correlation matrix, might be applied. The correlation matrix provides information on how correlated different bins are by relating an element of the correlation matrix to an element of the covariance matrix through the relation

$$\rho_{ij} = \frac{M_{ij}}{\sigma_i \sigma_j} \tag{7.5}$$

The ρ_{ij} is in the range [-1,1] for -1 being a completely anti-correlated value (increase in bin i and decrease in bin j), a value of 0 means the bins are completely uncorrelated, and a value of 1 means that the bins are entirely positively correlated (increase in bin i and increase in bin j).

Chi-squared calculations (χ^2), or goodness-of-fit tests, uses the covariance matrix's on and off-diagonal elements to compare the observed values to the values expected from the models in place. The chi-squared value is given by

$$\chi^2 = \sum_{i=1}^{N} \sum_{j=1}^{N} (x_i - y_i) M_{ij}^{-1} (x_j - y_j) \tag{7.6}$$

for N the number of bins, M_{ij}^{-1} the matrix element for the inverse of the covariance matrix, x_i the value of bin i in the covariance matrix, y_i the value of bin i for the model being compared against.

7.4 Systematic Uncertainty Summary

For plotting purposes, the systematic uncertainties are categorized into one of four major groups: flux, detector model, interaction model, and MINERvA tunes as seen in Figure 7.1 with a breakdown of each subcategory found in Table 7.1.

The resultant systematic uncertainties for the measured cross section enter at different steps of the cross section extraction. At the event selection level, data has no systematic uncertainties. Uncertainties pertaining to background processes are introduced at the background subtraction. Uncertainties related to detector responses typically enter at the unfolding step because the shifting of variables will introduce populations with more

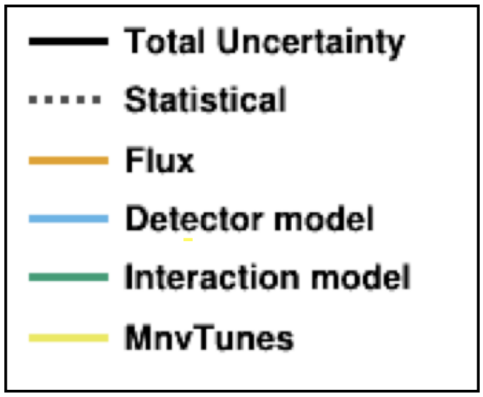


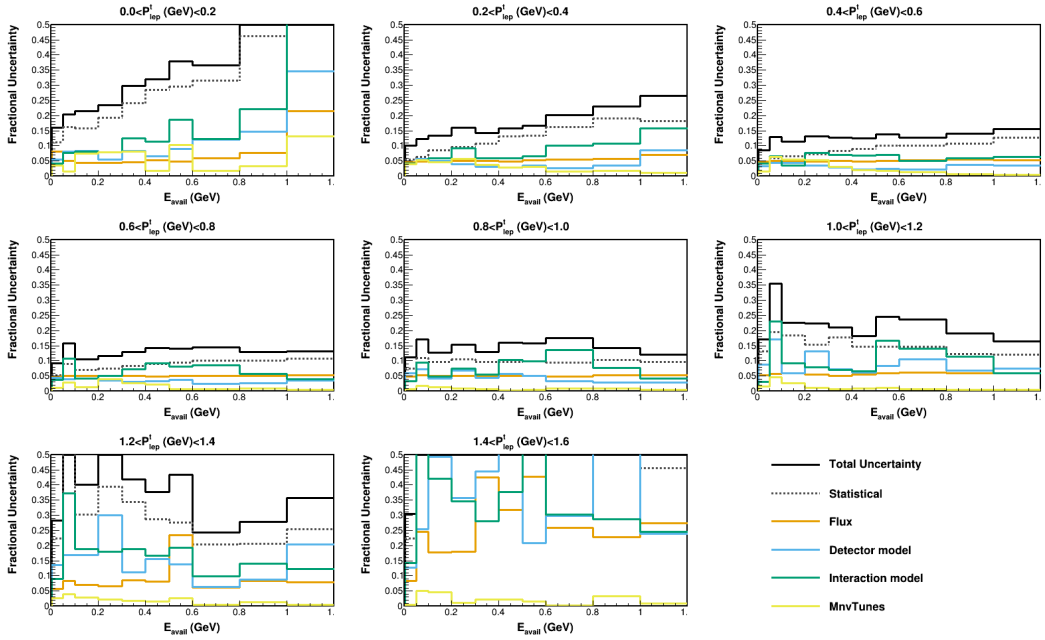
Figure 7.1: All systematic uncertainties fall under one of the categories.

Uncertainty Group Name	Uncertainty Name
Flux	Flux
Detector Model	Angular Resolution Beam Angle EM Energy Scale Particle Response Target Mass Leakage Uncertainty
Interaction Model	GENIE Model FSI Model
MnvTunes	RPA Low Recoil 2p2h High Q2 Pion Low Q2 Pion

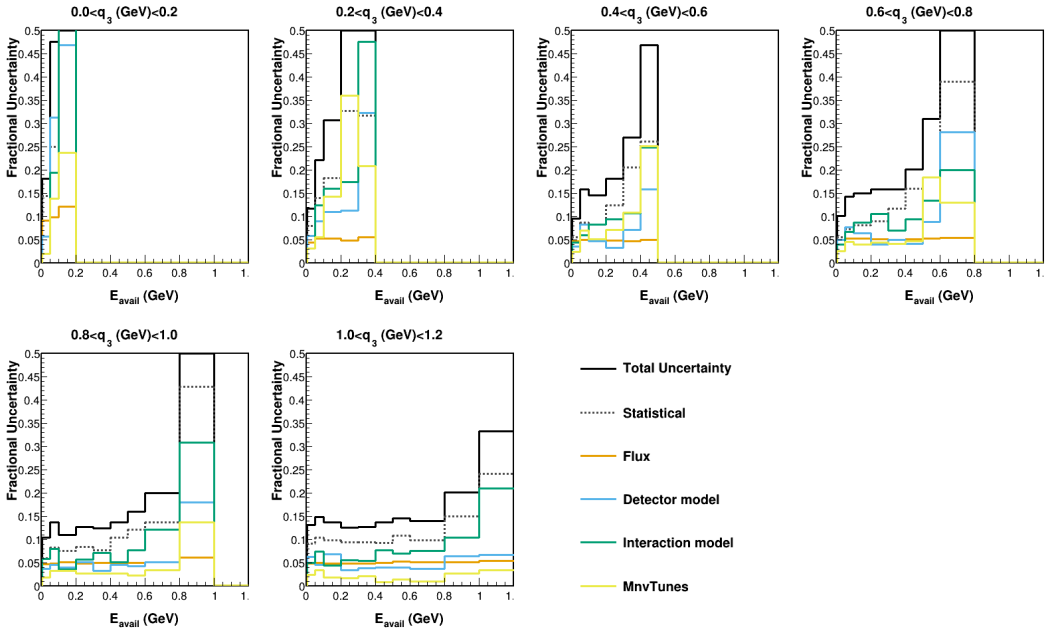
Table 7.1: Categorical breakdown of systematics uncertainties.

off diagonal entries, which will become noticeable during the unfolding procedure. The efficiency correction introduces uncertainties on models pertaining to signal events.

Figure 7.2 shows the total cross section error summary for E_{avail} vs. p_T and E_{avail} vs. q_3, broken down into the four main categories. The remainder of the chapter will discuss the leading contributors of systematic errors to each error group.



(a) E_{avail} vs. p_T cross section error summary



(b) E_{avail} vs. q_3 cross section error summary

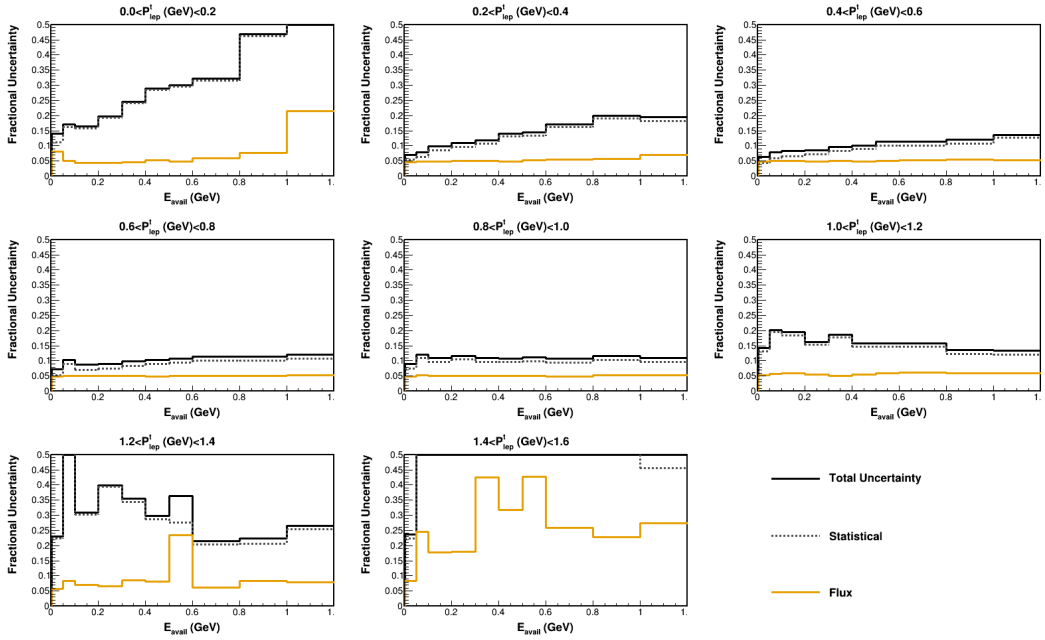
Figure 7.2: Total cross section error summary broken down into four major subgroups.

7.4.1 Flux Model Uncertainty

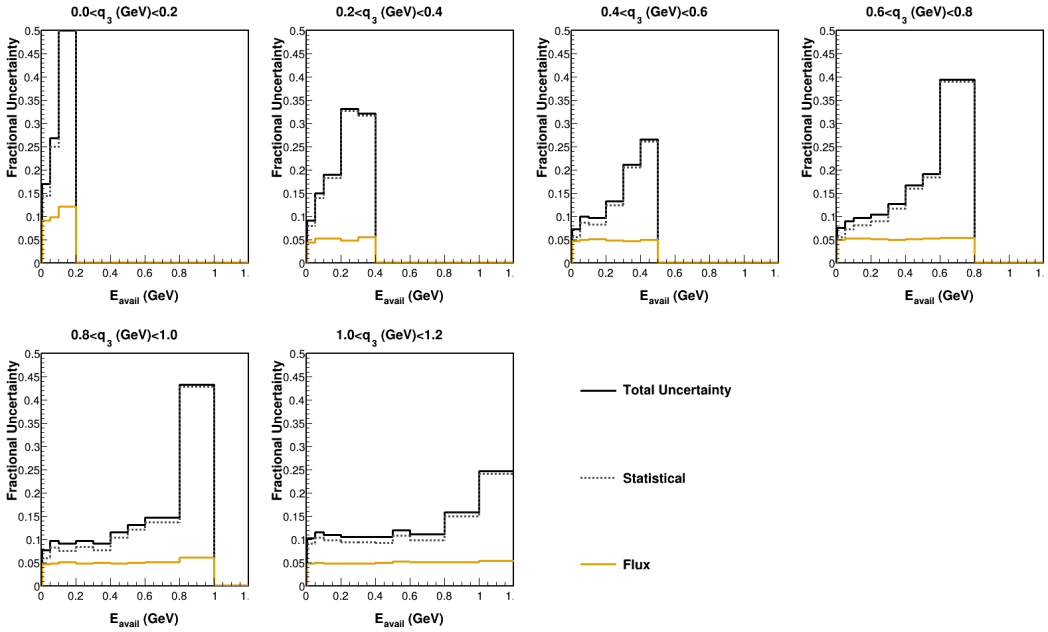
The simulated flux, as discussed previously in Section 2.3.2, has many integrated parameters and therefore many sources of uncertainty. The uncertainties related to the flux can be broken down into two major categories: focusing uncertainties associated with all components related to the NuMI beam and hadron production uncertainties related to the uncertainty of hadron production from the proton beam incident on the graphite target. All uncertainties under those two major categories are highly correlated; thus, a single parameter shift of 1 σ is insufficient to encompass those correlations.

To capture the correlations between the flux uncertainties, many different fluxes are simulated. The flux universes refer to some random variations in each parameter pulled from a gaussian distribution centered on the respective nominal parameter value with a width equivalent to a 1σ parameter uncertainty. In MINERvA, 200 flux universes are produced and the flux uncertainty is estimated by comparing each flux universe to the CV flux value. A weight is produced to adjust the likelihood of an event occurring given the true neutrino energy and flavor with the systematic uncertainty being the spread of the weighted distributions.

The flux uncertainty can be reduced by applying the neutrino flux constraint resulting from the $\nu - e$ and IMD results as discussed Section 2.3.2. The constrained flux is calculated through a posterior probability weighted average among an ensemble of flux universes, each characterized by the neutrino flux spectrum and the posterior probability. The best estimation of the flux after the constraint is about a 10% reduction of the nominal flux. Figure 7.3 shows the contribution of the flux uncertainty to the total cross section uncertainty. The subtraction of neutrino-electron scattering events that are used to constrain the flux is



(a) E_{avail} vs. p_T cross section error summary isolating the flux systematic error



(b) E_{avail} vs. q_3 cross section error summary isolating the flux systematic error

Figure 7.3: Cross section error summary showing flux contribution to total uncertainty.

the likely driver for the larger uncertainty in the lowest q_3 bin.

7.4.2 Interaction Model Uncertainty

The models implemented in MINERvA's Monte Carlo generator through GENIE version 2.12.6 each have parameters that carry their uncertainties. For the most part, GENIE uncertainties are evaluated using a re-weighting method where two universes are produced for each parameter in question using a $\pm\sigma$ shifted universe, referred to as "knobs". GENIE knobs can be broken down into two categories: GENIE model uncertainties and GENIE final state interaction uncertainties. Each knob has a specific name associated with it.

GENIE knob	**+1 σ**	**-1 σ**	**GENIE knob**	**+1 σ**	**-1 σ**
AGKYxF1pi	20%	20%	FrPiProd_N	20%	20%
AhtBY	25%	25%	FrPiPridπ	20%	20%
BhtBY	25%	25%	MFP_N	20%	20%
CCQEPauliSupViaKF	30%	30%	MFP_pi	20%	20%
CV1uBY	30%	30%	MaNCEL	25%	25%
CV2uBY	40%	40%	MaRES	20%	20%
EtaNCEL	30%	30%	MvRES	10%	10%
FrAbs_N	20%	20%	NormNCRES	20%	20%
FrAbs_pi	30%	30%	RDecBR1gamma	50%	50%
FrCEx_N	50%	50%	Rnv2pi	50%	50%
FrCEx_pi	50%	50%	Rnp2pi	50%	50%
FrElas_N	30%	30%	Theta_Delta2Npi	100%	100%
FrElas_pi	10%	10%	VecFFCCQEshape	100%	100%
FrInel_N	40%	40%			

Table 7.2: Systematic uncertainties for the GENIE universes based on $\pm\sigma$ two universe shifts.

