

ELECTRON NEUTRINO APPEARANCE IN THE MINOS  
EXPERIMENT

Anna Maria Holin  
University College London

Submitted to University College London  
in fulfilment of the requirements for the award  
of the degree of Doctor of Philosophy

February 2010

## **Declaration**

I, Anna Maria Holin, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

## Abstract

The MINOS experiment is a long-baseline neutrino oscillation experiment which sends a high intensity muon neutrino beam through two functionally identical detectors, a Near detector at the Fermi National Accelerator Laboratory in Illinois, 1km from the beam source, and a Far detector, 734km away, in the Soudan Mine in Minnesota.

MINOS may be able to measure the neutrino mixing angle parameter  $\sin^2 2\theta_{13}$  for the first time. Detector granularity, however, makes it very hard to distinguish any  $\nu_e$  appearance signal events characteristic of a non-zero value of  $\theta_{13}$  from background neutral current (NC) and short-track  $\nu_\mu$  charged current (CC) events. Also, uncertainties in the hadronic shower modeling in the kinematic region characteristic of this analysis are relatively large. A new data-driven background decomposition method designed to address those issues is developed and its results presented. By removing the long muon tracks from  $\nu_\mu$ -CC events, the Muon Removed Charge Current (MRCC) method creates independent pseudo-NC samples that can be used to correct the MINOS Monte Carlo to agree with the high-statistics Near detector data and to decompose the latter into components so as to predict the expected Far detector background. The MRCC method also provides an important cross-check in the Far detector to test the background in the signal selected region.

MINOS finds a 1.0-1.5  $\sigma$   $\nu_e$ -CC excess above background in the Far detector data, depending on method used, for a total exposure of  $3.14 \times 10^{20}$  protons-on-target. Interpreting this excess as signal, MINOS can set limits on  $\sin^2 2\theta_{13}$ . Using the MRCC method, MINOS sets a limit of  $\sin^2 2\theta_{13} < 0.265$  at the 90% confidence limit for a CP-violating phase  $\delta = 0$ .

## Acknowledgements

This thesis would not have been possible if not for the help of many people. In particular, I would like to thank my supervisor Prof. Jenny Thomas for her advice, guidance, and encouragement throughout my Ph.D. She has provided me with constant support and her wide-ranging practical and theoretical knowledge have been crucial to the progression of my research.

I would also like to thank Trish Vahle, who has given me invaluable help throughout my Ph.D. and has taught me a great deal. I am also grateful to Chris Smith, who has helped me during the early stages of my analysis. This thesis would also not have been possible were it not for my collaborators from the MINOS  $\nu_e$  group, Josh Boehm, Tingjun Yang, Pedro Ochoa, Greg Pawloski, Mayly Sanchez and others, who I have enjoyed working with over the years. Thank you also to all my other collaborators from MINOS and to my colleagues from the High Energy Physics group at UCL.

Finally, I would like to thank my family for their faith in me. My sister Olga has given me much support and always brought a smile back to my face. Last but not least, I would like to thank my husband Jack, who is my rock and never complained when I was working long days or was away, but would always inspire me to do the best I can. His faith in me never faltered and for this I cannot express my gratitude enough.

# Contents

|          |  |           |
|----------|--|-----------|
| <b>1</b> | <b>Introduction</b>  | <b>15</b> |
| <b>2</b> | <b>Neutrino Theory</b>   | <b>17</b> |
| 2.1      | A Short History of the Neutrino . . . . .                                  | 17        |
| 2.2      | The Standard Model . . . . .   | 22        |
| 2.3      | Neutrino Oscillations . . . . .  | 23        |
| 2.3.1    | Two-Flavour Oscillations . . . . .   | 24        |
| 2.3.2    | Three-Flavour Oscillations . . . . .                                       | 27        |
| 2.3.3    | Neutrino Mass Hierarchy . . . . .  | 28        |
| 2.4      | Neutrino Matter Effects . . . . .  | 29        |
| 2.5      | The Chooz experiment . . . . .   | 31        |
| 2.6      | Neutrino Interactions in MINOS . . . . .                                   | 32        |
| <b>3</b> | <b>The MINOS Experiment</b>  | <b>36</b> |
| 3.1      | The NuMI Beam . . . . .  | 36        |
| 3.2      | The MINOS Detectors . . . . .  | 40        |
| 3.2.1    | The MINOS Near Detector . . . . .  | 41        |
| 3.2.2    | The MINOS Far Detector . . . . .   | 42        |
| 3.3      | Calibration of the MINOS Detectors . . . . .                               | 43        |
| 3.4      | MINOS Results to Date . . . . .  | 46        |
| 3.4.1    | Muon Neutrino Disappearance . . . . .                                      | 46        |
| 3.4.2    | Active Neutrino Disappearance Search . . . . .                             | 48        |
| <b>4</b> | <b>The Electron Neutrino Appearance Analysis</b>                           | <b>51</b> |
| 4.1      | General Method . . . . .   | 52        |
| 4.2      | $\nu_e$ -Appearance Analysis Backgrounds . . . . .                         | 52        |
| 4.3      | Monte Carlo and Reconstruction . . . . .                                   | 55        |
| 4.4      | Monte Carlo Flux Weights . . . . .   | 56        |
| 4.5      | Analysis Selection Cuts . . . . .  | 57        |
| 4.5.1    | Data Quality Cuts . . . . .  | 57        |
| 4.5.2    | Analysis Preselection Cuts . . . . .                                       | 57        |
| 4.6      | Cross-Talk . . . . .   | 59        |
| 4.7      | $\nu_e$ Appearance Analysis Selection Algorithms . . . . .                 | 60        |
| 4.7.1    | A Neural Net Particle Identification Algorithm . . . . .                   | 61        |
| 4.7.2    | A Library Event Matching (LEM) Particle Identification Algorithm . . . . . | 69        |
| 4.8      | ND Data/Monte Carlo Spectra and Differences . . . . .                      | 70        |

|          |   |            |
|----------|---|------------|
| 4.9      | Near Detector Spectrum Decomposition Methods . . . . .                                  | 75         |
| 4.9.1    | The Horn On-Off Method . . . . .  | 75         |
| 4.9.2    | The Muon Removed Charged Current Method . . . . .                                       | 81         |
| <b>5</b> | <b>The Muon Removed Charged Current Method - MRCC</b>                                   | <b>83</b>  |
| 5.1      | Description of Removal Process . . . . .  | 84         |
| 5.1.1    | Generating the MRCC Sample . . . . .  | 86         |
| 5.1.2    | MRCC Analysis Selection . . . . .   | 87         |
| 5.2      | The MRCC Analysis Method . . . . .  | 93         |
| 5.2.1    | Making the Correction . . . . .   | 96         |
| 5.2.2    | Using Different Quantities to carry out the MRCC Correction . . .                       | 98         |
| <b>6</b> | <b>MRCC Discussion and Results</b>  | <b>99</b>  |
| 6.1      | Considerations of the MRCC Method . . . . .   | 99         |
| 6.1.1    | Differences in Kinematics between NC and CC Events . . . . .                            | 100        |
| 6.1.2    | Net Charge of the MRCC Showers . . . . .  | 102        |
| 6.1.3    | $\pi^0$ Content of MRCC Showers . . . . .   | 104        |
| 6.2      | Unknown Source of Data/MC Differences . . . . .   | 104        |
| 6.3      | MRCC Systematic Errors . . . . .  | 106        |
| 6.4      | Final Near Detector MRCC Result . . . . .   | 114        |
| 6.5      | MRCC Mock Data Challenge . . . . .  | 117        |
| 6.6      | Comparison of MRCC and HOO Results . . . . .  | 117        |
| <b>7</b> | <b>Pathway to a Final <math>\nu_e</math>-Appearance Result</b>                          | <b>120</b> |
| 7.1      | Far Extrapolation of ND Decompositions . . . . .  | 120        |
| 7.1.1    | Oscillating the FD MC . . . . .   | 120        |
| 7.1.2    | Extrapolation Method . . . . .  | 122        |
| 7.1.3    | Systematic Errors in the FD Extrapolation . . . . .                                     | 124        |
| 7.1.4    | Final Predictions . . . . .   | 125        |
| 7.2      | Sidebands . . . . .   | 126        |
| 7.2.1    | Anti-PID Sideband . . . . .   | 126        |
| 7.2.2    | MRCC Sideband . . . . .   | 128        |
| 7.3      | Discussion of the MRCC Sideband Results . . . . .                                       | 130        |
| 7.3.1    | Applying a Charge Sign Cut . . . . .  | 130        |
| 7.3.2    | MRCC Far Detector Distributions . . . . .   | 131        |
| 7.3.3    | MRCC Sideband Prediction in One Detector . . . . .                                      | 131        |
| 7.3.4    | MRCC Sideband by MINOS Data Run Periods . . . . .                                       | 132        |
| 7.4      | MRE Sideband . . . . .  | 133        |
| <b>8</b> | <b>Final <math>\nu_e</math>-Appearance Analysis Result</b>                              | <b>144</b> |
| 8.1      | Far Detector Final Box Opening . . . . .  | 145        |
| 8.2      | Cross-Checks carried out on the Final Data Set . . . . .                                | 146        |
| 8.2.1    | Far Detector Selection Numbers . . . . .  | 146        |
| 8.2.2    | Far Detector Distributions . . . . .  | 147        |
| 8.2.3    | Event Scanning . . . . .  | 150        |
| 8.2.4    | Breakdown of Events by MINOS Data Run . . . . .   | 150        |
| 8.3      | Final Contours for the $3.14 \times 10^{20}$ POT $\nu_e$ -Appearance Analysis . . . . . | 151        |

|          |   |            |
|----------|---|------------|
| 8.4      | Comments on Far Detector Data . . . . .                           | 153        |
| 8.5      | Future MINOS $\nu_e$ -Appearance Analysis . . . . .               | 154        |
| <b>9</b> | <b>Final Summary and Conclusions</b>                              | <b>161</b> |
|          | <b>Bibliography</b>   | <b>163</b> |
| <b>A</b> | <b>Appendix - Additional ND Distributions</b>                     | <b>168</b> |
| <b>B</b> | <b>Appendix - Oscillations in the <math>\nu_e</math> Analysis</b> | <b>198</b> |
| <b>C</b> | <b>Appendix - Additional MRCC FD Distributions</b>                | <b>200</b> |
| <b>D</b> | <b>Appendix - FD Distributions and Golden Event Displays</b>      | <b>207</b> |
| <b>E</b> | <b>Appendix - FD Data Event Lists</b>                             | <b>221</b> |

# List of Figures

|      |  |    |
|------|--|----|
| 2.1  | Oscillations Result of the Super-Kamiokande Experiment . . . . .                                 | 20 |
| 2.2  | Supernova 1987a Hubble Space Telescope Photograph . . . . .                                      | 22 |
| 2.3  | Number of Neutrino Species from LEP . . . . .  | 24 |
| 2.4  | $\nu_\mu \rightarrow \nu_x$ Oscillation Probability in MINOS . . . . .                           | 27 |
| 2.5  | Electron Neutrino Charged Current Scattering Interaction . . . . .                               | 29 |
| 2.6  | 90% Exclusion Limits generated by the Chooz Experiment . . . . .                                 | 34 |
| 2.7  | Neutrino Charged Current Event . . . . .   | 35 |
| 2.8  | Neutrino Neutral Current Event . . . . .   | 35 |
| 2.9  | Deep Inelastic Scattering Schematic . . . . .  | 35 |
| 3.1  | MINOS Experiment Design . . . . .  | 37 |
| 3.2  | Different Neutrino Beam Configurations . . . . .   | 38 |
| 3.3  | Plan and Elevation Views of the NuMI Beam Facility . . . . .                                     | 39 |
| 3.4  | Schematic of the MINOS Focusing Horns . . . . .  | 41 |
| 3.5  | Schematic of MINOS Ionization Chambers . . . . .   | 42 |
| 3.6  | Far Detector Alternative Planes . . . . .  | 43 |
| 3.7  | The MINOS Near Detector . . . . .  | 44 |
| 3.8  | The MINOS Far Detector . . . . .   | 45 |
| 3.9  | The MINOS Calibration Detector . . . . .   | 46 |
| 3.10 | MINOS Exposure in Protons-on-Target . . . . .  | 47 |
| 3.11 | MINOS Far Detector $\nu_\mu$ -CC Energy Spectrum for $3.36 \times 10^{20}$ POT . . . . .         | 48 |
| 3.12 | MINOS Best-Fit Contours in $\Delta m_{32}^2$ and $\sin^2 2\theta_{23}$ Parameter Space . . . . . | 49 |
| 3.13 | MINOS Far Detector Neutral Current Spectrum for $3.18 \times 10^{20}$ POT . . . . .              | 50 |
| 4.1  | Monte Carlo Simulated Far Detector Events . . . . .  | 53 |
| 4.2  | Monte Carlo Example of $\nu_e$ -Appearance Signal Spectrum . . . . .                             | 54 |
| 4.3  | Beam $\nu_e$ -CC Parent Hadrons at NuMI Target . . . . .   | 55 |
| 4.4  | $\nu_\mu$ -CC Spectra in the MINOS Near Detector . . . . .                                       | 56 |
| 4.5  | Effects of Long Track Rejection Cuts on FD MC Events . . . . .                                   | 59 |
| 4.6  | Effects of Shower and Energy Selection Cuts on FD MC Events . . . . .                            | 64 |
| 4.7  | Data/MC Comparison of Low PE Hits in the Near Detector . . . . .                                 | 65 |
| 4.8  | Data/MC Comparison of Low PE Hits in the Far Detector . . . . .                                  | 65 |
| 4.9  | Longitudinal and Transverse Energy Profiles of a FD $\nu_e$ -CC Event . . . . .                  | 66 |
| 4.10 | Signal/Background Separation of ANN PID Variables Part 1 . . . . .                               | 67 |
| 4.11 | Signal/Background Separation of ANN PID Variables Part 2 . . . . .                               | 68 |
| 4.12 | Signal/Background Separation of ANN PID . . . . .  | 69 |
| 4.13 | LEM PID - Examples of Library Matches . . . . .  | 70 |



|      |   |     |
|------|---|-----|
| 4.14 | Signal/Background Separation of LEM PID Variables . . . . .                                       | 71  |
| 4.15 | Signal/Background Separation of LEM PID . . . . .   | 72  |
| 4.16 | ND Data/MC Comparison for the ANN PID . . . . .   | 73  |
| 4.17 | ND Data/MC Comparison for the LEM PID . . . . .   | 74  |
| 4.18 | ND Hadronic Model Uncertainty for the PID Selections . . . . .                                    | 75  |
| 4.19 | ND Data and MC in Horn On-Off Beam Configurations . . . . .                                       | 76  |
| 4.20 | ND MC Backgrounds in Horn On-Off Beam Configurations . . . . .                                    | 77  |
| 4.21 | ND MC Horn Off/On Ratios for the NC Background . . . . .  | 78  |
| 4.22 | ND MC Horn Off/On Ratios for the $\nu_\mu$ -CC Background . . . . .                               | 79  |
| 4.23 | NC Horn Off/On Ratios for Data and MC . . . . .   | 80  |
| 4.24 | $\nu_\mu$ -CC Horn Off/On Ratios for Data and MC . . . . .  | 81  |
| 4.25 | Final ND Data Decomposition for the HOO Method - ANN PID . . . . .                                | 82  |
| 4.26 | Final ND Data Decomposition for the HOO Method - LEM PID . . . . .                                | 82  |
|      |   |     |
| 5.1  | Principle behind Muon Removed Charged Current Method . . . . .                                    | 84  |
| 5.2  | Muon Removed Charged Current Method Procedure . . . . .   | 85  |
| 5.3  | ND MRCC Data Reconstructed Energy Distribution . . . . .  | 89  |
| 5.4  | ND MRCC Data Pulse Height Weighted Purity Distribution . . . . .                                  | 90  |
| 5.5  | ND MRCC Data Pulse Height Weighted Completeness Distribution . . . . .                            | 90  |
| 5.6  | ND MRCC Data Original $\nu_\mu$ -CC PID Distribution . . . . .                                    | 91  |
| 5.7  | ND MRCC MC Original $\nu_\mu$ -CC PID Distribution . . . . .                                      | 92  |
| 5.8  | ND MRCC MC Reconstructed Energy Distributions . . . . .   | 93  |
| 5.9  | ND MRCC and Standard Data/MC Ratios - ANN PID Selection . . . . .                                 | 94  |
| 5.10 | ND MRCC and Standard Data/MC Ratios - LEM PID Selection . . . . .                                 | 95  |
| 5.11 | Near Detector NC MC before and after MRCC Correction . . . . .                                    | 97  |
|      |   |     |
| 6.1  | ND MRCC and Standard MC Kinematic Truth Distributions . . . . .                                   | 101 |
| 6.2  | MRCC Distributions for Different Original Track QP . . . . .                                      | 103 |
| 6.3  | Different MRCC Ratios depending on Charge Sign Cut . . . . .                                      | 104 |
| 6.4  | Pion Final State Distributions for MRCC and Standard Events . . . . .                             | 105 |
| 6.5  | Data/MC MRCC/Standard Comparisons for Shower Rise Parameter . . . . .                             | 107 |
| 6.6  | Data/MC MRCC/Standard Comparisons for Fraction of Pulse Height in<br>a Two-Plane Window . . . . . | 108 |
| 6.7  | Data/MC MRCC/Standard Comparisons for Shower Moliere Radius . . . . .                             | 109 |
| 6.8  | MRCC and Standard Data/MC Energy Distributions - ANN PID . . . . .                                | 112 |
| 6.9  | MRCC and Standard Data/MC Energy Distributions - LEM PID . . . . .                                | 113 |
| 6.10 | Results for the MRCC ND Background Decomposition . . . . .  | 115 |
| 6.11 | Final Result for MRCC Method - ANN PID Selection . . . . .  | 116 |
| 6.12 | Final Result for MRCC Method - LEM PID Selection . . . . .  | 116 |
| 6.13 | MRCC Method Mock Data Challenge Results . . . . .   | 118 |
| 6.14 | Final Comparison of ND Data Decomposition Methods - ANN PID . . . . .                             | 118 |
| 6.15 | Final Comparison of ND Data Decomposition Methods - LEM PID . . . . .                             | 119 |
|      |   |     |
| 7.1  | Far/Near Ratios for the ANN PID Selection . . . . .   | 134 |
| 7.2  | Far/Near Ratios for the LEM PID Selection . . . . .   | 135 |
| 7.3  | ND Muon Removed Electron Added PID Distributions . . . . .  | 136 |
| 7.4  | ND MRE and Standard FD Efficiencies - ANN PID selection . . . . .                                 | 136 |

|      |   |     |
|------|---|-----|
| 7.5  | Systematic Error Bands - ANN PID Selection . . . . .  | 137 |
| 7.6  | ND Background Decomposition for Anti-PID Sidebands . . . . .  | 137 |
| 7.7  | Comparison of ND Background Decomposition Methods for Anti-PID<br>Sidebands . . . . .   | 138 |
| 7.8  | FD Data and Prediction for ANN Anti-PID Sideband . . . . .  | 138 |
| 7.9  | FD Data and Prediction for LEM Anti-PID Sideband . . . . .  | 139 |
| 7.10 | ANN PID Distributions for the FD MRCC Preselected Events . . . . .  | 139 |
| 7.11 | LEM PID Distributions for the FD MRCC Preselected Events . . . . .  | 140 |
| 7.12 | FD MRCC Spatial Distributions after $\nu_e$ Preselection . . . . .  | 141 |
| 7.13 | MRCC Sideband Extrapolation within One Detector . . . . .   | 142 |
| 7.14 | FD MRE Preselected Event Sample PID Distributions . . . . .   | 143 |
|      |   |     |
| 8.1  | Shower Distributions for Preselected FD Data . . . . .  | 145 |
| 8.2  | ANN PID Distributions for Exposure of $3.14 \times 10^{20}$ POT . . . . .   | 146 |
| 8.3  | LEM PID Distributions for Exposure of $3.14 \times 10^{20}$ POT . . . . .   | 147 |
| 8.4  | FD Data and HOO Prediction Reconstructed Energy Spectra . . . . .   | 148 |
| 8.5  | FD Data and MRCC Prediction Reconstructed Energy Spectra . . . . .  | 149 |
| 8.6  | FD Data Events as a Function of Time . . . . .  | 155 |
| 8.7  | FD Data Event Vertices . . . . .  | 156 |
| 8.8  | FD Data Distributions after ANN PID Selection . . . . .   | 157 |
| 8.9  | Event Display for Golden FD Event . . . . .   | 158 |
| 8.10 | Final Limit Contours after ANN PID Selection . . . . .  | 159 |
| 8.11 | Final Limit Contours after LEM PID Selection . . . . .  | 159 |
| 8.12 | Projected 90% Confidence Limits in the Absence of $\nu_e$ -Appearance and<br>an Exposure of $7.0 \times 10^{20}$ POT . . . . .  | 160 |
| 8.13 | Projected 90% Confidence Limits in the Absence of $\nu_e$ -Appearance and<br>an Exposure of $10.0 \times 10^{20}$ POT . . . . . | 160 |
|      |   |     |
| A.1  | ND MRCC and Standard MC Kinematic Truth Distributions - ANN PID<br>Selection and Energy between 1 and 2 GeV . . . . .           | 169 |
| A.2  | ND MRCC and Standard MC Kinematic Truth Distributions - ANN PID<br>Selection and Energy between 2 and 3 GeV . . . . .           | 170 |
| A.3  | ND MRCC and Standard MC Kinematic Truth Distributions - ANN PID<br>Selection and Energy between 3 and 4 GeV . . . . .           | 171 |
| A.4  | ND MRCC and Standard MC Kinematic Truth Distributions - ANN PID<br>Selection and Energy between 4 and 5 GeV . . . . .           | 172 |
| A.5  | ND MRCC and Standard MC Kinematic Truth Distributions - ANN PID<br>Selection and Energy between 5 and 8 GeV . . . . .           | 173 |
| A.6  | ND MRCC and Standard MC Kinematic Truth Distributions - LEM PID<br>Selection and Energy between 1 and 2 GeV . . . . .           | 174 |
| A.7  | ND MRCC and Standard MC Kinematic Truth Distributions - LEM PID<br>Selection and Energy between 2 and 3 GeV . . . . .           | 175 |
| A.8  | ND MRCC and Standard MC Kinematic Truth Distributions - LEM PID<br>Selection and Energy between 3 and 4 GeV . . . . .           | 176 |
| A.9  | ND MRCC and Standard MC Kinematic Truth Distributions - LEM PID<br>Selection and Energy between 4 and 5 GeV . . . . .           | 177 |

|  |     |
|--|-----|
| A.10 ND MRCC and Standard MC Kinematic Truth Distributions - LEM PID   |     |
| Selection and Energy between 5 and 8 GeV . . . . .                     | 178 |
| A.11 Data/MC MRCC/Standard Comparisons for the ANN PID . . . . .       | 179 |
| A.12 Data/MC MRCC/Standard Comparisons for the LEM PID . . . . .       | 180 |
| A.13 Data/MC MRCC/Standard Comparisons for Shower Rise Parameter . . . | 181 |
| A.14 Data/MC MRCC/Standard Comparisons for Shower Fall Parameter . . . | 182 |
| A.15 Data/MC MRCC/Standard Comparisons for Shower Moliere Radius . . . | 183 |
| A.16 Data/MC MRCC/Standard Comparisons for Shower Lateral Spread . . . | 184 |
| A.17 Data/MC MRCC/Standard Comparisons for Fraction of Pulse Height in |     |
| a Two-Plane Window . . . . .   | 185 |
| A.18 Data/MC MRCC/Standard Comparisons for Fraction of Pulse Height in |     |
| a Four-Plane Window . . . . .  | 186 |
| A.19 Data/MC MRCC/Standard Comparisons for Fraction of Pulse Height in |     |
| a Six-Plane Window . . . . .   | 187 |
| A.20 Data/MC MRCC/Standard Comparisons for Fraction of Pulse Height in |     |
| Eight Biggest Hits . . . . .   | 188 |
| A.21 Data/MC MRCC/Standard Comparisons for Fraction of Pulse Height in |     |
| a Narrow Road . . . . .  | 189 |
| A.22 Data/MC MRCC/Standard Comparisons for Energy Weighted Sum of      |     |
| Distance of each Hit to Z-Axis . . . . .                               | 190 |
| A.23 Data/MC MRCC/Standard Comparisons for Minimum Distances between   |     |
| Larger than Average Hits . . . . .                                     | 191 |
| A.24 Data/MC MRCC/Standard Comparisons for Fraction of Best 50 LEM     |     |
| CC- $\nu_e$ Matches with $y > 0.9$ . . . . .                           | 192 |
| A.25 Data/MC MRCC/Standard Comparisons for Mean Fractional Charge of   |     |
| Best 50 LEM CC- $\nu_e$ Matches with $y > 0.9$ . . . . .               | 193 |
| A.26 Data/MC MRCC/Standard Comparisons for Average $y$ of Best 50 LEM  |     |
| CC- $\nu_e$ Matches with $y > 0.9$ . . . . .                           | 194 |
| A.27 Data/MC MRCC/Standard Comparisons for Number of Shower Hits . . . | 195 |
| A.28 Data/MC MRCC/Standard Comparisons for Energy of Biggest Hit in    |     |
| Primary Shower . . . . .   | 196 |
| A.29 Data/MC MRCC/Standard Comparisons for Energy of Second Biggest    |     |
| Hit in Primary Shower . . . . .  | 197 |
| C.1 FD MRCC Difference between MRCC and Original Vertex X . . . . .    | 200 |
| C.2 FD MRCC Difference between MRCC and Original Vertex Y . . . . .    | 201 |
| C.3 FD MRCC Difference between MRCC and Original Vertex Z . . . . .    | 201 |
| C.4 FD MRCC Pulse Height Weighted Completeness . . . . .               | 202 |
| C.5 FD MRCC Original Reconstructed Neutrino Energy . . . . .           | 202 |
| C.6 FD MRCC Original Reconstructed Hadronic $Y$ . . . . .              | 203 |
| C.7 FD MRCC Original Reconstructed $W^2$ . . . . .                     | 203 |
| C.8 FD MRCC Original Reconstructed $Q^2$ . . . . .                     | 204 |
| C.9 FD MRCC Number of Showers . . . . .                                | 204 |
| C.10 FD MRCC Number of Shower Hit Strips . . . . .                     | 205 |
| C.11 FD MRCC Shower Rise Parameter . . . . .                           | 205 |
| C.12 FD MRCC Shower Fall Parameter . . . . .                           | 206 |

|      |   |     |
|------|---|-----|
| D.1  | FD Data Number of Event Hit Planes . . . . .  | 207 |
| D.2  | FD Data Number of Event Hit Strips . . . . .  | 208 |
| D.3  | FD Data Number of Event Tracks . . . . .  | 208 |
| D.4  | FD Data Number of Track Hit Planes . . . . .  | 209 |
| D.5  | FD Data Number of Event Showers . . . . .   | 209 |
| D.6  | FD Data Number of Shower Hit Planes . . . . .   | 210 |
| D.7  | FD Data Number of Shower Hit Strips . . . . .   | 210 |
| D.8  | FD Data Shower Lateral Spread . . . . .   | 211 |
| D.9  | FD Data Shower Parameter A . . . . .  | 211 |
| D.10 | FD Data Shower Parameter B . . . . .  | 212 |
| D.11 | FD Data Shower Moliere Radius . . . . .   | 212 |
| D.12 | FD Data Energy Weighted Sum of Distance of each Hit to the Z-Axis . . . . .                       | 213 |
| D.13 | FD Data Minimum Distances between Larger than Average Hits . . . . .                              | 213 |
| D.14 | FD Data Fraction of Pulse Height in a Two-Plane Window . . . . .                                  | 214 |
| D.15 | FD Data Fraction of Pulse Height in a Four-Plane Window . . . . .                                 | 214 |
| D.16 | FD Data Fraction of Pulse Height in a Six-Plane Window . . . . .                                  | 215 |
| D.17 | FD Data Fraction of Pulse Height in Eight Biggest Hits . . . . .                                  | 215 |
| D.18 | FD Data Fraction of Pulse Height in a Narrow Road . . . . .                                       | 216 |
| D.19 | FD Data Fraction of Best 50 LEM CC- $\nu_e$ Matches with $y > 0.9$ . . . . .                      | 216 |
| D.20 | FD Data Mean Fractional Charge of the Best 50 LEM CC- $\nu_e$ Matches<br>with $y > 0.9$ . . . . . | 217 |
| D.21 | FD Data Average $y$ of the Best 50 LEM CC- $\nu_e$ Matches with $y > 0.9$ . . . . .               | 217 |
| D.22 | Event Display for Golden FD Event - Subrun 32687 . . . . .  | 218 |
| D.23 | Event Display for Golden FD Event - Subrun 37242 . . . . .  | 219 |
| D.24 | Event Display for Golden FD Event - Subrun 38197 . . . . .  | 220 |

# List of Tables

|      |   |     |
|------|---|-----|
| 2.1  | Leptons in the Standard Model . . . . .                                 | 23  |
| 4.1  | ND MC Numbers of Events for Different Selections . . . . .              | 72  |
| 4.2  | ND Data and MC Numbers of Events for Different Selections . . . . .     | 73  |
| 4.3  | ND Data Decomposition Numbers for the Horn On-Off Method . . . . .      | 80  |
| 5.1  | MRCC ND Data Event Numbers and Efficiencies . . . . .                   | 88  |
| 5.2  | MRCC ND Data Completeness and Purity . . . . .                          | 89  |
| 5.3  | MRCC ND MC Background Numbers . . . . .                                 | 92  |
| 5.4  | MRCC ND MC Background Percentages . . . . .                             | 93  |
| 5.5  | MRCC ND Data/MC Ratios . . . . .  | 93  |
| 6.1  | Final MRCC Decomposition Numbers for the ANN PID Selection . . . . .    | 111 |
| 6.2  | Final MRCC Decomposition Numbers for the LEM PID Selection . . . . .    | 114 |
| 6.3  | Summary of MRCC ND Background Decomposition Results . . . . .           | 117 |
| 7.1  | Average MRE Selection Efficiencies . . . . .                            | 123 |
| 7.2  | Final Systematic Errors for the HOO and MRCC Methods . . . . .          | 125 |
| 7.3  | Final Far Detector Predictions . . . . .                                | 126 |
| 7.4  | ND Data Decompositions for the Anti-PID Sidebands . . . . .             | 127 |
| 7.5  | FD MC and Prediction Numbers for the Anti-PID Sidebands . . . . .       | 127 |
| 7.6  | FD Prediction Numbers with Errors for the Anti-PID Sidebands . . . . .  | 128 |
| 7.7  | FD Anti-PID Sidebands Data/Prediction Comparisons . . . . .             | 128 |
| 7.8  | FD MRCC Sideband Predictions . . . . .                                  | 129 |
| 7.9  | FD MRCC Sideband Data/Prediction Comparisons . . . . .                  | 130 |
| 7.10 | FD MRCC Sideband Data/Prediction Comparisons - Neutrinos Only . . . . . | 131 |
| 7.11 | MRCC Sideband Extrapolation within One Detector . . . . .               | 132 |
| 7.12 | FD MRCC Sideband Numbers by MINOS Run Period . . . . .                  | 133 |
| 7.13 | Far Detector MRE Sideband Results for the two PIDs . . . . .            | 133 |
| 8.1  | Final FD Results . . . . .  | 145 |
| 8.2  | FD Event Numbers and Selection Efficiencies . . . . .                   | 150 |
| 8.3  | Final FD Results by MINOS Run Period . . . . .                          | 151 |
| 8.4  | Final FD Data Fit Results . . . . .                                     | 153 |
| E.1  | FD Data Event List - Selected by ANN PID . . . . .                      | 222 |
| E.2  | FD Data Event List - Selected by LEM PID . . . . .                      | 223 |
| E.3  | FD Data Event List - Selected by Both PIDs . . . . .                    | 224 |

E.4 FD Data Event List - Preselection . . . . . 229

# Chapter 1

## Introduction

The MINOS experiment is a long baseline neutrino experiment and is currently one of the experiments that are at the forefront of neutrino physics. Neutrinos are very weakly interacting elementary particles whose properties are very hard to measure, and since they were first postulated in the 1930s, neutrino science has come a long way and some surprising discoveries have been made. One of those discoveries is that neutrinos oscillate between flavours as they travel, so a muon neutrino may for example oscillate into an electron or a tau neutrino as it travels through space.

The MINOS experiment was built in order to make precision measurements of the neutrino oscillation parameters, in particular the parameters for the dominant oscillation mode of muon neutrinos into tau neutrinos. For MINOS, which uses a Near and a Far detector, this oscillation mode is evidenced by a smaller than predicted number of muon neutrino charged current interactions in the Far detector. This is because the appearing tau neutrinos do not possess enough energy to create tau neutrinos in charged current interactions. This main MINOS analysis is usually referred to as the muon neutrino disappearance analysis. If muon neutrinos however also oscillate into electron neutrinos, then MINOS can attempt to measure the parameters involved by trying to observe electron neutrino appearance resulting from this sub-dominant oscillation mode. Unfortunately, the MINOS electron neutrino appearance analysis is very difficult because the detectors were optimized to look for muon neutrino disappearance, or in particular for long muon tracks resulting from charged current muon neutrino interactions. Nevertheless, despite the difficulties, MINOS has attempted to look for electron neutrino appearance and this analysis is described in detail in this document.

First, a brief introduction into the theory and history of neutrinos is presented. Furthermore, a brief introduction into neutrino interactions in MINOS is also given. The MINOS experiment is then introduced in detail before focus is turned to the electron neutrino appearance analysis itself. The methods involved in this analysis are discussed extensively with particular attention paid to the muon removed charged current method, a data-driven analysis background estimation method. The results of the exciting and complex electron neutrino appearance analysis are then summarised and future prospects are discussed.



## Chapter 2

# Neutrino Theory

One of the most fascinating and most elusive particles ever observed is the neutrino. Even though the Standard Model of particle physics is highly successful in explaining elementary particle interactions, it no longer explains neutrino behaviour as we observe it. This intriguing particle could be pointing to new and unexplored physics beyond the Standard Model and its properties have far reaching implications for both fundamental physics and cosmology, as will be explained in this chapter.

### 2.1 A Short History of the Neutrino

Neutrinos are elementary particles that were first postulated by W. Pauli in 1930 in a letter<sup>1</sup> starting “Dear Radioactive Ladies and Gentlemen” in order to explain the apparent non-conservation of energy in beta decay. The phenomenon that this unconventional remedy, in the form of a new particle now called a neutrino, was to resolve was the spectrum observed for the outgoing electrons in beta decay. Naively, since the recoil nucleus is very large, one would expect the electron to take away most of the energy from the beta decay in the form of kinetic energy. Thus the spectrum of the beta decay electrons would be centered around a particular maximum energy  $KE_{max}$ . Instead, beta decay experiments were observing a continuous electron energy spectrum that peaked at much lower values and then declined, with very few electrons having an energy of  $KE_{max}$ . Thus, for most of the events, a large part of the kinetic energy was simply vanishing. After many unsuccessful attempts to explain the unexpected behaviour of the beta spectrum, the only reasonable solution remaining which avoided discarding the

---

<sup>1</sup>This letter was not an official publication but it has been reproduced in various books and publications since it was written [1].

law of conservation of energy, was to postulate a new particle which would carry away the excess energy [1]. Since the particle in question was not being observed at the time, it had to be very weakly interacting with other matter. This particle would also have no charge and half integer spin. The reason for the former is simple charge conservation. The reason for the latter can be explained if one considers the following example. If for example the initial nucleus had integer spin, then after the beta decay, it still has integer spin as the number of nucleons has not changed (even though one neutron is now a proton). So therefore, if only an electron (which has half integer spin) were emitted, then angular momentum would not be conserved. But if a neutrino has half integer spin, then the non-conservation of angular momentum problem is solved. Shortly after being proposed, the new particle became known as the neutrino. It may be interesting to note that when Pauli first proposed the neutrino as a new particle, he proposed that the particle have a small mass not more than 1% of the proton mass. In 1933, E. Fermi first wrote down the correct theory for beta decay [2]. In this, a neutron decays into a proton, an electron, and an anti-neutrino.

From the time neutrinos were first proposed by Pauli it took more than 20 years for them to be observed. Neutrinos were first detected in 1953 by F. Reines and C.L. Cowan [3]. They later confirmed their discovery with more data in [4]. They put their experiment next to a nuclear reactor and used the induced beta decay reaction:



where the positron is expected to take most of the energy to carry out their observation. They used a scintillator detector with photomultiplier tubes to observe the gamma rays emitted by the annihilation of the positron with an electron, and also to detect gammas delayed by about  $5\mu s$  resulting from neutron capture. After many checks, including turning off the reactor, they found that they were indeed seeing anti-neutrinos interacting within their detector.

In 1957, C.S. Wu published the results of a difficult experiment using polarised  $\text{Co}^{60}$  [5]. She observed an assymetry in the direction of electrons emitted in the beta decay of this isotope. This was proof that parity was not being conserved in beta decay. This result is closely linked to what Goldhaber, Grodzins and Sunyar found, also in 1957,

namely that neutrinos are left-handed [6]. They used  $\text{Eu}_{63}^{152}$  that has a spin of zero and odd parity. By K-orbital electron capture, this isotope decays into an excited state in the form of  $\text{Eu}_{62}^{152}$  and emits a neutrino in the process. This excited state then goes into a ground state by emitting a photon of the same helicity as the neutrino. Goldhaber, Grodzins and Sunyar measured the circular polarisation of those photons. Based on this measurement, they found that their results pointed to a 100% negative helicity of the neutrinos, that is all the neutrinos in their experiment were left-handed.

Even though neutrinos were generally thought to be massless at the time (this was the accepted wisdom during the early development of neutrino physics), Bruno Pontecorvo proposed the phenomenon of neutrino oscillations in 1957 in analogy with the oscillations observed in the neutral kaon  $K_0 - \bar{K}_0$  system [7, 8, 9]. Oscillations of neutrino flavours were then first introduced by Z. Maki, M. Nakagawa and S. Sakata in the form of two-flavour neutrino mixing [10]. An important feature of neutrino oscillations and neutrino mass mixing is that for those to exist, neutrinos have to have some mass (although it could be minuscule). Neutrino oscillations will be described in more detail later in this chapter.

The muon neutrino was first observed by Lederman, Schwartz, Steinberger and their colleagues in 1962 at the Brookhaven National Laboratory in the US [11]. They used a beam of protons from the Alternating Gradient Synchrotron at BNL to create the first muon neutrino beam. The neutrinos could then be detected downstream as they interacted and produced muons that were detected using a spark chamber. This experiment proved that there was a second type of neutrino in addition to the electron neutrino observed in beta decay: the muon neutrino.

In the 1960s, the Homestake experiment located in the Homestake Mine in South Dakota in the US and led by Davis was the first experiment to detect neutrinos from the sun. It used solar neutrino flux predictions calculated by Bahcall ([12]) to predict the number of solar neutrinos that should have been detected. The experiment itself was a huge tank of perchloroethylene which contains chlorine. Upon being struck by a neutrino, chlorine converts to argon. The argon particles were then extracted and counted and it was found over a period of many years that the number of neutrinos seen was about one third of the predicted number of neutrinos [13]. This observation became

known as the solar neutrino problem and was confirmed by a number of subsequent experiments [14, 15, 16, 17].

In the 1980s, the IMB experiment, and the Kamiokande II experiment, both large underground water cerenkov detectors, found a deficit of muon neutrinos produced as a result of cosmic ray interactions in the atmosphere [18, 19]. This became known as the atmospheric neutrino deficit. In 1998, the Super-Kamiokande collaboration presented data from 535 days of exposure and concluded that, as a function of angle and momentum, the observed deficit of muon neutrinos compared to atmospheric electron neutrinos is consistent with oscillations of muon neutrinos to tau neutrinos [20], thus resolving this observed deficit.

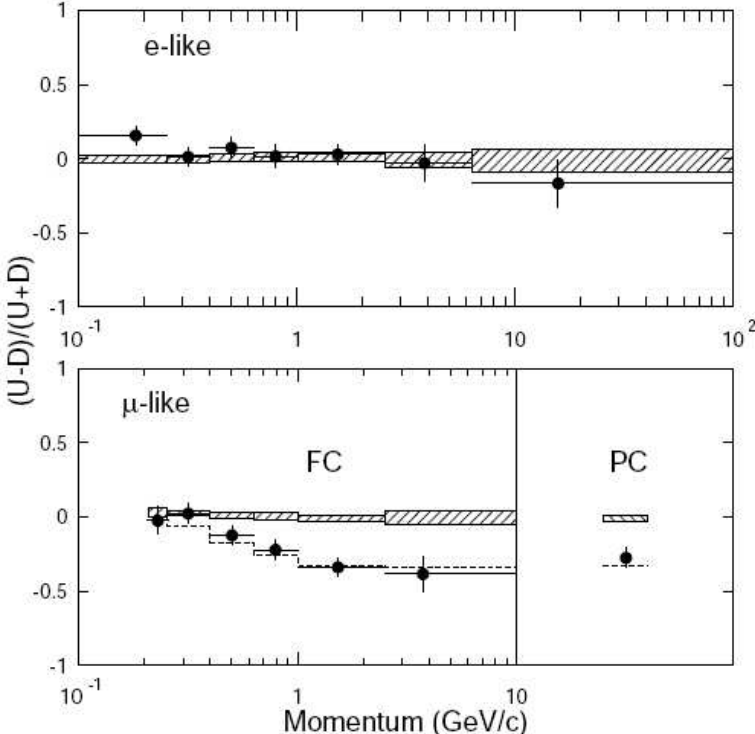


Figure 2.1: The  $(U-D)/(U+D)$  asymmetry of upward and downward going neutrinos should be zero in the case of no oscillations, but it was found to differ from zero as a function of momentum for fully contained (FC) and also for one energy bin whose mean was estimates at around 15GeV for partially contained (PC) events. The Monte Carlo simulation expectation without neutrino oscillations is shown in the hatched region, and the dashed line indicates  $\nu_\mu \rightarrow \nu_\tau$  neutrino oscillations for  $\sin^2 2\theta = 1.0$  and  $\Delta m^2 = 2.2 \times 10^{-3} eV^2$  (Source of Figure [20]).