GENIE FSI uncertainties beginning with the naming convention of "Fr"$\langle..\rangle$_{N, pi} refer to the different FSI probabilities governing hadrons and pions produced from the primary interaction. The different interactions include elastic "Elas" scattering where the

nucleus remains in the ground state and the hadron keeps the same charge, inelastic "Inel" scattering where the nucleus is broken apart, charge exchange "CEx" where the final state hadron has a different charge than it had in the original interaction, absorption "Abs" where the hadron is absorbed in the nucleus, and pion production "PiProd" where the nucleon contains enough energy to produce an additional particle in the interaction. The final state models refer to the probability for particles produced by the initial neutrino-nucleus interaction to re-interact during their progression through the nuclear medium. The models are simulated using the INTRANUKE-hA intra-nucleon hadron cascade package [104]. The mean free path for nucleons (MFP_N) and for pions (MFP_pi) each have a corresponding shift of $\pm$ 20%, and refer to the mean free path of hadrons produced within the nucleus. Typically, the shorter the mean free path, the higher the probability of re-interaction within the nucleus.

GENIE interaction model uncertainties control the types of neutrino-nucleus interactions that are allowed to occur. As mentioned in Section 3.6.1, these cover interactions like deep inelastic scattering, quasi-elastic scattering, pion production, and 2p2h. The MaRES and MvRES knobs refer to the GENIE model that is used to produce a charged current resonance event where a target nucleon gets excited to a resonance state after the neutrino-nucleon interaction and immediately decays to a nucleon and a pion. The resonant interaction is modeled using the Rein-Sehgal model which includes nucleon form factors akin to the Llewellyn-Smith CCQE model [105]. The knobs referring to the inelastic neutrino-nucleus scattering models using Bodek-Yang parameters include AhtBy, BhtBy and CV1uBy, CV2uBy are the parameters related to the vector form factors of the valence quarks [106]. The CCQEPauliSupViaKF knob requires that a final state nucleon have energy above the

Fermi energy to break the nuclear binding and exit the nucleus.

The MaCCQE knob refers to the uncertainties due to the axial mass M_A used in the axial form factor, F_A, in the CCQE scattering formalism. By default, the uncertainty is large for GENIE (-15%, +25%) and is one of the leading systematics in this analysis. A modification to the axial form factor has proven to reduce the magnitude. Originally the axial form factor used in the Llewellyn-Smith model for quasi-elastic scattering was derived from neutrino-deuteron scattering data using a dipole shape assumption [45]. Instead of using a dipole shape assumption, F_A can instead be expanded using a variable z for F_A written as

$$F_A(q^2) = \sum_k a_k z(q^2)^k$$

for a_k dimensionless numbers encoding nucleon structure information [37]. Referred to as the z-expansion, this approach is used to derive the axial form factor and has proven more rigorous resulting in a reduction in the systematic uncertainty of MaCCQE. In this analysis, the z-expansion is applied to evaluate the GENIE_MaCCQE systematic uncertainty which requires 100 separate F_A universes. All GENIE knobs and their corresponding shifts can be referred to in Table 7.2. Descriptions of each GENIE uncertainty can be found in Tables 7.3 through 7.6.

The remaining GENIE model uncertainties have little to no effect on the event count and are therefore negligibly small uncertainties. The hadronization model, AGKYxF1pi, is derived from a neutrino scattering off of a quark or gluon within the nucleus, resulting in a fragmented nucleus also known as a DIS event. The AGKYxF1pi knob is simulated in GENIE by varying the distribution of the pion Feynman parameter, x_F, in the 1-pion states of the model [45] [46]. Resonant scattering that produces a nucleon and pion modeled

using the Rein-Sehgal model is represented with the Theta_Delta2Npi uncertainty. The RDecBR1gamma uncertainty varies the branching ratio of photon decay to resonant states.

GENIE Model Uncertainty	**GENIE Knob**	**Description**	$\pm$ **1** σ
M_A (Elastic Scattering)	MaNCEL	Adjust M_A in elastic scattering cross section	$\pm$ 25%
M_A (CCQE Scattering)	MaCCQE	Adjust M_A in Llewellyn-Smith cross section, affecting shape and normalization	$\pm$ 25%
CCQE Normalization	NormCCQE	Adjusts CCQE Normalization	$\pm$ 25%
CCQE Vector Form factor model	VecFFCCQEshape	Changes BBBA to dipole, affecting shape only	$\pm$ 25%
CC Resonance Normalization	NormCCRES	Changes the normalization of CC Rein-Sehgal cross section	$\pm$20%
NC Resonance Normalization	NormNCRES	Changes the normalization of NC Rein-Sehgal cross section	$\pm$20%
M_A (Resonance Production)	MaRES	Adjusts M_A in Rein-Sehgal cross section, affecting shape and normalization	$\pm$20%
M_V (Resonance Production)	MvRES	Adjusts M_V in Rein-Sehgal cross section, affecting shape and normalization	$\pm$10%
1 π production from $\frac{\nu p}{\bar{\nu} n}$ non resonant interactions	Rvp1pi	Affects NC and CC production of single pion final states from non resonant inelastic scattering with νp,$\bar{\nu} n$ initial states	$\pm$50%

Table 7.3: Magnitude of systematic shifts for GENIE model universes.

GENIE Model Uncertainty	GENIE Knob	Description	1 σ
1 π production from $\frac{\nu n}{\bar{\nu} p}$ non resonant interactions	Rvn1pi	Affects NC and CC production of single pion final states from non resonant inelastic scattering with νn,$\bar{\nu} p$ inital states	$\pm 50\%$
2 π production from $\frac{\nu p}{\bar{\nu} n}$ non resonant interactions	Rvp2pi	Affects NC and CC production of single pion final states from non resonant inelastic scattering with νp,$\bar{\nu} n$ inital states	$\pm 50\%$
2 π production from $\frac{\nu n}{\bar{\nu} p}$ non resonant interactions	Rvn2pi	Affects NC and CC production of single pion final states from non resonant inelastic scattering with νn,$\bar{\nu} p$ inital states	$\pm 50\%$
Bodek-Yang parameter A_{HT}	AhtBY	Refines Bodek-Yang model parameter A_{ht} (shape and normalization effect)	$\pm 25\%$
Bodek-Yang parameter B_{HT}	BhtBY	Refines Bodek-Yang model parameter B_{ht} (shape and normalization effect)	$\pm 25\%$
Bodek-Yang parameter C_{V1u}	CV1uBY	Refines Bodek-Yang model parameter CV1u (shape and normalization effect)	$\pm 30\%$
Bodek-Yang parameter C_{V2u}	CV2uBY	Refines Bodek-Yang model parameter CV2u (shape and normalization effect)	$\pm 40\%$
DIS CC Normalization	NormDISCC	Adjusts overall normalization of the non-resonance inclusive cross section	

Table 7.4: Magnitude of systematic shifts for GENIE model universes.

FSI Uncertainty	GENIE Knob	Description	1 σ
Pion mean free path	MFP_pi	Adjusts mean free path for pions	$\pm 20\%$
Nucleon mean free path	MFP_N	Adjusts mean free path for nucleons	$\pm 20\%$
Pion - absorption	FrAbs_Pi	Adjusts absorption probability for pions, for given total rescattering probability	$\pm 30\%$
Pion - charge exchange	FrCEx_Pi	Adjusts charge exchange probability for pions, for given total rescattering probability	$\pm 50\%$
Pion - elastic	FrElas_Pi	Adjusts elastic probability for pions, for given total rescattering probability	$\pm 10\%$
Pion - inelastic	FrInel_Pi	Adjusts inelastic probability for pions, for given total rescattering probability	$\pm 40\%$
Pion - pion production	FrPiProd_Pi	Adjusts pion production probability for pions, for given total rescattering probability	$\pm 20\%$
Nucleon - charge exchange	FrICEx_NN	Adjusts charge exchange for nucleons, for given total rescattering probability	$\pm 50\%$
Nucleon - elastic	FrElas_N	Adjusts elastic probability for nucleons, for given total rescattering probability	$\pm 30\%$
Nucleon - inelastic	FrInel_N	Adjusts inelastic probability for nucleons, for given total rescattering probability	$\pm 40\%$
Nucleon - absorption	FrAbs_N	Adjusts absorption probability for nucleons, for given total rescattering probability	$\pm 20\%$
Nucleon - pion production	FrPiProd_N	Adjusts pion production probability for nucleons, for given total rescattering probability	$\pm 20\%$

Table 7.5: Magnitude of systematic shifts for GENIE FSI universes.

FSI Uncertainty	GENIE Knob	Description	1 σ
AGKY hadronization model for x_F distribution	AGKYxF1pi	Adjusts x_F distribution for low multiplicity N + pi DIS fs, produced by AGKY	$\pm 20\%$
AGKY hadronization model for p_T model	AGKYpT1pi	Adjusts p_T distribution for low multiplicity N + pi DIS fs, produced by AGKY	$\pm 3\%$
Delta decay angular distribution	Theta_Delta2Npi	Changes delta decay angular distribution	On/off

Table 7.6: Magnitude of systematic shifts for GENIE FSI universes.

Figure 7.4 shows the cross section error summary broken into the GENIE FSI model and GENIE interaction model uncertainties. The bin of the highest p_T and the bin of lowest p_T and high E_{avail} show very large systematic uncertainties. This is also true for the bin of highest E_{avail} and low q_3 bins. These bins have a low number of events and a systematic variation of them has a larger, more dramatic impact. This is due to events either migrating in and out of bins or inefficiencies in the bin. Additionally, large bin-to-bin variations indicate statistical uncertainty in the measurement of systematic error, as seen in the highest p_T bins. A further breakdown of the systematic errors from GENIE models are shown in Figures 7.5 and 7.6. The GENIE MaRES is the leading systematic uncertainty for most bins followed by GENIE MvRES and GENIE NormCCRES. As shown in Figures 6.19 and 6.22 in Chapter 6, the CC resonant pion production has a large contribution to the cross section measurement. Since the GENIE MaRES and MvRES knobs control resonant pion production it is not surprising that they are among the leading contributors to the uncertainty along with NormCCRES that is a knob for the CC resonance normalization.

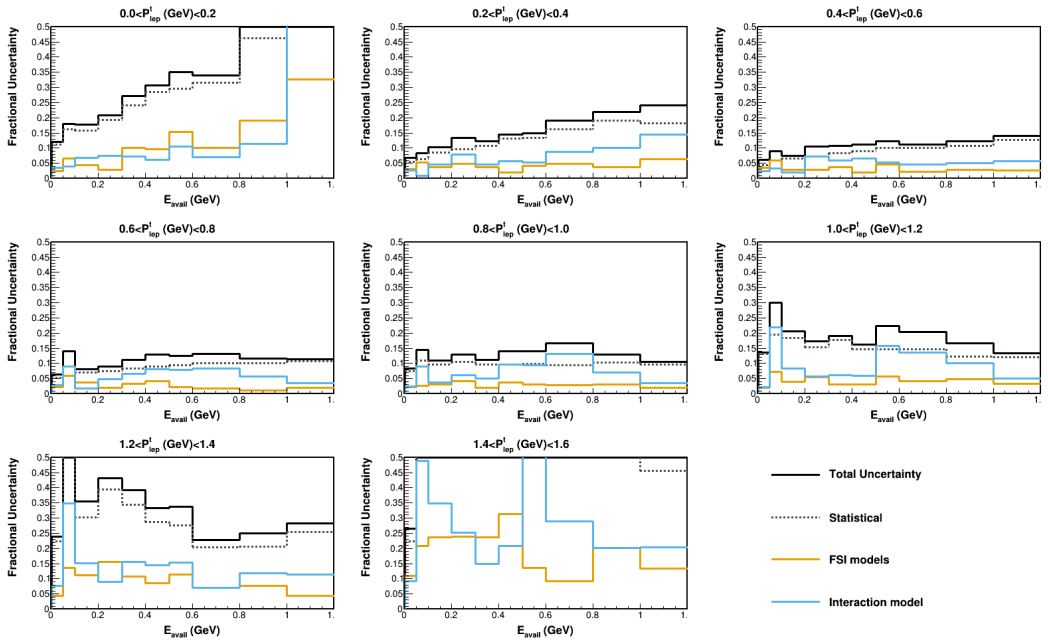
The GENIE FSI systematic error breakdown is shown in Figure 7.7. The GENIE FrInelas_N uncertainty contributes to the first bin of q_3 and highest E_{avail} bin with less contribution for higher q_3 bins. The GENIE FrInelas_N knob adjusts the inelastic probability for nucleons. Most likely this would include a neutron losing a large amount of energy in a collision, resulting with a proton in the final state. Interactions that have a change in the multiplicity of particles in the final state also impact the energy reconstruction.

7.4.3 MnvTune Uncertainties

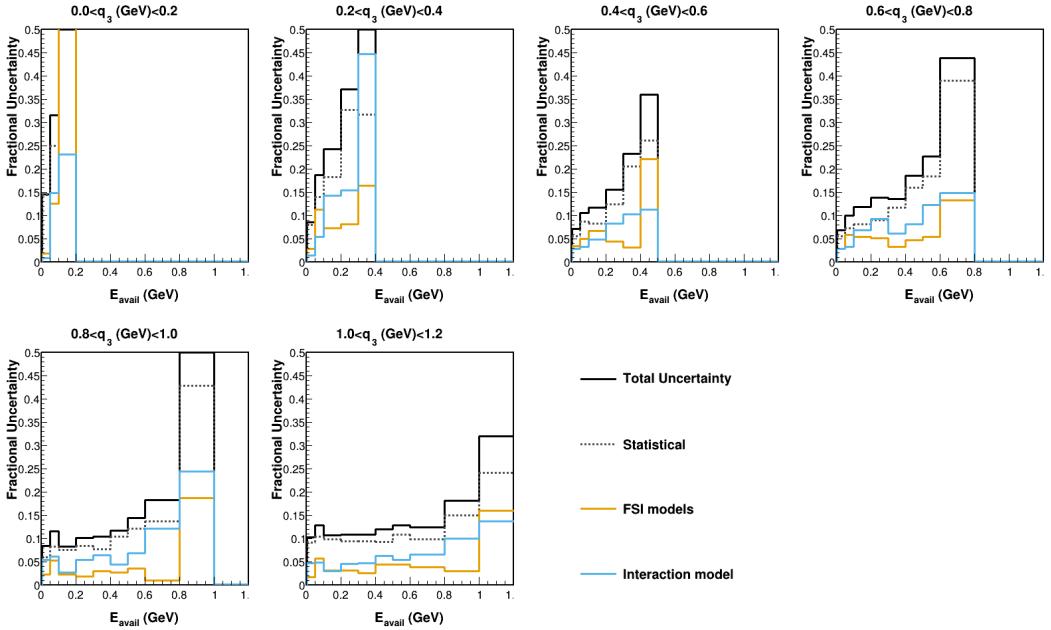
For each component that enters into the MnvGenie tune shown in Tables 3.4 and 3.5, there is an associated systematic uncertainty [55]. This analysis will have "MnvTune" shown as a single systematic uncertainty; however, the components that compose the uncertainty are discussed below.

RPA Uncertainty The random phase approximation (RPA) uncertainty considers the effect of long-range correlations between nucleons. The RPA uncertainty can be broken down into two sources of high and low Q^2 regions to account for different RPA effects. RPA weights are only applied to quasi-elastic events; therefore, the effect is seen in the low E_{avail} region [60]. The uncertainty in $q_0 - q_3$ phase space corresponds to four systematic universe shifts of $\pm 1\sigma$ for each of the high and low Q^2 regions.

Low Recoil 2p2h Uncertainty The 2p2h weight uncertainty is based on the three variations of the 2p2h model. The first is for the case of the neutrinos only scattering off of neutron-proton (np) pairs, the second is for neutrinos scattering only off of neutron-neutron (nn) or proton-proton (pp) pairs and the third is the data-MC disagreement seen at low-q_3



(a) E_{avail} vs. p_T cross section error summary isolating GENIE systematic errors divided into interaction model and FSI uncertainties.



(b) E_{avail} vs. q_3 cross section error summary isolating GENIE systematic errors divided into interaction model and FSI uncertainties.

Figure 7.4: Cross section error summary showing GENIE FSI and GENIE interaction model uncertainty contribution.

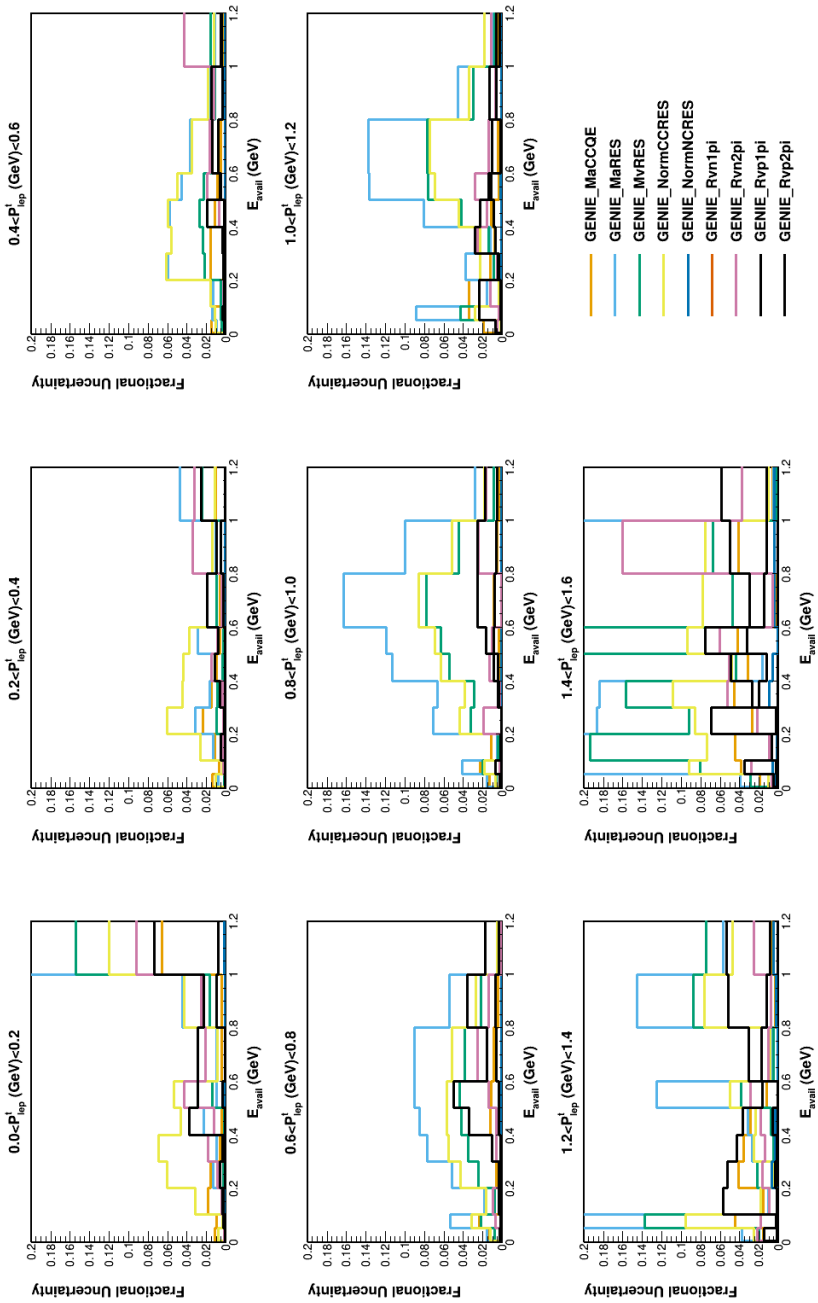


Figure 7.5: A decomposition of the cross section systematic errors broken into leading GENIE systematic error contributors for E_{avail} vs p_T. The leading systematic uncertainty is GENIE_MaRES (light blue), GENIE_MvRES (green) and GENIE_NormCCRES (yellow).

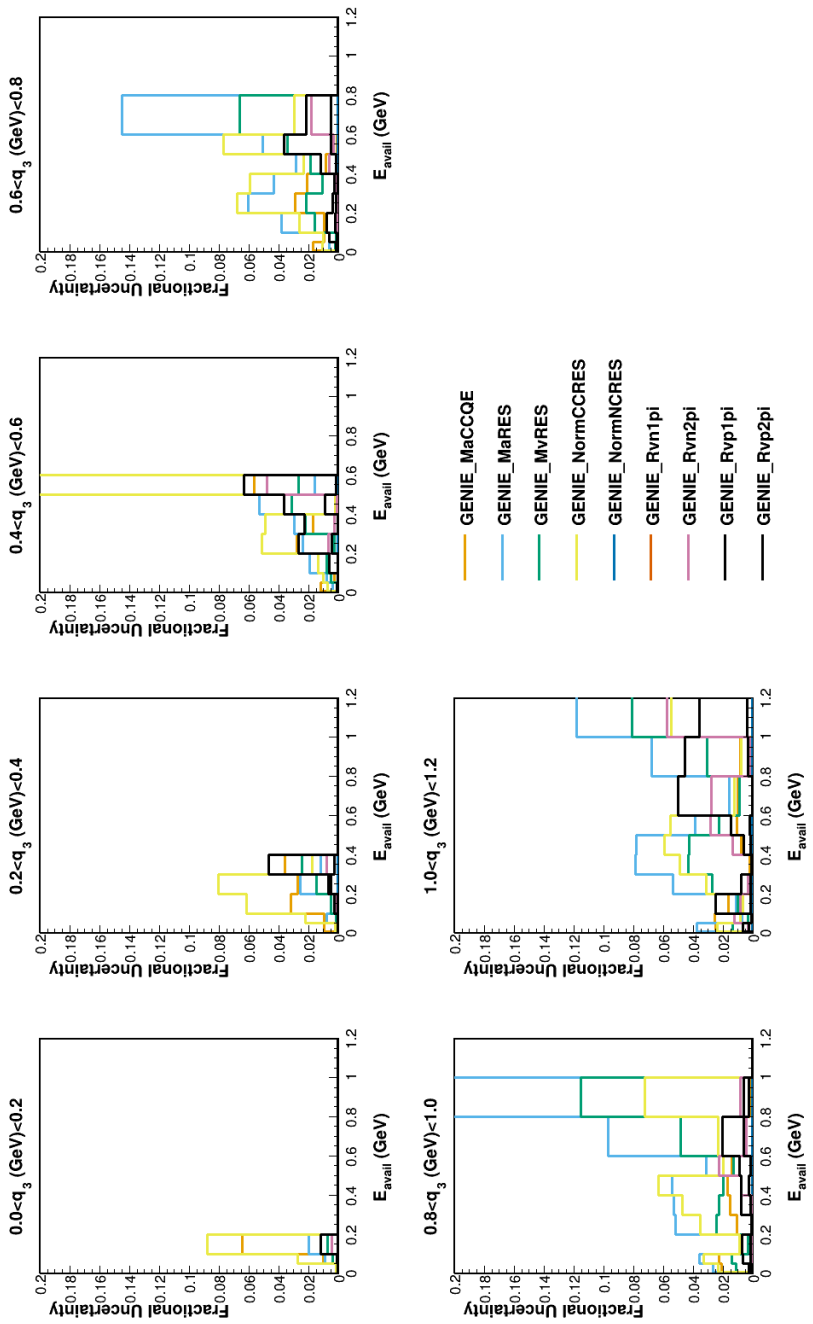
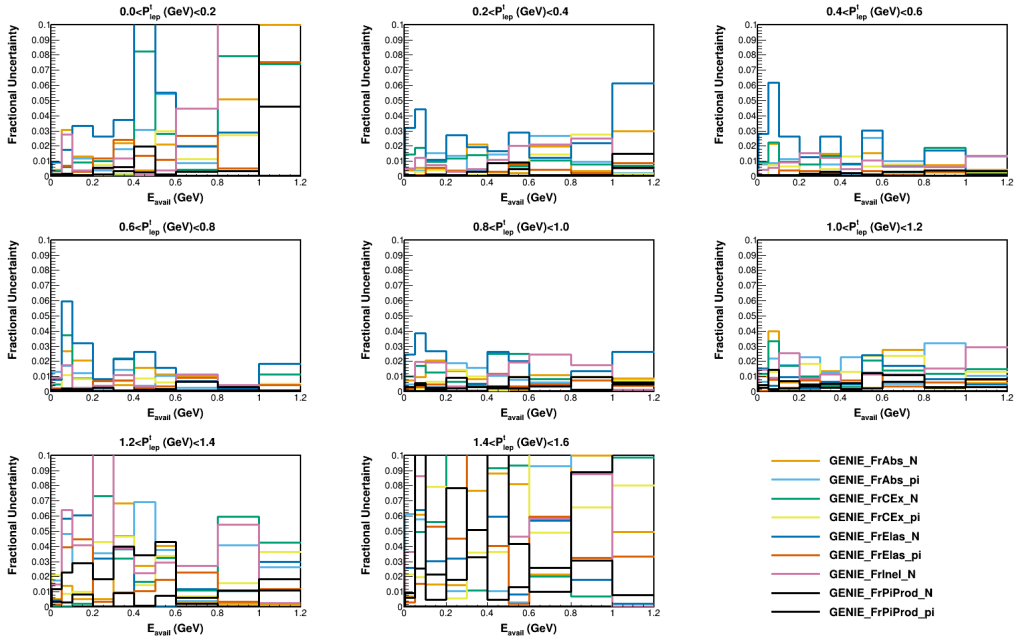
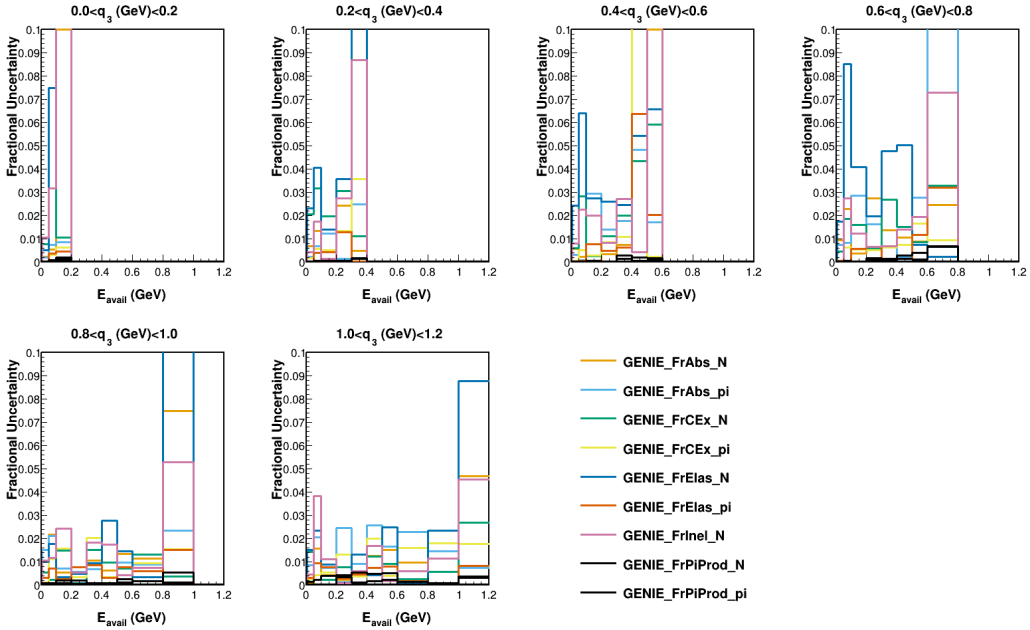


Figure 7.6: A decomposition of the cross section systematic error broken into leading GENIE systematic error contributors for E_{avail} vs q_3. The leading systematic uncertainty is GENIE_MaRES (light blue), GENIE_MvRES (green) and GENIE_NormCCRES (yellow).



(a) E_{avail} vs. p_T cross section error summary isolating GENIE systematic errors divided into interaction model and FSI uncertainties.



(b) E_{avail} vs. q_3 cross section error summary isolating GENIE systematic errors divided into interaction model and FSI uncertainties.

Figure 7.7: Cross section error summary showing GENIE FSI and GENIE interaction model uncertainty contribution.

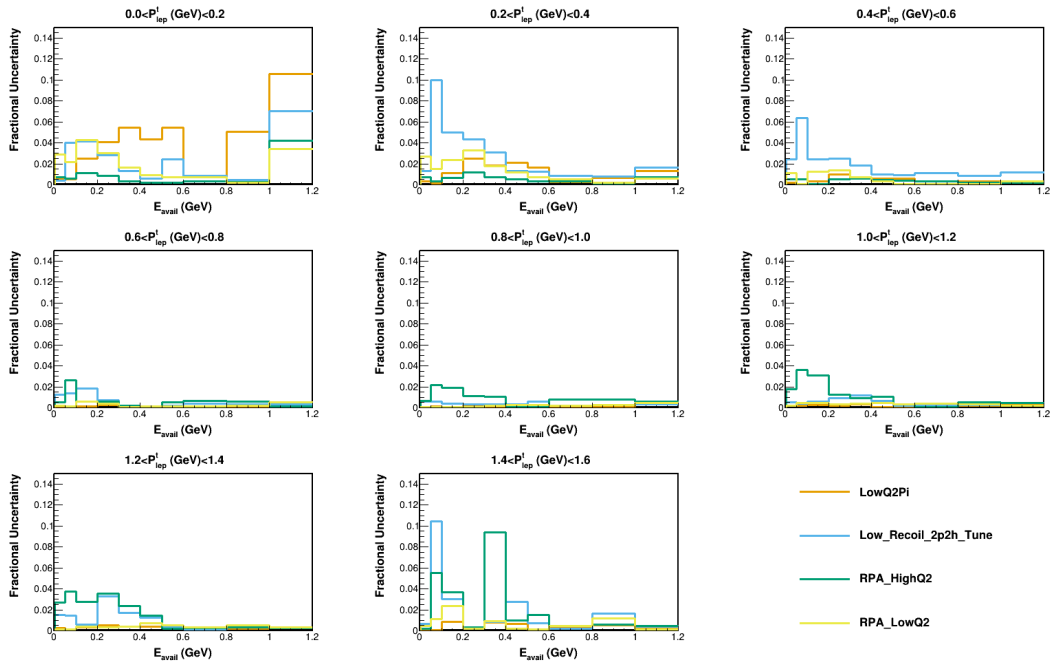
described as single-particle scattering. Each of the three variations has one corresponding universe and the RMS between the universes becomes the 2p2h uncertainty based on the assumption that neutrinos will interact on a variety of correlated nucleon pairs [99].

Non-resonant pion Uncertainty Charged current and neutral current non-resonant pion production is re-weighted based on the reanalysis of bubble chamber data. All CC and NC 1 π knobs are formed together for a single 1 σ uncertainty. The uncertainty is applied by default at 4.0% to the CV universe. The knobs corresponding to non-resonant pion production are GENIE_RvpCC1pi, GENIE_RvpNC1pi for νp and $\bar{\nu} p$ interactions and GENIE_RvnCC1pi, GENIE_RvnNC1pi for νn and $\bar{\nu} n$ interactions [63].