A momentous event in neutrino physics occurred in 1987 with the explosion of the 1987A supernova in the Large Magellanic Cloud on the outskirts of our galaxy. The ring-structure remaining after this supernova was photographed by the Hubble Space Telescope and Figure 2.2 shows one of the photographs of the supernova remnant. This supernova was an incredible stroke of luck and initiated a new era of neutrino astronomy, which is currently in its early stages but which has already spawned numerous experiments that are currently underway. When supernova 1987A happened, the IMB and the Kamiokande II experiments (both water Cerenkov detectors) were taking data. During a supernova explosion, neutrinos are copiously produced and in this case, both experiments observed neutrinos coming from the 1987A supernova. The IMB experiment observed 8 neutrino events within a time interval of six seconds and within visible energies of 20-40MeV [21]. The Kamiokande II experiment observed 11 events within a time window of 13 seconds and with a visible energy range of 7.5-36MeV [22]. It was possible to reconstruct the origin of two of the 11 events in the general direction of the Large Magellanic Cloud. Unfortunately, supernovae in our vicinity are relatively rare and are expected to occur no more than 2-3 times per century. Neutrino astronomy can however look for other neutrinos, like for example ultra-high-energy neutrinos, that would be messengers from astrophysical sources that cannot be investigated any other way due to their distance from the solar system.

It was only very recently (in 2001) that the SNO [17, 23] experiment in Sudbury, Canada, solved the solar neutrino problem. SNO used a heavy water Cerenkov detector located  $\sim 2000\text{m}$  underground to carry out its observations. This experiment solved the problem by detecting both solar electron neutrinos which created electrons when interacting in the heavy water of the detector, and elastic scattering interactions in which muon and tau neutrinos could also participate and so provided some sensitivity to non-electron neutrino flavours. When the data was analysed, it was found that the total solar neutrino flux could be calculated and that it agreed with the hypothesis that electron neutrinos oscillate away into other neutrino flavours on their way to the earth.

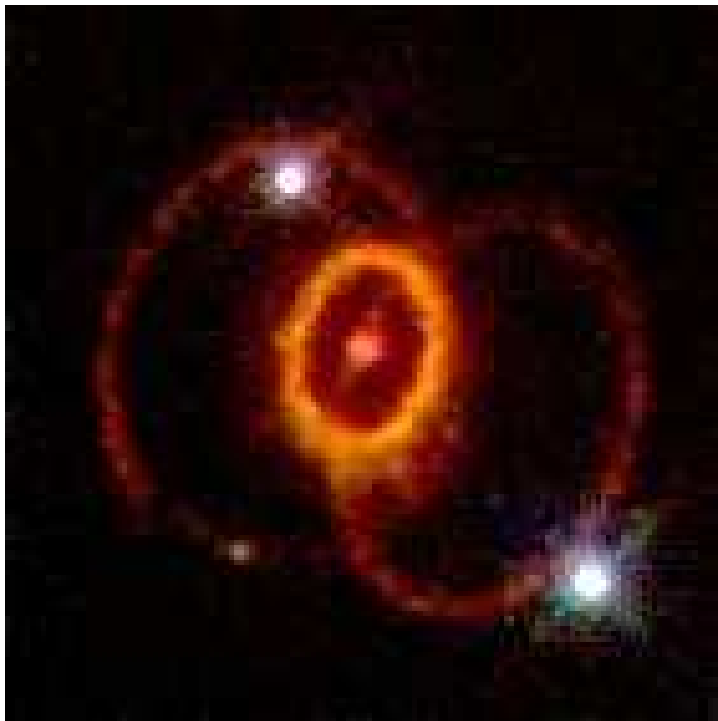


Figure 2.2: A photograph of the ring structure of the explosion remnant of supernova 1987A in the Large Magellanic Cloud. The main ring in the centre is due to gas and dust heated by the supernova. The two larger overlapping rings are somewhat mysterious and different explanations have been put forward. One such explanation is that the two rings actually lie in different planes, one behind the star remnant and one in front of it. In this explanation, the rings are explained as the result of 2 jets in opposite directions emanating from the supernova remnant, which is assumed to have some wobble in its rotation axis aligned towards us and thus creating the rings that we see. The 1987A supernova initiated the new era of neutrino astronomy which is currently in its infancy but holds much promise in investigating the universe and its origins (Source of image and information: NASA HST internet site).

## 2.2 The Standard Model

In the Standard Model of particle physics [24, 25, 26], there are three generations of quarks:

$$\begin{pmatrix} u \\ d \end{pmatrix}, \begin{pmatrix} c \\ s \end{pmatrix}, \begin{pmatrix} t \\ b \end{pmatrix}. \quad (2.2)$$

Similarly, there are three generations of leptons:

$$\begin{pmatrix} \nu_e \\ e \end{pmatrix}, \begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}, \begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix} \quad (2.3)$$

Finally, there are the gauge bosons that are the mediators of the forces between the

particles:

- photons are the exchange particle of the electromagnetic force,
- gluons are the exchange particles of the strong force,
- finally, the massive  $W^{+-}$  and  $Z^0$  bosons are the mediators of the weak force.

The weak force is responsible for known neutrino interactions<sup>2</sup>. Gravity is not included in the Standard Model as it stands now as no satisfactory theory of quantum gravity currently exists.

Data from LEP at CERN [27] have shown (see Figure 2.3) that there are in total three generations of light active neutrino<sup>3</sup>: the electron neutrino ( $\nu_e$ ), the muon neutrino ( $\nu_\mu$ ), and the tau neutrino ( $\nu_\tau$ ). Together with the charged electron, muon, and tau, these particles make up the six leptons in the Standard Model of particle physics.

The properties of the six leptons in the Standard Model are summarised in Table 2.1.

| Lepton     | Charge | Mass                          | Lepton Number |         |          |
|------------|--------|-------------------------------|---------------|---------|----------|
|            |        |                               | $L_e$         | $L_\mu$ | $L_\tau$ |
| $e^-$      | -1     | 0.511 MeV                     | 1             | 0       | 0        |
| $\nu_e$    | 0      | $< 2\text{eV}^4$              | 1             | 0       | 0        |
| $\mu^-$    | -1     | 105.66 MeV                    | 0             | 1       | 0        |
| $\nu_\mu$  | 0      | $< 0.19\text{ MeV}$           | 0             | 1       | 0        |
| $\tau^-$   | -1     | $1776.84 \pm 0.17\text{ MeV}$ | 0             | 0       | 1        |
| $\nu_\tau$ | 0      | $< 18.2\text{ MeV}$           | 0             | 0       | 1        |

Table 2.1: Leptons in the Standard Model (for more detailed particle listings see [31]).

## 2.3 Neutrino Oscillations

A lot of evidence has accumulated in recent years which shows that neutrinos oscillate as they travel through space. Neutrino oscillations require at least two of the neutrinos to have mass. The most essential feature of neutrino oscillation theory is that each observed neutrino flavour eigenstate is a superposition of mass eigenstates. Therefore, the flavour eigenstates govern the neutrinos' interactions with matter, however, the

<sup>2</sup>The neutrino is believed to interact only via the weak and the gravitational force.

<sup>3</sup>In some neutrino theories, other neutrinos like for example sterile neutrinos - i.e. neutrinos that do not interact via the weak force - are postulated.

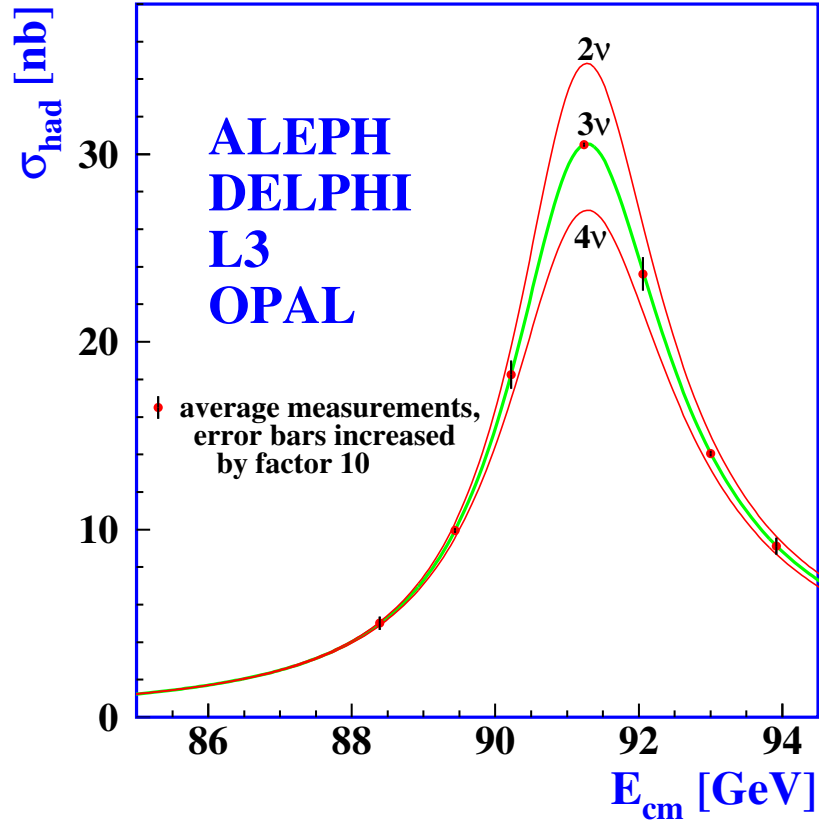


Figure 2.3: LEP measurements of the hadron production cross-section around the  $Z^0$  resonance. The curves indicate the predicted cross-section for two, three and four neutrino species with Standard Model couplings and negligible mass. The data show agreement with the three neutrino species prediction (source of Figure [27]). LEP measured the number of active neutrino species to be  $2.9840 \pm 0.0082$ .

mass eigenstates govern how the neutrinos propagate through space.

### 2.3.1 Two-Flavour Oscillations

In a simplified form, assuming only two species of neutrino, two flavour neutrino oscillations can be parametrized as follows:

$$\begin{pmatrix} \nu_\alpha \\ \nu_\beta \end{pmatrix} = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} \nu_1 \\ \nu_2 \end{pmatrix} \quad (2.4)$$



where  $\nu_\alpha$  and  $\nu_\beta$  are the neutrino flavour eigenstates, and  $\nu_1$  and  $\nu_2$  are neutrino mass eigenstates. We then have:

$$\begin{aligned} |\nu_\alpha\rangle &= \cos\theta|\nu_1\rangle + \sin\theta|\nu_2\rangle \\ |\nu_\beta\rangle &= -\sin\theta|\nu_1\rangle + \cos\theta|\nu_2\rangle \end{aligned} \tag{2.5}$$

If the propagation of the mass states in space with time is given by:

$$\begin{aligned} |\nu_1(t)\rangle &= |\nu_1\rangle e^{-iE_1 t} \\ |\nu_2(t)\rangle &= |\nu_2\rangle e^{-iE_2 t} \end{aligned} \tag{2.6}$$

then after time  $t$ , we have for the time evolution of the flavour states:

$$\begin{aligned} |\nu_\alpha(t)\rangle &= \cos\theta e^{-iE_1 t} |\nu_1\rangle + \sin\theta e^{-iE_2 t} |\nu_2\rangle \\ |\nu_\beta(t)\rangle &= -\sin\theta e^{-iE_1 t} |\nu_1\rangle + \cos\theta e^{-iE_2 t} |\nu_2\rangle \end{aligned} \tag{2.7}$$

Hence after time  $t$  we have the amplitude:

$$\begin{aligned} A_{\nu_\alpha \rightarrow \nu_\beta} &= \langle \nu_\alpha | \nu_\beta(t) \rangle \\ &= (\cos\theta \langle \nu_1 | + \sin\theta \langle \nu_2 |) (-\sin\theta e^{-iE_1 t} |\nu_1\rangle + \cos\theta e^{-iE_2 t} |\nu_2\rangle) \\ &= -\cos\theta \sin\theta e^{-iE_1 t} \langle \nu_1 | \nu_1 \rangle + \cos^2\theta e^{-iE_2 t} \langle \nu_1 | \nu_2 \rangle \\ &\quad - \sin^2\theta e^{-iE_1 t} \langle \nu_2 | \nu_1 \rangle + \sin\theta \cos\theta e^{-iE_2 t} \langle \nu_2 | \nu_2 \rangle \\ &= -\cos\theta \sin\theta e^{-iE_1 t} + \sin\theta \cos\theta e^{-iE_2 t} \\ &= \frac{\sin 2\theta}{2} (-e^{-iE_1 t} + e^{-iE_2 t}) \end{aligned} \tag{2.8}$$

The probability that the neutrino will be found in state  $\beta$  after time  $t$  is then:

$$\begin{aligned}
P_{\nu_\alpha \rightarrow \nu_\beta} &= |A_{\nu_\alpha \rightarrow \nu_\beta}|^2 \\
&= \frac{\sin^2 2\theta}{4} \left( -e^{-iE_1 t} + e^{-iE_2 t} \right) \left( -e^{iE_1 t} + e^{iE_2 t} \right) \\
&= \frac{\sin^2 2\theta}{4} \left( 1 - e^{-iE_1 t} e^{iE_2 t} - e^{-iE_2 t} e^{iE_1 t} + 1 \right) \\
&= \frac{\sin^2 2\theta}{4} \left( 2 - e^{it(E_2 - E_1)} - e^{-it(E_2 - E_1)} \right) \\
&= \frac{\sin^2 2\theta}{4} \left\{ 2 - \cos[t(E_2 - E_1)] - i \sin[t(E_2 - E_1)] \right. \\
&\quad \left. - \cos[-t(E_2 - E_1)] - i \sin[-t(E_2 - E_1)] \right\} \\
&= \frac{\sin^2 2\theta}{4} \left\{ 2 - 2 \cos[t(E_2 - E_1)] \right\} \\
&= \frac{\sin^2 2\theta}{2} \left\{ 1 - \cos[t(E_2 - E_1)] \right\}
\end{aligned} \tag{2.9}$$

Now, in the relativistic limit,  $|\vec{p}_i| = p_i \gg m_i$ . Hence, we approximately have:

$$E_i = \sqrt{p_i^2 + m_i^2} \simeq p_i + \frac{m_i^2}{2p_i} \simeq E + \frac{m_i^2}{2E} \tag{2.10}$$

Also, if  $t \simeq L/c = L$  in natural units ( $\hbar = c = 1$ ), then Equation 2.9 becomes:

$$\begin{aligned}
P_{\nu_\alpha \rightarrow \nu_\beta} &= \frac{\sin^2 2\theta}{2} \left\{ 1 - \cos \left\{ t \left[ \left( E + \frac{m_2^2}{2E} \right) - \left( E + \frac{m_1^2}{2E} \right) \right] \right\} \right\} \\
&= \frac{\sin^2 2\theta}{2} \left\{ 1 - \cos \left[ L \left( \frac{m_2^2 - m_1^2}{2E} \right) \right] \right\} \\
&= \frac{\sin^2 2\theta}{2} \left[ 1 - \cos \left( \frac{\Delta m^2 L}{2E} \right) \right] \\
&= \sin^2 2\theta \sin^2 \left( \frac{\Delta m^2 L}{4E} \right) \\
&= \sin^2 2\theta \sin^2 \left( \frac{1.27 \Delta m^2 L}{E} \right)
\end{aligned} \tag{2.11}$$

where the baseline  $L$  is measured in km,  $E$  is measured in GeV, and  $\Delta m^2 = m_2^2 - m_1^2$  is measured in  $eV^2/c^4$ . If  $P_{\alpha \rightarrow \beta}$  is the oscillation probability, then the survival probability is simply:

$$P_{\nu_\alpha \rightarrow \nu_\alpha} = 1 - \sin^2 2\theta \sin^2 \left( \frac{1.27 \Delta m^2 L}{E} \right) \tag{2.12}$$

In order to illustrate how neutrino oscillations affect the true muon neutrino energy spectrum, the neutrino oscillation probability function using the two-flavour mixing parametrisation is plotted in Figure 2.4 for the atmospheric oscillation parameters

currently measured by the MINOS experiment [47] and for a baseline  $L=735\text{km}$ .

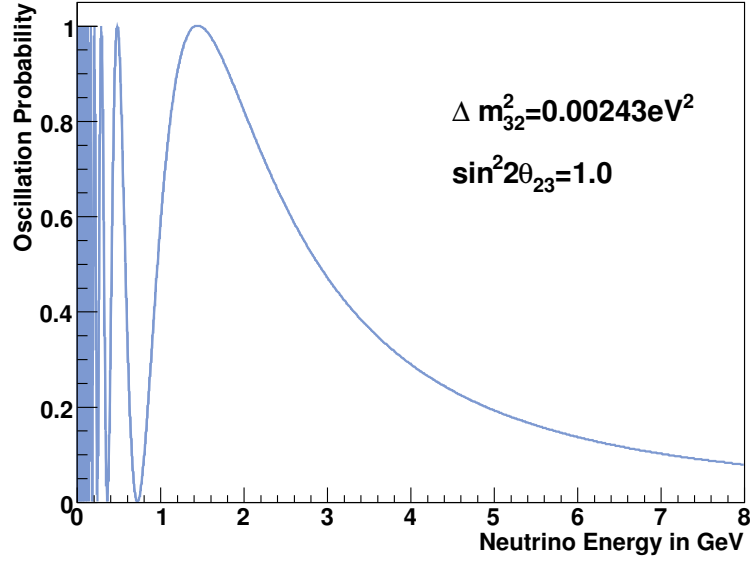


Figure 2.4:  $\nu_\mu \rightarrow \nu_x$  oscillation probability for the MINOS experiment, which has a baseline of 735km.

### 2.3.2 Three-Flavour Oscillations

In the more general case of three neutrino flavours, the neutrino mixing can be parametrised with the Pontecorvo-Maki-Nakagava-Sakata (PMNS) matrix:

$$U = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix} \quad (2.13)$$

In this parametrization,  $c_{ij} = \cos \theta_{ij}$ ,  $s_{ij} = \sin \theta_{ij}$ ,  $\theta_{ij}$  is the mixing angle, and  $\delta$  is

the CP-violating phase. The PMNS matrix can also be rewritten as follows:

$$U = \begin{pmatrix} 1 & & \\ & c_{23} & s_{23} \\ & -s_{23} & c_{23} \end{pmatrix} \times \begin{pmatrix} c_{13} & s_{13}e^{-i\delta} & \\ & 1 & \\ -s_{13}e^{i\delta} & & c_{13} \end{pmatrix} \times \begin{pmatrix} c_{12} & s_{12} & \\ -s_{12} & c_{12} & \\ & & 1 \end{pmatrix} \times \begin{pmatrix} 1 & & \\ & e^{i\alpha} & \\ & & e^{i\beta} \end{pmatrix} \quad (2.14)$$

This way of writing the matrix decomposes it into the different mixing regimes. Solar neutrino oscillations are associated with the (12) sector and the solar mass difference term  $\Delta m_{\odot}^2 = \Delta m_{21}^2$ . Atmospheric oscillations are associated with the (23) sector and with the mass difference term  $\Delta m_{atm}^2 = \Delta m_{32}^2$ . Finally, the (13) sector is associated with  $\nu_e$  flavour appearance and as can be seen from the matrix, it is a term closely linked to the CP-violating phase  $\delta$ . The mixing angle  $\theta_{13}$  has not been measured to date and if it is found to be zero, then all the terms containing the CP violating term will vanish and CP violation in the neutrino sector will be equivalent to zero (strictly speaking the CP-phase  $\delta$  could have any value then but it would have no effect). Neutrino CP violation is particularly important for various models of baryogenesis through leptogenesis in the early universe, so could have far reaching implications for cosmological models. If the CP-violating phase is found to be non-zero, this could explain the observed matter-antimatter imbalance in the universe. A non-zero  $\theta_{13}$  would be observable for example via electron neutrino appearance in a muon neutrino beam.

### 2.3.3 Neutrino Mass Hierarchy

If there are indeed only three neutrino types, then there are two possible mass hierarchies. These are the normal and the inverted mass hierarchies. The normal mass hierarchy is defined:

$$m_{\nu_1} < m_{\nu_2} < m_{\nu_3}$$

The inverted mass hierarchy corresponds to:

$$m_{\nu_3} < m_{\nu_1} < m_{\nu_2}$$

Oscillation experiments like MINOS cannot measure absolute neutrino masses because individual masses cannot be extracted from the  $\Delta m^2$  term in the oscillation propagation

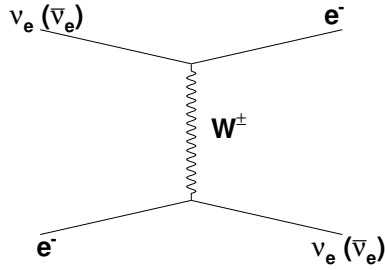


Figure 2.5: Electron (anti-) neutrinos will interact with matter and will scatter elastically via charged current interactions off electrons. The neutrino strikes an electron, exchanges a charged W-boson. The outgoing particles are an electron and a (anti-) neutrino.

formula (see Equation 2.11). Oscillation experiments are thus only able to measure and put limits on mixing angles and mass-squared differences between the mass eigenstates. Direct measurements of neutrino masses have so far only yielded limits (see Table 2.1) and it has not yet been possible to determine the mass of any of the neutrinos. It is even possible that the mass of the lightest neutrino is zero. Neutrino matter effects, which are discussed in the next section, should modify the neutrino oscillation probabilities to a large enough extent to potentially allow us to determine the neutrino mass hierarchy at future experiments.

## 2.4 Neutrino Matter Effects

When neutrinos traverse matter, their propagation is altered by coherent forward scattering from atoms in their way. These matter effects are described by the Mikheyev–Smirnov–Wolfenstein mechanism (MSW) [28, 29]. Neutrinos of all flavours will scatter off matter, but only electron neutrinos will scatter elastically via charged current interactions off electrons as shown in Figure 2.5. This affects oscillation probabilities. Matter effects are for example important for electron neutrinos produced in the sun as they traverse the electron dense layers of the sun.

The magnitude of the coherent forward scattering depends on the electron number density  $N_e$  and the Fermi Constant  $G_F$ . This extra potential is given by:

$$V_e = \sqrt{2}G_F N_e \tag{2.15}$$

For anti-neutrinos, the sign of  $V_e$  is reversed, therefore, matter effects distinguish between neutrinos from anti-neutrinos. Recall from Equation 2.4 that:

$$\begin{pmatrix} \nu_e \\ \nu_x \end{pmatrix} = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} \nu_1 \\ \nu_2 \end{pmatrix} \quad (2.16)$$

The time evolution of the mass eigenstates will be given by:

$$i \frac{d}{dt} \begin{pmatrix} \nu_1 \\ \nu_2 \end{pmatrix} = \begin{pmatrix} E_1 & 0 \\ 0 & E_2 \end{pmatrix} \begin{pmatrix} \nu_1 \\ \nu_2 \end{pmatrix} \quad (2.17)$$

The flavour basis of Equation 2.17 can be changed by rewriting Equation 2.16 as:

$$X = A Y$$

Then:

$$A^{-1} X = A^{-1} A Y = Y$$

Then, substituting  $A^{-1} X$  for  $Y$  in Equation 2.17, and multiplying both sides by  $A$  from the left, it can be rewritten as:

$$i \frac{d}{dt} \begin{pmatrix} \nu_e \\ \nu_x \end{pmatrix} = \begin{pmatrix} E_1 \cos^2 \theta + E_2 \sin^2 \theta & (E_2 - E_1) \sin \theta \cos \theta \\ (E_2 - E_1) \sin \theta \cos \theta & E_1 \sin^2 \theta + E_2 \cos^2 \theta \end{pmatrix} \begin{pmatrix} \nu_e \\ \nu_x \end{pmatrix} \quad (2.18)$$

Now, by subtracting a phase  $E_1 I$  which has no effect on the neutrino oscillations, and by adding to the hamiltonian the extra potential arising from matter effects, one obtains:

$$i \frac{d}{dt} \begin{pmatrix} \nu_e \\ \nu_x \end{pmatrix} = \begin{pmatrix} (E_2 - E_1) \sin^2 \theta \pm \sqrt{2} G_F N_e & (E_2 - E_1) \sin \theta \cos \theta \\ (E_2 - E_1) \sin \theta \cos \theta & (E_2 - E_1) \cos^2 \theta \end{pmatrix} \begin{pmatrix} \nu_e \\ \nu_x \end{pmatrix} \quad (2.19)$$

The eigenstates for Equation 2.19 are:

$$\begin{aligned} |\nu_{1m}\rangle &= \cos \theta_m |\nu_e\rangle + \sin \theta_m |\nu_x\rangle \\ |\nu_{2m}\rangle &= -\sin \theta_m |\nu_e\rangle + \cos \theta_m |\nu_x\rangle \end{aligned} \quad (2.20)$$

where the angle  $\theta_m$  is defined by:

$$\tan 2\theta_m = \frac{(E_2 - E_1) \sin 2\theta}{(E_2 - E_1) \cos 2\theta \mp \sqrt{2} G_F N_e} \quad (2.21)$$

Now, using the same derivation mechanism as for the two-flavour vacuum oscillations,

one obtains:

$$P_{\nu_e \rightarrow \nu_x} = \sin^2 2\theta_m \sin^2 \left( \frac{(E_{1m} - E_{2m})L}{2} \right) \quad (2.22)$$

where  $E_{im}$  are the eigenvalues derived from Equation 2.19 and their difference is given by:

$$E_{1m} - E_{2m} = \frac{\Delta m_{21}^2}{2E} \sqrt{\sin^2 2\theta + \left( \frac{\pm \sqrt{2} G_F N_e}{\Delta m_{21}^2 / 2E} - \cos 2\theta \right)^2} \quad (2.23)$$

and  $\sin^2 2\theta_m$  is given by:

$$\sin^2 2\theta_m = \frac{\sin^2 2\theta}{\sin^2 2\theta + \left( \cos 2\theta \mp \frac{\sqrt{2} G_F N_e}{\Delta m_{21}^2 / 2E} \right)^2} \quad (2.24)$$

If we then have:

$$\Delta M_{21}^2 = \Delta m_{21}^2 \sqrt{\sin^2 2\theta + \left( \frac{\pm \sqrt{2} G_F N_e}{\Delta m_{21}^2 / 2E} - \cos 2\theta \right)^2} \quad (2.25)$$

Then we can write the oscillation probability as:

$$P_{\nu_e \rightarrow \nu_x} = \sin^2 2\theta_m \sin^2 \left( \frac{1.27 \Delta M_{21}^2 L}{E} \right) \quad (2.26)$$

When the matter density is low, then  $N_e \rightarrow 0$ ,  $\Delta M_{21}^2 \rightarrow \Delta m_{21}^2$  and  $\sin^2 2\theta_m \rightarrow \sin^2 2\theta$ , so the oscillation probability  $P_{\nu_e \rightarrow \nu_x}$  reverts to the vacuum oscillation probability.

An interesting feature of matter effects is that it is possible to achieve resonant enhancement of the oscillation probability when:

$$\frac{\sqrt{2} G_F N_e}{\Delta m_{21}^2 / 2E} = \cos 2\theta \quad (2.27)$$

This maximal mixing occurs for neutrinos if  $m_2 > m_1$ , and for anti-neutrinos if  $m_1 > m_2$ , and is therefore dependent on the mass hierarchy.

## 2.5 The Chooz experiment

An important experiment with regard to the analysis described in this document is the Chooz experiment in France [30], which finished data taking in 1998 and currently has set the world's best limit on the value of  $\sin^2 2\theta_{13}$ , which is the parameter the electron neutrino appearance analysis discussed in this documents attempts to measure/set a better limit on. Chooz used a liquid scintillator detector instrumented with

photomultiplier tubes to detect electron anti-neutrinos generated by the running of the two nuclear reactors at the nuclear power station in Chooz, France. The detector was located about 1km away from the two nuclear reactors which were generating the anti-neutrinos. In particular, it compared data taken when the reactor was switched off, and when the reactor was on. The number of electron anti-neutrinos observed was consistent with predictions and no deficit was observed that would have indicated anti-neutrinos oscillating away into other flavours on their way to the detector. The Chooz experiment was therefore able to set limits on the oscillation mode  $\bar{\nu}_e \rightarrow \bar{\nu}_x$  as is shown in Figure 2.6. The limit contours are shown in atmospheric mass difference term  $\Delta m_{atm}^2$  versus  $\sin^2 2\theta_{13}$  parameter space. It is important to note that MINOS measures the former precisely and the most current results are given in the next chapter, which constrains the limit on the latter to  $\sin^2 2\theta_{13} < 0.15$ .

## 2.6 Neutrino Interactions in MINOS

There are two event types which neutrinos undergo. Those are charged current (CC) and neutral current (NC) interactions, as shown in Figures 2.7 and 2.8 respectively. The important difference between the two types of interaction is that for CC events, a charged lepton is produced, whereas for NC events, no charged lepton is produced. This allows CC events to be identified. For example in MINOS, the muon neutrinos strike iron nuclei in the detectors and muons are produced which leave long tracks that can be detected. Usually, a hadronic shower is also present in the events. For NC events, only a shower is produced since the outgoing lepton is a neutrino which doesn't leave a trace.

There are several categories of neutrino events that can occur in MINOS:

**Quasi-Elastic Events** - these are events that are usually characterised by an electron or muon track without a shower since the vast majority of the neutrino energy is transferred directly to the outgoing lepton

**Resonance Events** - these are events where the neutrino usually produces a  $\Delta$  resonance ( $\Delta^+$ ,  $\Delta^{++}$ ,  $\Delta^0$ ) in the final state which then produces a pion

**Coherent Pion Events** - these are events in which a coherent  $\pi^0$  is produced when the neutrino impacts the nucleus. Such events are characterised by the fact that the outgoing pion is very forward and the struck nucleus remains in its ground state



**Deep Inelastic Scattering (DIS) Events** - these are events where the incoming neutrino transfers part of its energy to the struck nucleus thus generating a particle shower, and part of its energy to the outgoing lepton (see Figure 2.9).

Most of the events occurring in the MINOS experiment are DIS events. The kinematic quantities needed to describe DIS events are shown in Figure 2.9. If

$$\nu = \frac{q \cdot P}{M} = E - E' \quad (2.28)$$

is the loss of energy of the lepton in the nucleon rest frame, then there are four important quantities:

1.  $Q^2$  is the four-momentum squared transferred to the nucleus by the incoming neutrino (lepton):

$$Q^2 = -q^2 \quad (2.29)$$

2.  $x$  is the fraction of the nucleon's momentum carried by the struck quark:

$$x = \frac{Q^2}{2M\nu} \quad (2.30)$$

3.  $y$  is the fraction of the lepton's energy lost in the nucleon' rest frame. In other words, this is the fraction of the original neutrino energy that creates a hadronic shower:

$$y = \frac{q \cdot P}{k \cdot P} = \frac{\nu}{E} \quad (2.31)$$

4.  $W^2$  is the invariant mass squared of the hadronic system recoiling against the scattered lepton:

$$W^2 = (P + q)^2 = M^2 + 2M\nu - Q^2 \quad (2.32)$$

More information about the above quantities can be found in [31]. They are Lorentz invariant and the four quantities together describe the whole interaction system. They are also very important as they determine the size and angle of the hadronic shower and the length and momentum of the lepton track seen in the detector. For example, a high hadronic  $y$  for a muon neutrino CC event will mean that the hadronic shower will be large, but the muon track will be very short and very hard to identify within the shower.

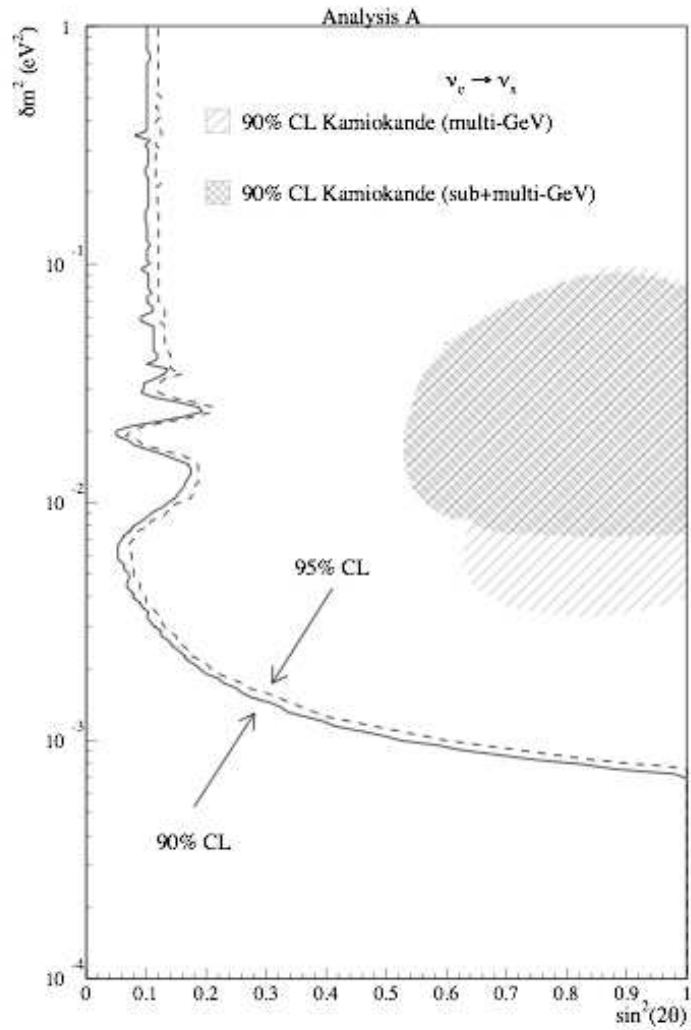


Figure 2.6: This figure shows the 90% exclusion limits generated by the Chooz experiment [30]. The exclusion limits are in the  $\Delta m_{atm}^2$  versus  $\sin^2 2\theta_{13}$  parameter space. For the  $\Delta m^2$  measured by the MINOS experiment, the limit set by the Chooz experiment is  $\sin^2 2\theta_{13} < 0.15$ .

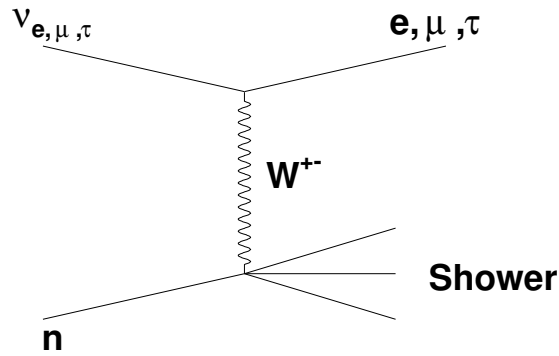


Figure 2.7: Neutrino charged current event. The neutrino strikes a nucleus, exchanges a charged  $W$ -boson, becomes a muon or an electron (or a tau). The struck nucleus produces a particle shower.

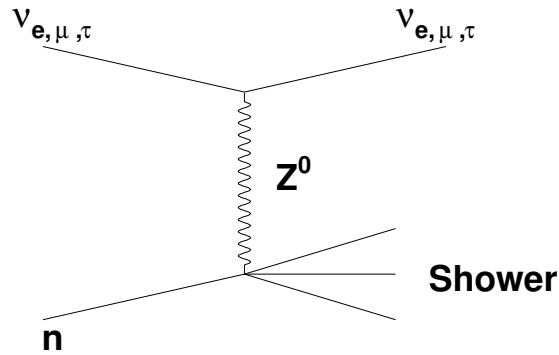


Figure 2.8: Neutrino neutral current event. The neutrino strikes a nucleus, exchanges a neutral  $Z$ -boson, and continues as a neutrino. The struck nucleus produces a particle shower.

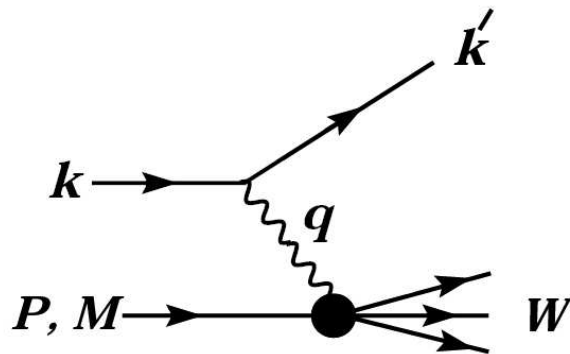


Figure 2.9: Diagram of the kinematic quantities involved in deep inelastic scattering. The four-momenta of the incoming and outgoing leptons are  $\mathbf{k}$  and  $\mathbf{k}'$ .  $\mathbf{q}$  is the momentum transferred to the struck nucleus.  $\mathbf{P}$  and  $M$  are the four-momentum and the mass of the struck nucleon, and  $\mathbf{q}$  is the mass of the recoiling system (Source [31]).

## Chapter 3

# The MINOS Experiment

The Main Injector Neutrino Oscillation Search (MINOS) is a long-baseline neutrino experiment designed to investigate the phenomenon of neutrino oscillations. MINOS is in a unique position to be able to distinguish neutrino oscillations from more exotic theories like neutrino decay. The primary goal of the MINOS experiment is to confirm the validity of the neutrino oscillations theory, and to determine, to a high precision, parameters associated with those oscillations.

In the MINOS experiment, a beam of protons from the Main Injector at the Fermi National Accelerator Laboratory (FNAL or Fermilab), Illinois, is incident on a fixed graphite target, eventually generating a high intensity / high purity beam of  $\nu_\mu$ -neutrinos. As Figure 3.1 shows, the neutrino beam first traverses the Near detector at Fermilab before the neutrinos have had a chance to oscillate, then passes through 734km of earth rock, and then arrives at the Far detector that is situated 710m underground in the Soudan Mine in Minnesota.

### 3.1 The NuMI Beam

To create the NuMI (Neutrinos at the Main Injector) beam [32, 33, 34], the booster accelerator produces protons of momentum 8GeV/c, which are then accelerated by the Main Injector to 120GeV/c. Those protons are extracted from the Main Injector in 5 to 7 batches per cycle with a cycle time of 2.2-2.4s and contain  $2.1\text{-}3.0 \times 10^{13}$  protons per pulse depending on running conditions (improvements by the NuMI facility engineers mean that this number can reach  $3.0 \times 10^{13}$  and more in 2009/2010). The protons pass



Figure 3.1: MINOS experiment design. A high intensity beam of muon neutrinos is generated at Fermilab in Illinois. It passes a Near detector and then travels 734km through the earth to a Far detector in the Soudan Mine in Minnesota.

through a collimating baffle which is made out of graphite and is 1.5m long and has an 11mm diameter bore at its centre along its full length. The baffle collimates the beam and serves to protect downstream elements (especially the magnetic horns) from exposure to misdirected proton pulses. After collimation, the protons are incident onto a target with a typical beam spot of 1.2mm. The NuMI target is a thin graphite rod composed of 47 individual 2cm long segments with a gap of 3mm between the individual segment (called fins), giving a total length of 94cm, which allows most of the protons in the beam pulse to interact. The target's width is 6.4mm, and its height 15mm. The target's transverse dimensions were determined so as to provide sufficient target integrity in the tough beam conditions, yet for the target to be thin enough to obtain a good flux of mesons and minimize secondary interactions within and reabsorption of the latter by the target. The target density is  $1.78\text{g/cm}^3$  and it can be moved up to 2.5m from the zero-position along the beam direction. This, together with the two focusing horns, results in NuMI's in-built ability to adjust the peak energy of the produced neutrinos.

The two NuMI focusing horns are connected in series and are pulsed to produce toroidal magnetic focusing fields of 3.0T. The horn fields allow the charge and momentum selection of the secondary particles produced by the impact of the beam protons on the target. The horns have a nominal current of 200kA, but the current can be varied, and so, in conjunction with the movable target, MINOS is able to vary the energy spectrum of the produced neutrinos. There are three main configurations: low energy, medium energy, and high energy - this can be seen in Figure 3.2. The low energy configuration yields a neutrino spectrum that peaks at 2-3GeV, whereas the high energy configuration produces a spectrum peaking at around 8-10GeV. Most of the data taken until now has been in the low energy LE-10 configuration as this is where MINOS has the most sensitivity to neutrino oscillations. Figure 3.3 is a schematic drawing of how the neutrino beam is generated.

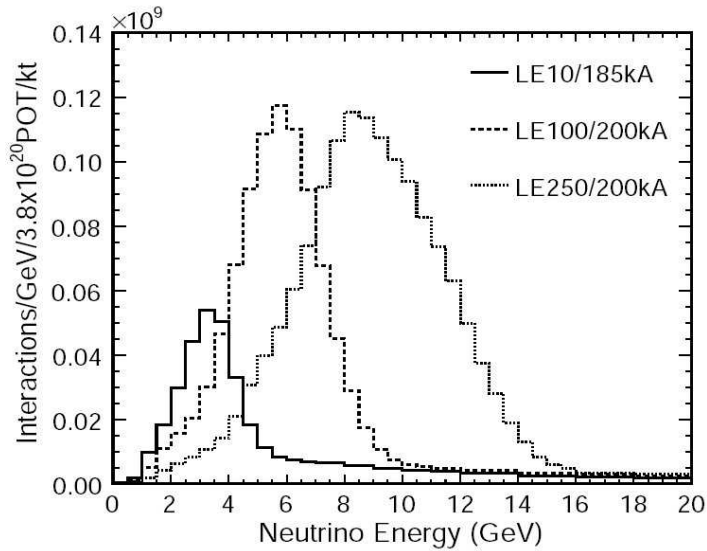


Figure 3.2: This figure shows the calculated rate of  $\nu_\mu$  charged current interactions in the MINOS Near detector for three different spectra. The LE10/185kA spectrum is for a horn current of 185kA and a target position of 0.1m from its zero-position along the beam axis (at zero, the target is furthest downstream and within horn I). Similarly, the LE100/200kA spectrum is for a target position of 1m and a horn current of 200kA, and the LE250/200kA spectrum is for the same current, but a maximal target position of 2.5m (Figure from [32]).