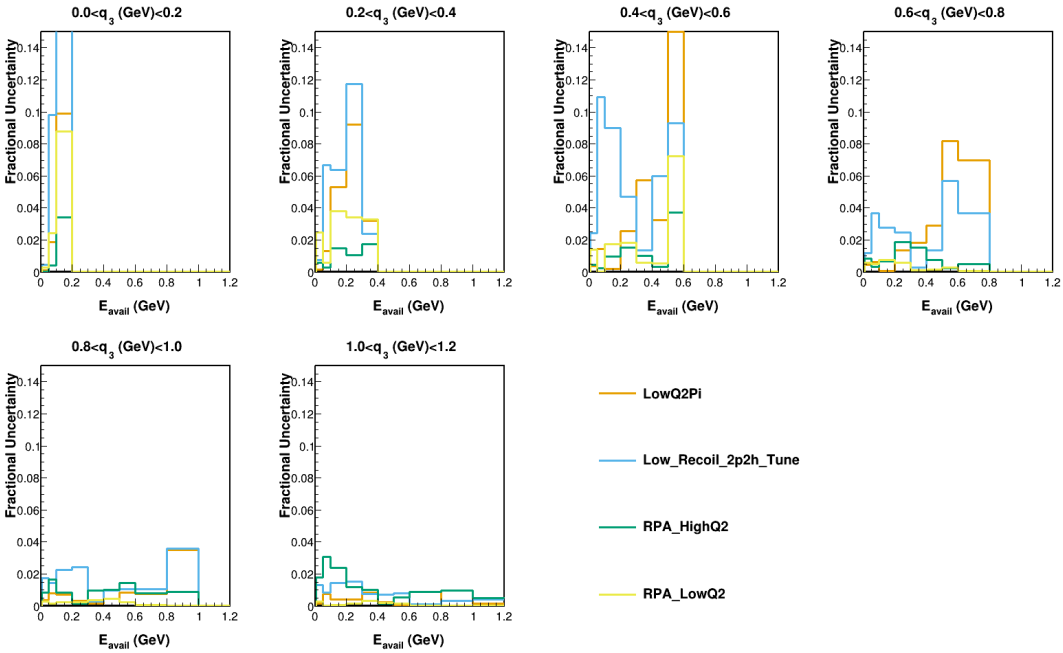
Figure 7.8 shows the systematic error breakdown for the MnvTunes. The low Q^2 tune affects the highest bin of E_{avail} for a given q_3 and p_T. This is expected because these are the regions in which the process is most dominant. The low recoil 2p2h tune has a larger systematic uncertainty for values of q_3 compared to p_T. The shifting of the 2p2h model impacts the E_{avail} distribution and the effects are seen more easily in q_3 due to the model dependency.

7.4.4 Detector Model Uncertainties

The detector model uncertainties consist of the uncertainties pertaining to the simulation of particle propagation through the detector, particle and kinematic reconstruction and the particle response of the detector. The detector model uncertainty for this analysis is broken up into two groups: detector response group including angular resolution, beam angle, EM energy scale, particle response, target mass, leakage uncertainties and GEANT group



(a) E_{avail} vs. p_T cross section error summary isolating Mnvtune uncertainties



(b) E_{avail} vs. q_3 cross section error summary isolating Mnvtune uncertainties

Figure 7.8: Cross section error summary showing MnvTune uncertainties.

uncertainties.

GEANT group uncertainties stem from MINERvA's reliance on GEANT4 to simulate final state particles in the detector. The uncertainty is produced by varying the hadron inelastic cross section of final state hadrons based on either test beam data (for protons and pions) or GEANT cross sections (neutrons).

Derived from studies of the test beam as reported in Ref. [103], the particle response systematic refers to the fractional uncertainty relating to each type of particle contributing to the hadronic recoil system. The universe shifts apply to the proton, meson, electromagnetic, and other categories (which serve to encompass other particles that typically do not account for a large fraction of the total energy). The uncertainties for the detector response are created by shifting the recoil energy a certain fraction for a given particle. In this case there are two universes, broken into an up shift and a down shift. A summary of the fractional uncertainties can be found in Table 7.7 [103], [45]. Additionally, cross-talk uncertainty is included in this systematic and was calculated using a cross-talk bench characterization of the PMTs prior to installation. A best-fit slope to the bench test vs. measured muons in the PMT was used to derive a 20% uncertainty.

The uncertainty on the beam angle is calculated based on an up and down shift in the "X" and "Y" directions with respect to the beam coordinate system. The shifts are 0.1 mrad in X and 0.09 mrad in Y. The ability to determine the angle of the electron is based on the resolution of the tracks in the detector. The estimation is that the position resolution is within 1 mm and the angular resolution is simulated to be within 1 mrad [40]. The uncertainty is evaluated from universes in which the electron angle is shifted by an amount randomly selected from a Gaussian centered at zero with width 1 mrad. The

Particle Type	Energy Range	Fractional Uncertainty
Proton	$E_{kinetic} < 50MeV$	$\pm 4\%$
Proton	$50 < E_{kinetic} < 100MeV$	$\pm 3.5\%$
Proton	$E_{kinetic} > 100MeV$	$\pm 3\%$
Pion	$400 < E_{kinetic}$	$\pm 5\%$
Pion	$400 < E_{kinetic} < 1900MeV$	$\pm 4\%$
Neutron	50 MeV $< E_{kinetic}$	$\pm 25\%$
Neutron	$50 < E_{kinetic} < 150MeV$	$\pm 10\%$
Neutron	$E_{kinetic} > 150MeV$	$\pm 20\%$
Electromagnetic		$\pm 3\%$
Other		$\pm 20\%$
Cross-talk		$\pm 20\%$

Table 7.7: Fractional uncertainties accounted for in the particle response systematic

uncertainty on the energy scale in the tracker and electromagnetic calorimeter is estimated by comparing energy of reconstructed π^0 candidates in charged current ν_μ events between data and simulation. The π^0 sample comparison leads to a 5.8% adjustment for energy deposits in the calorimeter with a 1.5% uncertainty in the electromagnetic response of the tracker and electromagnetic calorimeter. Additionally, based on early results of the MINERvA test bean, a 5% uncertainty is applied to the hadronic calorimeter [40] [34]. The detector mass has a fractional uncertainty of 1.4% [40].

The target mass uncertainty is the uncertainty in the measurement of the density for both pass and active targets. The uncertainty is evaluated by shifting the normalization of the simulated events by $\pm 1\sigma$.

A custom systematic uncertainty is required to correct the leakage of electron energy outside of the electron cone. The energy leakage outside the cone leads to an overestimation of the available energy. The energy leakage was estimated to be %8 of the electron energy as discussed in Section 4.5. The uncertainty is calculated from the available energy distribution of a specific subsection of neutrino-electron elastic scattering ($\nu + e \rightarrow \nu + e$) events. Since

the hadronic energy is 0 for neutrino-electron events by definition, this measurement is an *in-situ* measurement of the energy leakage from the electromagnetic shower. The subset of neutrino-electron events is formed by requiring cuts of Q^2_{QE} <0.02 GeV^2 and $E\theta^2 <$ 0.0032 GeV*Radian^2 for Q^2_{QE} calculated based on CCQE assumptions that use only lepton kinematics as implemented from MINERvA electron-neutrino analysis found in Reference [34]:

$$E_\nu^{QE} = \frac{M_n E_e - m_e^2/2}{m_n - E_n + p_e \cos\theta_e} \tag{7.7}$$

$$Q^2_{QE} = 2m_n(E_\nu^{QE} - E_e) \tag{7.8}$$

As seen in Figure 7.9 the mean value is calculated for data and MC in the region of higher statistics (values of $E_{avail} < 100$ MeV) and the uncertainty is estimated based on the difference of the mean between the first two bins of lepton energy. It is determined that the leakage needs to be corrected from the simulated value by 5 MeV and the uncertainty on the shift is ± 2 MeV.

The systematic uncertainties related to the detector are found in Figure 7.10. The leading uncertainty in q_3 bins with the highest E_{avail} is the leakage uncertainty. The highest p_T bin shows large systematic error values, similar to the GENIE error summary, due to the low number of events in that bin. The leakage uncertainty is the leading systematic uncertainty for the lower p_T bins. The error summary for GEANT systematic uncertainties is found in Figure 7.11. The leading systematic tends to be GEANT_Neutron, especially for the E_{avail} vs q_3 distribution. This is expected because the GEANT neutron uncertainty essentially takes the invisible energy, which has no contribution to E_{avail}, and adds energy from the neutron interacting in the detector. The q_3 distribution is more sensitive to this

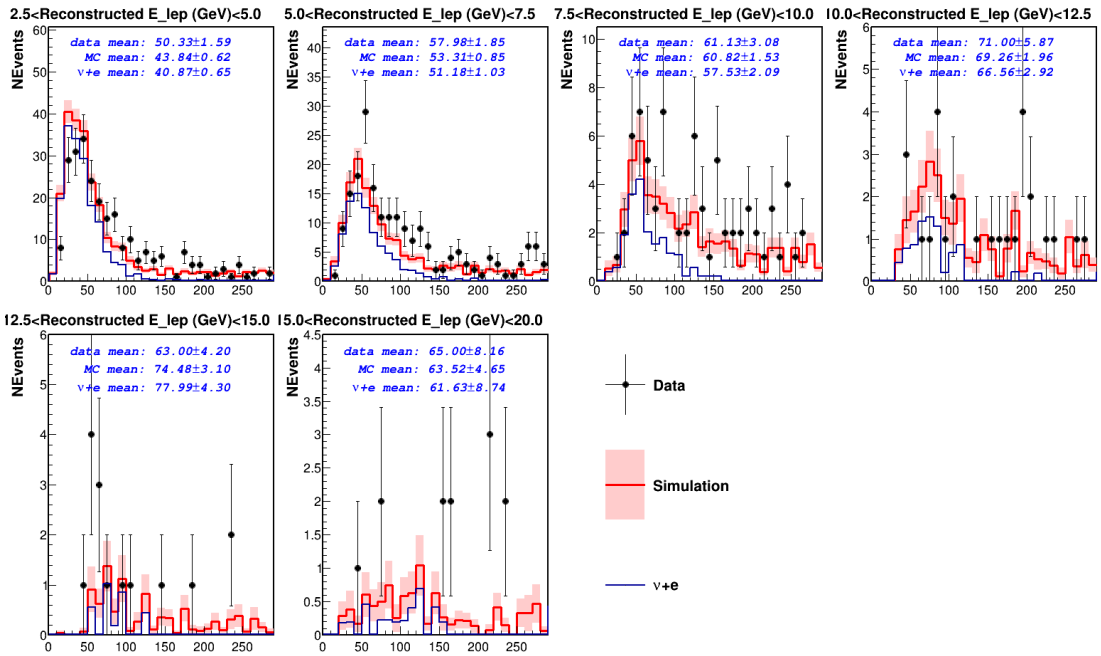
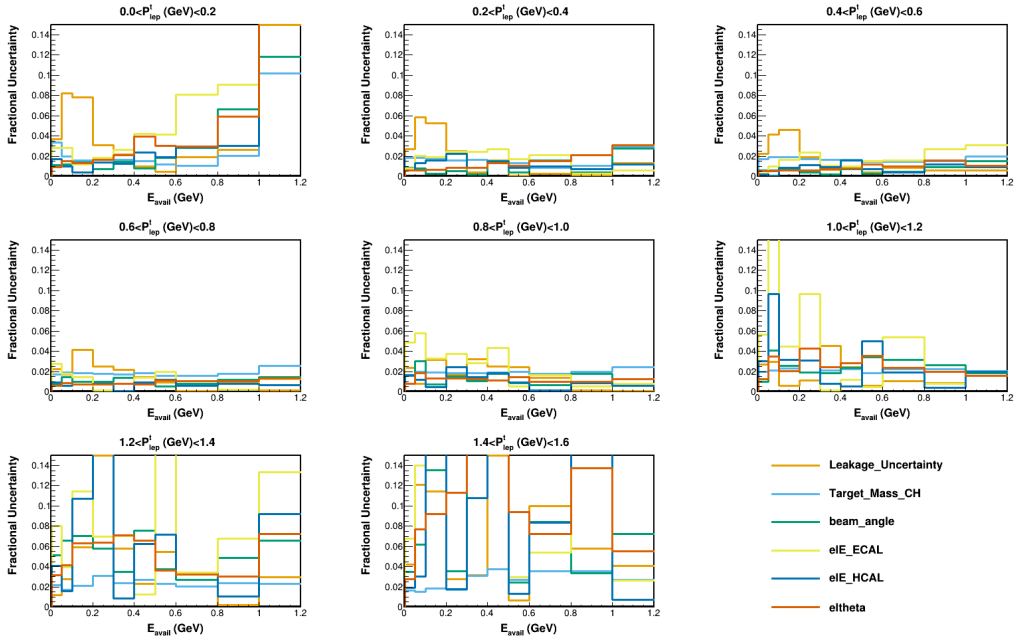
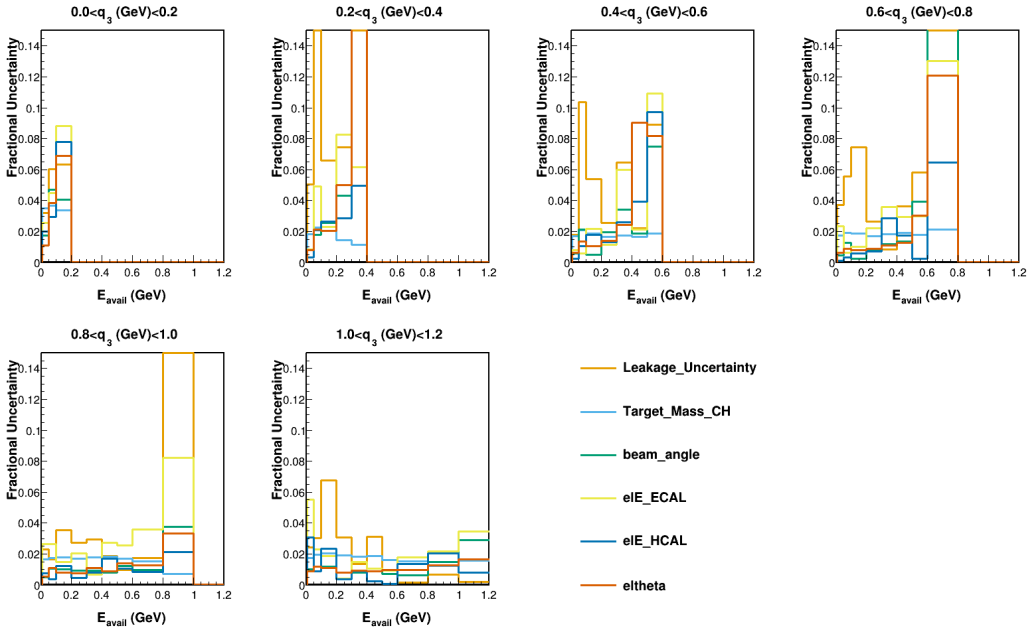


Figure 7.9: Distribution used to derive leakage systematic uncertainty value.

knob since E_{avail} contributes to the calculation of q_3.

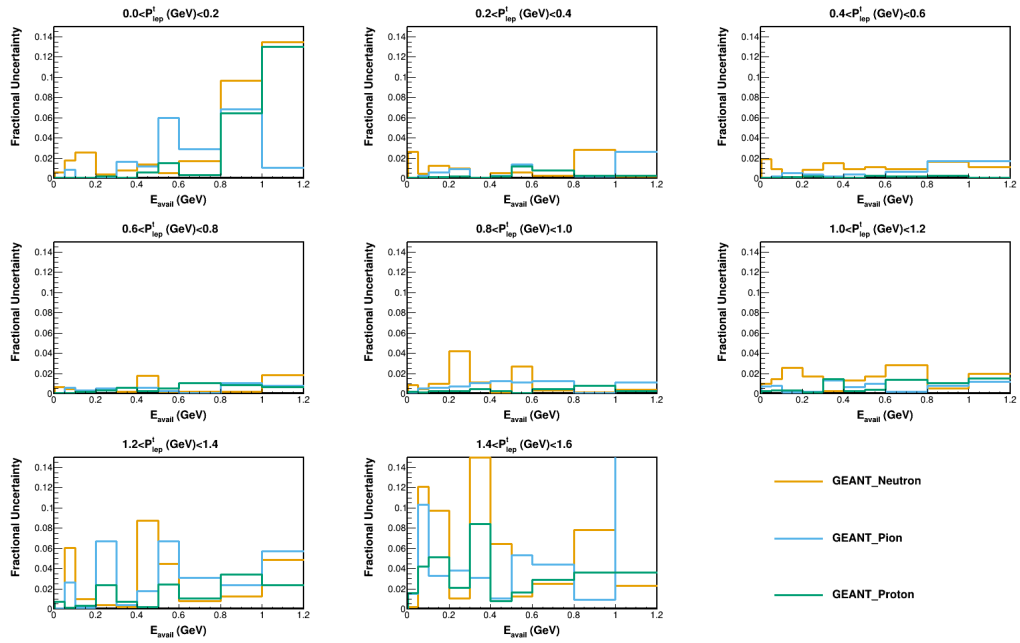


(a) E_{avail} vs. p_T cross section error summary isolating detector response uncertainties.

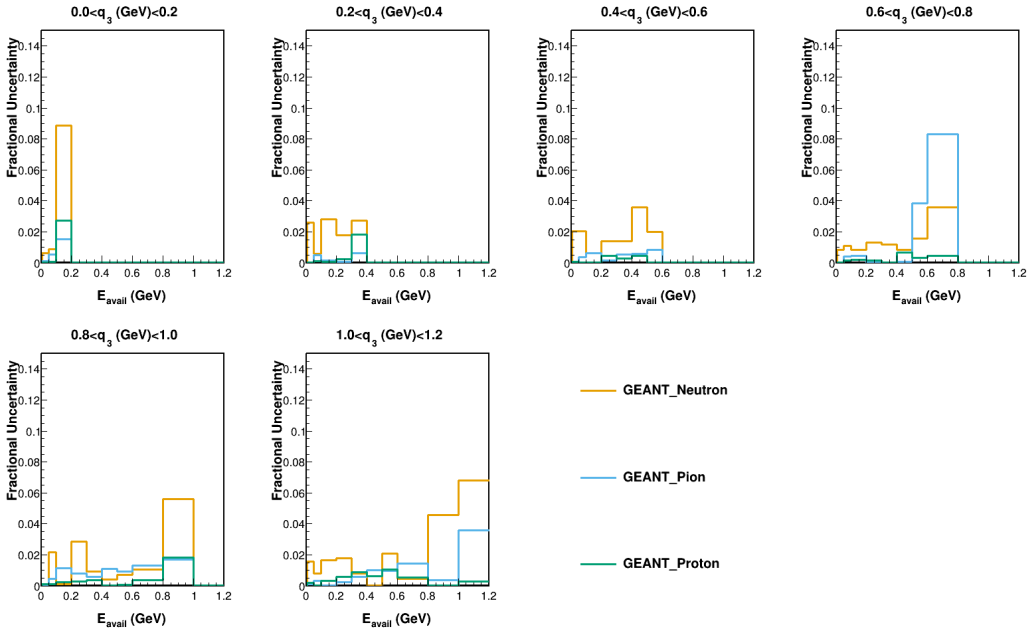


(b) E_{avail} vs. q_3 cross section error summary isolating detector response uncertainties.

Figure 7.10: Cross section error summary looking at the leading detector response uncertainties.



(a) E_{avail} vs. p_T cross section error summary isolating GEANT_Neutron, GEANT_Pion and GEANT_Proton uncertainties.



(b) E_{avail} vs. q_3 cross section error summary isolating GEANT_Neutron, GEANT_Pion and GEANT_Proton uncertainties.

Figure 7.11: Systematic uncertainties for GEANT.

8. Conclusion

8.1 Discussion

A discussion of the results can be broken into three major categories: excess of events at low E_{avail} values in the cross section result, model comparisons, and comparison to previous MINERvA results.

8.1.1 Excess of events at nearly zero E_{avail}

There is an excess of data events over the simulated prediction found in the first bin of E_{avail} seen in the cross section for both E_{avail} vs q_3 in Figures 6.13 and 6.14 and E_{avail} vs p_T as seen in Figures 6.16 and 6.17. A comparison between the measured and predicted cross section values in the first bin of E_{avail} is found in Table 8.1 and 8.2. The values presented in the table indicate the size of the data excess for events for the first bin of E_{avail} compared to the MC prediction.

	$0.0 < q_3(GeV) < 0.2$	$0.2 < q_3(GeV) < 0.4$
MC: $0.0 < E_{avail} < 0.04$	13.2	53.2
Data: $0.0 < E_{avail} < 0.04$	25.1 ± 5.0	63.6 ± 8.4
	$0.4 < q_3(GeV) < 0.6$	$0.6 < q_3(GeV) < 0.8$
MC: $0.0 < E_{avail} < 0.04$	77.5	62.0
Data: $0.0 < E_{avail} < 0.04$	109.5 ± 10.8	93.6 ± 9.7
	$0.8 < q_3(GeV) < 1.0$	$1.0 < q_3(GeV) < 1.2$
MC: $0.0 < E_{avail} < 0.04$	41.5	28.3
Data: $0.0 < E_{avail} < 0.04$	65.4 ± 7.2	38.8 ± 5.8

Table 8.1: Comparison of the cross section for data and MC in the first bin of E_{avail} for all bins of q_3. The table entries are $10^{-39}\frac{cm^2}{GeV^2}$.

	$0.0 < P^t_{lep}(GeV) < 0.2$	$0.2 < P^t_{lep} < 0.4$	$0.4 < P^t_{lep} < 0.6$
MC: $0.0 < E_{avail} < 0.04$	19.9	71.8	90.8
Data: $0.0 < E_{avail} < 0.04$	34.2 ± 6.1	94.5 ± 10.6	133.4 ± 12.6
	$0.6 < P^t_{lep} < 0.8$	$0.8 < P^t_{lep} < 1.0$	$1.0 < P^t_{lep} < 1.2$
MC: $0.0 < E_{avail} < 0.04$	63.8	36.5	15.5
Data: $0.0 < E_{avail} < 0.04$	99.4 ± 9.8	55.2 ± 6.6	21.3 ± 3.7
	$1.2 < P^t_{lep} < 1.4$	$1.4 < P^t_{lep} < 1.6$	
MC: $0.0 < E_{avail} < 0.04$	6.2	2.7	
Data: $0.0 < E_{avail} < 0.04$	7.7 ± 2.1	13.2 ± 3.6	

Table 8.2: Comparison of the cross section for data and MC in the first bin of E_{avail} for all bins of p_T. The table entries are $10^{-39}\frac{cm^2}{GeV^2}$.

Throughout this analysis, we know that the events populating the first bin of available energy tend to be events where the final state is neutral, typically final state neutrons. In Figures 6.19 and 6.22 we can see that the first bin of E_{avail} consists of charged current quasi-elastic events that account for $\sim$ 60-90% of the model prediction. The quasi-elastic events are expected to be the dominant contributor to the first E_{avail} bin because, in the absence of final state interactions, there is only a lepton and neutron in the final state.

Looking more closely at the q_3 bin of 0.4-0.6 GeV in Figure 6.19, there is a population

of inelastic events that leak into the first bin of E_{avail}. It is possible that some type of inelastic events with mostly neutrons in the final state is not being correctly simulated. It also could involve events where the final state pion doesn't have much energy and is absorbed within the nucleus, resulting in only final state neutrons. The last proposal to explain the excess of data events in the first E_{avail} bin is that the MC simulation predicts too many quasi-elastic events at higher values of E_{avail}. A larger population of quasi-elastic MC events near zero E_{avail} would help fill in the excess.

8.1.2 Model Comparison

The model comparisons seen in Figure 6.26 show a large enhancement for the MvnTune v1 at high E_{avail} values in q_3. As discussed in Section 3.7.4, the central value uses the MnvTune v2, which includes a low Q^2 pion suppression. A low Q^2 suppression equates to the highest E_{avail} values for a given q_3 for the events where p_T is low. Therefore, a large change in the pion production would manifest as a large change in MnvTune v1, as seen in the figure. As expected for the model comparisons in E_{avail} vs p_T shown in Figure 6.28, there is a large enhancement for MnvTune v1 at low values of p_T. Although there is a better agreement in the low p_T bins for MnvTune v1, this discrepancy is seen in all bins for high values of E_{avail}. This indicates that there might be a better agreement with data with an enhancement of pion production in all bins relative to MnvTune v2.

8.1.3 Comparison with Published $\bar{\nu}_\mu$ Measurements

The last step is to compare the result with MINERvA's low recoil LE $\bar{\nu}_\mu$ result [107]. Major differences in the two analyses make direct comparison between the two results difficult.

ME $\bar{\nu}_e$	LE $\bar{\nu}_\mu$
$0.0 < E_{avail}(GeV) < 0.04$	$0.0 < E_{avail}(GeV) < 0.03$
$0.04 < E_{avail}(GeV) < 0.08$	$0.03 < E_{avail}(GeV) < 0.07$
$0.8 < E_{avail}(GeV) < 0.12$	$0.07 < E_{avail}(GeV) < 0.17$
$0.12 < E_{avail}(GeV) < 0.16$	$0.17 < E_{avail}(GeV) < 0.27$
$0.16 < E_{avail}(GeV) < 0.24$	$0.27 < E_{avail}(GeV) < 0.35$
$0.24 < E_{avail}(GeV) < 0.32$	$0.35 < E_{avail}(GeV) < 0.5$
$0.32 < E_{avail}(GeV) < 0.4$	
$0.4 < E_{avail}(GeV) < 0.5$	

Table 8.3: Comparison between the E_{avail} binning used for the low recoil ME $\bar{\nu}_e$ analysis (left) and the low recoil LE $\bar{\nu}_\mu$ analysis (right). The $\bar{\nu}_e$ binning is truncated at 0.5 GeV for the comparison to LE but the results are reported up to 1.2 GeV.

The $\bar{\nu}_\mu$ measurement was made during MINERvA's low energy era for neutrino energies between 2 GeV and 6 GeV, while the reported $\bar{\nu}_e$ cross section measurement is in ME with peak a flux value of $\sim$6 GeV. Additionally, the neutrino flux spectrums of $\bar{\nu}_e$ and $\bar{\nu}_\mu$ are different.

There are also selection level differences between the two analyses, so the signal definitions are not identical. The $\bar{\nu}_\mu$ result requires a lepton momentum of greater than 1.5 GeV, while the $\bar{\nu}_e$ analysis requires lepton energy greater than 2.5 GeV to eliminate a large π^0 background at low electron energy. In addition, the $\bar{\nu}_e$ analysis has no scattering angle requirements while the $\bar{\nu}_\mu$ analysis requires the lepton scattering angle to be less than 20 degrees due to the difficulty of reconstructing high angle muons. The analyses are also reported using different binning. Table 8.3 shows a binning comparison between the ME and LE results for E_{avail} and Table 8.4 shows the comparison for q_3.

Lastly, the two analyses took different approaches in unfolding. The $\bar{\nu}_e$ unfolds using coarse binning and a large number of iterations and the $\bar{\nu}_\mu$ analysis unfolds using fine binning and a small number of iterations. This is due in large to the difference in observables.

ME $\bar{\nu}_e$	LE $\bar{\nu}_\mu$
$0.0 < q_3(GeV) < 0.2$	$0.0 < q_3(GeV) < 0.2$
$0.2 < q_3(GeV) < 0.4$	$0.2 < q_3(GeV) < 0.3$
$0.4 < q_3(GeV) < 0.6$	$0.3 < q_3(GeV) < 0.4$
$0.6 < q_3(GeV) < 0.8$	$0.4 < q_3(GeV) < 0.5$
$0.8 < q_3(GeV) < 1.0$	$0.5 < q_3(GeV) < 0.6$
$1.0 < q_3(GeV) < 1.2$	$0.6 < q_3(GeV) < 0.8$

Table 8.4: Comparison between the q_3 binning used for the low recoil ME $\bar{\nu}_e$ analysis (left) and the low recoil LE $\bar{\nu}_\mu$ analysis (right).

The ME $\bar{\nu}_e$ analysis has to account for the energy leakage outside the electron cone and into the available energy. Overall, the LE $\bar{\nu}_\mu$ has a much better energy resolution compared to the $\bar{\nu}_e$ analysis. With the consideration of the differences between the two analyses, the cross section result for the LE $\bar{\nu}_\mu$ is shown in Figure 8.1 and the relevant cross section bins for the $\bar{\nu}_e$ are shown in Figure 8.2.

There are similar features between the two results. As expected in the cross section model prediction, both cross section results have quasi-elastic events as the dominant contributor in the first bin of E_{avail}. There is a population of 2p2h events for values of low E_{avail}. The delta resonance becomes the dominant process at the higher values of E_{avail} in both predicted cross sections. There is a noticeable difference between the data results for values of $\sim E_{avail} > 0.2$ GeV. The $\bar{\nu}_\mu$ cross section prediction is consistently over-predicting the data while the $\bar{\nu}_e$ under-predicts the data.

Both cross section results contain many events that have no available energy. This creates a sharp peak at zero E_{avail} followed by a cross section that falls slowly compared to the size of the peak in the first E_{avail} bin. Therefore, to compare the first two E_{avail} bins for both results we assume that the peak at zero E_{avail} is a delta-like peak and the remaining cross section distribution is flat. To determine the magnitude of the delta function we

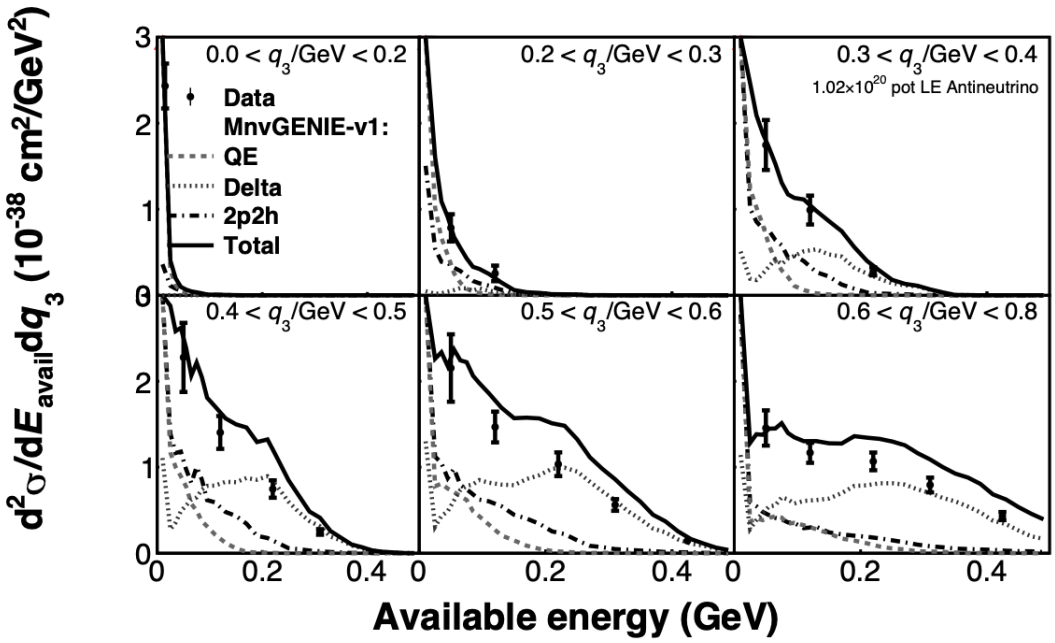


Figure 8.1: $d^2\sigma/dE_{avail}dq_3$ cross section per nucleon compared to the model with RPA and tune 2p2h components. Figure from Ref. [107]