After the beam passes the focusing horns, it travels along a 2m diameter steel pipe 675m in length, that has been evacuated to a pressure of 0.5 Torr to reduce particle

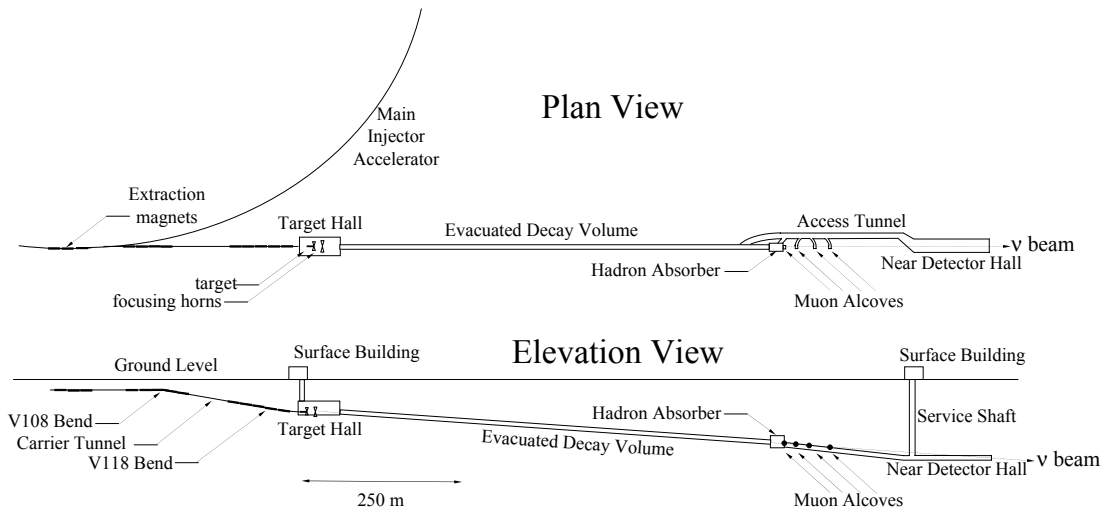


Figure 3.3: Plan and elevation views of the NuMI beam facility. A proton beam is directed onto a target, where the secondary mesons are focused into an evacuated decay volume via magnetic horns. Ionization chambers at the end measure the secondary hadron beam and tertiary muon beam. The hadron absorber and the earth rock traversed by the beam removes any remaining pions, kaons, and muons, thus leaving a pure beam of neutrinos heading for the Near and Far detectors (source of figure [32]).

scattering and absorption<sup>1</sup>. The decay pipe is surrounded by up to 3.5m of concrete shielding. The mesons decay along their path mainly via the channels (the branching ratios - BR - are taken from [31]):

$$\begin{aligned}
 \pi^+ &\rightarrow \mu^+ + \nu_\mu & (\text{or } \pi^- &\rightarrow \mu^- + \bar{\nu}_\mu) & (\text{BR of } \sim 100\%) \\
 K^+ &\rightarrow \mu^+ + \nu_\mu & (\text{or } K^- &\rightarrow \mu^- + \bar{\nu}_\mu) & (\text{BR of } 63.55 \pm 0.11\%) \\
 K^\pm &\rightarrow \pi^\pm + \pi_0 & & & (\text{BR of } 20.66 \pm 0.08\%) \\
 \mu^+ &\rightarrow e^+ + \nu_e + \bar{\nu}_\mu & & & (\text{BR of } \sim 100\%)
 \end{aligned}$$

The length of the NuMI decay pipe means that 75% of 10GeV/c pions and 100% of 10GeV/c kaons will decay by the end of the decay volume [35]. Due to the fact that the original beam is made of protons (hence positively charged secondary mesons are favoured over negatively charged ones), and also because the horn current is set so as to focus positively charged particles, the final neutrino beam consists mostly of muon neutrinos (92.9%). However, there is a small muon anti-neutrino component (5.8%) and a small electron neutrino intrinsic background component (1.3% - the latter will be described in more detail in Chapter 4). It is possible to reverse the horn current to focus

<sup>1</sup>In the most recent MINOS data run, usually referred to as run III, this decay pipe has been filled with helium so as to reduce the stress on the pipe entrance window caused by the aging of the aluminum. This change affected the neutrino beam slightly, however, the data from this run was not used in the results presented here.

negatively charged mesons rather than positively charged ones so as to obtain a muon anti-neutrino beam.

At the end of the decay pipe, the remaining mesons and protons are slowed down and absorbed from the beam by a hadron absorber. This is made of a water cooled core of steel and aluminum, surrounded by blocks of steel and concrete. After the absorber, the beam passes through another 240m of rock, which removes the vast majority of the remaining muons from the beam and leaves practically only neutrinos<sup>2</sup>. The neutrino beam comes out of the rock in the Near detector hall at a  $3.3^\circ$  angle and passes through the Near detector, and then continues traveling through the earth until it reaches the MINOS Far detector in Soudan, 734km away.

In order to align and monitor the beam there is different instrumentation along the beam. There are beam position monitors and other monitors to monitor the proton beam and the beam spot size [32]. Further downstream of the decay pipe, there is a hadron monitor which is an ionization chamber monitor that measures the flux and spatial profiles of the hadrons in the beam prior to the hadron absorber. Downstream of the hadron absorber, there are three muon monitors at intervals in the rock traversed by the beam which are used to measure the rates and spacial profiles of the muons remaining in the beam [36, 37]. The hadron and muon monitor positions can be seen better in the schematic in Figure 3.5.

## 3.2 The MINOS Detectors

The MINOS Near and Far detectors [39] are optimized to observe muons resulting from charged current interactions of muon neutrinos and are designed to be as similar as possible so as to cancel out systematic errors. Nevertheless, they differ in some respects like for example size and shape, but those differences are accounted for in the analysis of the data, so the detectors can be considered functionally identical. The MINOS detectors are steel-scintillator sampling calorimeters which are made of 2.54cm thick magnetised steel vertical planes covered on one side with planes of 4.1cm wide 1cm thick extruded

---

<sup>2</sup>Any muons that still remain in the beam after reaching the Near detector are removed at analysis time by applying data quality and fiducial volume selection cuts.



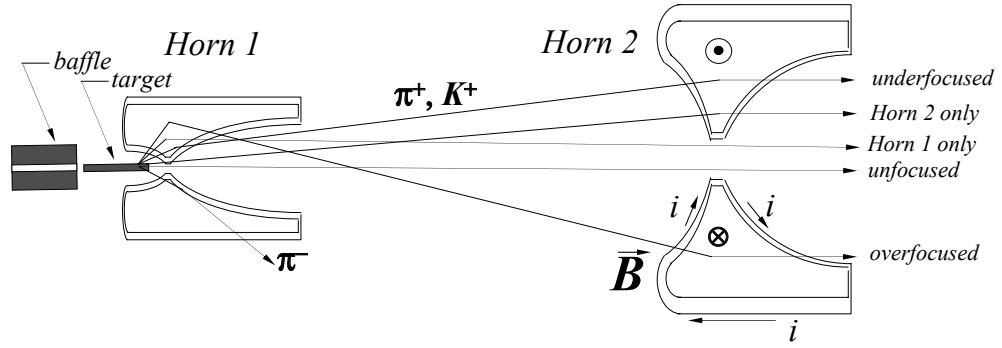


Figure 3.4: Schematic of the MINOS focusing horns. The hadrons produced by the impact of the proton beam onto the target are focused by the horns’ toroidal magnetic fields. The two horns are separated by 10m and the vertical scale is 4 times that of the horizontal scale (source of Figure [32]).

scintillator strips with embedded wavelength shifting optical fibers. Alternate scintillator planes are oriented at  $90^\circ$  to each other in  $u$  and  $v$  directions which correspond to the  $x$  and  $y$  directions being rotated by  $45^\circ$  - this can be seen in Figure 3.6. When particles interact in the detectors, the light is captured in the wavelength-shifting fibres and then clear optical fibres pass the light to multi-anode photo-multiplier tubes (PMTs). The magnetisation of the MINOS detectors allows the experiment to measure the momenta of muons from curvature. This complements the basic measurement of momentum from range. The detector magnetisation also allows the determination of the muon charge and thus whether the incoming particle is a neutrino or an anti-neutrino.

### 3.2.1 The MINOS Near Detector

The MINOS Near detector (ND) has a total mass of 0.98kT and comprises 282 steel planes, 153 of which are instrumented. The calorimeter section of the detector comprises the first 120 planes and is fully instrumented. The spectrometer section comprises planes above 120 and only every fifth plane is instrumented. The ND has dimensions of  $3.8\text{m} \times 4.4\text{m} \times 15\text{m}$  and uses Fast QIE electronics because of the high event rate due to its proximity to the target. The ND signal is read out by 64-anode Hamamatsu photomultiplier tubes (M64 PMTs). The ND cavern is located 100m underground at the end of the NuMi beamline at Fermilab in the US, approximately 1km from the target. The magnetic field strength, when averaged over the fiducial volume, is 1.28T in the ND. A

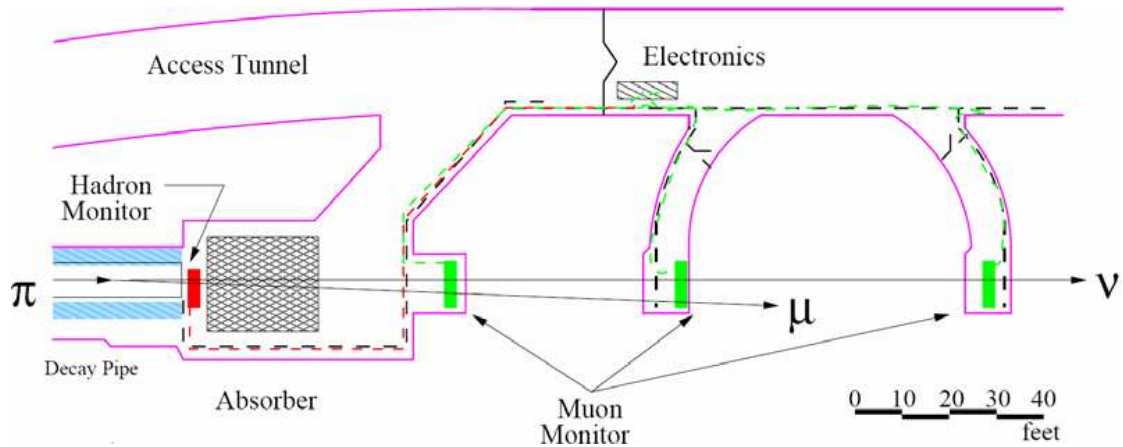


Figure 3.5: Schematic of the positions of the ionization chambers used to monitor the tertiary hadrons and muons with respect to the beam (source of Figure [37]).

picture of the detector is shown in Figure 3.7.

### 3.2.2 The MINOS Far Detector

The MINOS Far detector (FD) is located 710m underground in the Soudan Mine in Minnesota, US. It has a total mass of 5.4kT and consists of 486 octagonal steel/scintillator planes, but is divided into two supermodules (SM1 and SM2) of 249 and 237 planes respectively. The two supermodules are 14.78m and 14.10m respectively, with a gap of 1.15m between them, thus giving a total FD length of 30m. The detector is 8m high (and wide). The large mass of the FD is necessary because the FD event rate is much lower than for the ND. This is because over the 734km between the two detectors, the beam diverges significantly, as  $1:R^2$ , and is several km wide as it reaches the FD. Even with this large mass, the event rate at the FD is very low, just a few events each day. Because of the low event rate, it was possible to use VA electronics and  $8\times$  optical multiplexing in the FD signal readout. At their centre, throughout the length of the detector, the two supermodules each have a magnetic coil and the resulting magnetic field in the detector is 1.42T. GPS time stamping is used to synchronize the Far detector data to the NuMi beam and there is continuous readout of the whole detector. In order to be able to reduce background and identify cosmic muons, the FD also possesses a special veto shield above it. This shield can be seen above the detector in Figure 3.8.

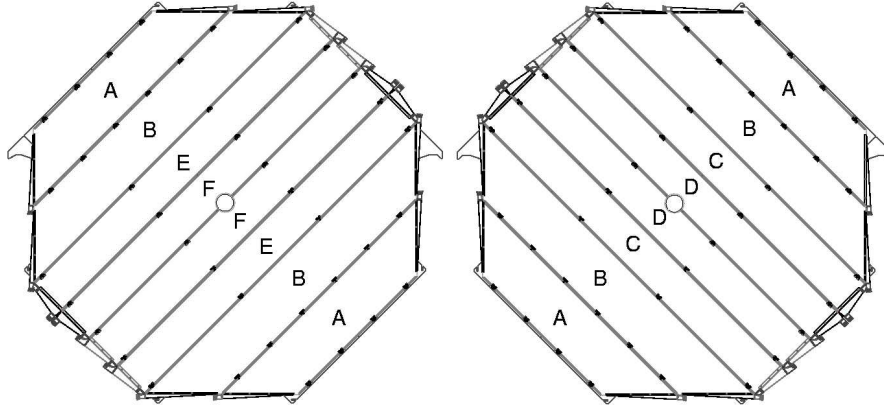


Figure 3.6: The MINOS Far detector schematic view of U and V planes, which alternate in the detector in order to allow the 3-D reconstruction of events. Module types A and B have 28 scintillator strips each, and the other, and C, D, E, and F have 20 scintillator strips each (source of Figure [39]).

### 3.3 Calibration of the MINOS Detectors

In order to be able to measure the energies of neutrino events reliably in both detectors, the MINOS detectors need to be calibrated accurately and reliably on a continuous basis - both the relative calibration between the two detectors and the absolute calibration need to be correct to within 2% and 5% respectively [40, 41, 42, 43]. Continuous calibration of the detectors is important because of effects like PMT gains changes over time (for example as temperature and humidity fluctuate in the detector halls) or scintillator light output changes due to scintillator aging. There are several steps which allow the MINOS detectors to be calibrated:

**Through-Going Cosmic Ray Muons** - in order to achieve strip-to-strip calibration within the two detectors and to account for attenuation in the optical fibers and connections, MINOS uses cosmic ray muons separately within each detector as standard candles (track angle variations are taken into account). The through-going cosmic ray muon data is able to reduce 30% strip-to-strip variations in light output to 8% ([32]).

**Test Bench Scan of Scintillator Modules** - in order to map out the response along each scintillator strip, a  $^{137}\text{Cs}$  radioactive source was used on a test bench setup prior to the installation of the modules. An empirical model of optical attenuation in the wavelength-shifting and clear optical fibers was used to parametrise the response data



Figure 3.7: The MINOS Near detector. This is a front view of the detector; the first steel plane that is seen is most upstream with respect to the neutrino beam.

so as to be able to correct the signals from cosmic-ray muons during the calculation of the uniformity calibration constants. The parametrisation also allows the correction of reconstructed shower energies based on hit strip positions by as much as 30% in the Far detector and 50% in the Near detector.

**Charge Injection** - in order to calibrate the electronics of the detectors, known quantities of electric charge are injected into the front end electronics and digitized which allows the precise determination (to within 1%) of the number of ADC counts for a given input charge. Pedestals are also continuously monitored and subtracted automatically.

**Light Injection System** - a light injection (LI) system is used for the calibration of the entire chain downstream of the scintillator. It uses 400nm blue LEDs which are fanned out into multiple optical fibres. The LED light is injected into the wavelength shifting fibers embedded in the scintillator strips. The LEDs are pulsed over a range of intensities thus allowing the mapping out of the PMTs' response curve once a month for each detector. The LEDs are also pulsed at one single intensity so as to provide a continuous gain calibration and drift corrections.

**Stopping Cosmic Ray Muons** - even though the mean energies of cosmic ray muons differ between the Near and Far detectors because of the depth of the detectors

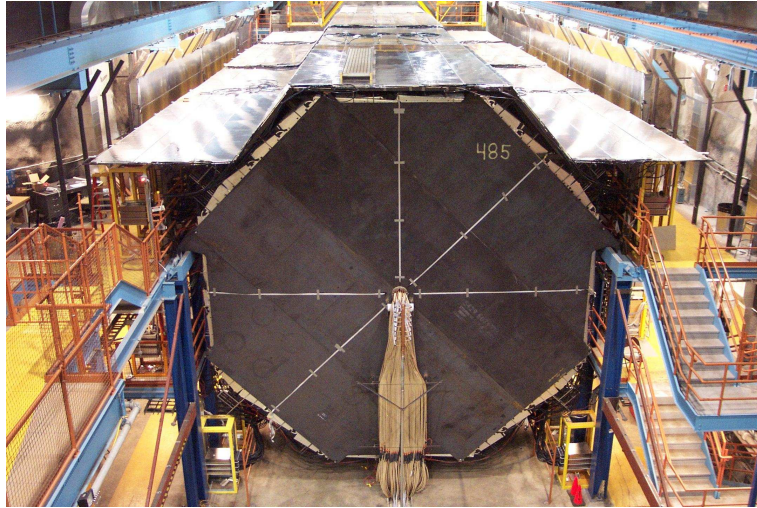


Figure 3.8: The MINOS Far detector. This front view of the detector is somewhat deceptive as it does not really convey the true size of it. In reality, the FD is 8m high and 30m long. The 'hat' on the detector is the veto shield which is used to help identify cosmic muon events.

in the earth, stopping cosmic ray muons can be used to provide an absolute energy scale calibration between the two detectors. In each detector, the portion of the muon track which corresponds to the energy between 0.5GeV and 1.1GeV is used to tabulate the detectors' response and to derive a single calibration constant for each detector. After this calibration step, the energy depositions in the scintillator strips in both detectors can be expressed in terms of Muon Energy Units (MEUs).

**Calibration Detector** - in order to achieve absolute calibration and relate the energies in MEUs to energies in GeV for different particles, a special Calibration Detector was constructed at the same time as the main MINOS detectors. The Calibration Detector (CalDet) was placed in a test beam at CERN between 2001 and 2003 and tested the response of the detector for electrons, muons, and hadrons of known momenta. The different particles could be identified via various cuts and a time-of-flight system [43, 44, 45, 46]. CalDet weighed 12 tons and consisted of 60 steel/scintillator planes that were of  $1\text{m} \times 1\text{m}$  size and had 24 scintillator strips each. This detector however was not magnetised unlike the Near and the Far detectors. Figure 3.9 shows a picture of this detector. In addition to providing an absolute energy scale for calibration, CalDet also allowed the testing and tuning of Monte Carlo (MC) simulations for low energy interactions (CalDet particle momenta were varied between 0.6GeV and around 10GeV).



Figure 3.9: The MINOS Calibration detector in a test beam at CERN.

### 3.4 MINOS Results to Date

The amount of exposure to neutrinos is measured in terms of the number of protons hitting the production target (POT). Since MINOS started taking data in 2005, it has accumulated approximately  $7 \times 10^{20}$  POT in forward horn current mode (neutrino mode) until the summer of 2009, when the third official run finished. The total exposure can be seen in Figure 3.10. The fourth official run has been underway since September 2009 in reverse horn current mode, which produces a beam of mostly muon anti-neutrinos, as explained earlier. In addition to the neutrino beam data, MINOS has also been taking cosmic data continuously while the experiment has been online. To date, several MINOS analyses have already carried out, and the main published results are summarised in the sections below.

#### 3.4.1 Muon Neutrino Disappearance

The main analysis goal of MINOS is the measurement of muon neutrino disappearance and the oscillation parameters associated with this measurement,  $\Delta m_{32}^2$  and  $\sin^2 2\theta_{23}$ . In order to select the final data sample, different cuts are applied to select the  $\nu_\mu$ -CC events, which are primarily recognised by the long muon tracks resulting from the CC interactions of muon neutrinos. This was a blind analysis which was carried out

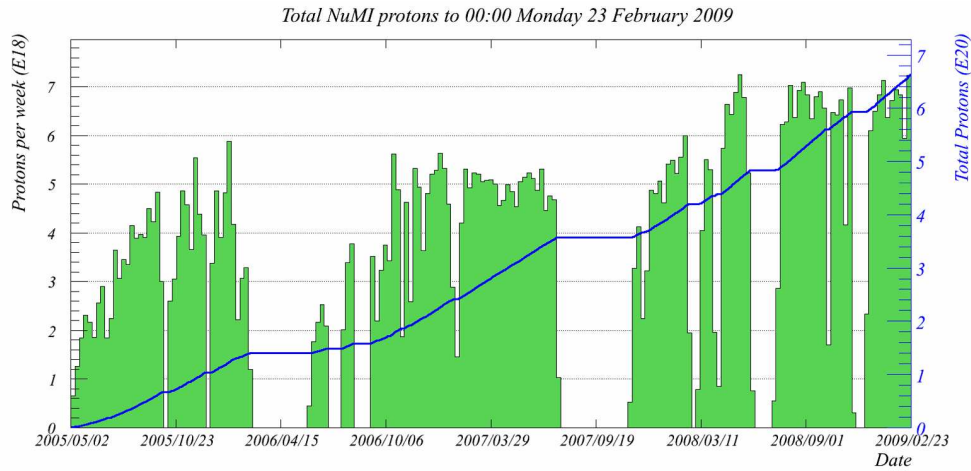


Figure 3.10: This Figure shows the MINOS exposure to neutrinos measured in terms of the number of protons hitting the production target (POT). The green histogram shows the weekly exposure in units of  $10^{18}$  POT. Three general run periods can be discerned in this Figure. They correspond to runs I, II, and III of the MINOS data (this Figure also contains some special data runs like for example in high energy mode at the beginning of Run II, but these special runs were short and are not pointed out on this graph). The blue line shows the accumulated total exposure in units of  $10^{20}$  POT.

for neutrinos only (that is the anti-neutrinos were cut out of the sample by applying a charge-sign cut based on the direction of the curvature of the muon track) and involves measuring the ND spectrum before oscillations have occurred, and then extrapolating this spectrum to the Far detector and comparing the prediction with no oscillations to the Far detector data set. The observed deficit of  $\nu_{\mu}$ -CC events is then interpreted as a neutrino oscillations signal and the oscillation parameters are extracted. Alternative neutrino disappearance theories like neutrino decay are also excluded. The most recent result [47] of this analysis, which superseded a first MINOS result published in 2006 [48], was for a total sample of  $3.36 \times 10^{20}$  POT and the following results were found:

$$\Delta m_{32}^2 = (2.43 \pm 0.13) \times 10^{-3} eV^2 \text{ at the 68\% confidence limit, and}$$

$$\sin^2 2\theta_{23} > 0.9 \text{ at the 90\% confidence limit.}$$

The value of  $\Delta m_{32}^2$  measured by MINOS currently represents the world's best measurement of this parameter. The best fit contour plot is shown in Figure 3.12 and the measured  $\nu_{\mu}$ -CC energy spectrum is shown in Figure 3.11. It can be seen that MINOS observes 848 events in the Far detector compared to a no-oscillations expectation of  $1065 \pm 60(\text{syst})$  events (those numbers include some higher energy beam data). The  $\nu_{\mu}$ -CC disappearance analysis is continuing and will soon be repeated with the new available data set of approximately  $7 \times 10^{20}$  POT exposure.

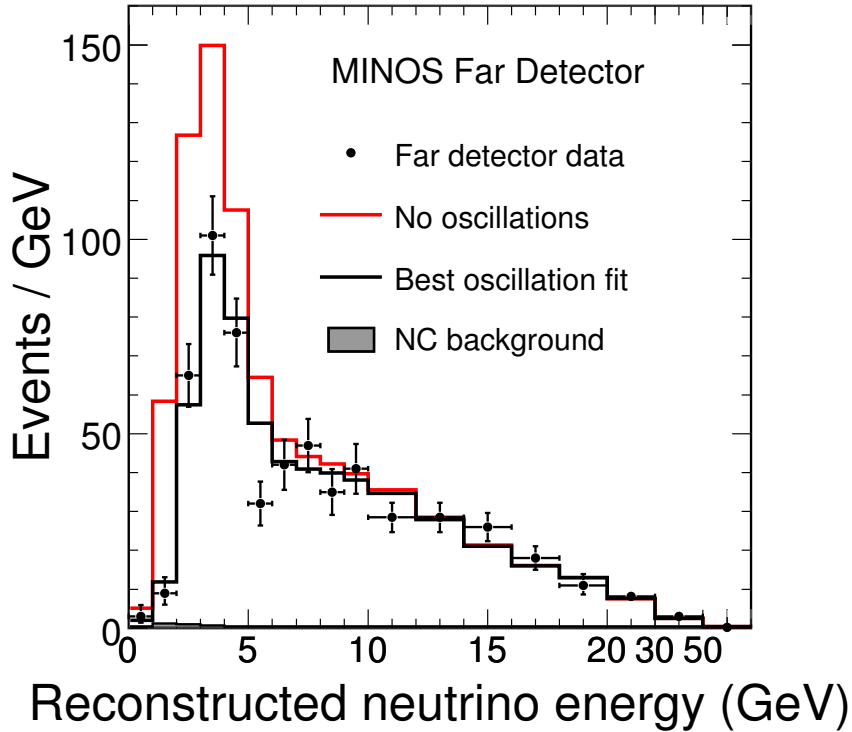


Figure 3.11: MINOS Far detector  $\nu_\mu$ -CC spectrum as a function of reconstructed neutrino energy. The dotted line is the expectation in the absence of neutrino oscillations. The solid line is the best oscillation fit, and the points are the data. A small neutral current background is expected in the data and is also shown as a shaded area (Figure can be found in [47]).

### 3.4.2 Active Neutrino Disappearance Search

A feature of neutrino oscillations is that even though  $\nu_\mu$  oscillate away into other flavours on their way to the FD, they should still interact in the detector via neutral current (NC) interactions. Therefore, any depletion in the expected FD NC spectrum would indicate the presence of at least one non-interacting or “sterile” neutrino in addition to the three active neutrino species. MINOS has carried out an analysis comparing the FD NC predicted spectrum with the selected data and found that for an exposure of  $3.18 \times 10^{20}$  POT no depletion was found in the number of neutral current like neutrino events [49, 50]. This can be seen in Figure 3.13 where the data is compared to two oscillation hypothesis - no  $\nu_e$  appearance and some  $\nu_e$  appearance and CP-violation.



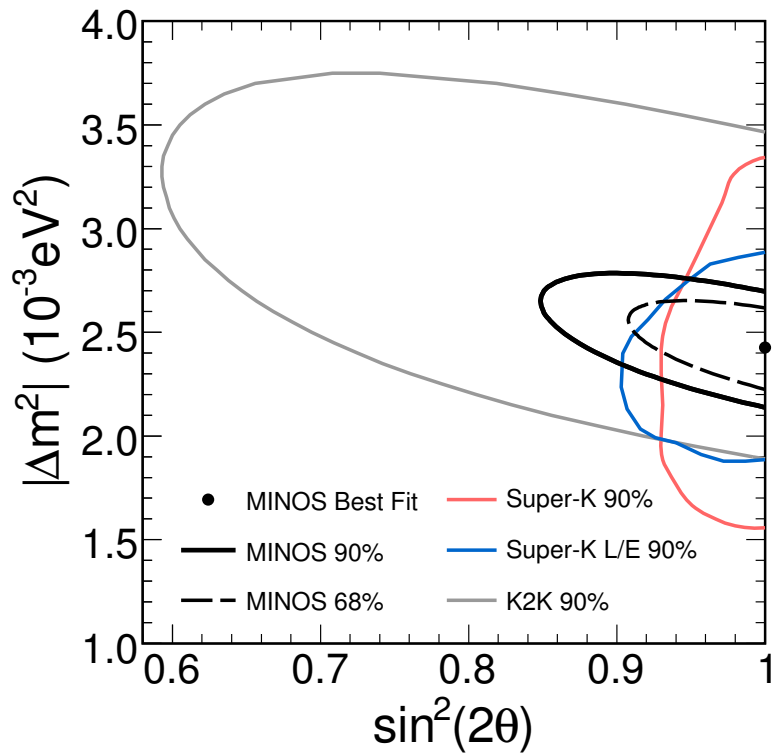


Figure 3.12: Contours in  $\Delta m_{32}^2$  and  $\sin^2 2\theta_{23}$  parameter space for the best neutrino oscillations fit to the  $\nu_\mu$ -CC data in the MINOS Far detector (Figure can be found in [47]).

The observed data was 388 events compared to a standard three-flavour neutrino model expectation of  $377 \pm 19.4(\text{stat.}) \pm 18.5(\text{syst.})$  events.

In addition to not observing depletion in the NC spectrum in the FD, MINOS has also set a limit on the fraction of active neutrinos oscillating away into  $\nu_{sterile}$  to be  $f_s < 0.52$  (0.55 for maximally allowed electron neutrino appearance) at a 90% confidence limit.

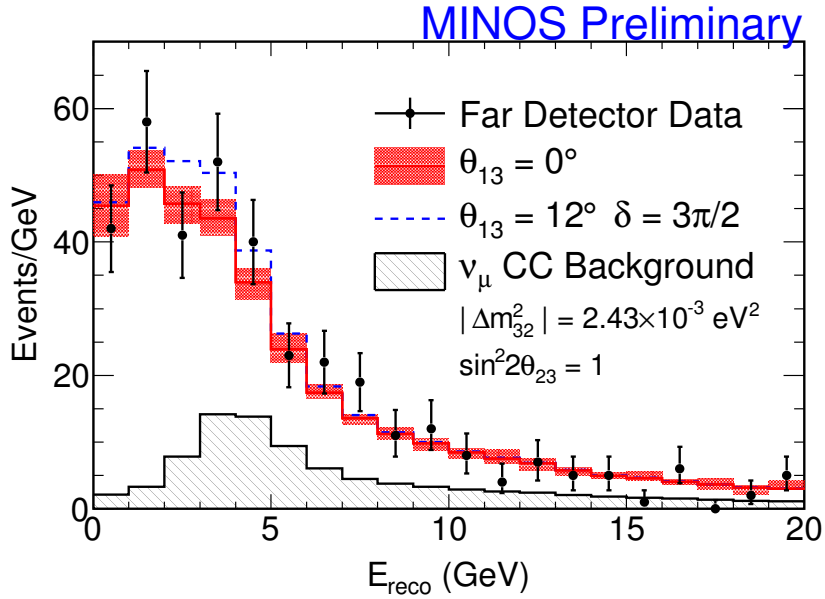


Figure 3.13: Far detector neutral current like event spectrum compared to predictions assuming no  $\nu_e$  appearance ( $\theta_{13} = 0$ ) and assuming some  $\nu_e$  appearance and CP-violation ( $\theta_{13} = 12^\circ$  and  $\delta = 3\pi/2$ ). The total exposure is  $3.18 \times 10^{20}$  POT. The data points are consistent with no  $\nu_\mu$  oscillating away into  $\nu_{sterile}$  neutrino species (Figure can be found in [50]).

## Chapter 4

# The Electron Neutrino Appearance Analysis

The MINOS  $\nu_e$ -appearance analysis attempts to measure / set a better limit on the  $\sin^2 2\theta_{13}$ , the oscillation term containing the neutrino mass mixing angle  $\theta_{13}$ , which has not been measured to date. This analysis is very difficult because the granularity of the MINOS detectors was optimized to detect muon tracks and is not very well suited to detecting electron energy deposits created by charged current  $\nu_e$  interactions, which would represent the electron neutrino appearance signal that MINOS is looking for. Figure 4.1 shows some typical events in the FD modeled by the Monte Carlo (MC) and one can easily see how difficult it can be to distinguish between signal  $\nu_e$ -CC events and other, background, events in MINOS. Even though an event may appear  $\nu_e$ -CC like, in reality, it is not possible to know for certain whether it is such an event. Apart from the difficulty in identifying signal  $\nu_e$ -CC events, there is also the added difficulty that most of the events specific to this analysis are in a kinematic region where events are very hard to model. This results in data/MC differences which need to be taken into account in order to predict the FD backgrounds correctly. Despite the above concerns, MC simulations show that it should be possible to carry out a MINOS  $\nu_e$ -appearance analysis by carefully selecting events so as to maximise the signal to background ratio.

## 4.1 General Method

The MINOS  $\nu_e$  appearance analysis is a blind analysis. This means that prior to looking at the FD data in the signal region, all procedures were finalised and the background predictions (what is expected in the absence of a signal) determined. Since MINOS has two detectors, it uses the ND, which has a high event rate, to measure the observed neutrino spectrum before oscillations have occurred. The observed ND spectrum consists purely of background, since there is no  $\nu_e$  appearance signal in the ND data. Unfortunately, the selected ND MC spectrum does not agree well with the ND data. The MC can thus not be relied upon to decompose the ND background into its individual components correctly. Because of this, two different data-driven methods have been developed to decompose the ND data spectrum. Those background spectra are then used in conjunction with Near/Far MC ratios to extrapolate them to the FD. During this procedure, the FD MC is oscillated<sup>1</sup> using the oscillation parameters measured in the main MINOS  $\nu_\mu$ -CC-Disappearance analysis (described in section 3.4.1). This allows the prediction of the spectrum in the FD in the absence of  $\nu_e$ -appearance and can then be compared to the FD selected data for the final result. A MC cartoon of what a signal could look like is shown in Figure 4.2. In order to make certain that there are no problems with the selection cuts or extrapolation procedure used, several FD sideband data sets - FD data not expected to be subject or insensitive to  $\nu_e$  appearance - are opened prior to unblinding the final selected FD data set.

## 4.2 $\nu_e$ -Appearance Analysis Backgrounds

Figure 4.1 gives an idea of the different event types seen in the MINOS detectors. In general, events are characterised by a track or a particle shower, or both. Since the ND is very close to the beam source, no neutrino oscillations have taken place yet so events selected for the  $\nu_e$  appearance analysis consist of backgrounds only. In the ND, there are thus three types of events:  $\nu_\mu$ -CC events, NC events, and intrinsic beam  $\nu_e$ -CC events. Beam  $\nu_e$ -CC events are a small component of the neutrino beam resulting mostly from the decay of secondary muons, especially in the low energy region which is of interest to the  $\nu_e$  appearance analysis. The following decays are responsible for the beam  $\nu_e$ -CC

---

<sup>1</sup>From a technical point of view, the MC is oscillated by using calculated oscillation weights at analysis time.

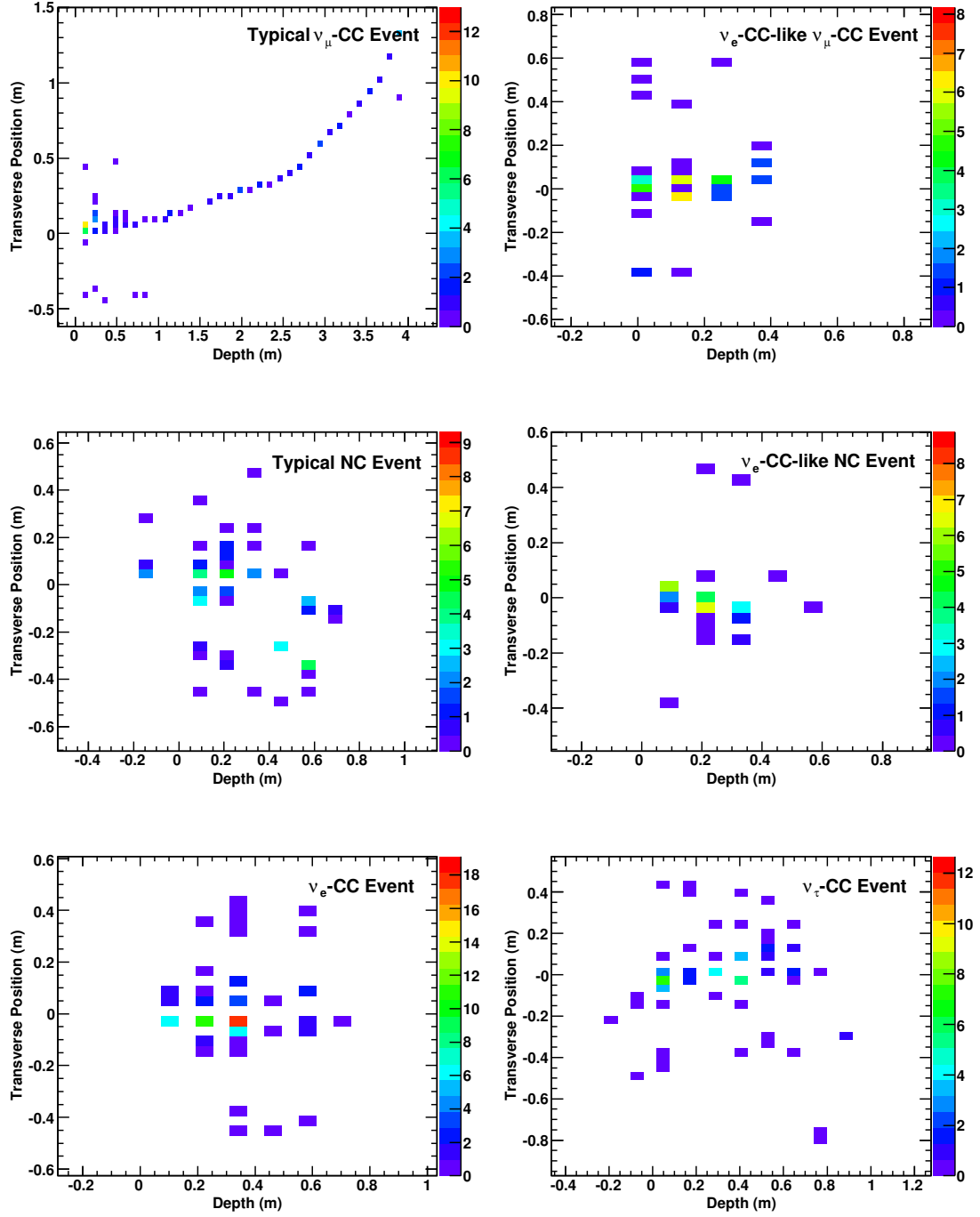


Figure 4.1: Example FD events modeled by the Monte Carlo. The top left plot shows a typical  $\nu_\mu$ -CC event with a long muon track and a hadronic shower deposition. The top right plot shows a  $\nu_\mu$ -CC event where most of the energy of the neutrino was transferred to the shower (high hadronic Y) and the muon track is very short, mimicking a  $\nu_e$ -CC signal event, an example of which is shown in the bottom left plot. The middle left plot shows a typical NC event which consists of a hadronic shower and possibly some short hadron tracks. The middle right plot shows a NC event that is  $\nu_e$ -CC-like. This is the largest background in the  $\nu_e$  appearance analysis. The bottom right plot is an example of a  $\nu_\tau$ -CC event which comes from  $\nu_\mu$  oscillating into  $\nu_\tau$  on their way to the FD; this type of event is a small background in this analysis.

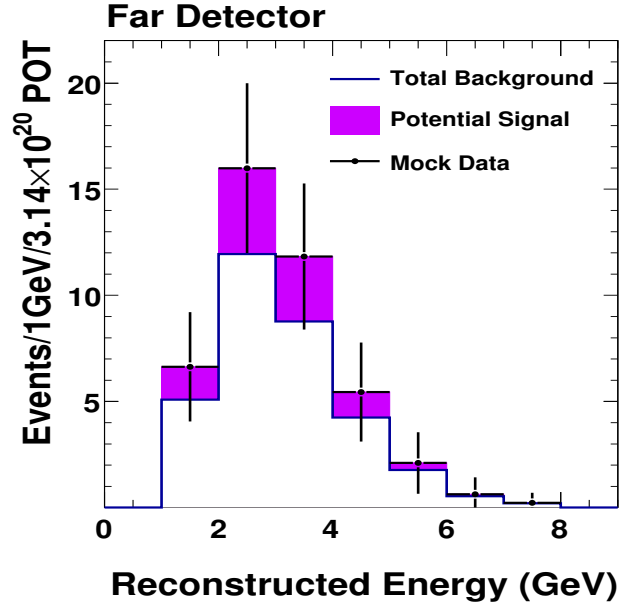


Figure 4.2: This Figure shows a Monte Carlo example of how the  $\nu_e$ -appearance signal could look for an exposure of  $3.14 \times 10^{20}$  POT. The mock data shows a signal oscillated with favourable oscillation parameters which were set at the Chooz limit ( $\sin^2 2\theta_{13} = 0.15$ ). Despite this, the signal is represented by only a relatively small excess above a large background. The latter mostly consists of misidentified neutral current events.

background overall:

$$\begin{aligned} \mu^+ &\rightarrow e^+ + \nu_e + \bar{\nu}_\mu && (\text{BR of } \sim 100\%) \\ K^+ &\rightarrow \pi^0 + e^+ + \nu_e && (\text{BR of } 5.07 \pm 0.04\%) \\ K_L &\rightarrow \pi^\mp + e^\pm + \nu_e(\bar{\nu}_e) && (\text{BR of } 40.55 \pm 0.12\%) \end{aligned}$$

The beam  $\nu_e$ -CC breakdown by decay parent can be seen in Figure 4.3.

On their way to the FD, the neutrinos oscillate. Different backgrounds are affected differently. The NC events do not oscillate, or strictly speaking, since all neutrino types interact via the NC interaction, the neutrino flavour is irrelevant and there will be no depletion or increase of events due to oscillations. The intrinsic beam  $\nu_e$ -CC background is also still present in the FD and any oscillations of it are too small to be visible. The  $\nu_\mu$ -CC events do oscillate on their way to the FD and this is taken into account. Finally, a new background appears in the FD arising from  $\nu_\mu$  oscillations into  $\nu_\tau$ . Those neutrinos represent a small  $\nu_\tau$ -CC background for the  $\nu_e$ -appearance analysis. Signal events, if they are present, are a result of  $\nu_\mu$  oscillations into  $\nu_e$  and are seen as  $\nu_e$ -CC in the FD.