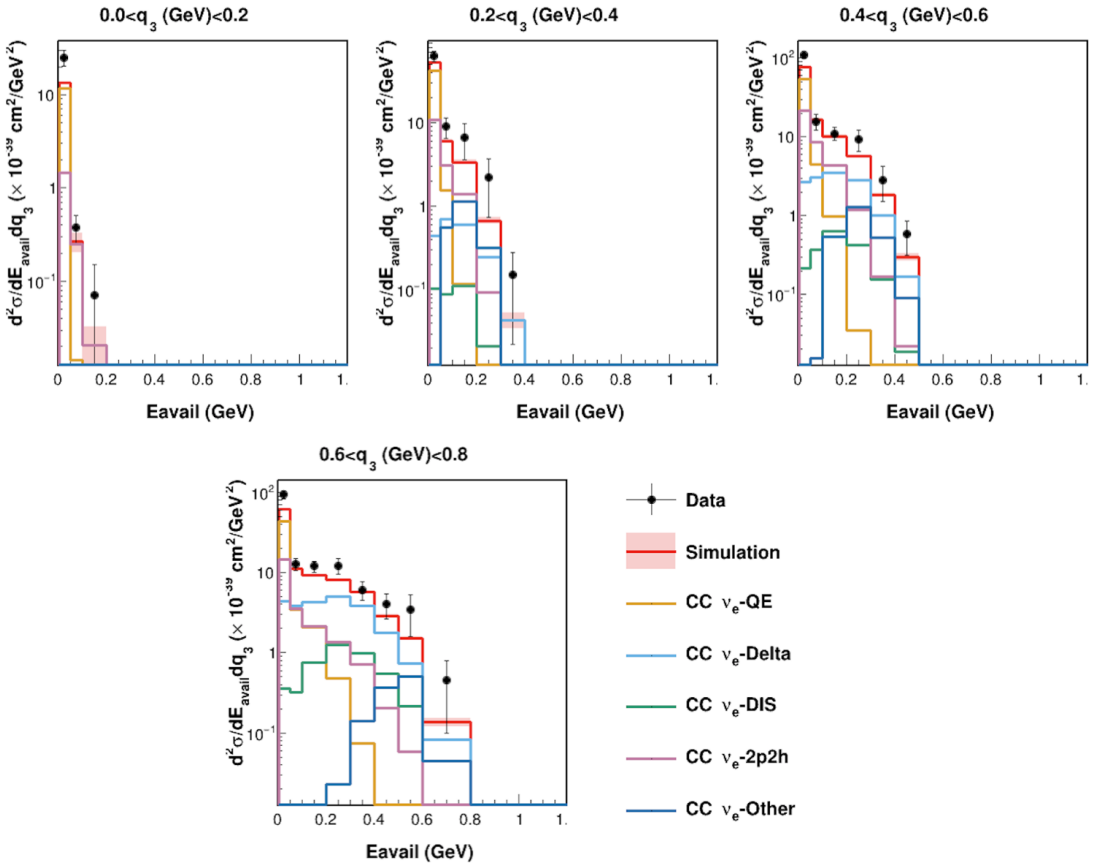


Figure 8.2: The $\bar{\nu}_e$ cross section result presented in this thesis truncated at 0.8 GeV on logy scale in q_3 for comparison to the LE result.

subtract the second E_{avail} bin from the first, or the flat distribution from the peak, leaving us with a $d\sigma/dE_{avail}$ value. We multiply $d\sigma/dE_{avail}$ by the bin width so that we end up with a cross section that is differential in each q_3 bin. This process is repeated for each result's data and MC values. The resultant bin combination for the two samples are 0.0 $< E_{avail}$(GeV) < 0.08 for the $\bar{\nu}_e$ cross section result and $0.0 < E_{avail}$(GeV) < 0.07 for the $\bar{\nu}_\mu$ cross section result. Tables 8.6 and 8.8 summarize the results. Tables 8.7 and 8.9 are the correlation matrices for the reported bins with the correlation matrix ordering defined in Table 8.5.

The conclusion drawn from the comparison between the data/MC peak at zero is that the ME $\bar{\nu}_e$ result is consistent with the LE $\bar{\nu}_\mu$ result in all q_3 bins except the first. The ME $\bar{\nu}_e$ result has a significant enhancement over the simulation.

	$0.0 < q_3(GeV) < 0.2$	$0.2 < q_3(GeV) < 0.4$
Estimated peak at zero	1	3
Subtraction size from first bin	2	4
	$0.4 < q_3(GeV) < 0.6$	$0.6 < q_3(GeV) < 0.8$
Estimated peak at zero	5	7
Subtraction size from first bin	6	8

Table 8.5: Correlation matrix ordering for both results. Note that the correlation matrix value for the data/MC peak is equivalent to the estimated peak at zero value.

	$0.0 < q_3(GeV) < 0.2$	Diagonal Unc.
Estimated data peak at zero	0.99	0.20
Data subtraction size from first bin	0.015	0.005
Estimated MC peak at zero	0.52	x
MC subtraction size from first bin	0.01	x
Data/MC peak at zero	1.90	0.38
	$0.2 < q_3(GeV) < 0.4$	Diagonal Unc.
Estimated data peak at zero	2.18	0.35
Data subtraction size from first bin	0.36	0.09
Estimated MC peak at zero	1.87	x
MC subtraction size from first bin	0.29	x
Data/MC peak at zero	1.17	0.19
	$0.4 < q_3(GeV) < 0.6$	Diagonal Unc.
Estimated data peak at zero	3.12	0.45
Data subtraction size from first bin	0.67	0.13
Estimated MC peak at zero	2.43	x
MC subtraction size from first bin	0.79	x
Data/MC peak at zero	1.53	0.18
	$0.6 < q_3(GeV) < 0.8$	Diagonal Unc.
Estimated data peak at zero	3.24	0.40
Data subtraction size from first bin	0.51	0.08
Estimated MC peak at zero	2.05	x
MC subtraction size from first bin	0.52	x
Data/MC peak at zero	1.58	0.19

Table 8.6: Summary of results for ME $\bar{\nu}_e$.

1.00	0.47	0.16	-0.03	0.36	0.04	0.36	0.15
0.47	1.00	-0.06	0.51	0.07	0.27	0.13	0.04
0.16	-0.06	1.00	-0.32	0.49	-0.30	0.51	-0.21
-0.03	0.51	-0.32	1.00	-0.30	0.60	-0.30	0.42
0.36	0.07	0.49	-0.30	1.00	-0.26	0.53	-0.26
0.04	0.27	-0.30	0.60	-0.26	1.00	-0.30	0.72
0.36	0.13	0.51	-0.30	0.53	-0.30	1.00	-0.14
0.15	0.04	-0.21	0.42	-0.26	0.72	-0.14	1.00

Table 8.7: Correlation matrix for ME $\bar{\nu}_e$ result for relevant bins.

	$0.0 < q_3(GeV) < 0.2$	Diagonal Unc.
Estimated data peak at zero	0.73	0.078
Data subtraction size from first bin	0.00	0.000
Estimated MC peak at zero	0.9	x
MC subtraction size from first bin	0.08	x
Data/MC peak at zero	0.74	0.08
	$0.2 < q_3(GeV) < 0.4$	Diagonal Unc.
Estimated data peak at zero	3.1	0.33
Data subtraction size from first bin	0.38	0.06
Estimated MC peak at zero	2.55	x
MC subtraction size from first bin	0.36	x
Data/MC peak at zero	1.15	0.12
	$0.4 < q_3(GeV) < 0.6$	Diagonal Unc.
Estimated data peak at zero	4.1	0.51
Data subtraction size from first bin	0.66	0.12
Estimated MC peak at zero	3.46	x
MC subtraction size from first bin	0.51	x
Data/MC peak at zero	1.16	0.14
	$0.6 < q_3(GeV) < 0.8$	Diagonal Unc.
Estimated data peak at zero	3.5	0.42
Data subtraction size from first bin	0.44	0.06
Estimated MC peak at zero	2.91	x
MC subtraction size from first bin	0.26	x
Data/MC peak at zero	1.20	0.15

Table 8.8: Summary of results for LE $\bar{\nu}_\mu$.

1.00	0.00	0.82	0.37	0.67	0.38	0.56	0.42
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.82	0.00	1.00	0.19	0.83	0.30	0.76	0.38
0.37	0.00	0.19	1.00	0.21	0.75	0.01	0.70
0.67	0.00	0.83	0.21	1.00	0.32	0.86	0.38
0.38	0.00	0.30	0.75	0.32	1.00	0.10	0.85
0.56	0.00	0.76	0.01	0.86	0.10	1.00	0.24
0.42	0.00	0.38	0.70	0.38	0.85	0.24	1.00

Table 8.9: Correlation matrix for the LE $\bar{\nu}_\mu$ result for relevant bins.

8.2 Ratio of $\bar{\nu}_\mu$ and $\bar{\nu}_e$ samples

As mentioned in the discussion about the LE $\bar{\nu}_\mu$ result in Section 8.1.3, muon analyses use different reconstruction algorithms, binning choices, and unfolding procedures. For a more direct comparison, we form a method that uses low recoil $\bar{\nu}_\mu$ and $\bar{\nu}_e$ samples at a reconstructed level with the intention of producing a double ratio comparison of $\bar{\nu}_\mu$ and $\bar{\nu}_e$ events of data and acceptance corrected Monte Carlo simulation or $\frac{\bar{\nu}_e^{data}/\bar{\nu}_\mu^{data}}{\bar{\nu}_e^{MC}/\bar{\nu}_\mu^{MC}}$. This effort is underway and only the preliminary studies are shown here.

There are differences in the cross section models between $\bar{\nu}_\mu$ and $\bar{\nu}_e$ but we expect those differences to be small when reporting in E_{avail}. In a neutrino interaction there has to be enough energy transfer to produce a muon but that energy doesn't appear in E_{avail}. Therefore, reporting in E_{avail} vs. q_3 or E_{avail} vs. p_T minimizes the cross section differences between the two neutrino flavors to an extent. In this sample comparison, the "electron cone" exclusion region utilized in the low recoil $\bar{\nu}_e$ result is also applied to the muon. Additionally, equivalent lepton kinematic cuts are applied between the two samples. This includes an energy range of $2 < E_{lep} < 20$ GeV and $\theta_{lep} < 20°$.

Since the fluxes of muons and electrons are not equivalent, the acceptance for the $\bar{\nu}_\mu$ must be corrected. The first step to producing an acceptance correction is to develop a weighting scheme to apply to the $\bar{\nu}_\mu$ sample. A neutrino energy estimator is formed using reconstructed available energy and reconstructed lepton energy, in a method similar to the one discussed in Section 5.2.1. Figure 8.3 shows that $E_e + E_{avail}$ is a good estimator of true neutrino energy. A ratio is formed from the RHC $\bar{\nu}_e$ flux to RHC $\bar{\nu}_\mu$ flux. The $\bar{\nu}_\mu$ sample is then weighted on an event-by-event by basis by the flux ratio using the energy estimator to select the weight in bins of true neutrino energy. Figure 8.4 shows a MC

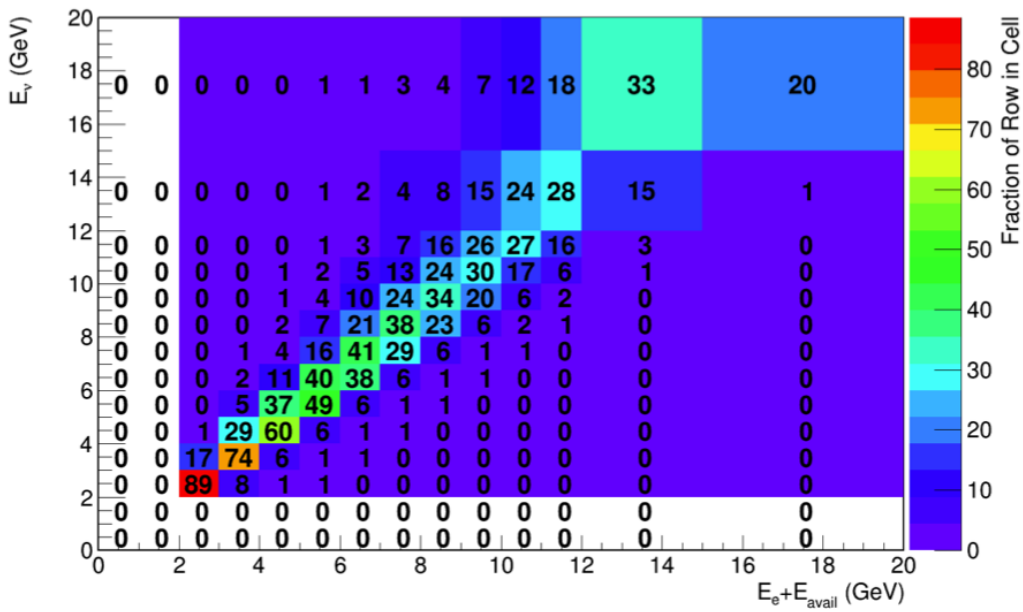


Figure 8.3: Neutrino energy estimator of electron energy plus available energy on the x-axis vs true neutrino energy on the y-axis.

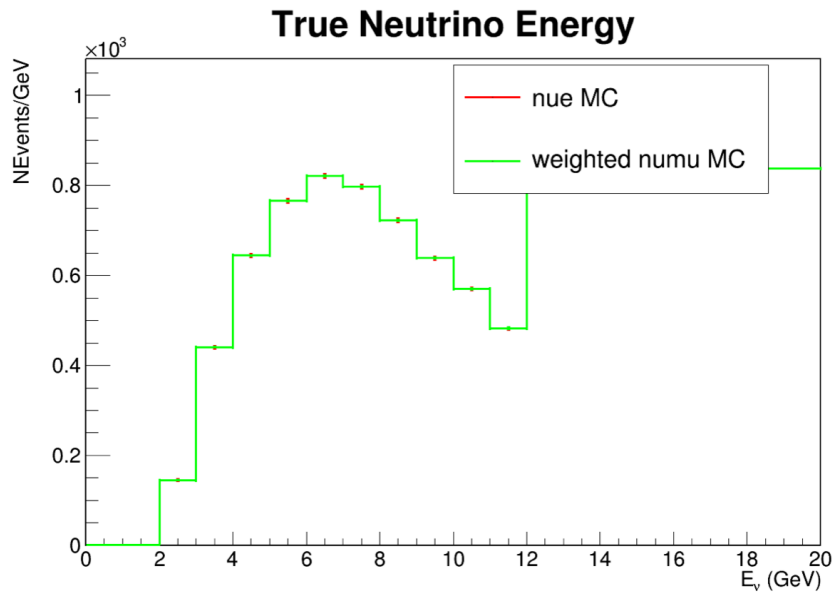


Figure 8.4: MC comparison of the $\bar{\nu}_e$ (red) and the flux weighted $\bar{\nu}_e$ (green) samples in true neutrino energy.

comparison between the $\bar{\nu}_e$ and the flux weighted $\bar{\nu}_e$ samples in bins of true neutrino energy. By construction, the weighting gives nearly identical agreement between the samples.

In the case of the low recoil $\bar{\nu}_e$ cross section result, the efficiency correction was defined as the ratio of selected true signal events to the number of true signal events simulated. In this case, we cannot apply an efficiency correction defined in only truth variables since it is a ratio of reconstructed quantities. Ideally, we would construct an efficiency correction in all reconstructed variables; however, you cannot have a reconstructed efficiency denominator by definition because there are events that will fail the reconstruction. For the purpose of this study we define two separate efficiency corrections. The first is a true efficiency which refers to the standard efficiency definition, as discussed in Section 6.3, used to extract the $\bar{\nu}_e$ cross section and a reconstructed efficiency defined as the ratio of accepted events of

reconstructed to true variables.

Figure 8.5 shows the true efficiency in bins of lepton p_T and true E_{avail}. There is a noticeably lower efficiency for the $\bar{\nu}_e$ selection at higher E_{avail} values. This is expected because we know that an electron neutrino interaction at high E_{avail} will have EM showers and hadronic activity, making it difficult for the electron reconstruction algorithm. The figure is used as a reference to confirm that the selection makes sense; however, it is not used as an efficiency correction. Figure 8.6 shows the reconstructed efficiency. There is a large migration of events out of the first E_{avail} bin which introduces a larger efficiency correction. This migration of events and the poor efficiency seen in the first E_{avail} bin still requires understanding and further studies. As seen in the figure, the reconstructed efficiencies are typically worse for high p_T muons because, even with an angle requirement of $< 20°$, many exit out the side of the MINERvA detector.

Figure 8.7 shows a comparison between the flux weighted and reconstructed efficiency corrected $\bar{\nu}_\mu$ sample to the reconstructed efficiency corrected $\bar{\nu}_e$ sample.[1] Figure 8.8 shows the ratio between the reconstructed efficiency corrected $\bar{\nu}_e$ MC sample to the flux weighted and reconstructed efficiency corrected MC $\bar{\nu}_\mu$ sample. After the efficiency corrections, the $\bar{\nu}_e$ and $\bar{\nu}_\mu$ samples have very good agreement. The next step in the ratio analysis is to interpret the data. Currently, we are performing studies to understand the systematic uncertainties on the samples. The final result will produce the ratios of $\bar{\nu}_e^{data}/\bar{\nu}_\mu^{data}$, $\bar{\nu}_e^{MC}/\bar{\nu}_\mu^{MC}$ and the double ratio, $\frac{\bar{\nu}_e^{data}/\bar{\nu}_\mu^{data}}{\bar{\nu}_e^{MC}/\bar{\nu}_\mu^{MC}}$.

[1]The figure shows the MC with a data overlay. The data is not yet well understood and no conclusions can be drawn from it.

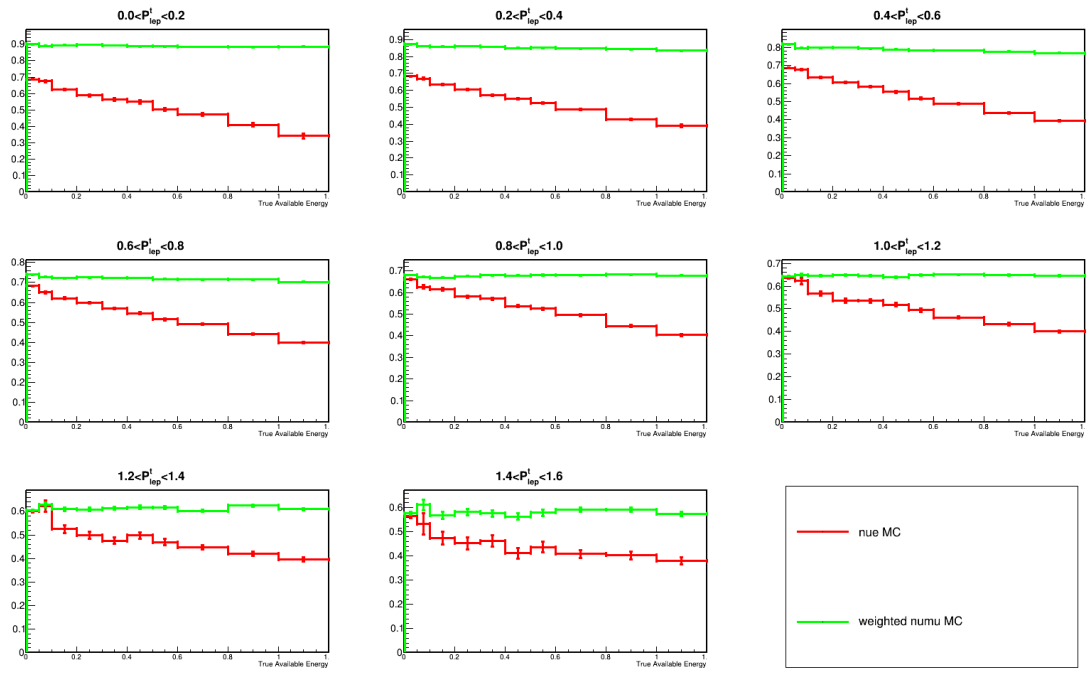


Figure 8.5: Comparison plot between the true efficiency of the flux weighted $\bar{\nu}_\mu$ events (green) and the unweighted $\bar{\nu}_e$ events (red).

!htp

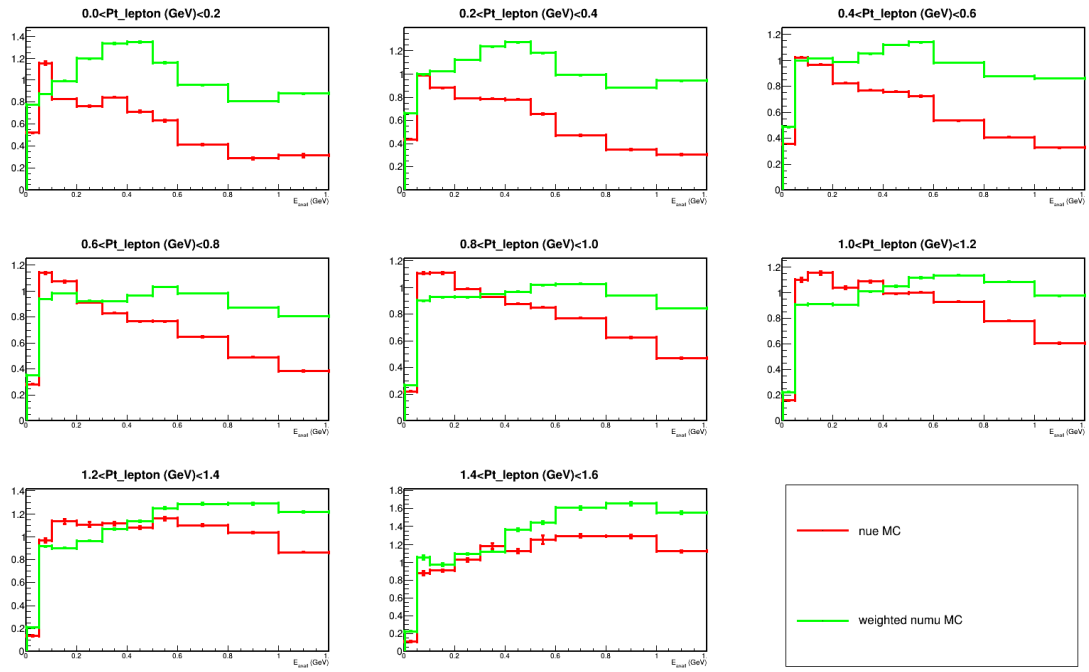


Figure 8.6: Comparison plot between the reconstructed efficiency of the flux weighted $\bar{\nu}_\mu$ events (green) and the unweighted $\bar{\nu}_e$ events (red).

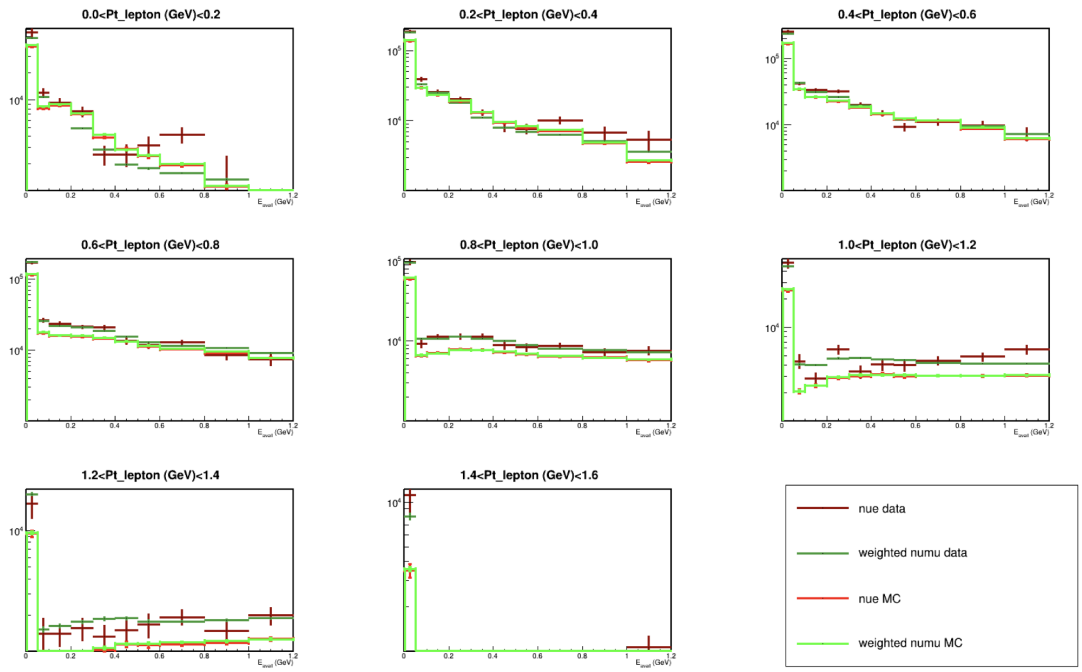


Figure 8.7: The flux weighted and reconstructed efficiency corrected $\bar{\nu}_\mu$ sample compared for MC (light green) and to the reconstructed efficiency corrected $\bar{\nu}_e$ sample for MC (light red) and data on a logy scale. The data has not yet been studied in detail and no comments can be made about its agreement to the MC.

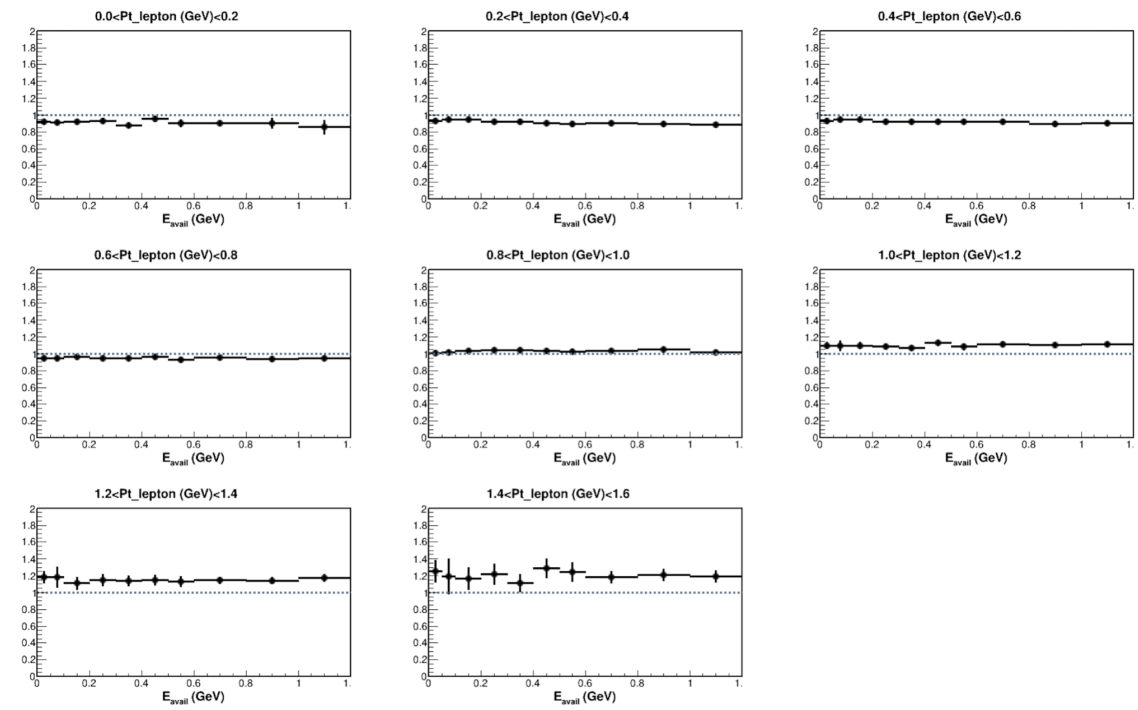


Figure 8.8: A MC ratio of the reconstructed efficiency corrected $\bar{\nu}_e$ sample to the the flux weighted and reconstructed efficiency corrected $\bar{\nu}_\mu$ sample.

Bibliography

[1] Wolfgang Pauli. “Pauli letter collection: Letter to Lise Meitner”. Typed copy (1930).

[2] James Chadwick. “The existence of a neutron. Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character”. 136.830 (1932).

[3] Roger Stuewer. “The Seventh Solvay Conference: Nuclear Physics at the Crossroads”. No Truth Except in the Details: Essays in Honor of Martin J. Klein 978-94-011-0217-9. (1995).

[4] Enrico Fermi. “An attempt of a theory beta radiation”. Z. Physics 88 (1934).

[5] H. R. Crane. Phys. Rev. 55, 501 (1939).

[6] F. Reines et al. C. L. Cowan Jr. Science 124, 103 (1956).

[7] DONUT Collaboration. “Observation of Tau Neutrino Interactions”. Phys.Lett. B504:218-224 (2001).

[8] Bruce T. Cleveland et al. “Measurement of the Solar Electron Neutrino Flux with the Homestake Chlorine Detector”. The Astrophysical Journal 496(1):505 (1998).

[9] Rolfs Haxton Parker. “Solar Hydrogen Burning and Neutrinos”. Nucl. Phys. 226-253 (2006).

[10] John N. Bahcall. Physical Review Letters 12, 300 (1964).

[11] Jr. Raymond Davis. Physical Review Letters 12, 303 (1964).

[12] ALICE Collaboration. “Unveiling the strong interaction among hadrons at the LHC”. Nature 588, 232-238 (2020).

[13] T. Kajita. “Atmospheric Neutrinos”. New Journal of Physics vol.6 (2004).

[14] Y. Fukuda et al. Physical Review Letters 81, 1562 (1998).

[15] SNO Collaboration. Physical Review Letters 87, 071301 (2001).

[16] A. Pich. "The Standard Model of Electroweak Interactions". arXiv:1201.0537 (2012).

[17] K. Altenmuller et al. M. Aker. "An improved upper limit on the neutrino mass from a direct kinematic method by KATRIN". Phys. Rev. Lett. 123.221802 (2019).

[18] Boris Kayser. "Neutrino Mass, Mixing, and Flavor Change". arXiv:0804.1497 (2008).

[19] Neutrino Oscillations. `https://en.wikipedia.org/wiki/Neutrino_oscillation-cite_note-MeszÃľna_2015-24`. Nov. 2021.

[20] K. Abe et al. [T2K]. Nucl. Instrum. Meth. A 659 (2011).

[21] D. S. Ayres et al. [NOvA]. "The NOvA Technical Design Report". doi:10.2172/935497 ().

[22] R. Acciarri et al. [DUNE]. "arXiv:1512.06148 [physics.ins-det]" ().

[23] DUNE Collaboration. "Deep Underground Neutrino Experiment (DUNE), Far Detector Technical Design Report". arXiv:2002.03005 (2020).

[24] "NuMI technical design handbook" (2002).

[25] E. L. Hubbard. "Booster Synchrotron". Technical report Fermilab (1973).