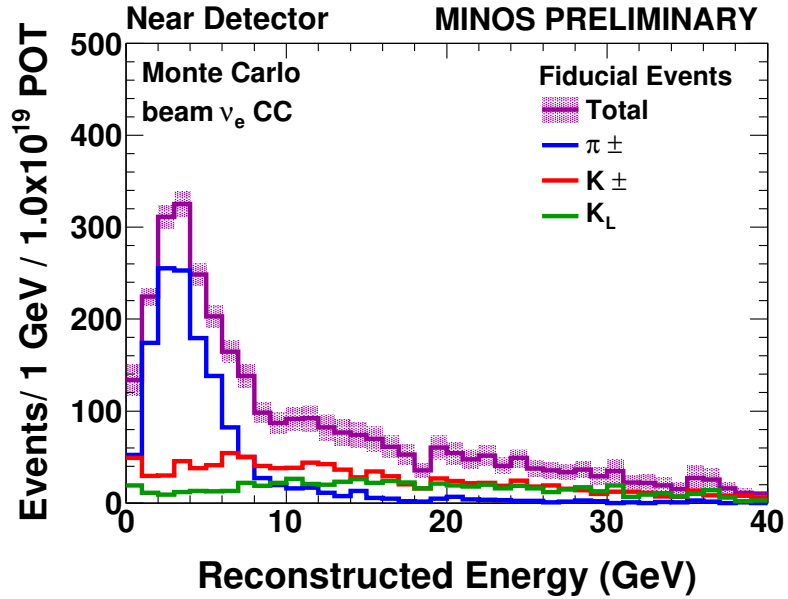


Figure 4.3: This ND Monte Carlo Figure shows the parent hadrons at the NuMI target of the intrinsic beam  $\nu_e$ -CC background as a function of reconstructed energy. The beam  $\nu_e$ -CC events are shown at fiducial volume cut level.

### 4.3 Monte Carlo and Reconstruction

The MINOS detectors and the beamline are simulated using GEANT3 [51]. To simulate the hadron production yields from the carbon target, FLUKA [52] is used in this analysis. Finally, NEUGEN3 [53] is used in order to simulate neutrino interactions and the re-interactions of the produced hadrons within the nuclei. The MINOS Cedar-phy version of the reconstruction is used to reconstruct the MC. The MINOS data is reconstructed using the MINOS reconstruction software version called Cedar-phy-bhcurve. This version includes an updated detailed mapping of the magnetic field in the detectors. It was not possible to use the new magnetic field mapping for the MC at the time due to time constraints on the analysis, however, it was shown that the new maps had a negligible effect on the  $\nu_e$ -appearance analysis selected events [54].

## 4.4 Monte Carlo Flux Weights

It became apparent in the early stages of the MINOS  $\nu_\mu$ -CC disappearance analysis that the ND MC  $\nu_\mu$ -CC spectrum does not agree exactly with the data. It was determined by comparing the data/MC spectra in different beam configurations that the main source of this discrepancy could be attributed to uncertainties in the neutrino flux at both detectors. Those in turn were caused by uncertainties in the yield of hadrons of the production target as a function of transverse and longitudinal (along the beam line) momentum. A method was developed [55] to constrain the beam flux calculations using data from different beam configurations. This tuning allowed the calculation of event-by-event weights for the MC which corrected the latter to agree with the data much better, as is shown in Figure 4.4. Since those weights are used to correct flux uncertainties, they were also incorporated into (and used by) the  $\nu_e$ -appearance analysis framework.

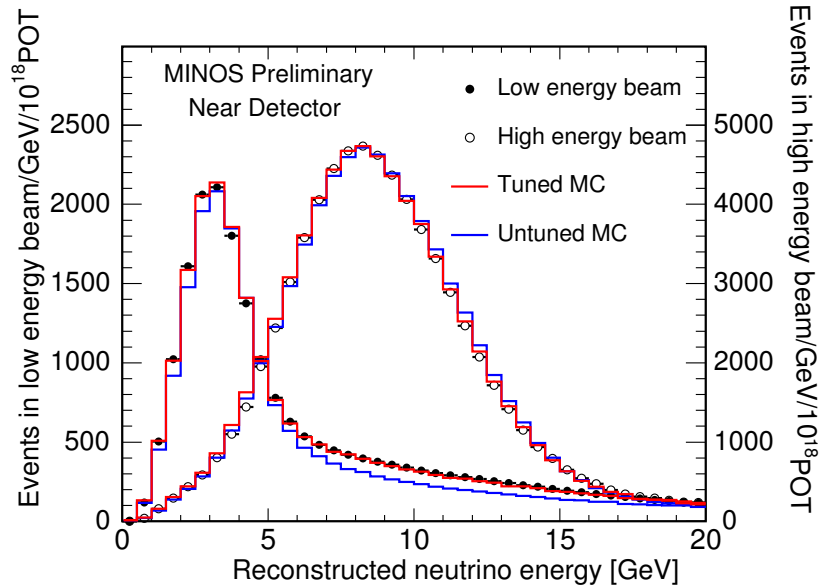


Figure 4.4: This Figure shows the  $\nu_\mu$ -CC spectra in the MINOS ND. Two beam configurations are used. The low energy beam is the standard configuration for MINOS data and was taken with the target set to a position of 10cm and the horn current set to 185kA. The high energy beam configuration was a special run during Run II of the MINOS data taking and was taken with a maximal target position of 250cm and a horn current of 200kA.. The points represent the data - black for low energy and white for high energy. The blue lines are MC before flux tuning, and red lines are MC after tuning. The tuning of the MC improves the data/MC agreement for both beam configurations.



## 4.5 Analysis Selection Cuts

Since the signal from  $\nu_e$ -appearance is expected to be very small, a number of selection cuts need to be applied in order to select as clean a signal as possible. This is achieved by a combination of cuts which include data quality, fiducial, preselection, and particle identification cuts [56].

### 4.5.1 Data Quality Cuts

There are several categories of data quality cuts and some of them differ for the two detectors. First a set of cuts is applied that removes data from consideration altogether which means the removed data is not counted in the total exposure. A cut requiring good beam is applied for both detectors. Then other cuts are applied which require the magnetic coils to be working correctly and in forward mode (the magnetic fields can also be reversed) and the current powering the coils to be within certain ranges for the detectors. For the FD, data taken when the GPS timing system was not working is excluded from the analysis.

A second set of data quality cuts does not affect the total exposure used but removes events that do not come from the neutrino beam. A cut which requires that an event not be a light injection (LI) event and be within a  $10\mu\text{s}$  window of the beam spill is applied. Cuts are also applied to reduce potential contamination from cosmic events. Those include a cut on the veto counter and cuts on other parameters like the shower and track angles [57, 58].

Finally, since the event rate in the FD from beam neutrinos is very low ( $\sim 2\text{-}3$  good neutrino events/day), a cut is applied to only accept the largest event in each spill. This removes small showers that could be artifacts of the reconstruction or just spurious hits.

### 4.5.2 Analysis Preselection Cuts

The first analysis cut that is applied is the fiducial volume cut. This cut is needed because it removes incomplete events at the edges of the detectors and events resulting from neutrinos interacting in the rock around the detectors (so-called rock muon events). Moreover, the reconstruction may not be working optimally at the edges of the detectors and the Monte Carlo may not be modeling effects in those detector areas very well.

Due to their different shapes and sizes, the fiducial volume cut is different for the two detectors. In addition to this, this cut is slightly different for the FD data and MC because of steel plane alignment. The fiducial volume cuts are as follows:

- ND volume -  $1.01080\text{m} < Z < 4.99059\text{m}$  and  $R < 0.8\text{m}$ , where  $R = \sqrt{X^2 + Y^2}$ ,  $X = X_{vertex} - 1.4885\text{m}$ , and  $Y = Y_{vertex} - 0.1397\text{m}$  (those two offsets are because in the ND, the magnetic coil is not within the fiducial volume). This fiducial volume does not follow the beam so it doesn't take into account the slight angle of the neutrino beam with respect to the ND.
- FD volume for MC -  $0.47692\text{m} < Z < 14.27860\text{m}$  and  $16.26470\text{m} < Z < 27.97240\text{m}$  (this is because of the gap between the two super-modules, and  $0.5\text{m} < R < \sqrt{14}\text{m}$ , where  $R = \sqrt{X_{vertex}^2 + Y_{vertex}^2}$  (no offsets because the FD is symmetrical around the magnetic coil).
- FD volume for data -  $0.49080\text{m} < Z < 14.29300\text{m}$  and  $16.27110\text{m} < Z < 27.98270\text{m}$ ,  $0.5\text{m} < R < \sqrt{14}\text{m}$ .

After the fiducial volume cut, preselection cuts are applied so as to select events that are  $\nu_e$ -CC-like and remove obvious backgrounds. Those cuts were optimised to cut out as much background as possible with minimal effect on the selected signal and their effects are shown in Figures 4.5 and 4.6. The preselection cuts are:

- if there is a track, its length needs to be less than 25 planes, defined as the absolute value  $|\text{Plane the track begins} - \text{plane the track ends}| < 25$  planes (Figure 4.5). This cut assures that any obvious  $\nu_\mu$ -CC events are removed since the latter are characterised by long muon tracks. In essence, this cut selects events that are more neutral current like.
- the number of track-like planes needs to be less than 16 (Figure 4.5). For a given track, the variable track-like planes is the number of planes that do not have non-track associated hits, like shower hits. This variable is hence the number of planes a track extends beyond the shower in the event. This cut is applied to eliminate CC- $\nu_\mu$  events, for which the track extends 15 planes or more out of the hadronic shower.
- the event has to have at least one shower (Figure 4.6). The vast majority of signal events have a shower, therefore it makes sense to cut out events that do not have one and are a priori very unlikely to be signal events.

- the number of contiguous event planes with at least a 1MeV scintillator strip deposition needs to be above 4 (Figure 4.6). This contiguous plane cut is used to select  $\nu_e$ -like events with a well-defined shower core<sup>2</sup>.
- the reconstructed energy should be more than 1GeV (Figure 4.6). This energy cut removes mostly small neutral current events.
- the reconstructed energy should be less than 8GeV (Figure 4.6). Hardly any signal is expected at energies higher than 8GeV, thus this cut serves to optimize the selection.

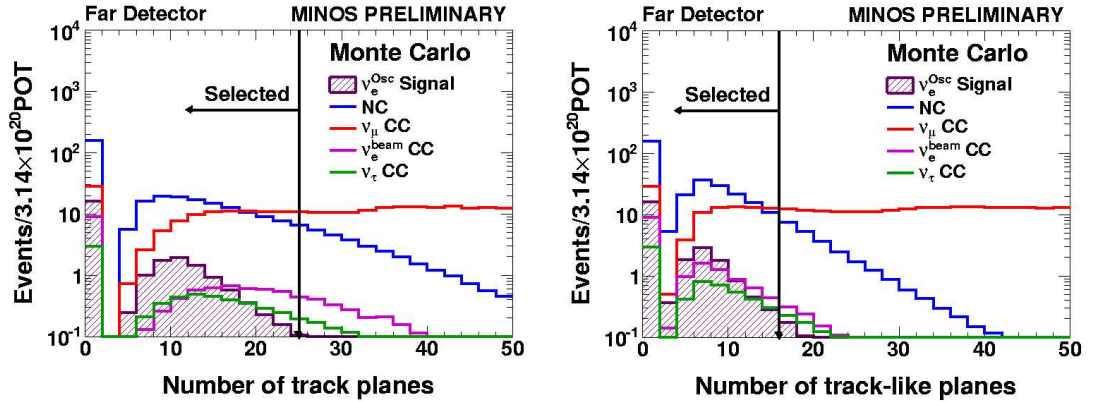


Figure 4.5: This Figure is made using FD Monte Carlo events after the application of the fiducial volume cut. The two plots show the effect of the track planes selection cut on the left, and of the track-like planes selection cut on the right. The signal  $\nu_e$ -CC events (oscillated at the Chooz limit) are shown in the shaded purple histograms. The blue lines show the neutral current background events, the red lines show the background  $\nu_\mu$ -CC events, the  $\nu_e$ -CC beam events (already present in the ND) are shown in pink, and the  $\nu_\tau$ -CC appearance background events are plotted in green. The respective selected regions for each cut are indicated by the vertical black lines and arrows.

## 4.6 Cross-Talk

During early stages of the analysis, it was found that the selection algorithms used at the time were sensitive to cross-talk hits, i.e. hits that arise for example from partial leakage of light into different PMT pixels. In addition to this, it was found that cross-talk hits below 2 photo-electrons (PE) were poorly modeled by the MC and there was a

<sup>2</sup>The minimum energy deposition is required to be 1MeV in scintillator strips for 4 consecutive planes, but the actual energy of a scintillator hit, when translated into shower energy by taking into account the energy which was deposited in the steel plane, is in the range of  $\sim 20$ MeV. Thus the contiguous plane cut is also effectively a cut on the reconstructed energy of the event.

large data/MC discrepancy for those hits. This can be seen in Figure 4.7 for the ND and in Figure 4.8 for the FD. As a consequence, a two-step approach was adopted. First, the cross-talk model was somewhat improved for both detectors to better agree with the data [59]. Second, a 2PE cut was adopted at the  $\nu_e$  analysis level. To clarify the latter it should be noted that when the events were reconstructed, all shower hits were used, so the hits below 2PE influenced the formation of showers. However, in order to calculate important shower topological variables for the purposes of the  $\nu_e$ -appearance analysis, the hits below 2PE were cut thus not directly feeding into the values of those variables<sup>3</sup>. As a result of removing the hits below 2PE, the analysis became less sensitive to cross-talk and poor modeling of those low hits.

## 4.7 $\nu_e$ Appearance Analysis Selection Algorithms

Due to the difficulty associated with identifying signal events, the  $\nu_e$  appearance analysis requires a careful selection of events with a high background rejection. The preselection cuts are good at rejecting the more obvious backgrounds, but at preselection level, it would be impossible to discern any potential signal as the signal to background ratio at this level is approximately 1 to 12 for  $\sim 230$  events in the final FD sample with an exposure of  $3.14 \times 10^{20}$  POT. Even at the Chooz limit, the signal events are so few, literally like a needle in a haystack, that simple cuts would never suffice to achieve adequate signal-background rejection. In order to provide the necessary level of background rejection, more sophisticated selection algorithms were developed and evaluated. Two algorithms were then chosen for the  $\nu_e$  appearance analysis to provide a cross-check for each other. A decision as to which of the two algorithms should be the primary was deferred until the opening of the analysis sidebands. In order to evaluate the performance of the selection algorithms, FD MC was used and oscillated with signal at the Chooz limit ( $\sin^2(2\theta_{13}) = 0.15$ ). A figure of merit (FOM) was calculated for each algorithm as a function of cut. It was found to be useful to define the FOM to include a parameter for the systematic error [60]:

$$FOM = \frac{signal}{\sqrt{\sigma_{stat}^2 + \sigma_{syst}^2}} \quad (4.1)$$

---

<sup>3</sup>As a consequence of the cross-talk studies, it was decided that for future MINOS analyses, a 2PE cut would be implemented at reconstruction level rather than just at analysis level.

where  $\sigma_{stat}$  is the statistical error, and  $\sigma_{syst}$  is the systematic error on the background. For this analysis, during the evaluation of the FOM and the best place to cut, a systematic error value of 10% was generally used as standard for comparison purposes since the systematic error in the FD was expected to be 10% or less (due to prior studies). The equation thus became:

$$FOM = \frac{signal}{\sqrt{background + 0.1^2 background^2}} \quad (4.2)$$

#### 4.7.1 A Neural Net Particle Identification Algorithm

The first particle identification algorithm (PID) used in this analysis is a neural net and is henceforth referred to as the ANN PID [61]. This neural network consists of 11 variables which are calculated on the whole event after hits below 2PE have been removed. A neural net is a decision making computational construct which takes in a number of variables and outputs a value which is indicative of whether an event is likely to be signal or background. Between the input layer and the final output, there are one or more hidden layers which link the input variables together and allow weights to be assigned to the different input variables as they are propagated through the nodes in the hidden layer(s). In order to assign the weights, the network has to be trained on a sample of signal and background MC for some time (this is called the number of epochs). The neural net is thus effectively learning and honing in on what combination of variable values means that an event is signal, and what combination means that it is background. The structure of the neural network used in this analysis was 11:6:6:1, that is 11 input variables, 2 hidden layers of 6 nodes each, and one output value, between -0.5 and 1.5 (the vast majority of values lie between 0 and 1), where a higher value signifies a higher likelihood of the event being a signal  $\nu_e$  event. Several of the variables included in the ANN PID rely on the mean longitudinal profile of the energy deposition in an electromagnetic cascade of a high energy electron or photon, which is approximately described by the following gamma distribution [31]:

$$\frac{dE}{dt} = E_0 b \frac{(bt)^{a-1} e^{-bt}}{\Gamma(a)}, \quad (4.3)$$

where  $E_0$ ,  $a$ , and  $b$  are parameters and the maximum occurs at:

$$t_{max} = \frac{a-1}{b} \quad (4.4)$$

The parameter  $b$  is directly related to the shower length  $X_s = X_0/b$ , where  $X_0$  is the radiation length (one steel plane in MINOS corresponds to approximately 1.4 radiation lengths, which is why the resolution of the detectors is not very good for EM showers).

When a neutrino interacts in the MINOS detectors and creates a shower, the longitudinal shower profile can be fit to Equation 4.3 and the parameters  $a$  and  $b$  can be extracted from the fit. An example of such a fit to the longitudinal shower profile of a good FD  $\nu_e$ -CC event is shown in Figure 4.9. Those extracted parameters are used with other variables in the ANN PID in order to separate electromagnetic-like showers from other hadronic showers and to hone in onto signal  $\nu_e$ -CC events. As a general rule,  $\nu_e$ -CC showers are more compact and forward than NC showers.

All 11 variables included in the neural network are calculated on the whole event, that is treating the event as one big shower (in practice, during the reconstruction process, some events are identified as having several (sub-)showers, one of which is identified as the primary shower). The ANN variables are shown in Figures 4.10 and 4.11 and are:

- shower fit parameter  $a$  - this comes from the fit to Equation 4.3 and constitutes the shower rise parameter in the longitudinal distribution of the shower, as can be seen in Figure 4.9
- shower fit parameter  $b$  - this also comes from the fit to Equation 4.3 and constitutes the shower fall parameter in the longitudinal distribution of the shower - see Figure 4.9
- minimal spanning tree summed weight - the minimum distances that join hits with pulse height bigger than the average for the event are calculated. The sum of those minimum distances is the total variable used in the ANN.
- the longitudinal projection of the event energy along the  $z$ -axis. More precisely, each hit forms a vector with the vertex. Those vectors projected onto the  $z$ -axis and multiplied by the hit pulse height. The total variable is the sum of those projections.
- UV RMS - this is the lateral shower spread, i.e. the RMS of the transverse energy deposition profile

- Moliere radius - this is the radius in number of strips that contains 90% of the event energy
- the fraction of the event energy deposited in a narrow, 3-strip wide road. The shower axis is found and all the hits that are within 1.5 strips of the shower axis are summed and the fraction of this sum to the total event pulse height is taken. This is a very powerful variable since EM showers usually tend to be narrower than hadronic showers.
- the fraction of the event energy in the 8 biggest hit strips divided by the total event energy
- the maximum fraction of the event energy in a n-plane sliding window where n=4,6,8 planes

The signal/background separation of the ANN PID is shown in Figure 4.12. The ideal cut was determined using Equation 4.2 and was placed at 0.7, that is in order to maximise the FOM, only events above this cut are selected. Using default oscillated FD MC at the Chooz limit of  $\sin^2 2\theta_{13} = 0.15$ , the FOM for the ANN PID is 1.57.

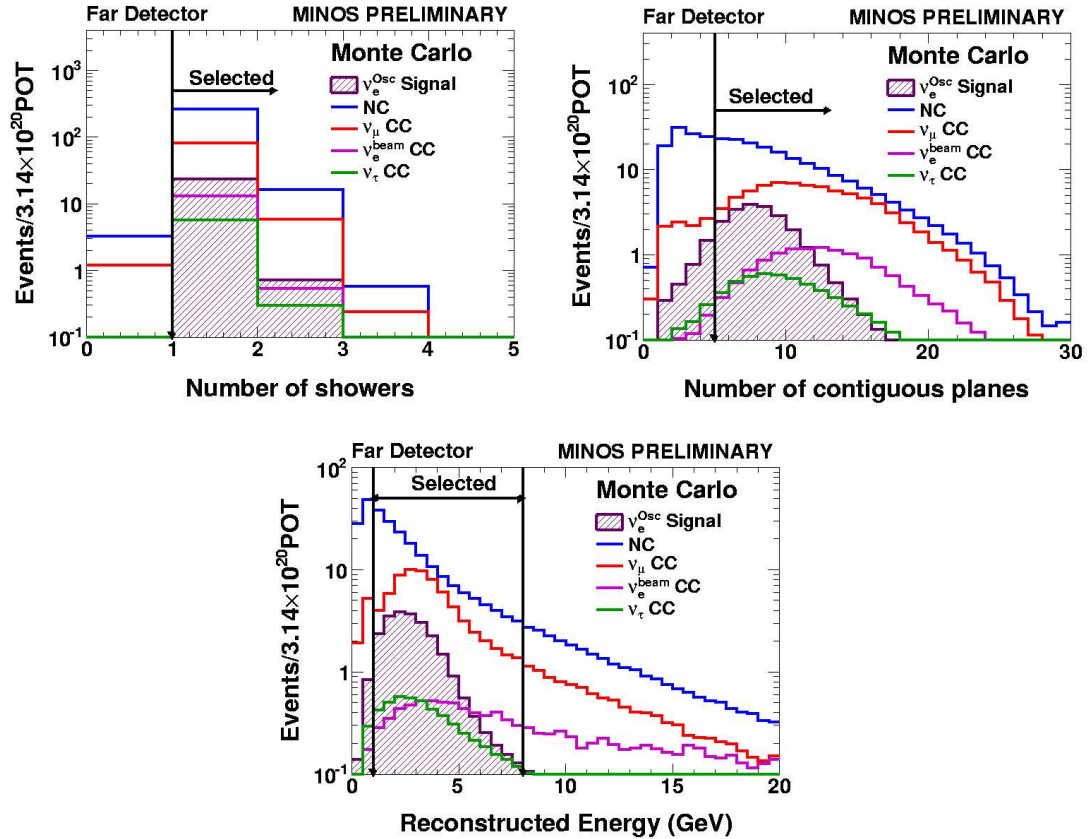


Figure 4.6: This Figure is made using FD Monte Carlo events after the application of the fiducial volume and both track selection cuts. The top left plot shows the effect of the number of showers selection cut, the top right plot shows the effect of the contiguous planes selection cut, and the bottom plot shows the effect of the reconstructed energy selection cuts. The signal  $\nu_e$ -CC events (oscillated at the Chooz limit) are shown in the shaded purple histograms. The blue lines show the neutral current background events, the red lines show the background  $\nu_\mu$ -CC events, the  $\nu_e$ -CC beam events (already present in the ND) are shown in pink, and the  $\nu_\tau$ -CC appearance background events are plotted in green. The respective selected regions for each cut are indicated by the vertical black lines and arrows.



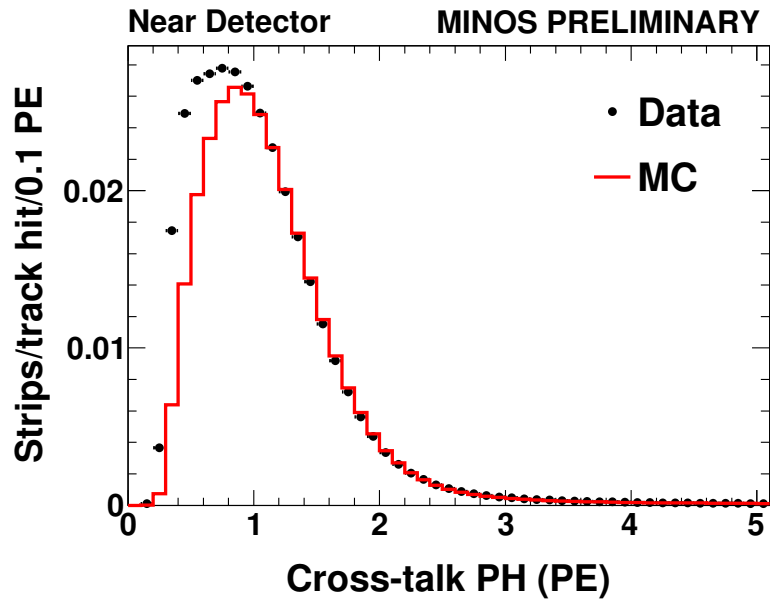


Figure 4.7: A comparison of data and MC for low PE cross-talk hits in the ND. It can be seen that below approximately 1PE, there is significant disagreement between the two samples.

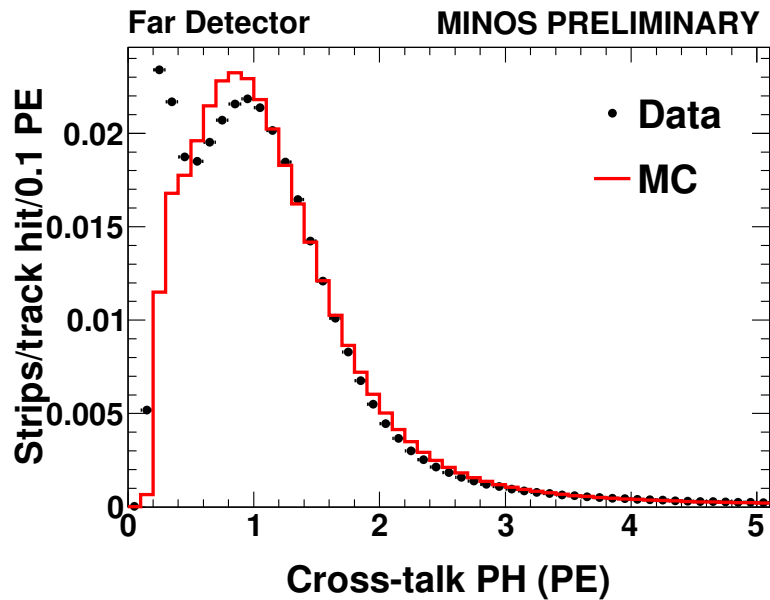


Figure 4.8: A comparison of data and MC for low PE cross-talk hits in the FD. It can be seen that below approximately 1PE, there is significant disagreement between the two samples.

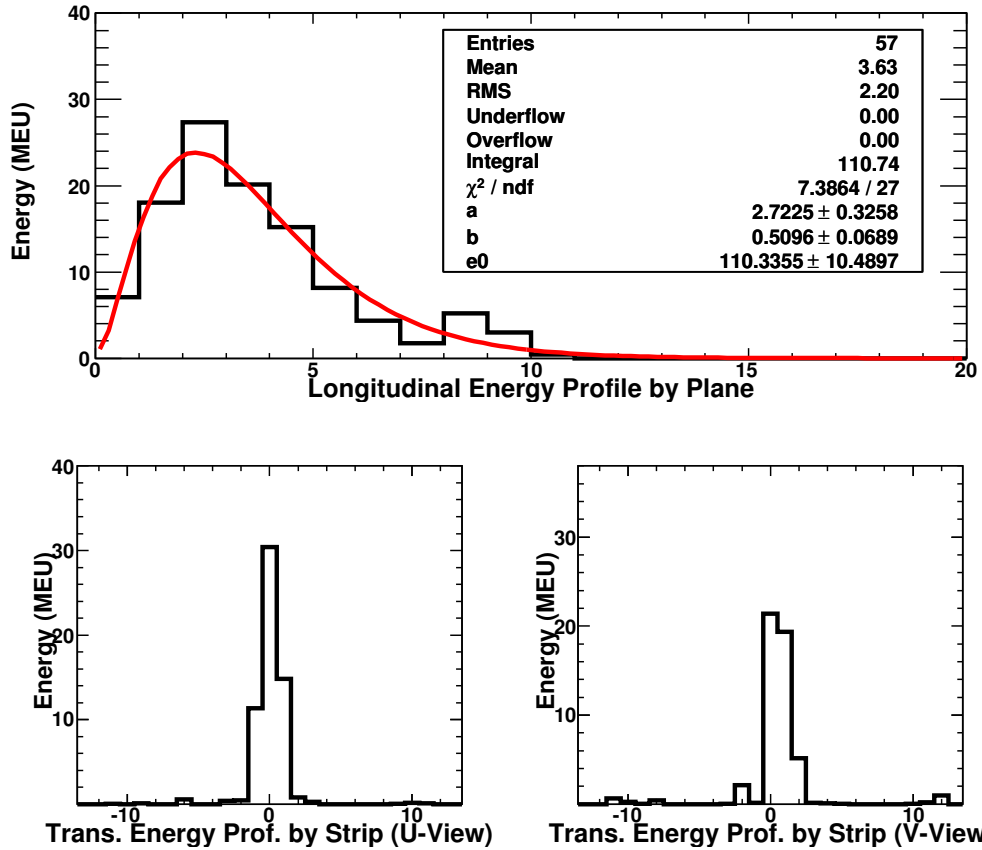


Figure 4.9: This Figure shows the longitudinal and transverse energy profiles of a FD  $\nu_e$ -CC event. The top plot is of particular interest and shows a fit of Equation 4.3 to the longitudinal energy profile and the extracted parameters a, b, and  $E_0$ . Parameter a is the shower rise parameter (upward slope of fit) and parameter b is the shower fall parameter (downward slope of fit). Those two parameters are used in the ANN PID.

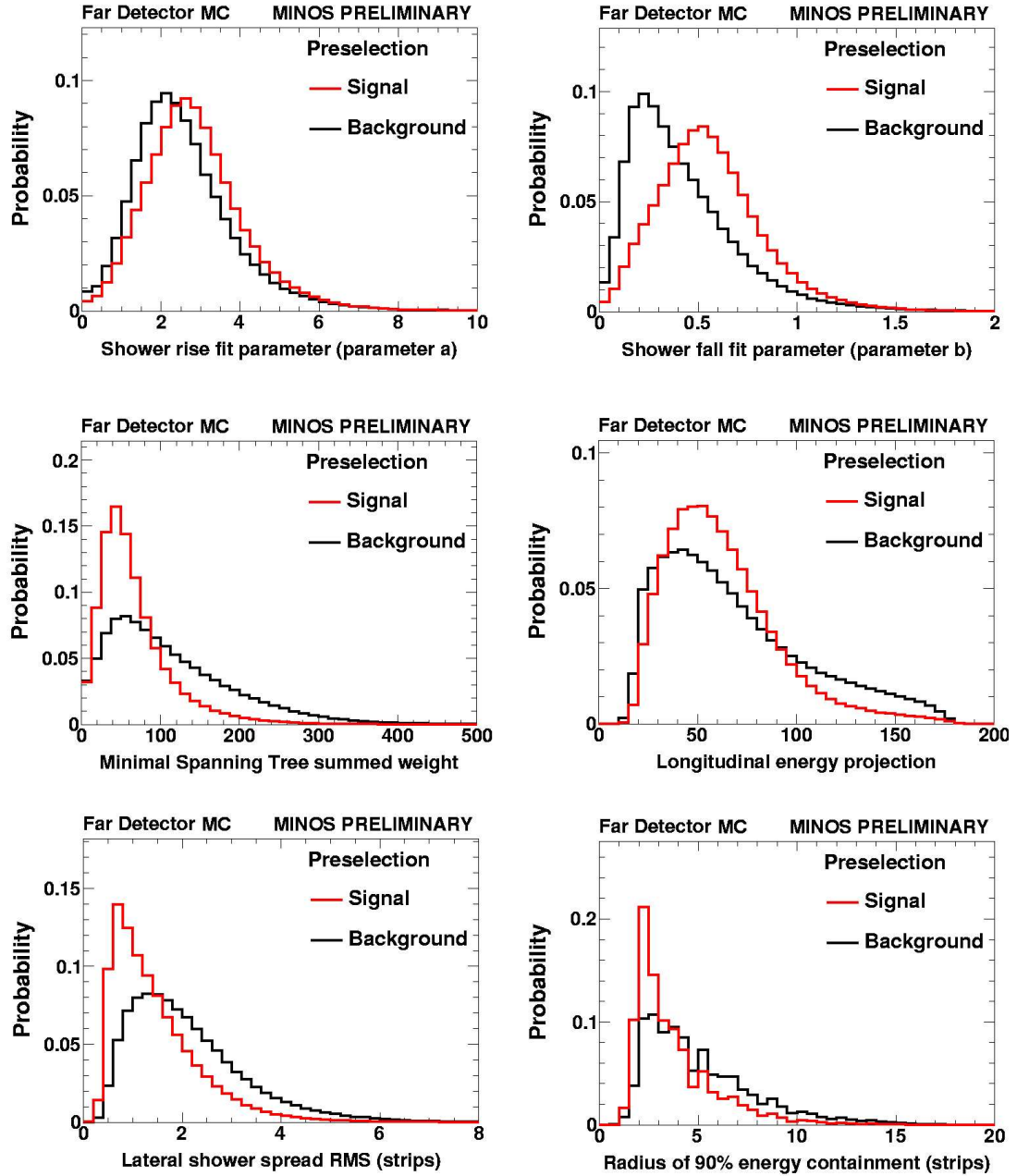


Figure 4.10: This Figure shows signal/background separation for the first 6 of the 11 variables used in the ANN PID. The full preselection has been applied. The signal is shown in red, and the background is shown in black. The histograms are area normalised to compare shape differences.

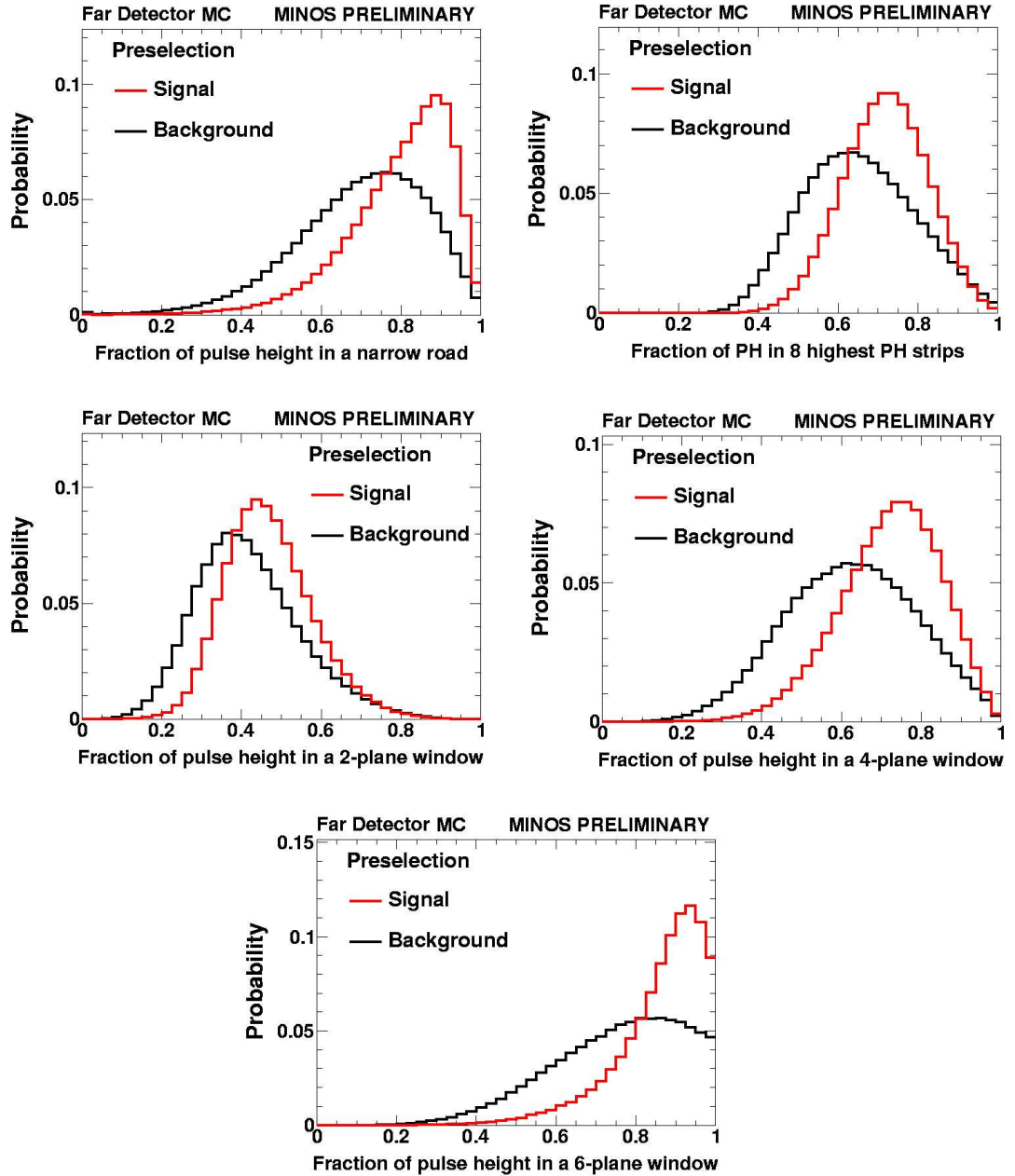


Figure 4.11: This Figure shows signal/background separation for the last 5 of the 11 variables used in the ANN PID. The full preselection has been applied. The signal is shown in red, and the background is shown in black. The histograms are area normalised to compare shape differences.

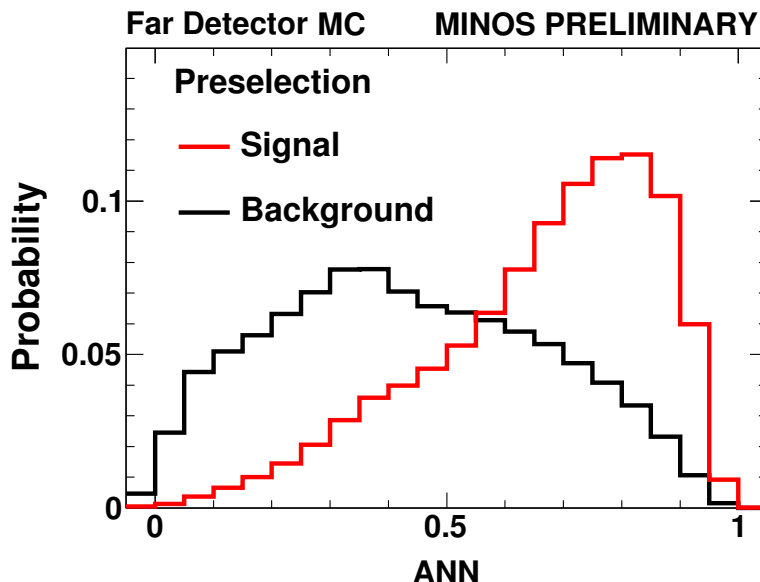


Figure 4.12: This Figure shows the signal/background separation of the ANN PID after the full preselection has been applied. The signal is shown in red, and the background is shown in black. In order to select signal events, a cut is placed at a PID value of 0.7 and only events above this cut are chosen for the final sample. The histograms are area normalised to compare shape differences.

#### 4.7.2 A Library Event Matching (LEM) Particle Identification Algorithm

The second particle identification algorithm used in the  $\nu_e$ -appearance analysis is the LEM PID [62], which constitutes, in principle, an ideal approach to selecting the signal from a large background. For this PID, large libraries of FD MC  $\nu_e$ -CC (10 million) and NC (20 million) events are constructed. Those library events are simplified by grouping hits away from the shower core into single hits and thus patterns are built up for signal and background events. In essence, the LEM PID is a MC nearest neighbour matching technique and for each candidate signal event, the 50 best matches are selected (this number was optimised to achieve the best results) and discriminants are constructed from the properties of those 50 best matches. An example of an event with a good and a bad library match is shown in Figure 4.13. Because the PID needs to be used in both the Near and the Far detectors, but the libraries were only available for FD MC (due to the very long time it takes to produce such libraries, in general more than half a year), events from both detectors needed to be set on an equal footing. To this end, when

the matching is carried out, each hit is corrected for attenuation and both FD and ND light levels are scaled to the levels of FD detector edge events. This is necessary because of Near/Far differences like scintillator strip and hence wavelength-shifting fibre length. Finally, hits below a charge of 3PE are not considered in the library matching. The three variables that are included in the final LEM PID are shown in Figure 4.14. Those variables are:

- $f_{50}$  - the fraction of the best 50 matches that are  $\nu_e$ -CC with  $y < 0.9$
- $y_{50}$  - the mean  $y$  of those best 50  $\nu_e$ -CC matches with  $y < 0.9$
- $q_{50}$  - the matched mean fractional charge of the best 50  $\nu_e$ -CC matches with  $y < 0.9$

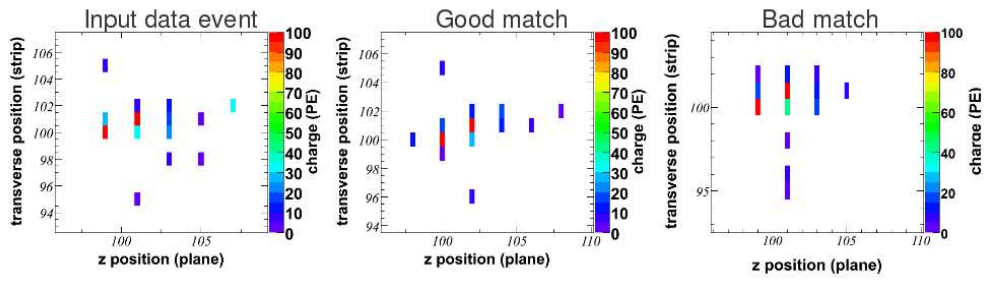


Figure 4.13: This Figure shows an example event (left plot) that is compared with a good (middle plot) and a bad (right plot) library match using the event hit pattern.