[26] A. Bashyal. "DUNE and MINERvA Flux Studies and a Measurement of the Charged Current Quasielastic Antineutrino Scattering Cross Section with Eν 6GeV on CH Target". Thesis (2021).

[27] S. Agostinelli et al. "GEANT4: A Simulation toolkit. Nucl. Instrum. Meth". A506:250303 (2003).

[28] C. Alt et al. "(NA49)". Eur. Phys. J. C49, 897 (2007).

[29] J. M. Paley et al. "MIPP". Phys. Rev. D 90, 032001 (2014).

[30] Systematic uncertainties from Feynman scaling applied to NNA49. `http://minerva-docdb.fnal.gov:8080/cgi-bin/ShowDocument?docid=8628`. 2013.

[31] R. Aliaga et al. "Neutrino Flux Predictions for the NuMI Beamline". Phys. Rev. D 94, no.9,092005 (2016).

[32] Richart Hill Tomalak Oleksandr. "Theory of elastic neutrino-electron scattering". Phys. Rev. D 101 (3) (2020).

[33] J. Park et al. (MINERvA). "Measurement of neutrino flux from neutrino-electron elastic scattering". Phys. Rev. D 93 (11) (2016).

[34] E. Valencia et al. (MINERvA). "Constraint of the MINERvA medium energy neutrino flux using neutrino-electron elastic scattering". Phys. Rev. D 100, (9), 092001 (2019).

[35] L. Lyons. "Statistics for nuclear annd particle physics" (1986).

[36] D. Ruterbories et al. "Constraining the NuMI neutrino flux using inverse muon decay reactions in MINERvA". arXiv: 2107.01059 [hep-ex] (2021).

[37] J. Devan et al. "Measurements of the inclusive neutrino and anti-neutrino charged current cross sections in MINERvA using the low-ν method". Phys. Rev. D. 94 (11) (2016).

[38] R. Fine et al. "Measurements of the Medium Energy NuMI Flux Using the Low-ν and High-ν Methods at MINERvA". Thesis oai:inspirehep.net:1823712 (2020).

[39] A. Bashyal et al. "Use of Neutrino Scattering Events with Low Hadronic Recoil to Inform Neutrino Flux and Detector Energy Scale". arXiv: 2104.05769 (2021).

[40] L. Aliaga *et al*. "Design, Calibration, and Performance of the MINERvA Detector". Nucl. Instrum. Meth. A.743, 130 (2014).

[41] P. Adamson et al. "Neutrino and antineutrino inclusive charged-current cross section measurements with the MINOS near detector". Phys. Rev. D. 81, 072002 (2010).

[42] Hamamatsu. "Photomultiplier Tubes". Basics and Applications Edition 3a (2007).

[43] C. Grab R. Luchsinger. Comp. Phys. Comm. 76, 262 (1992).

[44] R. Fine. "MINERvA DocDB Entry" (15047).

[45] C. Patrick. "Measurement of the Antineutrino Double-Differential Charged-Current Quasi-Elastic Scattering Cross Section at MINERvA". Ph.D. Dissertation, Northwestern University (2016).

[46] C.Andreopoulos et al. "The GENIE Neutrino Monte Carlo Generator". Nucl. Instr. Meth. A.614 (2010). DOI: `https://hep.ph.liv.ac.uk/~costasa/genie/index.html`.

[47] J. A. Formaggio and G. P. Zeller. "From ev to eev: Neutrino cross sections across energy scales." Rev. Mod. Phys. 84:1307–1341 (2012).

[48] R. Smith and E. Moniz. "Neutrino Reactions on Nuclear Targets". Nucl. Phys. B.43, 605 (1972).

[49] A. Bodek and J.L. Ritchie. "Fermi-motion effects in deep-inelastic lepton scattering from nuclear targets". Phys. Rev. D.23, 1070 (1981).

[50] J. E. Amaro et al. "Neutrino Interactions Importance to Nuclear Physics". AIP Conference Proceedings 1189, 24 (2009).

[51] A. M. Ankowski and J. T. Sobczyk. Phys. Rev C.77, 044311 (2008).

[52] Nadia et al. Fomin. "New Results on Short-Range Correlations in Nuclei". Annual Review of Nuclear a 67.1, pp. 129-159 (2017).

[53] C. Llewellyn Smith. "Neutrino reactions at accelerator energies". Physics Reports 3(5):261 – 379 (1972).

[54] Ruiz S. Nieves J. and Vacas V. "Inclusive Charged–Current Neutrino–Nucleus Reactions". Phys. Rev. C 83 (2011).

[55] P. A. Rodrigues et al [MINENRvA]. "Identification of Nuclear Effects in Neutrino-Carbon interactions at Low Three Momentum Transfer". Phys. Rev. Lett. 116, 071802 (2016).

[56] Gran et al. "Neutrino-nucleus quasi-elastic and 2p2h interactions up to 10 gev". Phys. Rev. D. 88:113007 (Dec. 2013).

[57] M. Valverde J. Nieves J. E. Amaro. "Inclusive quasielastic charged- current neutrino-nucleus reactions". Phys. Rev. C. 72: 019902 (July 2005).

[58] Bodek, Park, and Yang. "Improved low Q2 model for neutrino and electron nucleon cross sections in the few GeV region". Nuclear Physics B - Proceedings and Supplements 139:113 – 118 (2005).

[59] T. Yang. "AGKY transitions between KNO-based and JETSET". In AIP Conf. Proc. Volume 967, pages 269–275 (2007).

[60] R. Gran. "Model Uncertainties for Valencia RPA Effect for MINERvA" (2017).

[61] J. Schwehr, D. Cherdack, and R. Gran. "GENIE implementation of IFIC Valencia model for QE-like 2p2h neutrino-nucleus cross section". arXiv: 1601.02038 (2016).

[62] D. Rein. Nucl. Phys. B 278, 61 (1986).

[63] P. Rodrigues, C. Wilkinson, and K. McFarland. "Constraining the GENIE model of neutrino-induced single pion production using reanalyzed bubble chamber data". Eur. Phys. J. C 76(8):474 (2016).

[64] O. *et al.* Altinok. "Measurement of ν_μ charged-current single π^0 production on hydrocarbon in the few-GeV region using MINERvA." Phys. Rev. D 96, 072003. (2017).

[65] C. L. et al. McGivern. "Cross sections for ν_μ and $\bar{\nu}_\mu$ induced pion production on hydrocarbon in the few-GeV region using MINERvA". Phys. Rev. D 94, 052005. (2016).

[66] Eberly *et al.* "Charged Pion Production in ν_μ Interactions on Hydrocarbon at E_ν = 4.0 GeV". Phys. Rev. D 92, 092008. arXiv: 1406.6415 (2015).

[67] T. *et al.* Le. "Single Neutral Pion Production by Charged-Current $\bar{\nu}_\mu$ Interactions on Hydrocarbon at E_ν = 3.6 GeV". Phys. Lett. B. 749, 130–136. (2015).

[68] P. Stowell et al. "NUISANCE: a neutrino cross-section generator tuning and comparison framework". Journal of Instrumentation 12(01):P01016–P01016 (2017).

[69] P. Stowell et al MINERvA. "Tuning the GENIE pion production model with MINERvA data". Phys. Rev. D 100, no.7 072005 (2019).

[70] MINERvA Collaboration. MINERvA Publication List. 2021. URL: `https://minerva.fnal.gov/recent-minerva-results/` (visited on 01/02/2022).

[71] Wilkinson C Rodrigues P and McFarland K. "Constraining the GENIE model of neutrino-induced single pion production using reanalyzed bubble chamber data". arXiv 1601.01888 [hep-ex] (2016).

[72] Amaro J. Nieves J. and Valverde M. "Inclusive quasi-elastic neutrino reactions". Phys. Rev. C 70 (2005).

[73] K.A. Olive *el al.* "Particle Data Group". Chin. Phys. C 38, 090001 (2014).

[74] A. Lechner. "Particle Interactions in Matter". CERN Yellow Reports: School Proceedings Vol. 5/2018, CERN-2018-008-SP (2018).

[75] D. H. Perkins. "Introduction to High Energy Physics". Third Edition, page 59 (1987).

[76] T. Ferbel. "Techniques and Concepts of High Energy Physics X". NATO Science Series Series C: Mathematical and Physics Sciences Vol 534 (1998).

[77] Atomic Nuclear Properties. April 2018. URL: `http://pdg.ge.infn.it/2018/AtomicNuclearProperties/`.

[78] V.S. Kaushik. "Electromagnetic Showers and Shower Detectors". University of Texas at Arlington (2002).

[79] J. Wolcott. Measurement of Charged Current quasi-elastic cross-section for Electron Neutrinos on a h Ph.D Thesis University of Rochester (2016).

[80] R. Fruhwirth. Nucl.Instrum.Meth. A262, 444 (1987).

[81] J. Wolcott. "ν_e charged-current quasi-elastic (CCQE) scatting in the MINERvA experiment". August 4 (2015).

[82] Friedman Hastie Tibshirani. "The Elements of Statistical Learning 2nd Ed". Springer, NY (2009).

[83] D. Rein and L. M. Sehgal. Annals of Physics 133, 79 (1981).

[84] K.S. Lackner. Nucl. Phys. B153, 526 (1979).

[85] Observation of Neutrino and Antineutrino induced coherent neutral pion production off Al^{27}. Physics Letters B Volume 125 (1983).

[86] D. Rein and L. M. Sehgal. Nucl. Phys. B233, 29 (1983).

[87] A. A. Aguilar-Arevalo et al. Physics Letters B 664, 41 (2008).

[88] C. Baltay et al. Physics Letters 57, 2629 (1986).

[89] C. T. Kullenberg et al. Physics Letters B 686, 177 (2009).

[90] H. J. Grabosch et al. Physics Letters B 686, 177 (2009).

[91] F. Bergsma et al. Physics Letters B 157, 469 (1985).

[92] H. Faissner et al. Physics Letters B 125, 230 (1983).

[93] A. Higuera et al. MINERvA Collaboration. Physics Review Letters 113, 261802 (2014).

[94] G. D Agostini. "A Multidimensional unfolding method based on Bayes' theorem". Nucl. Instrum. Meth. A362, 487 (1995).

[95] G D'Agostini. "Improved iterative Bayesian unfolding". arXiv: 1010.0632 (2010).

[96] Webb G.I. "Bayes' Rule. In: Sammut C., Webb G.I." Encyclopedia of Machine Learning and Data Min (2017).

[97] J.E. Amaro et al. "Neutrino-Nucleus scattering the the SuSA model". arXiv:2106.02857 [hep-ex] (2021).

[98] R. Gran. "SuSA 2p2h and QE in more detail." MINERvA DocDB 25468 (2020).

[99] P.A Rodrigues et al. "Empirical constraint on the multi-nucleon component in neutrino-carbon reactions". arXiv:1906.10576v1 [hep-ex] (2019).

[100] M. Kabirnezhad. "Single pion production in neutrino-nucleon interactions." Phys. Rev. D 97:013002 (2018).

[101] L.A. Harewood and R. Gran. "Elastic hadron-nucleus scattering in neutrino-nucleus reactions and transverse kinematics measurements". arxiv 1906.10576v1 (2019).

[102] X. Lu. "MINERvA Low-Recoil Tune and its Implications". NuInt 18 (2018).

[103] L. Aliaga *et al*. "MINERvA neutrino detector response measured with test beam data". Nucl. Instrum. Meth. A.789, 28 (2015).

[104] S. Dytman. "INTRANUKE-hA". AIP Conference Proceedings 896 (2007).

[105] Dieter Rein and Lalit Sehgal. "Neutrino Excitation of Baryon Resonances and Single Pion Production". Annals Physics 133:79-153 (1981).

[106] A. Bodek and U. Yang. "Axial and Vector Structure Functions for Electron and Neutrino Nucleon Scattering Cross Sections at all Q^2 using Effective Leadking order Parton Distribution Functions". arXiv 1011.6592 (2010).

[107] Gran et al MINERvA. "Antineutrino Charged-Current Reactions on Hydrocarbon with Low Momentum Transfer". arXiv:1803.09377v2 [hep-ex] (2019).

A. The MINERvA Collaboration

Active members of the MINERvA collaboration as of January 2022

Faiza Akbar, Mohammad Sajjad Athar, Huma Haider, Vaniya Ansari, Prameet Gaur, Sayeed Akhter
Aligarh Muslim University, Aligarh, India

Gian Caceres, Helio da Motta, Adrian Lozano, Gilson Silva
Centro Brasileiro de Pesquisas Físicas, Rio de Janeiro, Brazil

Omar Miranda, Luis Montano, Marco Torres
Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional, Mexico City, Mexico

Jennica LeClerc, Christian Nguyen, Heather Ray
University of Florida, Gainsville, Florida

Leo Bellantoni, Minerba Betancourt, Robert Hatcher, Deepika Jena, Jorge Morfin, Gabe Perdue
Fermilab, Batavia, Illinois

Eric Christy
Hampton University, Hampton, Virginia

Alessandro Bravar
University of Geneva, Geneva, Switzerland

Diego Andrade, Jose Luis Bonilla, Julian Felix
Universidad de Guanajuato, Guanajuato, Mexico

Anežka Klustová, Abbey Waldron
Imperial College London, London, UK

Nishat Fiza, Satyajit Jena
Indian Institute of Science Education and Research, Mohali, India

Emily Maher
Massachusetts College of Liberal Arts, North Adams, Massachusetts

Laura Fields
University of Notre Dame, Notre Dame, IN

Amit Bashyal, Mateus Carneiro, Sean Gilligan, Maggie Greenwood, Heidi Schellman, Noah Vaughan
Oregon State University, Corvallis, Oregon

Nathaniel Tagg
Otterbein College, Westerville, Ohio

Komninos-John Plows, David Wark
University of Oxford, Oxford, United Kingdom

Steve Dytman, Ben Messerly, Donna Naples, Vittorio Paolone, Hang Su

University of Pittsburgh, Pittsburgh, Pennsylvania

Marvin Ascencio Sosa, Alberto Gago
Pontificia Universidad Catolica del Peru, Lima, Peru

Arie Bodek, Howard Budd, Tejin Cai, Gonzalo Diaz, Rob Fine, Sarah Henry, Jeffrey Kleykamp, Steve Manly, Kevin McFarland, Aaron McGowan, Miriam Moore, Andrew Olivier, Daniel Ruterbories, Mehreen Sultana, Clarence Wret
University of Rochester, Rochester, New York

Ronald Ransome
Rutgers University, New Brunswick, New Jersey

Hugh Gallagher, Trung Le, Tony Mann, Vladyslav Syrotenko
Tufts University, Medford, Massachusetts

Rik Gran, Emily Haase, Kyle Jacobsen, Ishmam Mahbub, Brandon Reed, Asit Srivastava
University of Minnesota at Duluth, Duluth, Minnesota

Ivancovish Julian Huancco Mamani, Carlos Solano Salinas
Universidad Nacional de Ingeniería, Lima, Peru

Jorge Chaves, Keith Kroma-Wiley, David Last, Christopher Mauger, Alejandro Ramirez Delgado
University of Pennsylvania, Philadelphia, Pennsylvania

William Brooks, Anushree Ghosh, Jonathan Miller, Barbara Yaeggy
Universidad Técnica Federico Santa María, Valparaiso, Chile

Steve Boyd, Xianguo Lu
University of Warwick

Jan Sobczyk
University of Wroclaw

Zubair Ahmad Dar, Amy Filkins, Michael Kordosky, Oscar E Moreno, Jeffrey Nelson, Edgar Valencia, Luis Zazueta
College of William and Mary, Williamsburg, Virginia

Deborah Harris, Maria Mehmood
University of York, Toronto, Ontario, Canada