Those three variables are then combined in a simple likelihood which represents the final LEM PID. The signal/background separation of this PID is shown in Figure 4.15 and the optimal cut was determined to be at 0.8 where only events with a LEM PID above 0.8 are selected for the final sample. Further to opening the FD sidebands however, the optimal LEM cut had to be relaxed and was set at a new value of 0.65, as will be described in Chapter 7. Because of this, all plots and numbers in this document use the LEM cut value of 0.65, unless otherwise indicated. Using default oscillated FD MC at the Chooz limit of  $\sin^2 2\theta_{13} = 0.15$ , the FOM for the LEM PID with a cut at 0.65 is 1.67.

## 4.8 ND Data/Monte Carlo Spectra and Differences

The events selected by the  $\nu_e$ -appearance analysis depend on the correct modeling of the neutrino event showers. The MINOS MC shower models are tuned using external

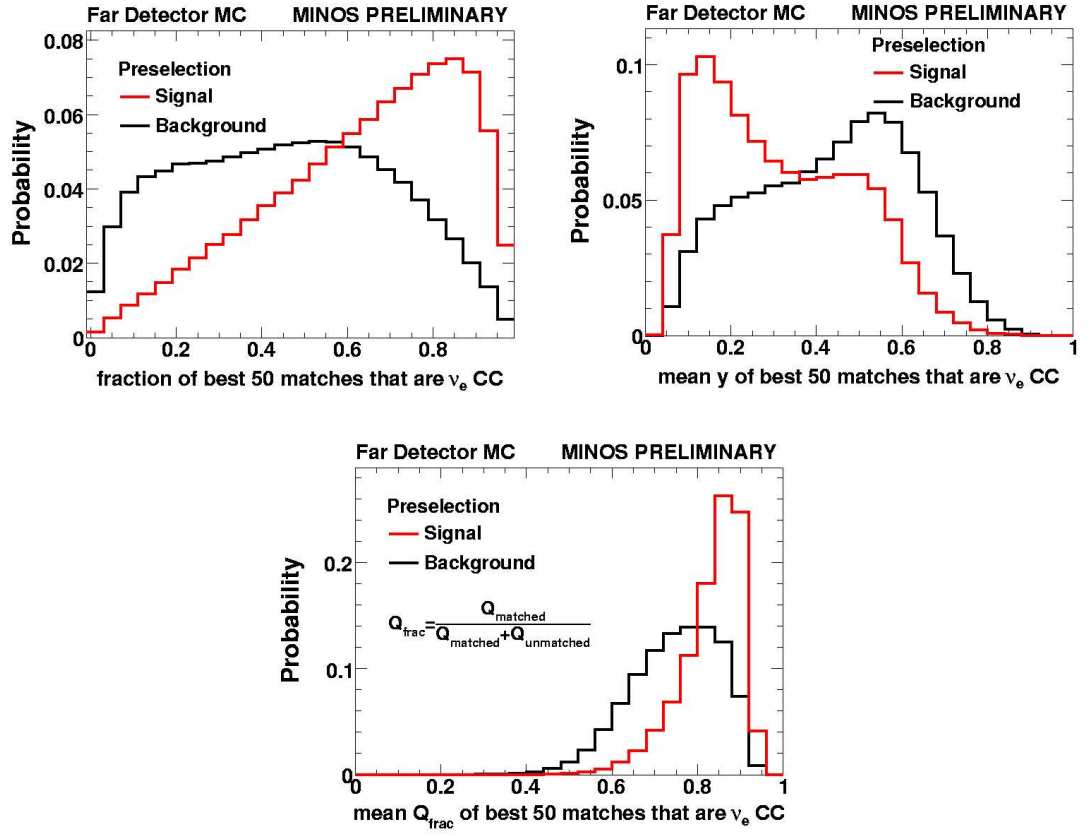


Figure 4.14: This Figure shows signal/background separation for the three variables used in the LEM PID. The full preselection has been applied. The signal is shown in red, and the background is shown in black. The histograms are area normalised to compare shape differences.

data and unfortunately there is very little external data available in the kinematic region of interest to the MINOS  $\nu_e$ -appearance analysis. Because of this, differences between the data and the MC can be observed in the ND. The ND data is very useful as it is a high statistics sample due to its proximity to the beam source. It can thus be used to compare the data and the MC and to derive meaningful conclusions from those comparisons. It can also be used to decompose the background to the  $\nu_e$  analysis into individual components. The composition of the nominal ND MC normalised to  $10^{19}$  protons on target (POT) can be seen in Table 4.1. All the  $\nu_e$ -appearance analysis preselection cuts are applied in this Table.

In the ND, the data and the MC show differences as shown in Figures 4.16 and 4.17 for the ANN PID and LEM PID selections respectively. Event numbers normalised to  $10^{19}$  POT are given in Table 4.2. From those it can be seen that the discrepancy between

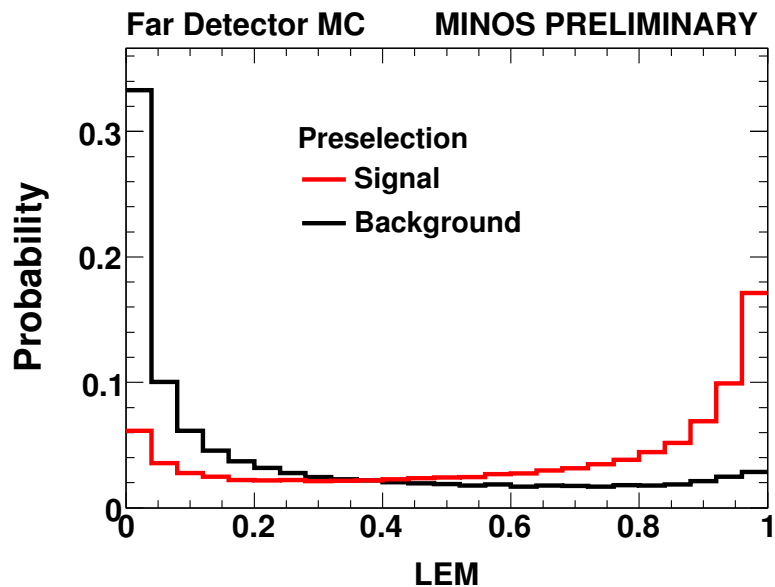


Figure 4.15: This Figure shows the signal/background separation of the LEM PID after the full preselection has been applied. The signal is shown in red, and the background is shown in black. In order to select signal events, a cut is placed at a PID value of 0.8 and only events above this cut are chosen for the final sample. The histograms are area normalised to compare shape differences.

| Selection    | MC ALL  | MC NC   | MC CC- $\nu_\mu$ | MC BEAM CC- $\nu_e$ |
|--------------|---------|---------|------------------|---------------------|
| Presel.      | 41943.1 | 20845.1 | 19799.5          | 1298.5              |
| ANN PID>0.7  | 6626.2  | 4284.9  | 1727.4           | 613.8               |
| LEM PID>0.65 | 5859.8  | 3639.7  | 1650.2           | 569.9               |

Table 4.1: ND MC numbers of events for different selections. All preselection cuts are applied and the PID cuts are at 0.7 and 0.65 for the ANN and LEM PIDs respectively. Numbers are normalised to a total exposure of  $10^{19}$  POT.

data and MC is 17% for the ANN PID and 40% for the LEM PID.

The up to 40% differences are quite large, however, it is important to note that since the hadronization model in MINOS suffers from large uncertainties [63] due to a lack of global data in the kinematic region of interest, the MC model is consistent with the data within the modeling uncertainties, as can be seen in Figure 4.18.



| Selection    | DATA    | MC ALL  | DATA/MC RATIO % |
|--------------|---------|---------|-----------------|
| Presel.      | 39224.7 | 41943.1 | 93.5            |
| ANN PID>0.7  | 5524.5  | 6626.2  | 83.4            |
| LEM PID>0.65 | 3528.2  | 5859.8  | 60.2            |

Table 4.2: ND data and MC numbers of events for different selections. All preselection cuts are applied and the PID cuts are at 0.7 and 0.65 for the ANN and LEM PIDs respectively. Numbers are normalised to a total exposure of  $10^{19}$  POT.

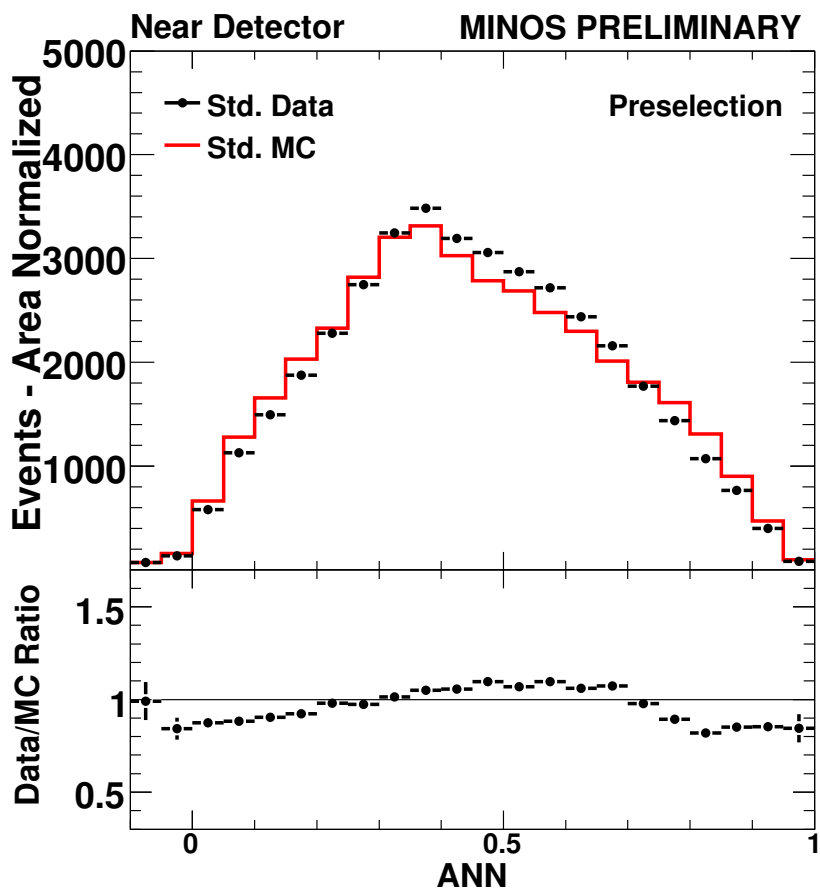


Figure 4.16: This Figure shows the preselected data versus MC in the ND for the ANN PID (the ANN cut is at 0.7). The red is the MC, and the black is the data. The distributions are normalised to the same area. The lower plot shows the data/MC ratios.

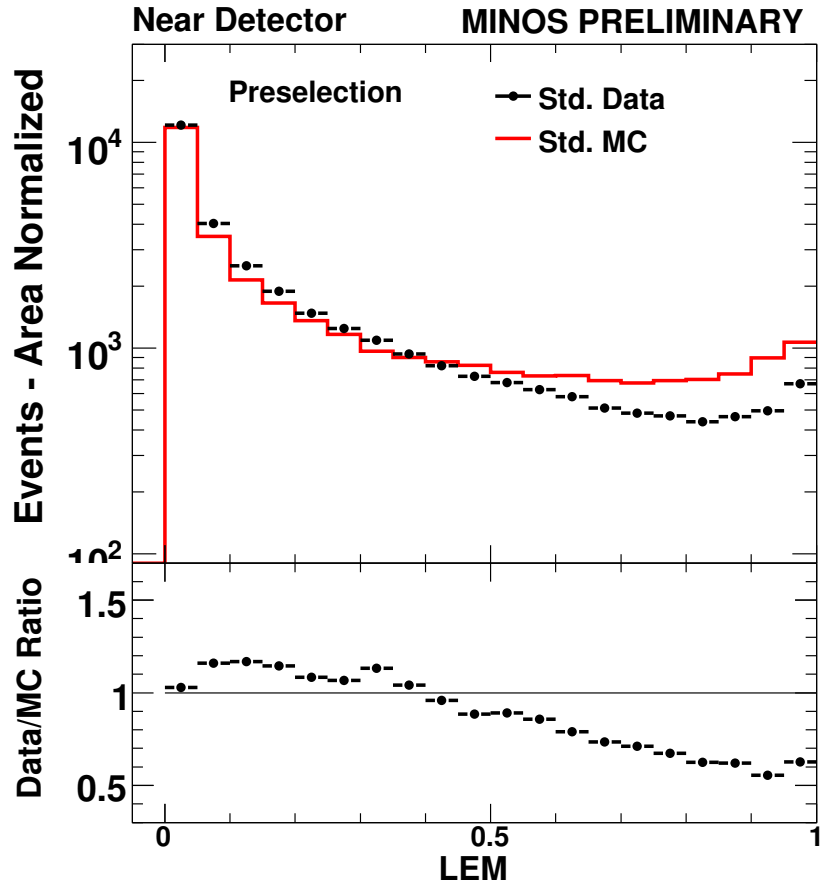


Figure 4.17: This Figure shows the preselected data versus MC in the ND for the LEM PID (the LEM cut is at 0.65). The red is the MC, and the black is the data. The distributions are normalised to the same area. The lower plot shows the data/MC ratios.

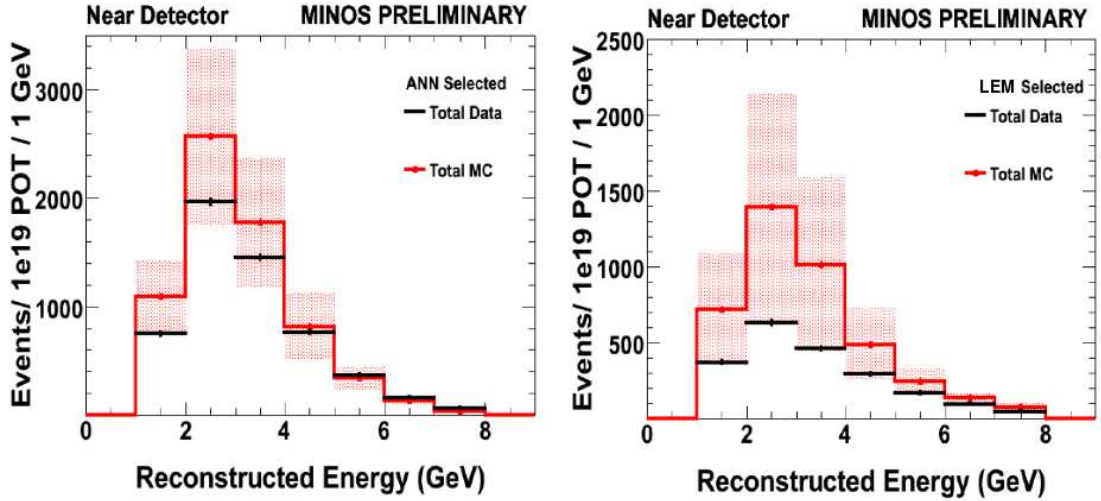


Figure 4.18: This Figure shows the uncertainty in the hadronic model of the MC for both PID selections. The data is in black, the MC is in red, and the shaded area corresponds to the uncertainty in the model.

## 4.9 Near Detector Spectrum Decomposition Methods

The data-MC differences are clearly of concern for the prediction of the  $\nu_e$  backgrounds in the FD and so two data-driven methods were developed to estimate the analysis backgrounds instead of relying on MC. The horn on-off (HOO) method is one such data-driven background estimation method, the muon removed charged current (MRCC) method is the other method and will be described in detail in Chapters 5 and 6.

### 4.9.1 The Horn On-Off Method

The horn on-off method (HOO) [64] of decomposing the background relies on a special short data run taken in the horn off mode, that is when the secondary pions and kaons produced by the impact of the protons onto the carbon target are not focused by the horns because the horns are not powered up. This means that the focusing peak at low energy vanishes because only the mesons that were originally along the beam direction contribute to the neutrino beam spectrum. The horn-on and horn-off spectra can be seen for both data and MC in Figure 4.19. It can be seen that at higher energies the two spectra are the same - no focusing by the horns is needed since the spectrum

comes from mesons that were originally along the beam direction. At low energies, the focusing peak vanishes in the horn off spectrum since many neutrino parent mesons that would normally have been focused by the horns are not. This results in the horn-off data set having a different background composition from the horn-on data. A MC truth comparison of the two spectra at  $\nu_e$  selection level can be seen in Figure 4.20, where it can be seen that for the horn off spectrum, the NC background is enhanced in comparison with the  $\nu_\mu$ -CC background. These differences in the two spectra can be leveraged to decompose the ND horn-on spectrum into backgrounds with the help of the respective MCs by solving what is essentially a set of equations with the background numbers as unknowns.

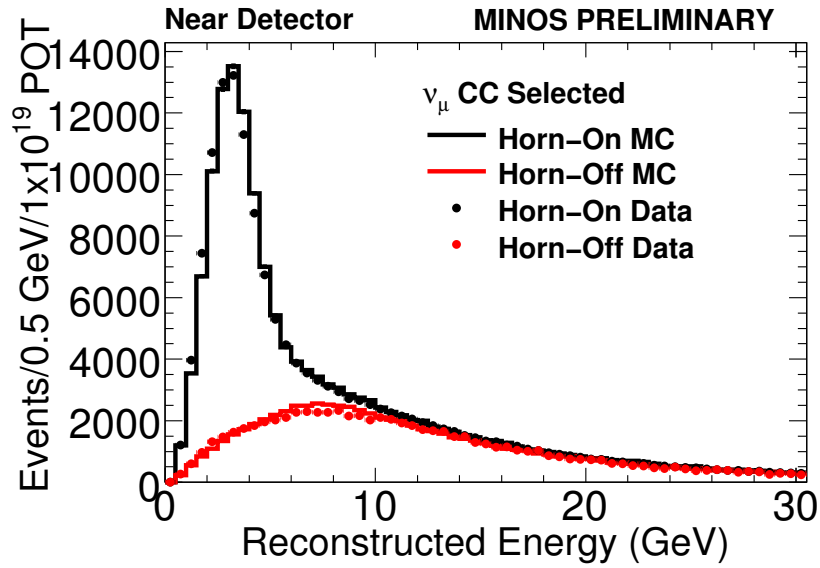


Figure 4.19: This Figure shows the  $\nu_\mu$ -CC selected horn on (black) and horn off (red) spectra for both data and MC. It can clearly be seen that when the horns are off, there is no focusing peak in the energy spectrum.

The spectra can be parametrised as follows:

$$N_i^{ON} = N_{i,NC}^{ON} + N_{i,CC}^{ON} + N_{i,\nu_e}^{ON} \quad (4.5)$$

$$N_i^{OFF} = N_{i,NC}^{OFF} + N_{i,CC}^{OFF} + N_{i,\nu_e}^{OFF} \quad (4.6)$$

where  $i$  denotes the bin number in reconstructed energy;  $N_i^{OFF}$  and  $N_i^{ON}$  are the total data numbers for a particular bin of energy; the total data numbers are sums of their respective three backgrounds: neutral current (NC),  $\nu_\mu$ -CC, and intrinsic beam

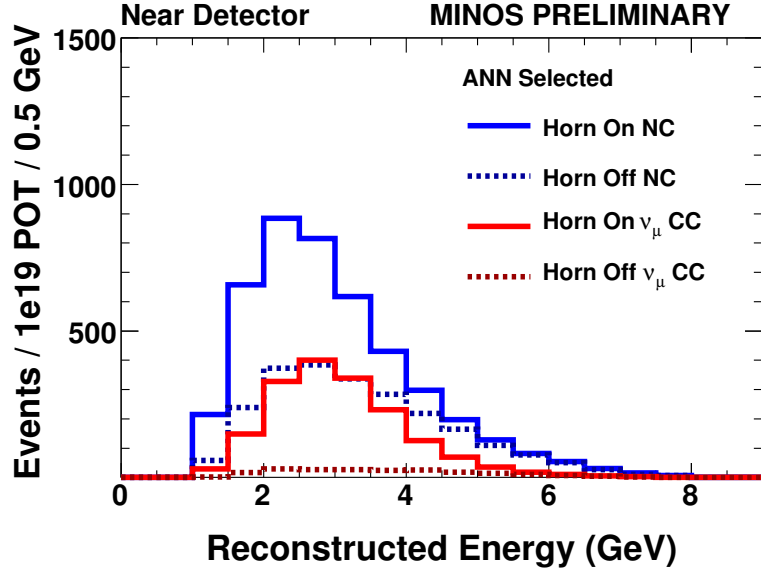


Figure 4.20: This Figure shows a MC comparison of the  $\nu_\mu$ -CC and NC components after the ANN PID selection for the horn on and horn off samples. The horn on NC sample is shown in solid blue, the horn off NC is shown in dashed blue. The horn on  $\nu_\mu$ -CC sample is shown in solid red, and the horn off  $\nu_\mu$ -CC is shown in dashed red. It can be seen that in the horn off sample, the NC component is much larger in comparison to the  $\nu_\mu$ -CC component.

$\nu_e$ -CC. Those background components parameters are at this stage unknowns. In order to extract them, Equation 4.6 can be rewritten as:

$$N_i^{OFF} = r_{NC}N_{i,NC}^{ON} + r_{CC}N_{i,CC}^{ON} + r_{\nu_e}N_{i,\nu_e}^{ON} \quad (4.7)$$

where:

$$r_x = \frac{N_x^{OFF}}{N_x^{ON}}. \quad (4.8)$$

The ratios  $r_x$  as defined by Equation 4.8 need to be taken from MC. Because of this, there could be some small mismodeling effects, however, in practice, such uncertainties will cancel to first order since the  $r_x$  parameters are ratios.

Equations 4.5 and 4.7 have now essentially become a set of two equations with three unknowns representing the data background components of the horn on sample. As the HOO method does not have a handle on the beam  $\nu_e$ -CC background, this is taken from horn on MC. This now means that the set of two equations has only two unknowns and so they can be solved and yield the following solutions for the NC and the  $\nu_\mu$ -CC

backgrounds:

$$N_{CC}^{ON} = \frac{r_{NC}N^{ON} - N^{OFF} + (r_{\nu_e} - r_{NC})N_{\nu_e}^{ON}}{r_{NC} - r_{CC}}. \quad (4.9)$$

$$N_{NC}^{ON} = \frac{N^{OFF} - r_{CC}N^{ON} - (r_{\nu_e} - r_{CC})N_{\nu_e}^{ON}}{r_{NC} - r_{CC}}. \quad (4.10)$$

The ratios  $r_x$  for the MC for  $\nu_\mu$ -CC events and NC events are shown in Figures 4.21 and 4.22 as a function of cut level. It can be seen in those plots that the ratios are quite stable, even as tighter and tighter cuts are applied.

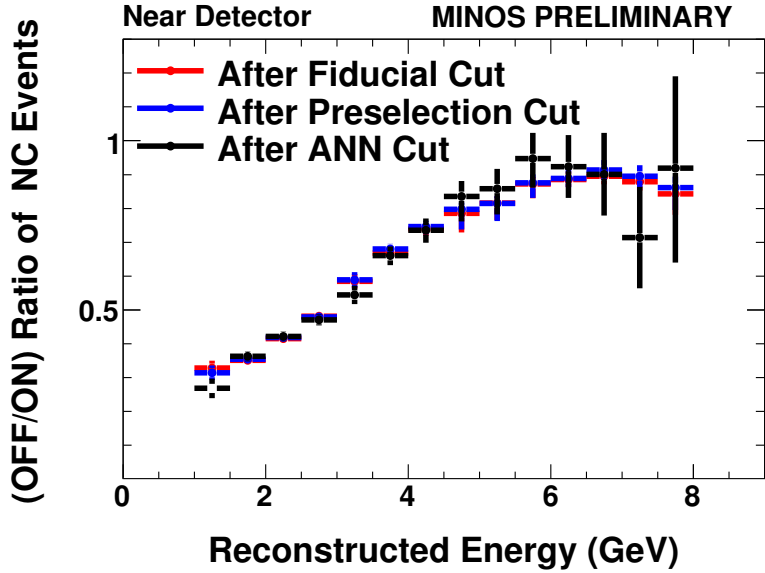


Figure 4.21: This Figure shows the MC NC horn off/on ratios as a function of cut level. The fiducial volume cut level is shown in red, the  $\nu_e$  preselection is shown in blue, and the ANN PID selected ratios are shown in black. It can be seen that the ratios change relatively little as the cut levels progress.

It is impossible to compare truth level selections in the data, and so in the  $\nu_e$  selected sample it is not possible to distinguish between the different background types. However, it is possible to compare the horn off/on ratios after the fiducial volume selection by applying the MINOS  $\nu_\mu$ -CC PID selection (this will select events that have a long muon track and yield a high purity sample of  $\nu_\mu$ -CC events). The fact that the data and the MC ratios agree after these cuts supports the premise that the horn off/on ratios are well modeled by the MC and will agree at deeper cut levels including the  $\nu_e$ -selection. The comparisons of the data and MC horn off/on ratios can be seen in Figures 4.23 and 4.24 for the NC (i.e. anti-selection of  $\nu_\mu$ -CC) and  $\nu_\mu$ -CC respectively.

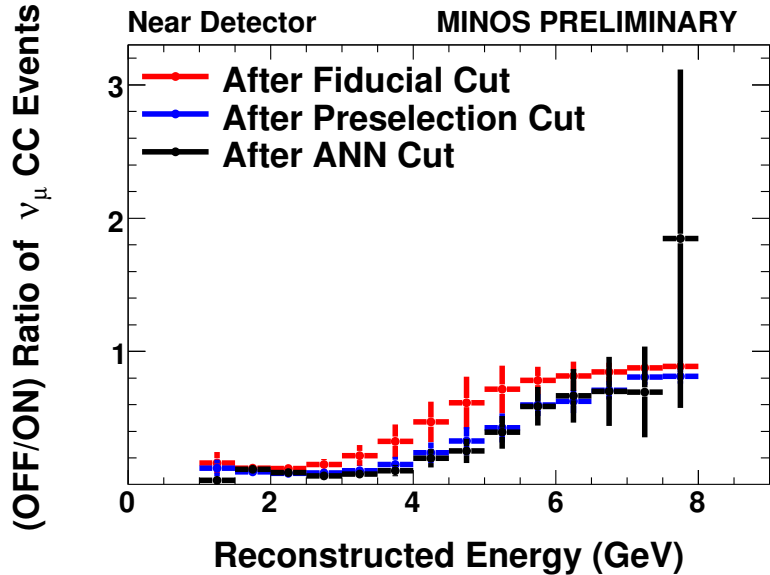


Figure 4.22: This Figure shows the MC  $\nu_\mu$ -CC horn off/on ratios as a function of cut level. The fiducial volume cut level is shown in red, the  $\nu_e$  preselection is shown in blue, and the ANN PID selected ratios are shown in black. It can be seen that the ratios change relatively little as the cut levels progress.

The systematic errors for the method are assigned by two methods. The first part of the error is assigned by comparing the  $r_x$  parameters for data and MC (the amount by which the MC ratios differ from the data ratios represents the systematic uncertainty). The second set of errors is assigned by comparing the MC  $r_x$  parameters at the  $\nu_e$  selection level with the  $r_x$  parameters at fiducial volume level. In order to account for the fact that the horn on beam  $\nu_e$ -CC background component is taken directly from MC, a conservative error of 30% is assigned to this component.

Since the relative background contributions in the horn on-off method result from solving a set of equations where the input numbers are data numbers that have finite statistics (especially in the case of the horn off sample), some bins in reconstructed energy may for example yield backgrounds event numbers that may be negative. It thus becomes necessary to adjust a few of the bins to make the decomposition numbers physical and also agree with the data. In order to achieve this, the number of events in a background bin that has a negative number of events (this only happens at the edges of the spectrum, where statistic may impact the horn off/on ratios) is set to zero and the total for the other unknown background is adjusted such that the total number plus the

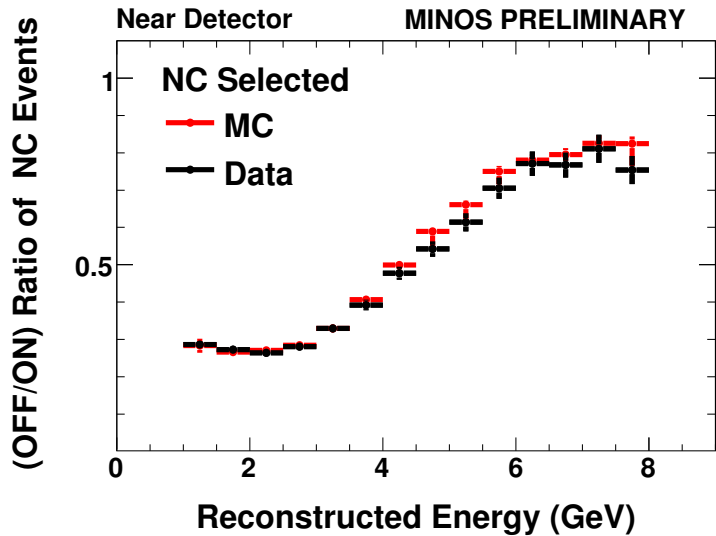


Figure 4.23: This Figure shows the NC horn off/on ratios for both data (black) and MC (red). It can be observed that the data ratios are well modeled by the MC.

number of beam  $\nu_e$ -CC equals the data in that bin.

The final results of the horn on-off ND background decomposition method can be seen for the ANN and LEM PIDs respectively in Figures 4.25 and 4.26. The final decomposition numbers are given in Table 4.3.

| Selection    | DATA    | NC                                   | $\nu_\mu$ -CC                        | Beam $\nu_e$ |
|--------------|---------|--------------------------------------|--------------------------------------|--------------|
| ANN PID>0.7  | 5524±35 | 3150 <sup>+292</sup> <sub>-273</sub> | 1781 <sup>+366</sup> <sub>-302</sub> | 593±178      |
| LEM PID>0.65 | 3528±28 | 2073 <sup>+260</sup> <sub>-258</sub> | 865 <sup>+351</sup> <sub>-216</sub>  | 590±177      |

Table 4.3: Final ND data decomposition numbers for the horn on-off Method. All preselection cuts are applied and the PID cuts are at 0.7 and 0.65 for the ANN and LEM PIDs respectively. Numbers are normalised to a total exposure of  $10^{19}$  POT.

Concerns were raised whether it is correct to use showers coming from a different kinematic region to determine the backgrounds in the horn-on beam. In particular, the showers in the unfocused horn off beam come from higher energy neutrinos than the showers in the focused horn-on beam. This can for example affect shower angle, which in theory the PIDs could be sensitive to. An additional study was carried out into the kinematics of using horn-off events selected by the  $\nu_e$ -appearance analysis and it found that the method is relatively unaffected by those kinematic issues and that it does have a sound physical basis [65].



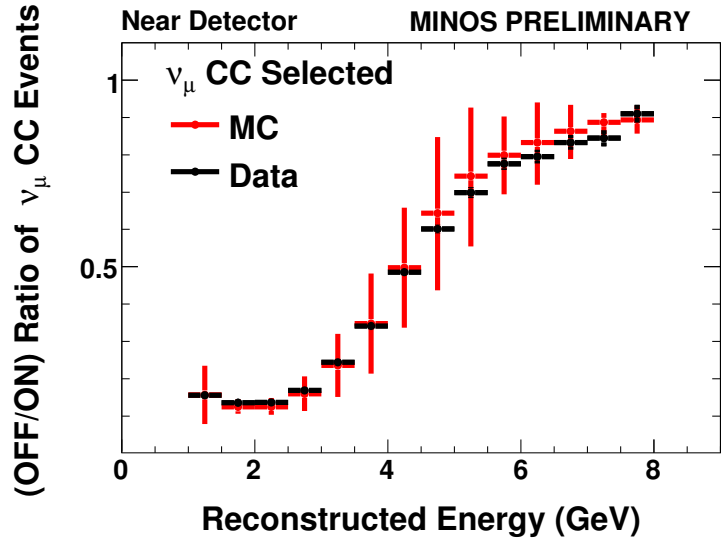


Figure 4.24: This Figure shows the  $\nu_\mu$ -CC horn off/on ratios for both data (black) and MC (red). It can be observed that the data ratios are well modeled by the MC.

#### 4.9.2 The Muon Removed Charged Current Method

The muon removed charged current (MRCC) method of decomposing the ND data spectrum uses a completely different concept to the HOO method, and therefore the two methods provide a very useful cross-check for each other. The MRCC method uses  $\nu_\mu$ -CC showers as if they were NC showers in order to correct the neutral current ND MC background component. The MRCC method is discussed in detail in the following two chapters.

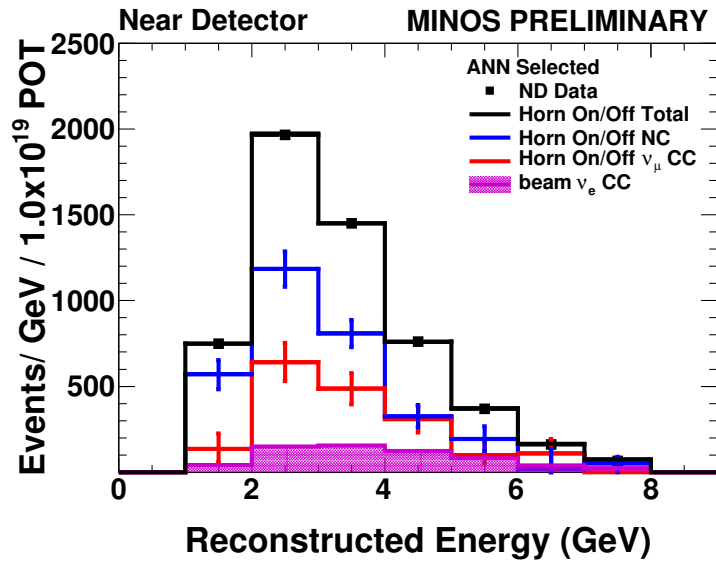


Figure 4.25: This Figure shows the final ND data decomposition for the horn on-off method for the ANN PID selection. The data is shown as black points, the total of all decomposed backgrounds is shown as black lines; the NC background in blue, the  $\nu_\mu$ -CC background in red, and the beam  $\nu_e$  background in purple.

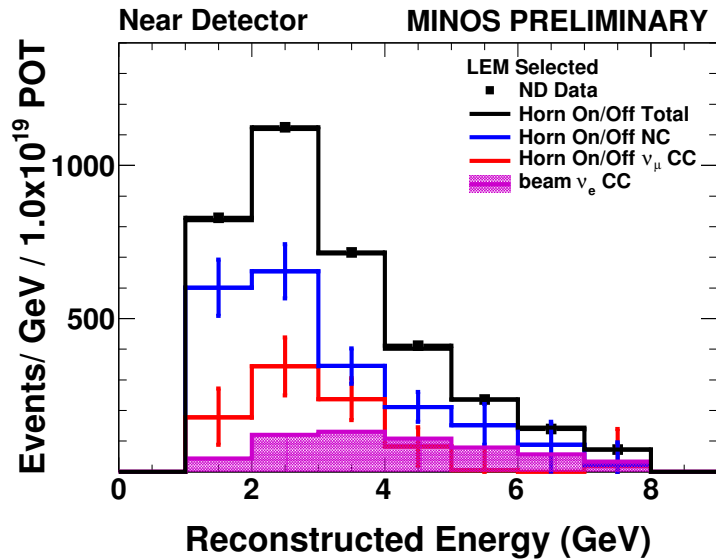


Figure 4.26: This Figure shows the final ND data decomposition for the horn on-off method for the LEM PID selection. The data is shown as black points, the total of all decomposed backgrounds is shown as black lines; the NC background in blue, the  $\nu_\mu$ -CC background in red, and the beam  $\nu_e$  background in purple.

## Chapter 5

# The Muon Removed Charged Current Method - MRCC

The Muon Removed Charged Current (MRCC) method [66, 67] provides a way to measure the ND NC background component in a data-driven way. It is one of the two ND background decomposition methods used in order to carry out the  $\nu_e$  appearance analysis, the other being the HOO method which has already been discussed in section 4.9.1. Charged current and neutral current events are very similar in that a neutrino strikes a nucleus, a lepton goes out and a shower is created. This idea is demonstrated in Figure 5.1. It is possible to take a sample of MINOS  $\nu_\mu$ -CC events with long tracks and remove the long tracks leaving only the showers. The resulting showers can then be essentially treated as an independent pseudo-NC sample since the muon tracks are no longer present. From now on, data and MC that have not been subjected to muon removal (MR) will be referred to as standard data and MC, whereas samples that have been subjected to muon removal will be referred to as MRCC samples. Of course, in reality, charged current events are not neutral current events. This will be investigated in more detail in Chapter 6. However, in the case of the MINOS detectors, showers from NC and from  $\nu_\mu$ -CC events look very similar, which is due to the granularity of the detectors.

The goal of the MRCC background estimation method is to obtain data-driven ND background component estimates, in particular for the NC and CC- $\nu_\mu$  event backgrounds, which can then be extrapolated to the FD in order to predict the expected number of observed events in the absence of  $\nu_e$  appearance.

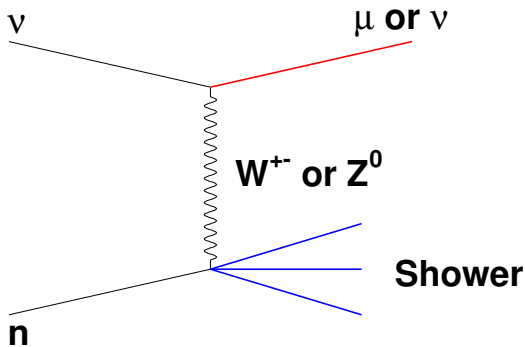


Figure 5.1: This Figure demonstrates the general idea behind the MRCC method in MINOS. The neutrino strikes a nucleus, exchanges a boson, and produces a muon. The struck nucleus produces a particle shower. Essentially, this means that charged current and neutral current events should be very similar if one removes the long muon track in the charged current event.

As will be shown later, there is evidence from the MRCC method that the discrepancy between the standard ND data and MC is due to imperfect modeling of hadronic showers to the level required to achieve good data-MC agreement for the very small and specific samples characteristic of the  $\nu_e$  analysis. Because of this, the MRCC method can be used to both carry out ND background decompositions, and also for one of the FD sidebands as will be described in Chapter 7.

## 5.1 Description of Removal Process

From a technical perspective, the MRCC samples are created by taking reconstructed event files (both data and MC), removing the longest tracks from the events, and then re-reconstructing the remnant showers as if they were NC events (MRCC events can also mimic high- $y$  CC- $\nu_\mu$  events. The latter are characterised by most of the neutrino energy being transferred to the shower and therefore resulting in a very short muon track that often does not extend beyond the shower and thus makes the event look NC-like). This means that there are two reconstruction passes, one prior, and one after the removal of the muon track. The MRCC process is illustrated in Figure 5.2 and is described in more detail in the subsequent sections.

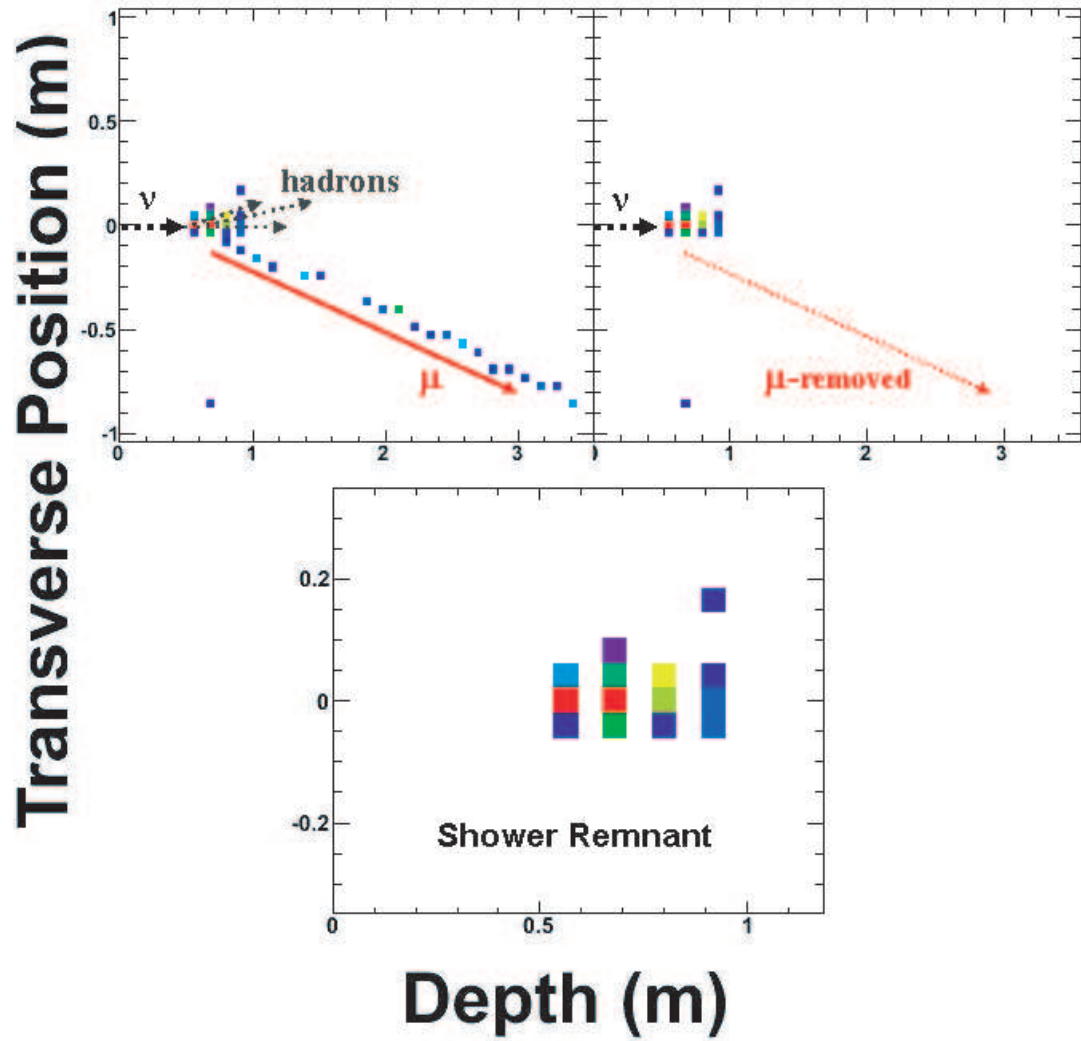


Figure 5.2: MRCC procedure example. The muon neutrino interacts in the detector and leaves a muon track and a hadronic shower. The muon track is removed, which leaves the shower remnant to be re-reconstructed as a new pseudo-NC event.

### 5.1.1 Generating the MRCC Sample

When the MRCC events are created, standard events have their first (usually longest) tracks removed, which is done by removing the hits along the track, and also, if a hit is both on the track and in the shower, 1 MIP<sup>1</sup> is subtracted from each hit that is in common to both track and shower. After the primary tracks have been removed, the resulting samples containing the shower remnants are re-reconstructed to obtain the MRCC samples. This means that the re-reconstruction stage attempts to reproduce any reconstruction issues that may be present for standard NC events, where showers may be split or may be too small to be reconstructed as an event.

During the MRCC process, certain quantities of the original event are retained and written out in the new MRCC event files for analysis and quality checking purposes. Examples of such quantities include the original reconstructed  $\nu_\mu$ -CC event PID classification (the original  $\nu_\mu$ -CC event PID classifications are described in more detail in [32] and [47]), the original x, y and z vertices of the event, the original x, y, and z momenta of the muon candidate, the original shower energy, and original track charge (q/p). Some of this retained information is used to apply selection cuts to the MRCC samples based on information from the events prior to the muon removal (before the second reconstruction pass). Those cuts will be described in more detail in the section 5.1.2.

In addition to the retained original event quantities, the completeness and purity of the muon removal process are recorded. The completeness is calculated as the fraction of the number of shared event hits between the original and the MRCC event and the number of original event hits (after the track was removed); this therefore reflects how many hits from the original shower were kept and are still in the new shower. The purity is calculated as the fraction of the number of shared event hits (between the original and the MRCC event) and the new re-reconstructed MRCC event hits; the purity therefore reflects how many hits may have migrated into the shower remnant during the re-reconstruction process. The purity and completeness are used as quality control to verify that the muon removal process worked as expected. In fact it can be seen in Figures 5.4 and 5.5 that both the MRCC purity and completeness are very high, mostly about 1.0 for the final samples selected by the PIDs. After all selection cuts, the vast

---

<sup>1</sup>A MIP is an abbreviation for the energy deposited by a minimum ionising particle and is the typical energy deposited by muons in the MINOS detectors.

majority of events in the ND have a completeness or purity above 0.9, as can be seen in Table 5.2. High purity and completeness indicate that the removal of the muon is working well and that most shower hits from the original event are still present in the 'new' event.

After the MRCC files are created, they are used to make  $\nu_e$ -analysis files that contain all the above information and also normal event information like the  $\nu_e$  PIDs and other  $\nu_e$ -analysis specific variables for the new events. The reconstructed energy of MRCC events is shown in Figure 5.3. The MRCC events can be treated as completely independent from the standard NC events since originally they were  $\nu_\mu$ -CC events.

### 5.1.2 MRCC Analysis Selection

The MRCC process is only meaningful for  $\nu_\mu$ -CC events with long tracks, but sometimes NC events with long pion tracks may have their tracks removed as well, while other, shower-like events, will be left intact and will still be present after the re-reconstruction process. An MRCC specific preselection is therefore important in order to make sure that the chosen sample is  $\nu_\mu$ -CC only. The MRCC selection cuts are applied to both data and MC and include:

- Muon removal completeness  $> -10$  - cuts out the events that did not have a track removed. It removes any events that were originally only showers, which are given a default MRCC completeness of -9999 at processing time.
- An expanded fiducial volume cut on the original event vertices of radius  $< 1.2\text{m}$  and  $0.5\text{m} < Z < 5.5\text{m}$  in order to cut out CC- $\nu_\mu$  events that were originally outside this fiducial volume. This is a quality control cut and its effects are small, of the order of 4%. The reason for applying it is to make sure that only showers from CC- $\nu_\mu$  events whose vertices are within a reasonable fiducial volume in the ND are included in the final MRCC sample.
- A track fit pass cut on the original track to make sure that the removed track was a good track. The effects of this cut are small - at the 3% level.
- A cut on the original  $\nu_\mu$ -CC event PID classification of  $\nu_\mu$ -CC PID  $> 0.3$  (as used in [47]). This cut is applied to select only events that were originally  $\nu_\mu$ -CC events

| Cut                  | Events | Total Eff. | Eff. to Prev. Cut |
|----------------------|--------|------------|-------------------|
| Std. Fid. Vol.       | 222138 | 100%       | 100%              |
| MRCC Completeness    | 176183 | 79.31%     | 79.31%            |
| MRCC Fid. Vol.       | 169143 | 76.14%     | 96.00%            |
| MRCC Track Fit Pass  | 164121 | 73.88%     | 97.03%            |
| MRCC Orig. CC PID    | 117533 | 52.91%     | 71.61%            |
| Std. $\nu_e$ Presel. | 34391  | 15.48%     | 29.26%            |
| ANN PID              | 5322.2 | 2.40%      | 15.48%            |
| LEM PID              | 2954.8 | 1.33%      | 8.59%             |

Table 5.1: MRCC ND data event numbers and efficiencies for progressive cuts ( $10^{19}$  POT, after data quality cuts)

with long tracks in order to make the selected shower sample completely independent from the standard NC events. The distribution of the  $\nu_\mu$ -CC PID is shown in Figure 5.6.

After the MRCC preselection, over 98% of events selected are true CC- $\nu_\mu$  events. Only  $\sim 1.5\%$  of NC events and  $\sim 0.1\%$  of beam CC- $\nu_e$  events remain in the sample. The remaining events are events that originally had a long track, like perhaps a pion track, and therefore were a natural background to the  $\nu_\mu$ -CC event selection. After the MRCC preselection, a high purity independent pseudo NC sample has been obtained. A detailed breakdown of the background numbers is shown for the  $\nu_e$ -analysis PIDs in Table 5.3.

After the MRCC specific cuts have been applied, in order to carry out the MRCC background estimation procedure, the same  $\nu_e$  preselection cuts need to be applied to the MRCC sample as are applied to the standard data and MC. It is important to stress that those cuts are applied on the new re-reconstructed event variables, not the original event quantities. Those cuts were described in Chapter 4 and include a fiducial volume cut, track length cuts, a requirement that there be at least one shower, a contiguous planes count cut, reconstructed energy cuts, and finally, a PID cut (ANN or LEM PID since the analysis was carried out for both PIDs). The progression of cuts for the MRCC data is shown in Table 5.1 and in Figures 5.3, 5.4, 5.5, and 5.6. The MRCC purity and completeness for the different cut levels are shown in Table 5.2.



| Cut                  | Pur. Mean | Comp. Mean | Pur. > 0.9 | Comp. > 0.9 |
|----------------------|-----------|------------|------------|-------------|
| Std. Fid. Vol.       | 94.56%    | 97.45%     | 70.83%     | 75.25%      |
| MRCC Comp.           | 94.56%    | 97.45%     | 89.31%     | 94.88%      |
| MRCC Fid. Vol.       | 95.72%    | 97.81%     | 90.78%     | 95.63%      |
| MRCC TrackFitPass    | 95.75%    | 97.81%     | 90.84%     | 95.64%      |
| MRCC Orig. CC PID    | 95.39%    | 97.75%     | 89.86%     | 95.49%      |
| Std. $\nu_e$ Presel. | 97.88%    | 98.65%     | 96.08%     | 98.04%      |
| ANN PID              | 98.52%    | 98.93%     | 97.81%     | 98.86%      |
| LEM PID              | 98.38%    | 98.87%     | 97.32%     | 98.83%      |

Table 5.2: MRCC ND data MRCC completeness and purity means and percentage of events with purity or completeness above 0.9 for progressive cuts ( $10^{19}$  POT, after data quality cuts)

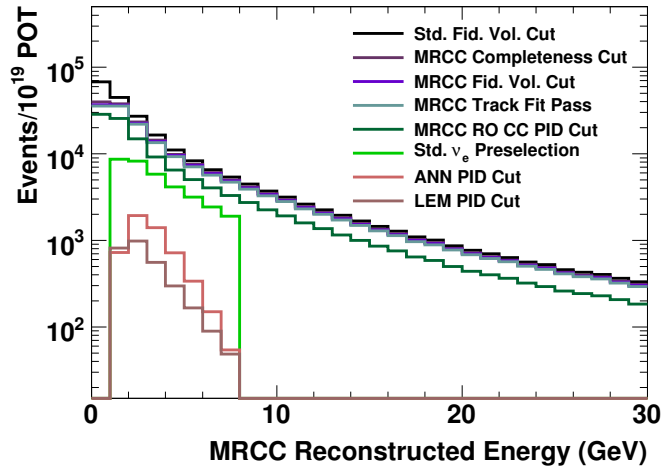


Figure 5.3: Reconstructed energy for  $10^{19}$  POT in the ND for MRCC data. The colours correspond to progressive selection cuts, where the standard fiducial volume cut was applied first. The two final PID selections are also shown.

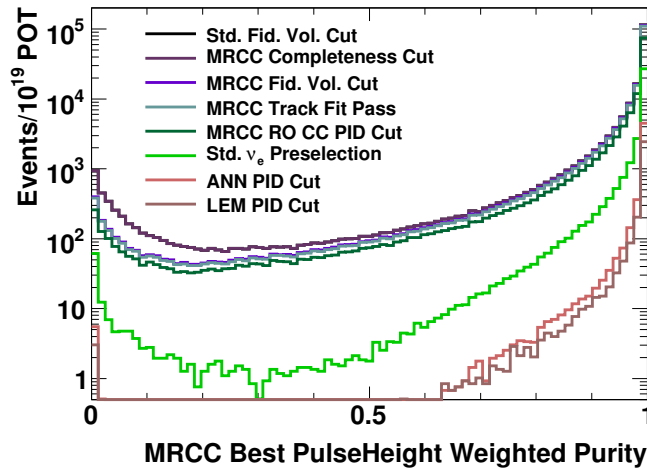


Figure 5.4: MRCC pulse height weighted purity for  $10^{19}$  POT in the ND for MRCC data. The colours correspond to progressive selection cuts, where the standard fiducial volume cut was applied first. The two final PID selections are also shown. Note that the lines for standard fiducial volume cut and the best completeness  $> -10$  cut are the same. This is because the purity is a MRCC quantity, which is only filled with a value between 0 and 1 if the event is a muon removed event, so only events that are already MRCC are shown in this Figure.

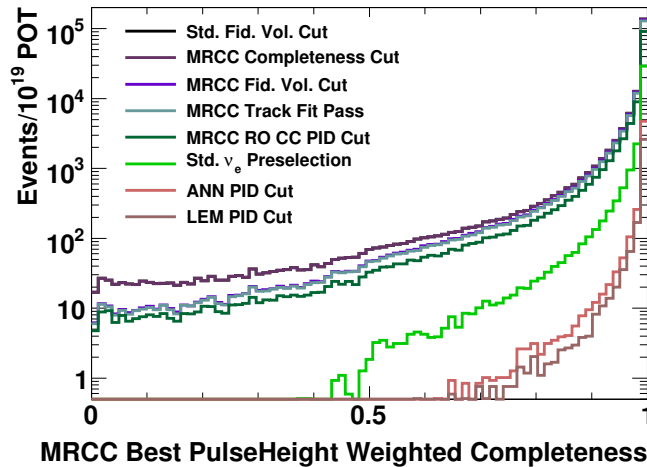


Figure 5.5: MRCC pulse height weighted completeness for  $10^{19}$  POT in the ND for MRCC data. The colours correspond to progressive selection cuts, where the standard fiducial volume cut was applied first. The two final PID selections are also shown. Note that the lines for standard fiducial volume cut and the best completeness  $> -10$  cut are the same. This is because the completeness is a MRCC quantity, which is only filled with a value between 0 and 1 if the event is a muon removed event, so only events that are already MRCC are shown in this Figure.

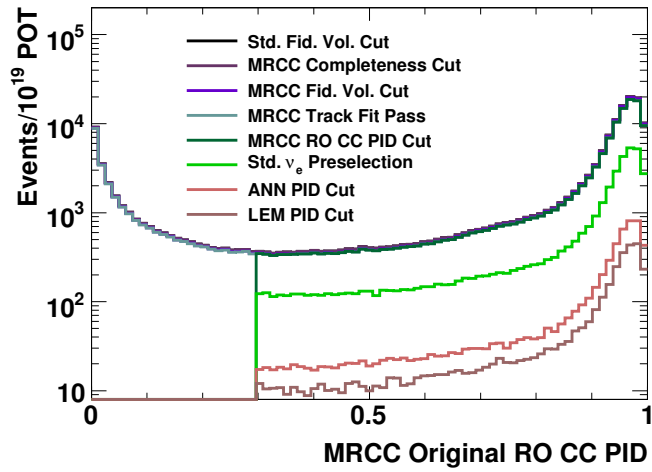


Figure 5.6: Original  $\nu_\mu$ -CC PID for  $10^{19}$  POT in the ND for MRCC data. The colours correspond to progressive selection cuts, where the standard fiducial volume cut was applied first. The two final PID selections are also shown. Note that the lines for standard fiducial volume cut and the best completeness  $> -10$  cut are the same. This is because the original  $\nu_\mu$ -CC PID is a MRCC quantity, which is only filled with a value between 0 and 1 if the event is a muon removed event, so only events that are already MRCC are shown in this Figure.

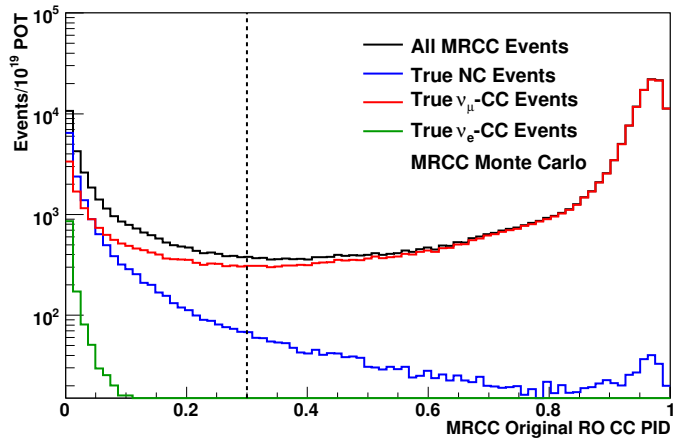


Figure 5.7: Original  $\nu_\mu$ -CC PID for  $10^{19}$  POT in the ND for MRCC MC. Only the standard (not MRCC-specific) fiducial volume cut is applied. The black line shows the total MRCC MC, the blue line shows the true NC background, the green line is the true beam  $\nu_e$ -CC background, and the true  $\nu_\mu$ -CC, which are providing the independent MRCC sample, are shown in red. Only events above a CC PID cut of 0.3 are selected by the MRCC preselection and this is indicated by the vertical dotted line.

Figure 5.7 shows the MC breakdown of the individual event types (NC,  $\nu_\mu$ -CC and beam  $\nu_e$ -CC) in the MRCC sample as a function of the original CC PID. It can be seen in this Figure that there is a background of true NC events (and a very small background of beam  $\nu_e$ -CC) that had tracks. Those events are present in the MRCC sample prior to the MRCC preselection cuts and then get removed by applying the CC PID cut. This is the main reason why it is so important to keep the information about the event prior to the muon removal.

| CUT                  | MC ALL | MC NC  | MC CCNUMU | MC BEAM NUE |
|----------------------|--------|--------|-----------|-------------|
| MRCC Presel.         | 125614 | 1621.2 | 123914    | 79.0        |
| Std. $\nu_e$ Presel. | 38047  | 561.5  | 37457     | 28.2        |
| ANN PID              | 6174.7 | 111.4  | 6053.8    | 9.5         |
| LEM PID              | 4985.1 | 91.8   | 4885.9    | 7.5         |

Table 5.3: MRCC ND MC background numbers of events for progressive cuts (based on MC truth -  $10^{19}$  POT)

The effects of the  $\nu_e$  selection on the MRCC selected sample are shown in Figure 5.8. It can clearly be seen here and in Table 5.3 that when all the selection cuts are applied, the result is a very pure sample - over 98% - of originally CC- $\nu_\mu$  events. Finally, Table 5.5 gives a summary of the  $10^{19}$  POT normalised event numbers for the MRCC data

| CUT                  | MC NC % | MC CCNUMU % | MC BEAM NUE % |
|----------------------|---------|-------------|---------------|
| MRCC Presel.         | 1.29    | 98.65       | 0.06          |
| Std. $\nu_e$ Presel. | 1.48    | 98.45       | 0.07          |
| ANN PID              | 1.80    | 98.04       | 0.15          |
| LEM PID              | 1.84    | 98.01       | 0.15          |

Table 5.4: MRCC ND MC background percentages for progressive cuts (based on MC truth -  $10^{19}$  POT)

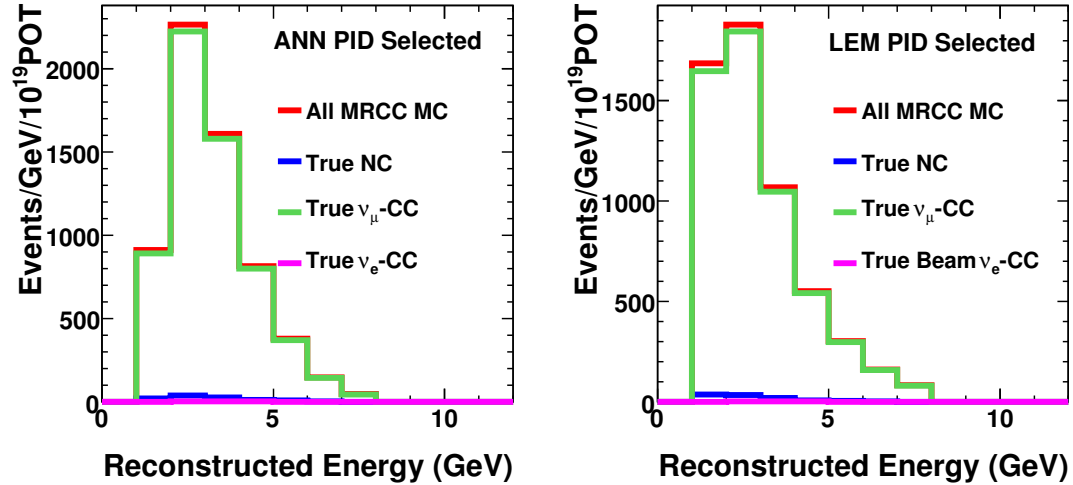


Figure 5.8: Reconstructed energy for MRCC MC normalised to  $10^{19}$  POT after all selection cuts including the ANN PID cut or LEM PID cut. The NC and CC- $\nu_e$  event contamination of the MC is shown for both PIDs, and the CC- $\nu_\mu$  purity is 98%.

and MC and the respective total data/MC ratios.

| PID                  | Data   | MC ALL | MRCC Data/MC | STD Data/MC |
|----------------------|--------|--------|--------------|-------------|
| Std. $\nu_e$ Presel. | 34391  | 38047  | 90.4 %       | 93.5 %      |
| ANN PID              | 5322.2 | 6174.7 | 86.2 %       | 83.4 %      |
| LEM PID              | 2954.8 | 4985.1 | 59.3 %       | 60.2 %      |

Table 5.5: MRCC ND data and MC numbers of events by PID ( $10^{19}$  POT) and their respective MRCC data/MC ratios. The standard data/MC Ratios are also shown for comparison purposes.

## 5.2 The MRCC Analysis Method

Figures 5.9 and 5.10 show the data/MC ratios as a function of reconstructed energy for both standard and MRCC events after final ANN and LEM PID selections. It can

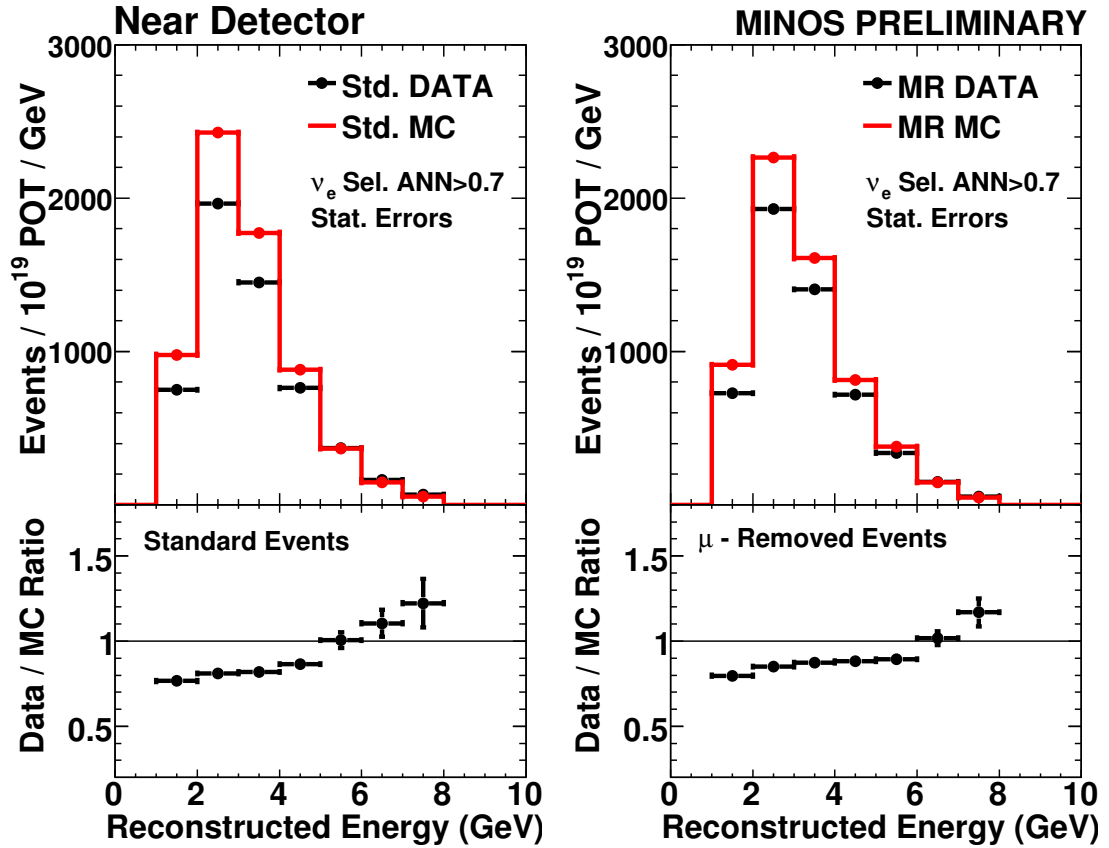


Figure 5.9: ND MRCC and standard data/MC ratios after  $\nu_e$  selection cuts and ANN PID greater than 0.7 as a function of reconstructed event energy (normalised to an exposure of  $10^{19}$  POT). It can be seen that the data/MC difference for MRCC events follows the data/MC difference in the standard events well, especially in the peak regions of the distributions.

be seen that the data/MC ratios are very similar for the standard and the MRCC events therefore the observed differences must be common to both the NC showers and the  $\nu_\mu$ -CC showers that the MRCC events originally come from. If the main differences between ND data and MC can be attributed to the hadronic shower modeling, one should be able to use the MRCC samples to derive a data-driven correction for the ND MC NC background since the MRCC events will suffer from the same modeling uncertainties as the standard NC events. There could be other reasons for the data/MC discrepancy (for example detector effects) but the shower modeling is the primary suspect given the current knowledge of the MINOS MC. Section 6.2 makes the case for shower modeling as the source of the observed data/MC differences in more detail.

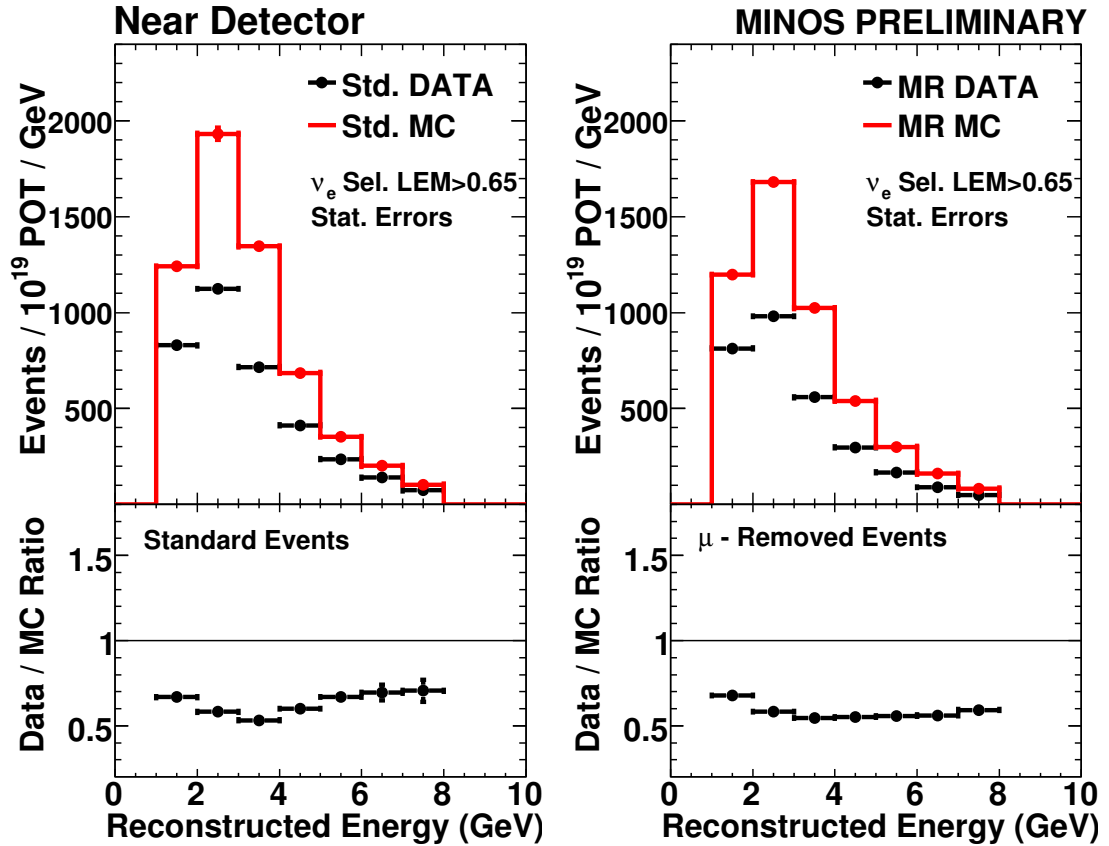


Figure 5.10: ND MRCC and standard data/MC ratios after  $\nu_e$  selection cuts and LEM PID greater than 0.65 as a function of reconstructed energy (normalised to an exposure of  $10^{19}$  POT). It can be seen that the data/MC difference for MRCC events follows the data/MC difference in the standard events well, especially in the peak regions of the distributions.

### 5.2.1 Making the Correction

In order to cancel potential systematic effects, the MRCC correction uses MRCC data/MC ratios and is given by the formula:

$$NC_{(i)}^{corrected} = \frac{MRCC_{(i)}^{Data}}{MRCC_{(i)}^{MC}} \times NC_{(i)}^{MC} \quad (5.1)$$

where  $i$  is the bin number in reconstructed energy. From this equation it can be seen that the NC background is corrected on a bin-by-bin basis by the ratio of MRCC data to MRCC MC as seen in Figures 5.9 and 5.10<sup>2</sup>.

The  $\nu_\mu$ -CC background can be estimated in two ways. It is possible to use the same correction as for the ND background:

$$\nu_\mu CC_{(i)}^{corrected} = \frac{MRCC_{(i)}^{Data}}{MRCC_{(i)}^{MC}} \times \nu_\mu CC_{(i)}^{MC} \quad (5.2)$$

This is because  $\nu_\mu$ -CC showers and NC showers are similar in MINOS due to detector granularity, thus it is reasonable to assume that any mismodeling of the NC showers will also affect high- $y$   $\nu_\mu$ -CC showers. Despite this, the method suffers from the flaw that the final total number of selected events does not agree with the total ND data by default. Also, there is an additional concern that the MRCC events and the  $\nu_e$ -selected  $\nu_\mu$ -CC background events come from a different space in hadronic  $Y$ . The MRCC events originally had a long track and hence a lower  $Y$ , whereas the latter events are shower only (the muon tracks are extremely short) and thus come from a high- $y$  region (this can also be seen in Figure 6.1 in the next chapter). Since the NC events are corrected using the MRCC ratios, and the MRCC events mimic real NC events, it is thus reasonable to assign the remaining data/MC difference to the  $\nu_\mu$ -CC background as follows:

$$\nu_\mu CC_{(i)}^{corrected} = DATA_{(i)} - NC_{(i)}^{corrected} - \nu_e CC_{(i)}^{MC} \quad (5.3)$$

where  $NC_{(i)}^{MC}$ ,  $\nu_\mu CC_{(i)}^{MC}$  and  $\nu_e CC_{(i)}^{MC}$  are the standard MC background event numbers.

---

<sup>2</sup>During the development of the MRCC method, various methods and variables were tried to carry out the MRCC correction. This meant that the MRCC data/MC ratios were not always well defined for certain bins. If for certain bins it was not possible or not reasonable to calculate a ratio due to the data or MC being equal to zero (because of limited statistics), the ratio was set to 1 by default, so as to have no effect. However, in the final method, in the range of energies used by the  $\nu_e$ -analysis (between 1.0 and 8.0GeV), all MRCC ratios are well defined.



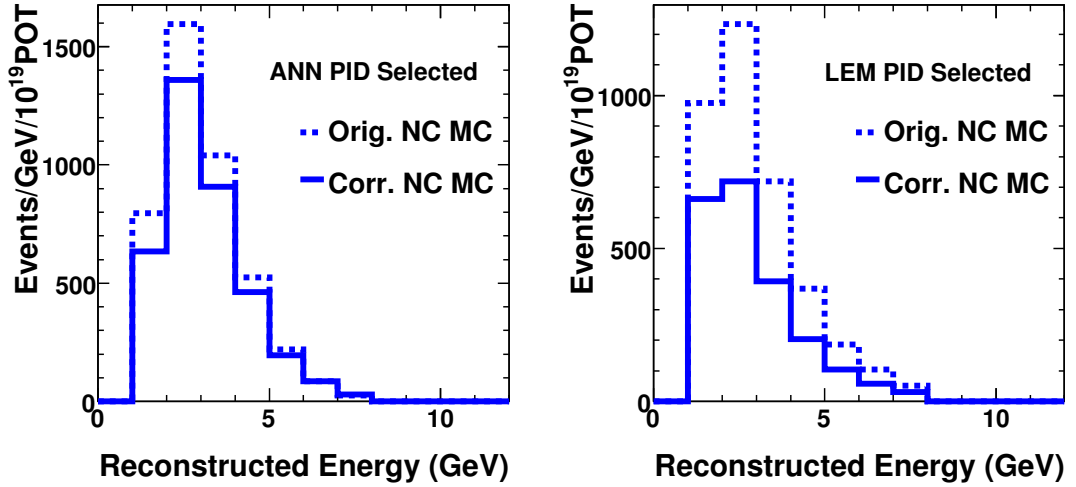


Figure 5.11: Near detector NC MC before and after MRCC correction (for  $10^{19}$  POT). It can be seen that the MRCC correction reduces the MC NC background, as needed to achieve better agreement with the data.

Even though the advantage of this method is clearly that it allows the use of the ND data to constrain the total backgrounds, it does have larger systematic uncertainties as a result of the uncertainties in the  $CC-\nu_e$  background. In addition, it can occasionally yield negative  $\nu_\mu$ -CC event numbers for bins at higher energies (above 8GeV - outside the region of interest) where the event statistics are low.

While the first method (see Equation 5.2) has smaller systematic errors, it is not constrained by the ND standard data (except indirectly by being impacted upon by the MRCC data). Also, as the first method uses  $\nu_\mu$ -CC events with a low  $y$  to correct  $CC-\nu_\mu$  events with a high  $y$ , the second method (see Equation 5.3) of estimating the  $\nu_\mu$ -CC event background is the official MRCC method, but the first method can provide an important check that the obtained NC background is reasonable.

Figure 5.11 shows the MRCC correction of the standard MC NC background. The MRCC data/MC ratios are calculated as a function of the reconstructed event energy. Since the MRCC correction is on a bin-by-bin basis, each event of the standard NC sample is weighted by the MRCC ratio corresponding to its reconstructed event energy (for instance an event that has a reconstructed energy of 3.2GeV will be weighted down by the MRCC ratio value of 0.87 for the ANN PID).

### 5.2.2 Using Different Quantities to carry out the MRCC Correction

The standard MRCC method can be modified to carry out an event-by-event MRCC correction using any reconstructed shower topology variable(s) to create the MRCC ratios as opposed to the correction in reconstructed event energy. It is even possible to use 2D MRCC ratios in order to create the correction weights so as to use more information. Many shower topological variables were tested during the development of the MRCC background decomposition method. It was found however that the results did not differ significantly from the standard correction in reconstructed energy. In a way, the use of the latter to create the MRCC data/MC ratios is the natural thing to do since it is the reconstructed energy that is being extrapolated to the FD. Another reason for not using a shower topological variable is that if one were used, a systematic error associated with the hadronic energy scale would not cancel out thus leading to a larger overall error. Despite this, using a shower shape variable for the correction is not strictly wrong and it tunes into the underlying hadronic shower modeling issues. Nevertheless, as there was no reason to assume that one weighting variable is better than another, and due to the above explained reasons, the MRCC correction is carried out as an ad-hoc correction in reconstructed energy at present.

# Chapter 6

## MRCC Discussion and Results

A number of considerations related to the MRCC background decomposition method need to be addressed prior to presenting the final results. These are discussed in this chapter. Systematic errors are also evaluated and final MRCC ND background decomposition results follow, including the results of a mock data challenge which was designed to test the validity of the MRCC method.

### 6.1 Considerations of the MRCC Method

The fundamental issue concerning the validity of the MRCC method is the difference between the NC showers and the shower component of the  $\nu_\mu$ -CC events. There are different kinematics and processes involved in those two types of events and the most important concerns are:

1. NC events are mediated by the neutral  $Z^0$  boson, whereas CC events are mediated using the charged  $W^\pm$  bosons. This means that there could be a difference in the kinematic distributions between the standard NC and the MRCC events because of the masses of the different bosons and leptons.
2. The differences in the mediating bosons also mean that the net charges of the resulting showers will be different for NC and muon removed events. A standard NC event will always transfer zero charge to the shower, however, since MRCC showers are caused by CC events, the transferred charge will be  $\pm 1$ .

3. The  $\pi^0$  content of the MRCC showers may be different from the  $\pi^0$  content of the standard NC events, therefore the electromagnetic component of the showers (which the  $\nu_e$  analysis selects on) may be different for MRCC and NC events.

There is also an issue not related to the NC/ $\nu_\mu$ -CC event differences but which could affect the validity of using the MRCC method to correct the NC showers. The issue is that the data/MC discrepancy for the standard ND samples could possibly come from something else and not imperfect shower modeling. One such possibility is for example detector modeling but the agreement between data and MC for other types of events (like long-track  $\nu_\mu$ -CC events) suggests that the detector modeling is unlikely to be the cause of the data/MC differences.

### 6.1.1 Differences in Kinematics between NC and CC Events

Addressing the first issue in section 6.1, it is possible to compare kinematic distributions for MRCC and standard NC and  $\nu_\mu$ -CC events after the  $\nu_e$  selections using MC truth variables. Comparisons have been done in slices of reconstructed energy and are shown for the ANN PID selection and the energy range of 3-4GeV in Figure 6.1. More comparisons for other energy slices and the LEM PID are included in Appendix A. The comparisons are made for MC truth distributions of the kinematic variables  $W^2$ ,  $Q^2$ ,  $x$  and  $y$  as described in section 2.6. These MC distributions show that the differences are of the order of 0-20% in the peak regions of the true  $W^2$ , true  $Q^2$  and true  $x$  distributions. In fact, it can be seen that even though the MRCC events originally come from  $\nu_\mu$ -CC events, the kinematic phase space of those events matches somewhat more closely true standard NC events than standard  $\nu_\mu$ -CC events selected by the  $\nu_e$  analysis. For true  $y$ , the differences seen in Figure 6.1 are well understood. By construction, the selected  $\nu_\mu$ -CC events come from a different phase-space in  $y$  for MRCC and standard events. Naturally, the  $\nu_e$  selection will select a background of high  $y$  standard  $\nu_\mu$ -CC events which are shower-like and have a very short muon track. On the other hand, MRCC events need to originally have had a track, so their high- $y$  kinematic region will consequently be depleted. It is reassuring though that the MRCC and the standard NC events are quite similar in  $y$  (except of course for the high- $y$  region as expected).

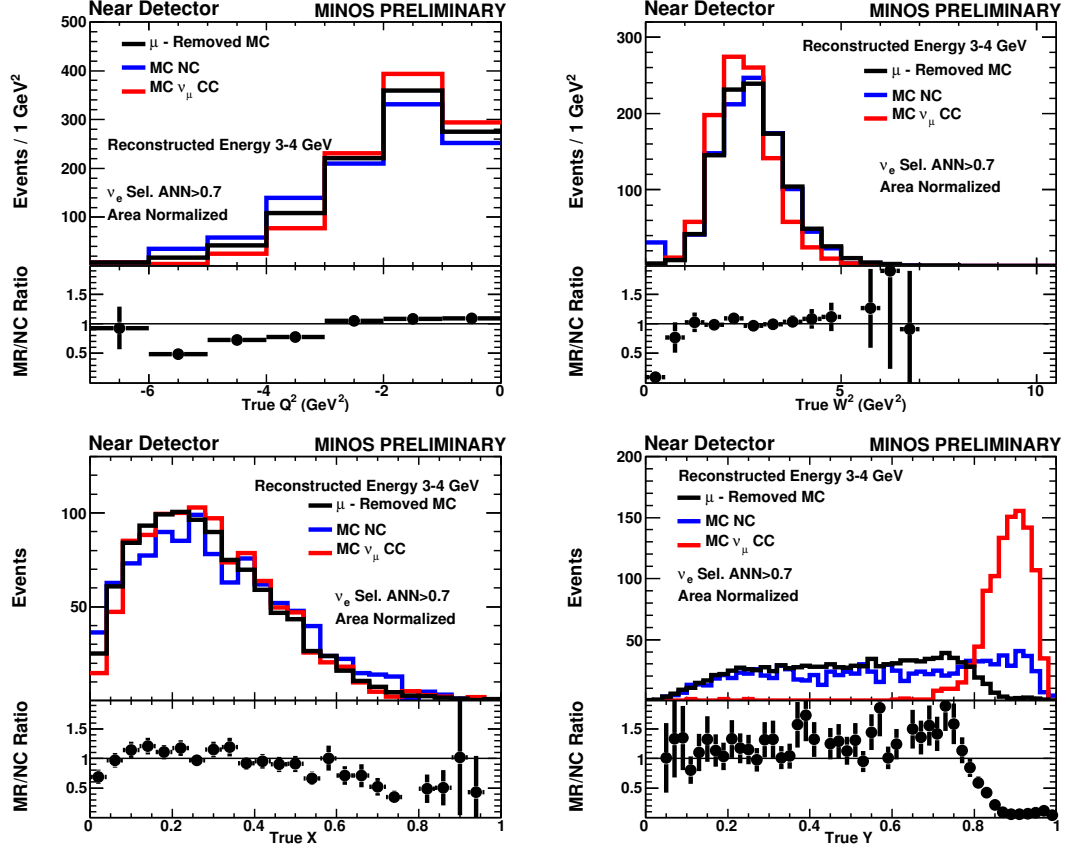


Figure 6.1: Kinematic truth distributions after  $\nu_e$  selection (ANN PID) for ND MRCC and standard MC. All distributions are for an energy slice between 3 and 4 GeV. In the four sets of plots, the top plots show the distributions of MRCC events in black, standard MC NC in blue, and standard MC  $\nu_\mu$ -CC in red. The top plots are area normalised to reveal shape differences. The bottom plots show the ratios of the MRCC to the standard MC NC distributions. The top left set of plots shows true  $Q^2$  - the four-momentum squared transferred by the neutrino to the struck nucleon. The top right set of plots shows true  $W^2$  - the mass squared of the system recoiling against the scattered neutrino. The bottom left set of plots shows true  $x$  - the fraction of the nucleon's momentum carried by the struck quark. Finally, the bottom right set of plots shows true  $y$  - the fraction of the neutrino's energy lost in the nucleon rest frame.

### 6.1.2 Net Charge of the MRCC Showers

In order to understand the second question in section 6.1 better, that is how much the net charge of the hadronic shower may affect the results, we can use the fact that the MRCC data and MC samples contain a mixture of both  $\nu_\mu$ -CC events and a smaller sample of anti- $\nu_\mu$ -CC events. It was suggested in [68] to split the MRCC events into neutrinos and anti-neutrinos by applying an additional muon sign selection. This is done by cutting on original muon event charge over momentum  $q/p$  (abbreviated as QP - the track charge sign) as follows:  $QP \leq 0$  for neutrinos, and  $QP > 0$  for anti-neutrinos. It was then suggested that one evaluate whether those samples differ from each other in their MRCC ratios. The NC showers lie in terms of shower charge between the neutrino and anti-neutrino CC showers, thus comparing neutrino and anti-neutrino MRCC events (charge difference of  $q=2$ ) should give a good estimate of how much the shower charge may impact the MRCC ratios.

MRCC events are a mixture of muon neutrino (around 90%) and muon anti-neutrino (10%) events. This can be seen in Figure 6.2. The corresponding data/MC ratios can be seen in Figure 6.3. Since the differences in the MRCC ratios are small between the neutrino and anti-neutrino samples, the conclusion can be drawn that the MINOS detectors have little sensitivity to differences in shower charge.

The visible differences for the neutrino and anti-neutrino samples can be used to estimate the systematic error associated with charge differences between NC and  $\nu_\mu$ -CC event showers. This is done by running the correction with the  $\nu_\mu$  and anti- $\nu_\mu$  MRCC ratios separately and using the differences as an estimate of this systematic error. During the development of the MRCC background decomposition method, an attempt was made to use QP charge information so as to cancel out the contributions from the  $\nu_\mu$  and anti- $\nu_\mu$  in the hope of making the net shower charge closer to the real standard NC events. Even though it appeared to be possible to calculate such weighted MRCC ratios, the difference from the normal (all QP) MRCC ratios was marginal. Also, there would be additional systematic errors associated with using a charge canceling approach that might remove any possible benefit. So as to take these effects into account however, the differences between the neutrino and anti-neutrino MRCC ratios were used as a systematic error for the MRCC method. This will be discussed in more detail in Section 6.3.

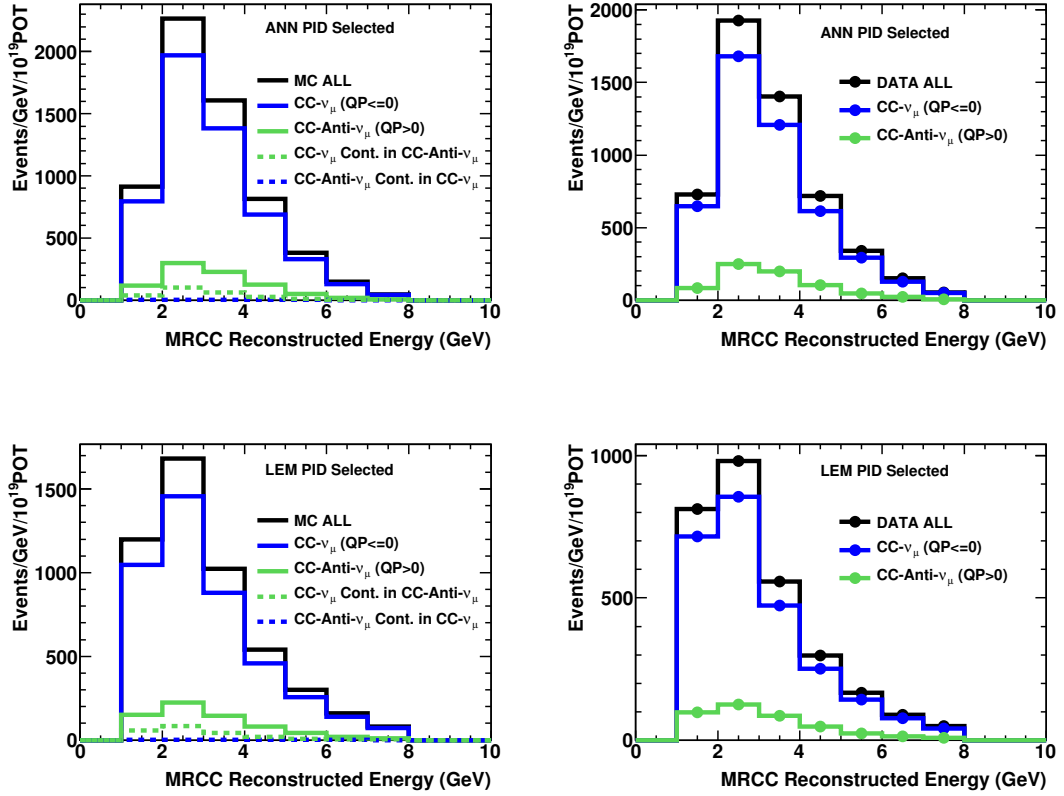


Figure 6.2: MRCC distributions for different original track charge sign cut QP - after ANN selection for the top set of plots, and after LEM selection for the bottom set of plots. The MC truth contamination distributions (dotted lines) show that there is a vanishing anti- $\nu_\mu$  contamination in the  $QP \leq 0$   $\nu_\mu$  sample, but that there is a non-vanishing  $\nu_\mu$  contamination in the  $QP > 0$  anti- $\nu_\mu$  sample (at the 20% level).

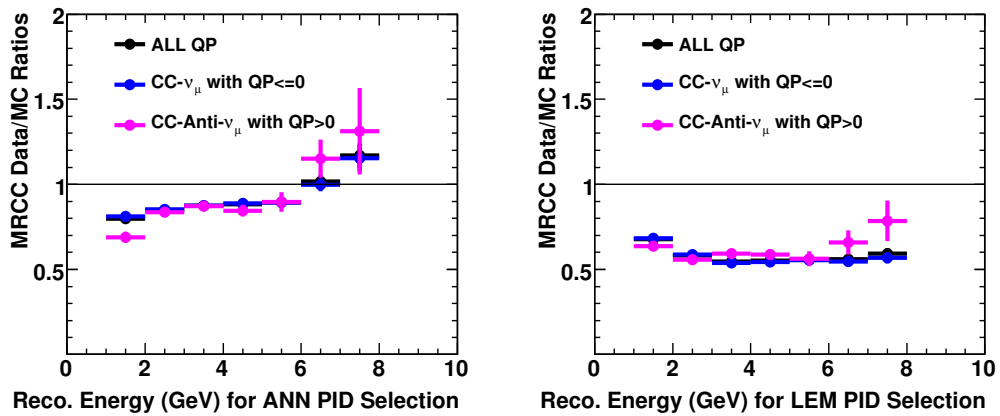


Figure 6.3: MRCC ratios depending on track charge sign cut QP (after ANN or LEM PID selection). The ratios for no QP cut, which contain both neutrinos and anti-neutrinos, lie between the neutrino and anti-neutrino ratios, but closer to the former, exactly as would be expected.

### 6.1.3 $\pi^0$ Content of MRCC Showers

In order to address the third issue in section 6.1, that is whether the  $\pi^0$  content of the MRCC showers is different to the standard NC events, the truth mean numbers and energies of the pions in MRCC and standard NC and  $\nu_\mu$ -CC events were plotted. This is shown in Figure 6.4. This Figure shows that according to MC, the  $\pi^0$  content of the MRCC showers is in fact very similar to the  $\pi^0$  content of NC events. However, this conclusion applies to MC and so, the situation could be somewhat different for data.

## 6.2 Unknown Source of Data/MC Differences

The last issue raised in section 6.1 is important because we need to be assured that most of the data/MC differences seen by the  $\nu_e$  appearance analysis indeed arise from imperfect shower topology modeling. Figures 6.5, 6.6, and 6.7 show distributions of several shower topological variables normalised by area to the data integral. These Figures show that data/MC differences are very similar for both MRCC and standard events. In addition, the double ratios show agreement within  $\sim 5$ -10% in the peaks of the distributions. Additional shower topological variables have been plotted in a similar way



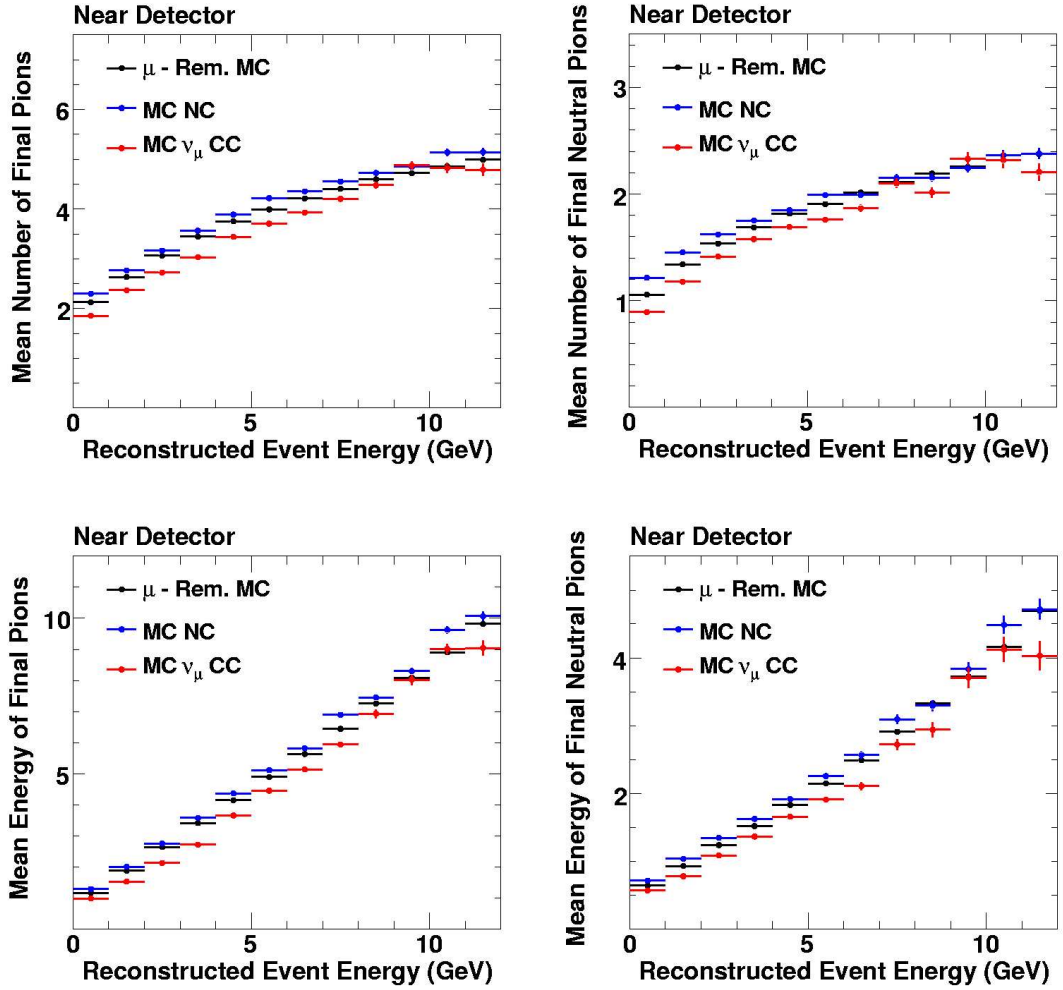


Figure 6.4: Mean numbers and energies of pions in the final states of events in 1GeV slices of reconstructed event energy for MRCC versus standard MC after  $\nu_e$  preselection. The top left plot shows the total average numbers of all pions ( $\pi^+$ ,  $\pi^-$ , and  $\pi^0$ ), the top right shows the average numbers of neutral pions ( $\pi^0$ ) only. The bottom left plot shows the total average energies of all pions ( $\pi^+$ ,  $\pi^-$ , and  $\pi^0$ ), the bottom right shows the average energies of neutral pions ( $\pi^0$ ) only. The black points represent MRCC MC events, the red points standard MC  $\nu_\mu$ -CC events, and the blue points represent standard MC NC events.

and are included in Appendix A. As the MRCC/Std double ratios agree consistently for so many shower topological variables (at the  $\sim 5\text{-}10\%$  level), it is clear that whatever is the source of the data/MC differences for the standard events must also be the source of the differences for the MRCC events. As those differences manifest themselves in shower topology for both the MRCC and standard events in the same way, the data-MC differences can indeed be explained as coming from imperfect shower modeling. Or in other words, the MRCC MC showers differ from the MRCC data showers in a similar manner as the standard MC showers differ from the standard data showers, and hence the most logical explanation for the data/MC differences is imperfect shower modeling.

### 6.3 MRCC Systematic Errors

The systematic errors which are expected to impact the MRCC background estimation method and have been evaluated are:

1. Cross-section errors - axial vector mass for quasi-elastic production processes [69]  
-  $M_A(QE) \pm 15\% - \delta_{M_A(QE)}$
2. Cross-section errors - axial vector mass for resonance production processes -  $M_A(RES)$   
 $\pm 15\% - \delta_{M_A(RES)}$
3. Cross-section errors - relative fraction of resonance to DIS - KNO parameters error  
 $\pm 50\% - \delta_{KNO}$
4. Flux Errors -  $\delta_{Flux}$
5. Hadronic energy scale  $\pm 11\% - \delta_E$
6. Charge of the MRCC shower -  $\delta_{QP}$
7. Additional selection error on the CC- $\nu_e$  beam background events of 30% -  $\delta_{\nu_e}$
8. Inherent differences between the MRCC events and standard events -  $\delta_{MR}$

The systematic errors 1-5 were calculated by changing both the MRCC MC sample and the standard MC sample and rerunning the correction for each systematic error. In particular, for errors 1-4, a special MINOS framework was used to weigh the MC events

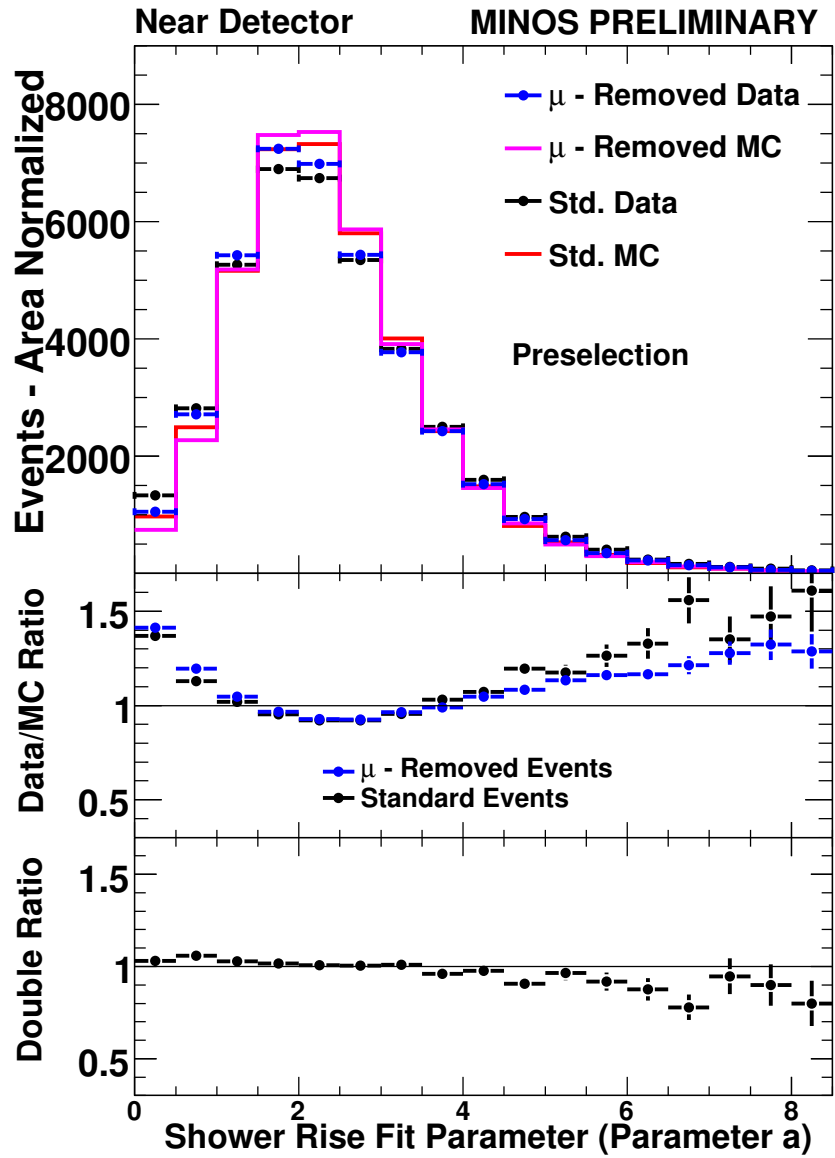


Figure 6.5: ND data/MC MRCC/standard comparisons after  $\nu_e$  preselection cuts. Distributions are normalised to the  $10^{19}$  POT standard data area to show shape differences. The variable shown here is the parameter associated with the rise of the shower.

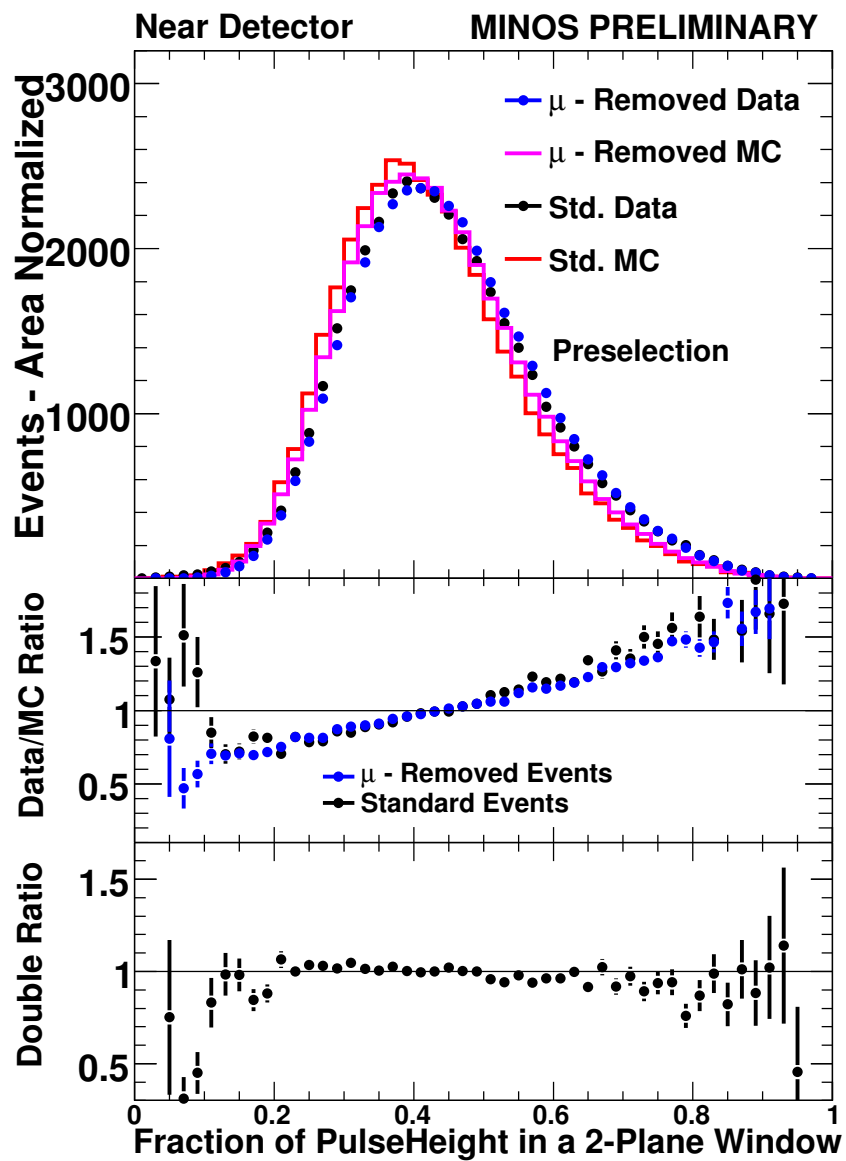


Figure 6.6: ND data/MC MRCC/standard comparisons after  $\nu_e$  preselection cuts. Distributions are normalised to the  $10^{19}$  POT standard data area to show shape differences. The variable shown here is the fraction of pulseheight in a 2-plane-window of the event.

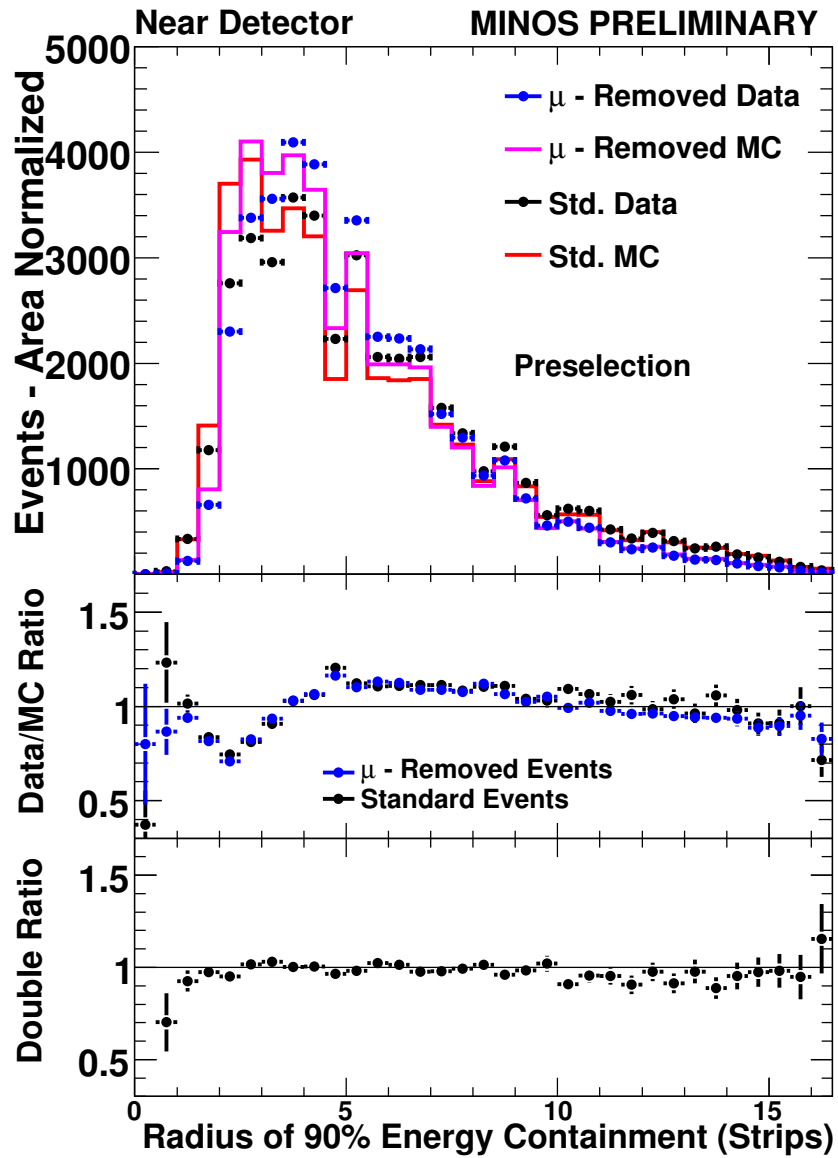


Figure 6.7: ND data/MC MRCC/standard comparison after  $\nu_e$  preselection cuts. Distributions are normalised to the  $10^{19}$  POT standard data area to show shape differences. The variable shown here is the Moliere radius of the shower.

up or down for a given systematic error. Since both MCs were changed by the same percentages, the differences cancel to a large extent resulting in a 5-10% agreement in most bins of energy.

In order to calculate the hadronic energy scale systematic, the event energy of both MCs was varied up or down by 11% and the MRCC decomposition method was run. Most of this systematic error canceled since the change was applied consistently to both standard and MRCC MC.

In order to calculate the systematic error resulting from the charge of the MRCC showers, the MRCC data/MC ratios for neutrinos ( $QP \leq 0$ ) and anti-neutrinos ( $QP > 0$ ) as shown in Figure 6.3 were used to run the MRCC decomposition method. The difference between those two results was divided by two and taken as the systematic error on the MRCC shower charge, since the difference between the neutrino versus anti-neutrino CC showers is two units of charge, while the NC/MRCC difference should only be one unit.

An additional selection error on the beam  $\nu_e$ -CC background events was taken into account by simply adding (subtracting from) to the CC- $\nu_e$  event sample 30% when carrying out the MRCC decomposition. This error was needed to cover uncertainties in the beam  $\nu_e$ -CC flux.

In order to obtain final bin-by-bin systematic uncertainties for errors 1-7, the up-down differences were divided by two for each bin and systematic error.

In order to calculate the systematic error resulting from inherent differences between the MRCC events and standard events, a different procedure was used. The MRCC data, MRCC MC, standard data and standard MC reconstructed energy distributions were area normalised and plotted for both PID selections in a similar fashion as in Figures 6.5, 6.6, and 6.7. The ANN PID selection is shown in Figure 6.8, and LEM PID selection is shown in Figure 6.9. The deviations of the double ratios from 1 were used as the systematic errors on a particular bin. Thus this systematic error corresponds to  $|(1 - \text{double ratio})|$  for each bin, therefore assuming that if the MRCC method were perfect, the double ratios would be equal to one. This way of calculating this systematic error may evolve for a future analysis as it may not be optimal and may even over estimate this error.

Since all the systematic errors are independent of each other, all eight were added in quadrature on a bin-by-bin basis to obtain a total systematic error for each bin  $i$  as follows:

$$\delta_{SYST}(i) = \sqrt{\frac{\delta_{QEMA}^2(i) + \delta_{ResMA}^2(i) + \delta_{KNO}^2(i) + \delta_{Flux}^2(i) + \delta_E^2(i)}{+\delta_{QP}^2(i) + \delta_{MR}^2(i) + \delta_{\nu_e}^2(i)}} \quad (6.1)$$

This allowed the calculation of total bin errors by adding the statistical and systematic errors in quadrature for each bin:

$$\delta_{TOTAL}(i) = \sqrt{\delta_{SYST}^2(i) + \delta_{STAT}^2(i)} \quad (6.2)$$

In order to quote the total final error for a particular systematic, and in order to be conservative and take into account possible correlations between the errors, the individual systematic uncertainties were summed linearly for all bins and then added in quadrature for the final number.

The systematic and statistical errors and their effects are listed in Tables 6.1 and 6.2 for the MRCC results for the ANN and LEM PIDs respectively.

|              |        |      |        |       |               |       |             |       |
|--------------|--------|------|--------|-------|---------------|-------|-------------|-------|
| Error        | Data   | +/-  | NC     | +/-   | $\nu_\mu$ -CC | +/-   | CC- $\nu_e$ | +/-   |
| Total        | 5524.4 | 34.9 | 3674.4 | 191.6 | 1236.2        | 281.3 | 613.8       | 185.1 |
| Stat.        | 5524.4 | 34.9 | 3674.4 | 45.0  | 1236.2        | 59.8  | 613.8       | 18.4  |
| Syst.        | 5524.4 | 0.0  | 3674.4 | 186.3 | 1236.2        | 274.9 | 613.8       | 230.2 |
| Shw.E.       |        |      |        | 128.9 |               | 120.2 |             | 113.5 |
| Flux         |        |      |        | 22.6  |               | 73.9  |             | 51.2  |
| QEMA         |        |      |        | 2.6   |               | 26.1  |             | 24.4  |
| RESMA        |        |      |        | 14.3  |               | 51.9  |             | 38.2  |
| Charge       |        |      |        | 84.1  |               | 84.1  |             | 0.0   |
| MRCC         |        |      |        | 99.7  |               | 99.7  |             |       |
| KNO          |        |      |        | 18.7  |               | 36.8  |             | 38.8  |
| $\nu_e$ Sel. |        |      |        | 0.0   |               | 184.1 |             | 184.1 |

Table 6.1: Final numbers and errors for MRCC decomposition for  $10^{19}$  POT and for reconstructed energy  $1.0\text{GeV} < \text{Reco. Energy} < 8.0\text{GeV}$  and ANN PID  $> 0.7$ . The quoted total error for the beam  $\nu_e$ -CC background (first line of Table) is only taking into account the 30% systematic error (last line), but when the  $\nu_\mu$ -CC background errors are determined, the beam  $\nu_e$ -CC component will vary as well and contribute to the  $\nu_\mu$ -CC uncertainty, as shown in the Table.

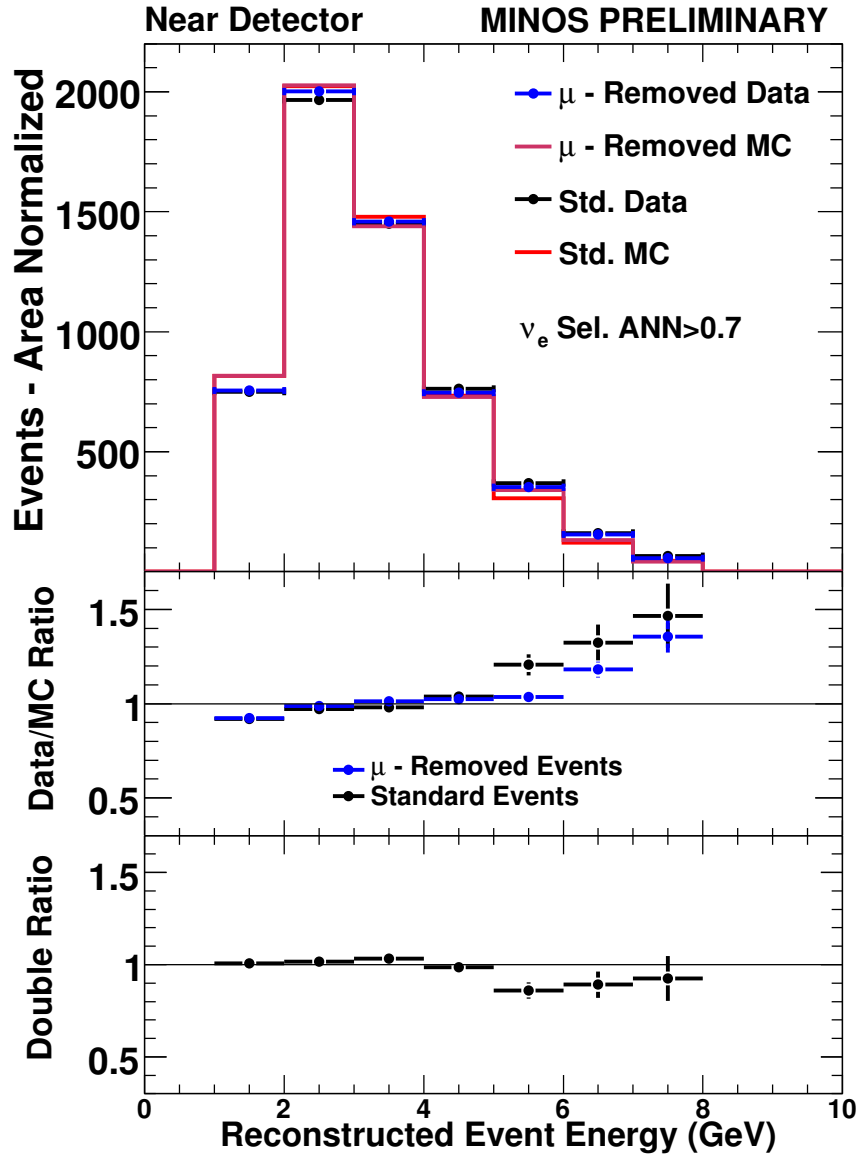


Figure 6.8: MRCC data and MC and standard data and MC distributions for reconstructed energy. Normalised to  $10^{19}$  POT MRCC data area to show inherent shape differences - after ANN PID Selection. Data/MC ratios for both MRCC and standard files are compared, and the resulting double ratios are used as the systematic error arising from from the MRCC process.



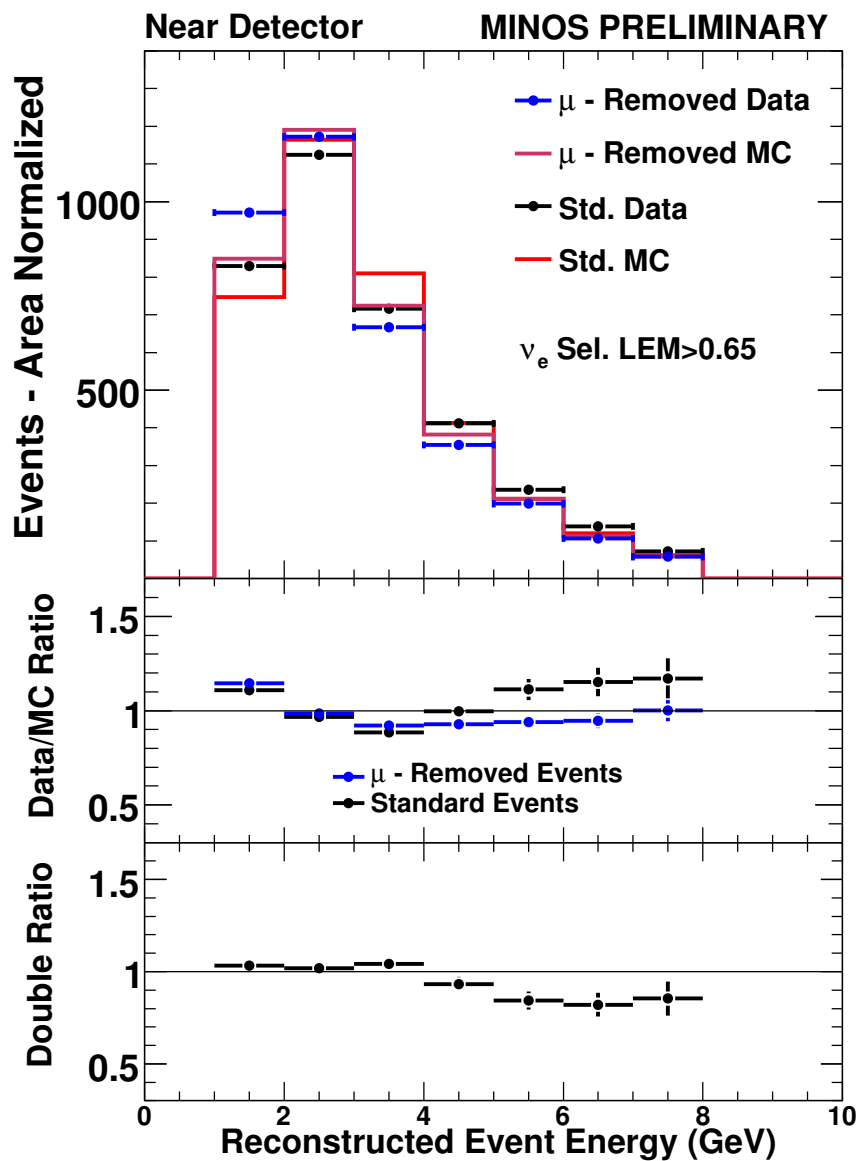


Figure 6.9: MRCC data and MC and standard data and MC distributions for reconstructed energy. Normalised to  $10^{19}$  POT MRCC data area to show inherent shape differences - after LEM PID Selection. Data/MC ratios for both MRCC and standard files are compared, and the resulting double ratios are used as the systematic error arising from from the MRCC process.

| Error        | Data   | +/-  | NC     | +/-   | $\nu_\mu$ -CC | +/-   | CC- $\nu_e$ | +/-   |
|--------------|--------|------|--------|-------|---------------|-------|-------------|-------|
| Total        | 3528.2 | 27.9 | 2169.8 | 135.8 | 788.6         | 243.7 | 569.9       | 171.9 |
| Stat.        | 3528.2 | 27.9 | 2169.8 | 34.6  | 788.6         | 47.8  | 569.9       | 17.7  |
| Syst.        | 3528.2 | 0.0  | 2169.8 | 131.3 | 788.6         | 239.0 | 569.9       | 208.7 |
| Shw.E.       |        |      |        | 63.6  |               | 85.3  |             | 93.0  |
| Flux         |        |      |        | 10.3  |               | 60.8  |             | 50.5  |
| QEMA         |        |      |        | 0.6   |               | 25.0  |             | 25.4  |
| RESMA        |        |      |        | 12.3  |               | 48.4  |             | 36.6  |
| Charge       |        |      |        | 81.0  |               | 81.0  |             | 0.0   |
| MRCC         |        |      |        | 77.8  |               | 77.8  |             | 0.0   |
| KNO          |        |      |        | 18.1  |               | 36.7  |             | 34.1  |
| $\nu_e$ Sel. |        |      |        | 0.0   |               | 171.0 |             | 171.0 |

Table 6.2: Final numbers and errors for MRCC decomposition for  $10^{19}$  POT and for reconstructed energy  $1.0\text{GeV} < \text{Reco. Energy} < 8.0\text{GeV}$  and LEM PID  $> 0.65$ . The quoted total error for the beam  $\nu_e$ -CC background (first line of Table) is only taking into account the 30% systematic error (last line), but when the  $\nu_\mu$ -CC background errors are determined, the beam  $\nu_e$ -CC component will vary as well and contribute to the  $\nu_\mu$ -CC uncertainty, as shown in the Table.

## 6.4 Final Near Detector MRCC Result

The final results for the MRCC method using the two  $\nu_e$  PID cuts are shown in Figures 6.10, 6.11, and 6.12. Figure 6.10 shows how much the MRCC method corrects the original MC NC for two different methods of carrying out the correction, as described in section 5.2. The first method shown is not the chosen one, but it is interesting because it scales the NC and  $\nu_\mu$ -CC backgrounds by the MRCC ratios equally, so by default contains a residual data/MC discrepancy after the correction. The size of this residual difference is potentially an indicator of whether the MRCC method is working well. The second method shown is the official MRCC method and relies on correcting the NC background and calculating the  $\nu_\mu$ -CC background from data minus corr.NC minus beam  $\nu_e$ -CC. This method agrees with the data by construction.

Both methods give similar results within errors. The final MRCC decomposition results are shown in Figures 6.11 and 6.12. and the totals are given in Table 6.3.

It has been demonstrated that the MRCC background estimation method is robust in that it allows correction of the nominal MC to much more closely resemble the data when both the NC and  $\nu_\mu$ -CC backgrounds are scaled equally by the MRCC data/MC ratios. This is irrespective of PID used. In fact, using this non-ideal MRCC method, the residual data-MC discrepancy is around 5% for both PIDs based on the total number

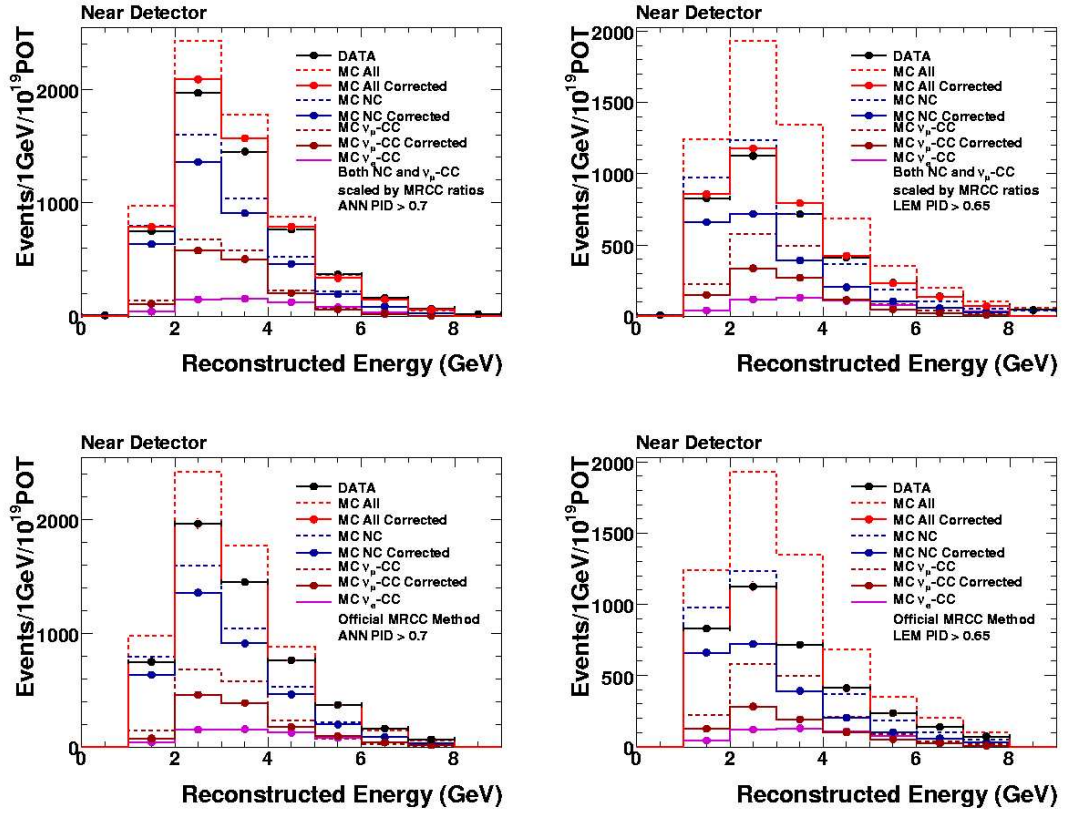


Figure 6.10: Results for the MRCC ND background decomposition of the standard ND data. The dotted lines are before MRCC correction and the solid lines are after correction. The lighter red lines are total MC, the blue lines are the NC background, the dark red lines are the  $\nu_\mu$ -CC background, the purple is the beam  $\nu_e$ -CC background, and the black points are the ND data. The left hand plots are the results for the ANN PID selection, and the right hand plots are the results for the LEM PID selection. The top plots are the results when both the NC and  $\nu_\mu$ -CC backgrounds are scaled by the MRCC data/MC ratios. Those results do not have to agree with the ND data by default, but they almost do. The bottom plots show the results of the official MRCC decomposition method, that is where only the NC background is scaled by the MRCC ratios, and the  $\nu_\mu$ -CC background is obtained by subtracting the beam  $\nu_e$ -CC and corrected NC backgrounds from the ND data. The official MRCC method agrees with the ND data by construction. Errors shown in this Figure are statistical only.

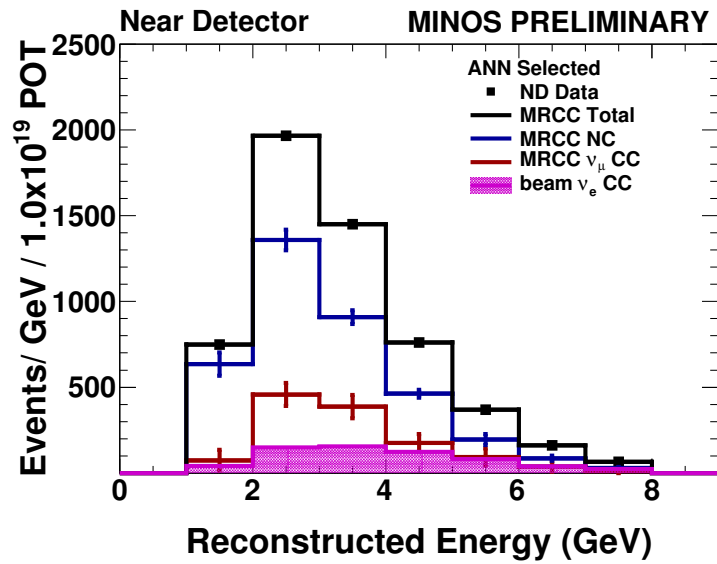


Figure 6.11: Final result for MRCC background decomposition method including systematic and statistical errors in the ND. The ANN PID selection is applied and distributions are normalised to an exposure of  $10^{19}$  POT. Reconstructed event energy is used to derive the MRCC correction ratios.

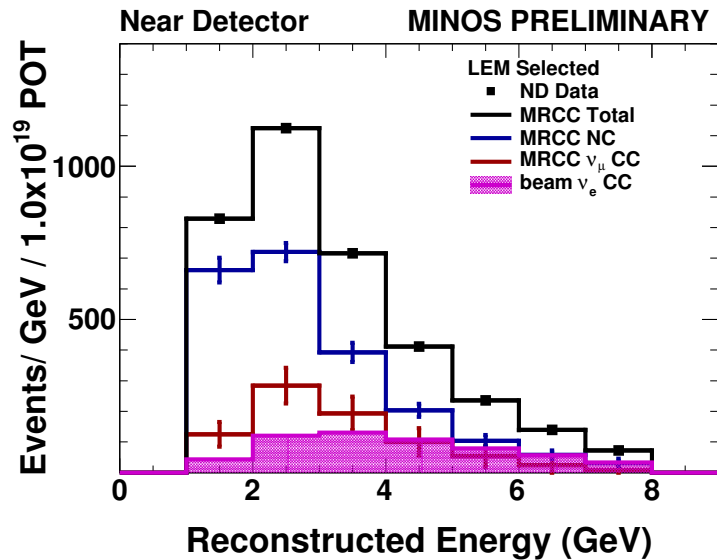


Figure 6.12: Final result for MRCC background decomposition method including systematic and statistical errors in the ND. The LEM PID selection is applied and distributions are normalised to an exposure of  $10^{19}$  POT. Reconstructed event energy is used to derive the MRCC correction ratios.

of calculated background events. This is given discrepancies of  $\sim 16\%$  and  $\sim 40\%$  before the correction. This result gives confidence that the official method (which agrees with data by construction) is correct within systematic errors since the CC and NC showers selected by the  $\nu_e$  analysis as background are expected to be very similar.

| PID | Data   | +/-  | NC     | +/-   | $\nu_\mu$ -CC | +/-   | CC- $\nu_e$ | +/-    |
|-----|--------|------|--------|-------|---------------|-------|-------------|--------|
| ANN | 5524.4 | 34.9 | 3674.4 | 191.6 | 1236.2        | 281.3 | 613.8       | 185.0  |
| LEM | 3528.2 | 27.9 | 2169.8 | 135.8 | 788.6         | 243.7 | 569.9       | 171.88 |

Table 6.3: Final numbers and errors for the MRCC decomposition for  $10^{19}$  POT and for reconstructed energy  $1.0\text{GeV} < \text{Reco. Energy} < 8.0\text{GeV}$

## 6.5 MRCC Mock Data Challenge

In order to verify that the MRCC method works as expected, a mock data challenge was carried out during which a MC with a modified intra-nuclear scattering model was used as data and the MRCC prediction was compared to the data 'truth'. The results of this data challenge can be seen in Figure 6.13. The results show that the MRCC prediction is indeed very close to the mock data truth information and that it appears to scale the backgrounds correctly.

## 6.6 Comparison of MRCC and HOO Results

The MRCC background decomposition method has been shown to work regardless of PID and a mock data challenge was completed successfully. Another important step was to compare the results of the MRCC and HOO background decomposition methods and it was found that both methods agree with each other within errors. This can be seen in Figures 6.6 and 6.6. This means that the ND backgrounds can now be extrapolated to the FD to create the FD predictions against which the FD data will be compared.

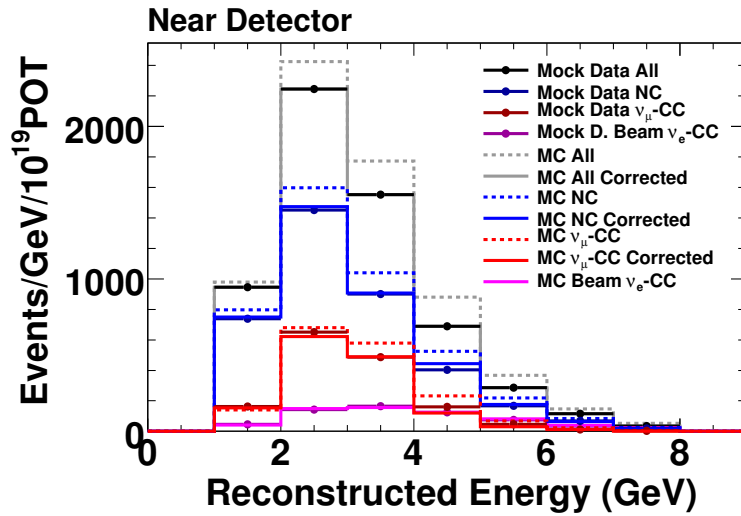


Figure 6.13: MRCC background decomposition method mock data challenge results for the ANN PID Selection. The points are the truth for the mock data, the dotted lines are the uncorrected MC, and the solid lines are the MC corrected using the MRCC decomposition method. The mock data are shown in black (and the darker colours for the individual backgrounds), the total MC is shown in gray, the NC are shown in blue, the  $\nu_\mu$ -CC are shown in red, and the beam  $\nu_e$  are shown in magenta. The truth spectra agree very well with the corrected MC using the MRCC decomposition method.

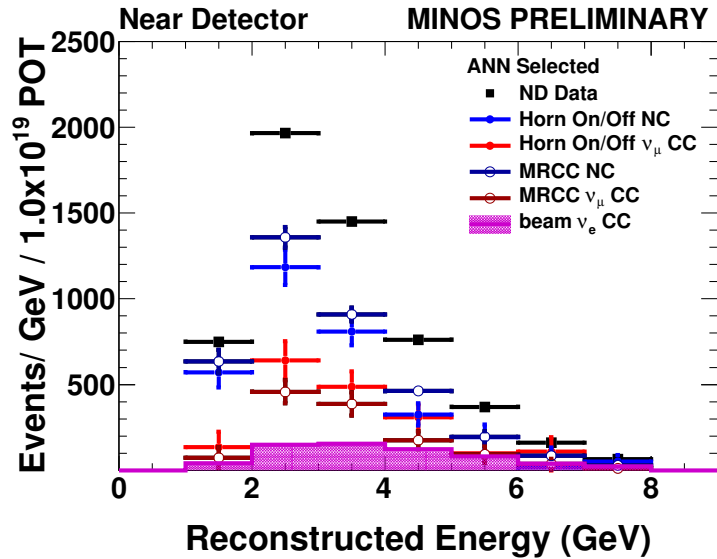


Figure 6.14: Final comparison of ND data decomposition methods for the ANN PID selected sample. Data is shown in black, the beam  $\nu_e$  are shown in purple, the NC are shown in blue, and the  $\nu_\mu$ -CC are shown in red. The HOO method result is shown in the lighter shades of red and blue, and the MRCC result is shown in the darker shades of red and blue. The two methods agree well with each other within errors.

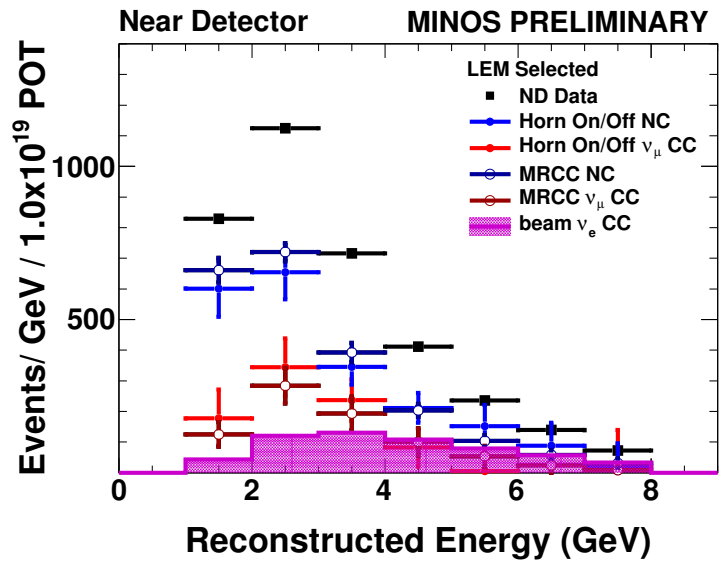


Figure 6.15: Final comparison of ND data decomposition methods for the LEM PID selected sample. Data is shown in black, the beam  $\nu_e$  are shown in purple, the NC are shown in blue, and the  $\nu_\mu$ -CC are shown in red. The HOO method result is shown in the lighter shades of red and blue, and the MRCC result is shown in the darker shades of red and blue. The two methods agree well with each other within errors.

## Chapter 7

# Pathway to a Final $\nu_e$ -Appearance Result

After obtaining the final ND decompositions, they can be extrapolated to the FD in order to predict the background in the absence of  $\nu_e$ -appearance. A number of systematic errors need to be considered in the process, and those will be described in more detail in this chapter. Before opening the blinded data, an important cross-check is carried out in the form of FD sidebands. The FD sidebands are parts of the FD data which are not expected to be subject or are insensitive to  $\nu_e$ -CC-appearance and are unblinded so as to check their compatibility with predictions. They also provide an important cross-check of the background extrapolation mechanisms.

### 7.1 Far Extrapolation of ND Decompositions

In the ND there are three components that need to be extrapolated to the FD. Those are the NC, the  $\nu_\mu$ -CC, and the intrinsic beam  $\nu_e$ -CC backgrounds. The extrapolation also involves taking care of  $\nu_\mu$ -CC oscillations resulting in the appearance of  $\nu_\tau$ -CC events and potentially  $\nu_e$ -CC appearance signal events.

#### 7.1.1 Oscillating the FD MC

In order to predict the FD spectrum, the MC needs to be oscillated. A special  $\nu_e$  analysis oscillation framework [70] is used to calculate oscillation weights for individual



FD MC events. Because the angle  $\theta_{13}$  is expected to be small, a two-flavour approximation cannot be used to calculate oscillation probabilities. Instead, a full three-flavour framework is needed. In an approximation, the probability of  $\nu_e$  appearance in a  $\nu_\mu$  beam in the absence of matter effects is given by [71]:

$$\begin{aligned}
P(\nu_\mu \rightarrow \nu_e) &\simeq \sin^2 2\theta_{13} \sin^2 \theta_{23} \sin^2 \Delta \\
&+ \alpha \Delta \cos \theta_{13} \sin 2\theta_{13} \cos \delta \sin 2\theta_{12} \sin 2\theta_{23} \sin \Delta \cos \Delta \\
&- \alpha \Delta \cos \theta_{13} \sin 2\theta_{13} \sin \delta \sin 2\theta_{12} \sin 2\theta_{23} \sin \Delta \sin \Delta
\end{aligned} \tag{7.1}$$

where:

$$\alpha \equiv \frac{\Delta m_{21}^2}{\Delta m_{31}^2} \tag{7.2}$$

and

$$\Delta \equiv \frac{\Delta m_{31}^2 L}{4E} \tag{7.3}$$

Once matter effects are included, the probability of  $\nu_e$  appearance in a  $\nu_\mu$  beam becomes:

$$\begin{aligned}
P(\nu_\mu \rightarrow \nu_e) &\simeq \sin^2 2\theta_{13} \sin^2 \theta_{23} \frac{\sin^2(A-1)\Delta}{(A-1)^2} \\
&+ 2\alpha \sin \theta_{13} \cos \delta \sin 2\theta_{12} \sin 2\theta_{23} \frac{\sin A\Delta \sin(A-1)\Delta}{A(A-1)} \cos \Delta \\
&- 2\alpha \sin \theta_{13} \sin \delta \sin 2\theta_{12} \sin 2\theta_{23} \frac{\sin A\Delta \sin(A-1)\Delta}{A(A-1)} \sin \Delta
\end{aligned} \tag{7.4}$$

where:

$$A \equiv \frac{G_f N_e L}{\sqrt{2}\Delta} \approx \frac{E}{11 \text{GeV}} \tag{7.5}$$

The oscillation framework used in the  $\nu_e$  analysis includes matter effects and the approximations used to oscillate the events in the  $\nu_e$  analysis are expanded to be precise to  $O(\alpha^\infty \sin \theta_{13})$  and  $O(\alpha^1 \sin \theta_{13}^\infty)$ . Those approximations are included in Appendix B and agree with an exact solution to within 0.1% or better. They assume a uniform electron density in the rock that the neutrino beam passes through of  $2.75 \text{g/cm}^3$ . The values for the various oscillation parameters are taken from the world's best measurements, including the muon neutrino disappearance results from MINOS ( $\Delta m_{32}^2 = 2.43 \times 10^{-3} \text{eV}^2$  and  $\sin^2 2\theta_{23} = 1$ ).

### 7.1.2 Extrapolation Method

In order to predict the FD spectrum, the Far/Near extrapolation method is used [72, 73]. This involves several steps for the different backgrounds. The NC,  $\nu_\mu$ -CC, and  $\nu_e$ -CC backgrounds are extrapolated using Far/Near ratios in bins of reconstructed energy  $i$  and for a particular background  $\alpha$  as follows:

$$N_{i,\alpha}^{FD,predicted} = N_{i,\alpha}^{ND,decomposed} \frac{N_{i,\alpha}^{FD,MC}}{N_{i,\alpha}^{ND,MC}} \quad (7.6)$$

As the beam  $\nu_e$ -CC background in the ND are taken directly from MC, and the FD beam  $\nu_e$ -CC background is also taken from MC, no Far/Near ratio needs to be constructed for the beam  $\nu_e$ -CC background. The NC and  $\nu_\mu$ -CC backgrounds are used as decomposed by the respective ND decomposition methods - the HOO and MRCC methods. Carrying out the Far extrapolation in bins of reconstructed energy reflects the best knowledge of various effects between the Near and the Far detectors, like event rate, energy smearing, efficiencies as a function of reconstructed energy, beam angle differences, fiducial volume differences, relative energy scale and electronics and data readout differences. Using Far/Near ratios also protects against flux errors, cross-section errors, and hadronic model uncertainties as they will cancel out between the two detectors. The Far/Near ratios with systematic errors can be seen in Figures 7.1 and 7.2.

As the  $\nu_e$ -CC appearance analysis is essentially a counting experiment because of low statistics in the FD, and also because the Far/Near ratios for the backgrounds are relatively flat, extrapolating the selected  $\nu_\mu$ -CC background in reconstructed energy only (instead of both true and reconstructed energy) does not impact the results and using true energy would not improve the sensitivity of the final result.

Along with the three backgrounds that are already present in the ND, the FD specific samples need to be predicted. As both the  $\nu_\tau$ -CC background and the  $\nu_e$ -CC signal events start out as  $\nu_\mu$  in the ND, their oscillation probabilities will depend on the latter. The  $\nu_e$ -CC and  $\nu_\tau$ -CC only appear in the FD, and the  $\nu_\mu$  oscillation probabilities need to be fed through during the prediction process, therefore true  $\nu_\mu$  energies become important during the  $\nu_e$ -CC and  $\nu_\tau$ -CC prediction procedure. During this procedure, Far/Near ratios are constructed for the  $\nu_\mu$ -CC events that have long tracks (are selected

by the  $\nu_\mu$ -CC disappearance analysis and are not  $\nu_e$ -CC like). In order to construct those ratios, Equation 7.6 is modified as follows:

$$N_{i,\nu_\mu CC}^{FD,predicted}(E_{reco,i}, E_{true}) = \frac{N_{i,\nu_\mu CC}^{ND,Data}(E_{reco,i})}{N_{i,\nu_\mu CC}^{ND,MC}(E_{reco,i})} N_{i,\nu_\mu CC}^{FD,MC}(E_{reco,i}, E_{true}) \quad (7.7)$$

where  $i$  denotes the event as this extrapolation is performed on an event-by-event basis as a function of both reconstructed and true energy. This prediction is corrected by the efficiency and purity of the  $\nu_\mu$ -CC-like selection, and oscillated and reweighted for  $\nu_\mu$ -CC and  $\nu_{\tau\mu}$ -CC (or  $\nu_e$ -CC) cross-section differences in bins of true energy. The oscillation probability is applied and generates true energy distributions for  $\nu_\tau$ -CC background and for  $\nu_e$ -CC signal events. Those are then converted to reconstructed energy spectra using a 2-D matrix of true versus reconstructed energy. The final step is to correct for selection efficiencies. For the  $\nu_\tau$ -CC background, the selection efficiency is taken from MC. For the  $\nu_e$ -CC signal events however, a special sample called MRE is used to obtain the selection efficiency. MRE [74, 73] (Muon Removed, Electron added) essentially creates a pseudo-signal sample from MRCC events where the original muon has been replaced by a simulated electron of the same four-momentum and the event has been re-reconstructed. This process can be carried out for both data and MC, just like the MRCC. The MRE samples allow the extraction of the efficiency of selecting  $\nu_e$ -CC signal events by correcting the FD standard MC efficiency by the data/MC MRE efficiency ratios in the ND as follows:

$$\epsilon_{(Std,i)}^{FD,pred.} = \frac{\epsilon_{(MRE,i)}^{ND,Data}}{\epsilon_{(MRE,i)}^{ND,MC}} \times \epsilon_{(Std,i)}^{FD,MC} \quad (7.8)$$

The MRE PID distributions in the ND are shown in Figure 7.3 and the ND MRE and FD standard efficiencies are shown in Figure 7.4. As can be seen from the latter, the MRE correction is small, especially in the peak region of the distributions. For the LEM PID it is a little larger on a bin-by-bin basis than for the ANN PID. The results are shown in Table 7.1.

| Cut     | ND MRE Data | ND MRE MC | FD Standard MC | FD Predicted |
|---------|-------------|-----------|----------------|--------------|
| ANN PID | 42.2%       | 44.3%     | 41.5%          | 41.4±1.5%    |
| ANN PID | 44.4%       | 42.2%     | 47.7%          | 45.2±1.6%    |

Table 7.1: Average MRE selection efficiency for the energy range 1-8GeV [73].

### 7.1.3 Systematic Errors in the FD Extrapolation

Since MINOS has two functionally identical detectors, many systematic errors will cancel out between the two. A number of systematic errors were evaluated during the analysis ([75]) but most of them proved to be small, especially considering that the  $\nu_e$  analysis is statistics limited rather than systematics limited at this stage. The systematic errors were calculated by using modified MC samples for both detectors and evaluating the Far/Near ratios using those samples. These systematic errors are the extrapolation systematic errors and are separate from the systematic errors that were evaluated for the ND background decomposition methods. Extrapolation systematic uncertainties included:

- cross-section and flux uncertainties
- hadronic shower uncertainties like intranuclear rescattering, hadron multiplicity and hadronic modeling
- calibration uncertainties:
  - absolute energy scale
  - hadronic energy scale
  - relative energy calibration
  - photomultiplier tube gains uncertainties
  - strip-to-strip uncertainties
  - attenuation uncertainties
- errors from the  $\nu_\tau$ -CC and signal  $\nu_e$ -CC extrapolation, for example the hadronic shower energy uncertainty affecting  $\nu_\mu$ -CC events
- detector modeling such as cross-talk between the readout channels and low pulse height hit modeling
- other errors are overall normalisation, beam intensity, preselection reconstruction effects [76], MRE errors

In order to calculate the different systematic errors, several methods were used. One method was to apply weights to MC in order to simulate changes. This was used for example for cross-section errors and hadronic modeling errors. A second method was

to shift the reconstructed quantities by a certain amount for both Near and Far MC - one such quantity was for example the event energy. Finally, the last method consisted of generating completely new MC that had for example cross-talk turned off. For each systematic error, the modified MC replaced the standard MC in the Far/Near extrapolation and the resulting shift in the total number of predicted events was determined and assigned to be the error on a given systematic.

The four biggest systematic errors were found to be relative energy, PMT gains, cross-talk, and normalisation. This can be seen in Figure 7.5.

After calculating the extrapolation errors, a total systematic error was calculated by combining them with the systematic errors from the ND background decomposition methods. Table 7.2 gives a summary of the systematic errors for the ANN PID, for both the HOO and the MRCC decomposition methods.

| PID:<br>Method:                                | ANN<br>HOO | ANN<br>MRCC | LEM<br>HOO | LEM<br>MRCC |
|--|------------|-------------|------------|-------------|
| Stat. Error from ND Separation                 | 2.3%       | 1.0%        | 3.2%       | 1.2%        |
| Syst. Error from ND Separation                 | 2.9%       | 3.3%        | 4.4%       | 4.1%        |
| FD Extrapolation                               | 6.4%       | 6.2%        | 10.6%      | 10.6%       |
| Total Sum in Quadrature                        | 7.4%       | 7.1%        | 12.0%      | 11.4%       |
| Syst. Error on Signal                          | 7.7%       | 7.7%        | 9.0%       | 9.0%        |
| Simple Stat. Error on<br>Background Prediction | 19.4%      | 18.9%       | 21.6%      | 21.3%       |

Table 7.2: Final systematic errors for both PIDs for the HOO and MRCC background decomposition methods [73]. Normalised to an exposure of  $3.14 \times 10^{20}$  POT.

#### 7.1.4 Final Predictions

The predictions of the FD extrapolation can be found in Table 7.3 for both background decomposition methods and both PIDs. It is interesting to note that the reason the MRCC method predicts slightly higher event numbers in the FD than the HOO method is that its ND decomposition has more NC events compared to the HOO method. NC interactions are insensitive to neutrino flavour, thus less of the background disappears on its way to the FD as the MRCC method predicts less  $\nu_\mu$ -CC background. Despite the small differences, the predictions of the two methods agree with each other well within errors.

| Selection and Method | Total Prediction | Stat. Error | Syst. Error | NC   | $\nu_\mu$ -CC | Beam $\nu_e$ | $\nu_\tau$ -CC |
|----------------------|------------------|-------------|-------------|------|---------------|--------------|----------------|
| ANN PID MRCC         | 28.1             | 5.30        | 2.14        | 21.3 | 3.6           | 2.1          | 1.0            |
| ANN PID HOO          | 26.5             | 5.15        | 1.91        | 18.2 | 5.2           | 2.1          | 1.0            |
| LEM PID MRCC         | 22.0             | 4.69        | 2.35        | 15.5 | 2.8           | 2.7          | 1.0            |
| LEM PID HOO          | 21.5             | 4.64        | 2.45        | 14.8 | 2.9           | 2.7          | 1.0            |

Table 7.3: Final FD predictions using the two different data decomposition methods and the two different PIDs. Normalised to an exposure of  $3.14 \times 10^{20}$  POT.

## 7.2 Sidebands

Of particular concern in the extrapolation are potential Far/Near differences that have not been accounted for by the MC Far/Near ratios. If one looks at the Far/Near extrapolation method in Equation 7.6, this method of predicting the FD backgrounds is equivalent to essentially scaling the oscillated FD MC by the ND data/MC ratios after decomposition. Mismodeling differences between the Near and the Far detectors could introduce errors. In order to carry out checks associated with this it was decided to look at three different sideband samples in the FD. Sideband samples are selections of the FD data that have very little or no signal, but are able to give information about the expected FD signal and backgrounds without looking at the final FD data sample. The sidebands that were looked at were the anti-PID sideband - events not selected by the PIDs, the MRCC sideband, and the MRE sideband<sup>1</sup>. Prior to unblinding the sidebands, it was decided that if the total number of events seen in the FD is within  $2\sigma$  of the predictions, then it will be deemed that one should proceed with the final data unblinding.

### 7.2.1 Anti-PID Sideband

The anti-PID sideband consisted of standard events after preselection, but below a cut of 0.55 for both the ANN and the LEM PIDs. This cut value was chosen so as to minimize any potential signal within this sideband [77]. Events which did not acquire a PID value (so were set to a default value of -9999) were removed from the anti-PID sideband. The anti-PID sideband was not a strong probe of data/MC differences between the ND and the FD because it was concentrated on PID regions where the data and

<sup>1</sup>Sidebands using data from the high energy beam configuration were considered in the early stages of the  $\nu_e$  analysis, but it was shown that there were too few events (single digit numbers) in those samples to be able to draw any useful conclusions about the expected FD backgrounds.

MC agree fairly well. This sideband was, however, very important to exercise the whole analysis chain. First, the HOO and MRCC ND decompositions had to be carried out, then those were extrapolated to the FD using the extrapolation method described in Section 7.1. Only the most important systematic errors were included in the interest of time (absolute / relative energy, gains, attenuation, strip-to strip calibration, linearity, normalisation, low PH strips, cross-talk).

| PID | Method | Total    | NC                                      | $\nu_\mu$ -CC                           | Beam $\nu_e$ -CC |
|-----|--------|----------|---|---|------------------|
| ANN | MC     | 28273    | 12686                                   | 15148                                   | 439              |
|     | HOO    | 26245±76 | 10801 <sup>+2392</sup> <sub>-1632</sub> | 15005 <sup>+1735</sup> <sub>-2390</sub> | 439±132          |
|     | MRCC   | 26245±76 | 11122±408                               | 14684±418                               | 439±132          |
| LEM | MC     | 27427    | 12809                                   | 14139                                   | 479              |
|     | HOO    | 27533±78 | 12594 <sup>+1990</sup> <sub>-1968</sub> | 14460 <sup>+2048</sup> <sub>-1945</sub> | 479±144          |
|     | MRCC   | 27533±78 | 12432±350                               | 14622±365                               | 479±144          |

Table 7.4: ND data decompositions / MC event numbers for the anti-PID sidebands. The uncertainties on the total numbers are statistical only, whereas the errors on the individual components include systematic errors as well.

As can be seen in Table 7.4 and Figure 7.6, the ND decomposition methods did not have to correct much for the anti-PID sidebands. The results of the two decomposition methods naturally agreed well - this is shown in Figure 7.7 - and consequently, both FD predictions were very similar. This can be seen in Table 7.5.

| PID | Method | Total      | NC       | $\nu_\mu$ -CC | Beam $\nu_e$ | $\nu_\tau$ -CC | Signal  |
|-----|--------|------------|----------|---------------|--------------|----------------|---------|
| ANN | MC     | 141.0      | 85.6     | 50.6          | 2.3          | 2.6            | 4.5     |
|     | HOO    | 131.8±8.4  | 72.7±3.7 | 54.5±3.5      | 2.2±0.4      | 2.4±1.2        | 4.9±0.3 |
|     | MRCC   | 130.2±7.5  | 77.0±4.5 | 48.6±3.2      | 2.2±0.4      | 2.4±1.2        | 4.9±0.3 |
| LEM | MC     | 157.2      | 99.4     | 52.7          | 2.4          | 2.9            | 6.4     |
|     | HOO    | 156.8±12.6 | 96.9±7.4 | 54.9±4.2      | 2.4±0.5      | 2.7±1.4        | 6.8±0.6 |
|     | MRCC   | 154.0±10.9 | 95.3±7.5 | 53.6±3.9      | 2.4±0.5      | 2.7±1.4        | 6.8±0.6 |

Table 7.5: FD default MC and data prediction event numbers for the anti-PID sidebands. The signal prediction was obtained using standard oscillation parameters ( $\Delta m_{32}^2 = 2.4 \times 10^{-3} eV^2$ ,  $\sin^2(2\theta_{23}) = 1.0$ ,  $\sin^2(2\theta_{13}) = 0.15$ ), and all numbers are scaled to the full FD data set of  $3.14 \times 10^{20}$  POT. The errors shown are systematic but do not include errors on the oscillation parameters.

An additional systematic error was evaluated on the FD prediction due to the uncertainties on the oscillation parameters, which directly contribute to the uncertainties on the large  $\nu_\mu$ -CC background in the anti-PID samples. Those additional errors are summarised in Table 7.6.

| PID | Method | Total | Syst. Error | Error on $\delta(\Delta m_{32}^2)$ | Error on $\delta(\sin^2(2\theta_{32}))$ | Total Syst. Error |
|-----|--------|-------|-------------|------------------------------------|---|-------------------|
| ANN | HOO    | 131.8 | $\pm 8.4$   | $\pm 2.5$                          | +2.8                                    | $\pm 9.2$         |
|     | MRCC   | 130.2 | $\pm 7.5$   | $\pm 2.5$                          | +2.8                                    | $\pm 8.4$         |
| ANN | HOO    | 156.8 | $\pm 12.6$  | $\pm 3.0$                          | +2.4                                    | $\pm 13.2$        |
|     | MRCC   | 154.0 | $\pm 10.9$  | $\pm 3.0$                          | +2.4                                    | $\pm 11.6$        |

Table 7.6: FD data prediction systematic errors summary for the anti-PID sidebands. The total systematic errors were obtained by adding all three contributions in quadrature.

After opening the FD anti-PID sideband, the observed data was compared with the predictions, as can be seen in Table 7.7. The data/prediction comparisons as a function of PID can be seen in Figures 7.8 and 7.9. For the ANN PID, 146 events were selected after all data quality and  $\nu_e$ -analysis preselection cuts, which was 0.97 standard deviations away from the expectation of  $131.8 \pm 11.5(\text{stat.}) \pm 9.2(\text{syst.})$  events. For the LEM PID, 176 events were selected, which was 1.06 standard deviations away from the expectation of  $156.8 \pm 12.5(\text{stat.}) \pm 13.2(\text{syst.})$  events. The results were consistent with the expectation within  $2\sigma$  and so were within the parameters defined prior to opening this sideband. Nevertheless, it should be remembered that the anti-PID sideband does not probe the PID selected region directly, so it only allows the identification of problems with the extrapolation method or the preselection of events in the FD.

| PID | Data | Method | Prediction | Stat. Error | Syst. Error | Total Error | Diff. in $\sigma$ |
|-----|------|--------|------------|-------------|-------------|-------------|-------------------|
| ANN | 146  | MC     | 141.0      | $\pm 11.9$  |             |             | +0.42(stat.)      |
|     | 146  | HOO    | 131.8      | $\pm 11.5$  | $\pm 9.2$   | 14.7        | +0.97             |
|     | 146  | MRCC   | 130.2      | $\pm 11.4$  | $\pm 8.4$   | 14.2        | +1.11             |
| LEM | 176  | MC     | 157.2      | $\pm 12.5$  |             |             | +1.50(stat.)      |
|     | 176  | HOO    | 156.8      | $\pm 12.5$  | $\pm 13.2$  | 18.2        | +1.06             |
|     | 176  | MRCC   | 154.0      | $\pm 12.4$  | $\pm 11.6$  | 17.0        | +1.29             |

Table 7.7: FD data versus background prediction event numbers for the anti-PID sidebands. All numbers are scaled to the full FD data set of  $3.14 \times 10^{20}$  POT. The statistical and systematic errors are added in quadrature for the final comparison.

## 7.2.2 MRCC Sideband

The second sideband that was opened was the MRCC sideband. It is important because it probes the behaviour of the data in the selected sample region of the PIDs,



where the data/MC discrepancy is higher. The MRCC FD sideband is unique because the high-PID sample is made purely of background events not containing any signal  $\nu_e$ -CC events. It is the only sideband that allows any judgement about potential background issues that may be encountered in the high PID region in the FD standard data.

The extrapolation procedure for the MRCC sideband is very simple. Essentially, the ND MRCC data and MC samples are used in conjunction with a sample of FD MRCC MC to predict the number of FD MRCC data events on a bin-by-bin basis according to the formula:

$$DATA_i^{FD} = MC_i^{FD} \left( \frac{DATA_i^{ND}}{MC_i^{ND}} \right) \quad (7.9)$$

where  $i$  is the bin number (in reconstructed energy) and all samples are MRCC samples as explained above. It was decided to open several MRCC sideband samples for both PIDs. Those were: the preselected sample, events above an ANN PID of 0.5, events above an LEM PID of 0.5, events above an ANN PID of 0.7, and events above an LEM PID of 0.8 (originally the best value for the LEM selection cut). In order to calculate systematic errors for the MRCC sideband predictions, it was assumed that the systematic error envelopes would be similar to the NC background envelopes from the standard Far/Near extrapolation. This is reasonable because the statistical errors are much larger than the systematic errors. The NC background error envelopes are shown in Figures 7.1 and 7.2. Unfortunately, they were not available at preselection level. In order to also take into account systematic errors in the  $\nu_\mu$  oscillation parameters, the latter were varied within their errors. Finally, all systematic errors were added in quadrature. The numbers of expected events in the FD are shown in Table 7.8.

| Cut     | MC    | Prediction | Stat. Error | Syst. Error | $\Delta m_{32}^2$ Error | $\sin^2(2\theta_{23})$ Error | Total Syst. Error | Total Error |
|---------|-------|------------|-------------|-------------|-------------------------|------------------------------|-------------------|-------------|
| Presel  | 254.3 | 233.3      | 15.3        | NA          | 3.8                     | +2.4                         | 4.5               | 15.9        |
| ANN>0.5 | 95.1  | 89.0       | 9.4         | 8.0         | 1.7                     | +1.0                         | 8.2               | 12.5        |
| LEM>0.5 | 46.1  | 30.5       | 5.5         | 2.7         | 0.9                     | +0.5                         | 2.9               | 6.2         |
| ANN>0.7 | 33.1  | 28.6       | 5.3         | 2.3         | 0.5                     | +0.3                         | 2.4               | 5.8         |
| LEM>0.8 | 14.5  | 7.4        | 2.8         | 0.8         | 0.1                     | +0.1                         | 0.8               | 2.9         |

Table 7.8: FD MRCC sideband predictions for different cut levels. All numbers are scaled to the full FD data set of  $3.14 \times 10^{20}$  POT.

The final results of the FD MRCC sideband are shown in Table 7.9. It can be seen that even though for the ANN PID above 0.7, the numbers agree within  $2\sigma$ , for the

LEM PID at an ideal cut of 0.8, they do not. As a result of the sideband results, it was decided that the ANN PID would be the primary  $\nu_e$  analysis PID. The LEM PID cut was relaxed to a value of 0.65 (the results for this cut are also shown in Table 7.9), which was less optimal (impacting the LEM PID sensitivity), but which pushed the LEM PID sideband discrepancy below  $2\sigma$  and still allowed to use the LEM PID as a cross-check for the results of the ANN PID.

| Cut      | MC    | Prediction | Stat. Error | Syst. Error | Total Error | Data | Difference in $\sigma$ |
|----------|-------|------------|-------------|-------------|-------------|------|------------------------|
| Presel   | 254.3 | 233.3      | 15.3        | 4.5         | 15.9        | 225  | -0.52                  |
| ANN>0.5  | 95.1  | 89.0       | 9.4         | 8.2         | 12.5        | 84   | -0.40                  |
| LEM>0.5  | 46.1  | 30.5       | 5.5         | 2.9         | 6.2         | 36   | 0.88                   |
| ANN>0.7  | 33.1  | 28.6       | 5.3         | 2.4         | 5.8         | 39   | 1.78                   |
| LEM>0.8  | 14.5  | 7.4        | 2.8         | 0.8         | 2.9         | 16   | 3.03                   |
| LEM>0.65 | 28.7  | 16.9       | 4.1         | 1.6         | 4.4         | 25   | 1.84                   |

Table 7.9: FD MRCC sideband results for different cut levels. The results for the new LEM PID cut of 0.65 are also given. All numbers are scaled to the full FD data set of  $3.14 \times 10^{20}$  POT.

The comparison of the MRCC FD prediction with the FD MRCC data can be seen in Figures 7.10 and 7.11. No obvious issues can be seen in those figures since the statistics of the samples are too small. There does however appear to be a trend towards the data being high at high LEM PID values.

## 7.3 Discussion of the MRCC Sideband Results

Even though the discrepancy for the MRCC sideband was less than  $2\sigma$  for the ANN PID selected events, for the optimal LEM selection, it was discrepant by  $3\sigma$ . Because of this, many investigations were carried out to determine whether this discrepancy was systematic in nature. These checks are explained in more detail in this section.

### 7.3.1 Applying a Charge Sign Cut

One of the checks carried out was to apply a charge sign cut on the original MRCC tracks to select neutrino events only instead of both neutrino and anti-neutrino events as for the default MRCC sideband. The results are summarized in Table 7.10 and show

that cutting on the charge sign did not help to resolve the problem with the sideband excess.

| Cut     | MC    | Prediction | Stat. Error | Syst. Error | Total Error | Data | Difference in $\sigma$ |
|---------|-------|------------|-------------|-------------|-------------|------|------------------------|
| Presel  | 223.2 | 206.0      | 14.4        | NA          | 14.4        | 205  | -0.07                  |
| ANN>0.5 | 83.5  | 78.6       | 8.9         | 7.0         | 11.3        | 78   | -0.05                  |
| LEM>0.5 | 40.6  | 26.9       | 5.2         | 2.4         | 5.7         | 34   | 1.25                   |
| ANN>0.7 | 29.3  | 25.4       | 5.0         | 2.0         | 5.4         | 36   | 1.95                   |
| LEM>0.8 | 12.8  | 6.4        | 2.5         | 0.7         | 2.6         | 14   | 2.90                   |

Table 7.10: FD MRCC sideband results for different cut levels for neutrinos only (original QP<=0). All numbers are scaled to the full FD data set of  $3.14 \times 10^{20}$  POT.

### 7.3.2 MRCC Far Detector Distributions

Apart from the PID distributions, other distributions were examined in the FD MRCC in order to gauge what could be responsible for the observed discrepancy. Four of those distributions are shown in Figure 7.12. Those distribution are the X, Y, and Z event vertices, and the  $\cos \theta$  of the original track angle with respect to the Z-axis after preselection. More distributions are shown in Appendix C. Unfortunately, after looking at many distributions, no obvious problems could be identified via this investigative route. Since the statistics of the samples were very small, it was hard to judge whether differences were purely statistical or were systematic in nature, as no obvious trends were found.

### 7.3.3 MRCC Sideband Prediction in One Detector

As the observed discrepancy for the LEM PID above 0.8 was  $3\sigma$ , the question arose whether the MRCC sideband extrapolation method described in Equation 7.9 was valid. As a test of this MRCC sideband prediction method, a test was carried out within one detector to eliminate Far/Near differences. To obtain a second MRCC sample, the horn off data and MC samples where processed through the MRCC and yielded low statistics samples within the ND. The prediction equation thus became:

$$DATA_i^{HornOff} = MC_i^{HornOff} \left( \frac{DATA_i^{HornOn}}{MC_i^{HornOn}} \right) \quad (7.10)$$

where  $i$  is the bin (in reconstructed energy) and all samples are MRCC samples in the ND. The investigation was carried out for both PIDs. The results of this verification can be seen in Figure 7.13. It can clearly be seen that within one detector, there is no tendency for this extrapolation method to overestimate (or underestimate) the predicted number of events, as is also shown in Table 7.11. Therefore, whatever the source of the discrepancy, it is more likely to be linked to a mismodeled Far/Near difference than a problem with the method itself.

| Cut      | MC    | Prediction | Data  | Stat. Error | Data/MC | Data/Prediction |
|----------|-------|------------|-------|-------------|---------|-----------------|
| Presel   | 22682 | 21045      | 20849 | 144         | 0.92    | 0.99            |
| ANN>0.7  | 3658  | 3191       | 3167  | 56          | 0.87    | 0.99            |
| LEM>0.65 | 2612  | 1513       | 1602  | 40          | 0.61    | 1.06            |

Table 7.11: MRCC sideband extrapolation within one detector results for different cut levels using the horn off ND MRCC data in place of the FD MRCC data. All numbers are scaled to a ND exposure of  $10^{19}$  POT.

### 7.3.4 MRCC Sideband by MINOS Data Run Periods

Another check of the MRCC FD sideband consisted of breaking down the data into different MINOS run periods. This was for two reasons. First, to check whether the observed excess could be due to a statistical fluctuation rather than an unknown systematic effect. Second, to look at a portion of newly taken Run III data (all the sidebands and the final box contained only data from Run I and Run II only by default) so as to verify whether any trends persist into the new data taking period. The breakdowns of the FD MRCC data by run period can be seen in Table 7.12.

It was established that, at least for the ANN PID, the sideband discrepancy was probably due to a statistical effect. This is because for most run periods, the sideband data agreed well with the prediction within a few percent, except for run period IIa, which was high for all cut levels. Because of low event numbers, it is very hard to draw conclusions about run IIa but the most logical conclusion is a statistical fluctuation. The additional run III data supported this conclusion for the ANN PID as well.

For the LEM PID however, it was hard to reach the same conclusion since the sideband data was higher than the prediction for all run periods, thus possibly pointing to a systematic effect. Adding the newest run III data for the LEM PID still showed that the

|                     | Run<br>POT $\times 10^{20}$ | Run I<br>1.21 | Run IIa<br>1.22 | Run IIb<br>0.71 | Run III<br>0.99 | Total<br>4.12 |
|---------------------|-----------------------------|---------------|-----------------|-----------------|-----------------|---------------|
| Prediction          | Preselection                | 89.8          | 90.9            | 52.7            | 73.2            | 306.4         |
|                     | ANN>0.7                     | 11.0          | 11.1            | 6.5             | 9.0             | 37.6          |
|                     | LEM>0.65                    | 6.5           | 6.6             | 3.8             | 5.3             | 22.2          |
| Data                | Preselection                | 73            | 97              | 55              | 74              | 299           |
|                     | ANN>0.7                     | 11            | 21              | 7               | 8               | 47            |
|                     | LEM>0.65                    | 8             | 12              | 5               | 10              | 35            |
| Ratio<br>Data/Pred. | Preselection                | 0.81          | 1.07            | 1.04            | 1.01            | 0.98          |
|                     | ANN>0.7                     | 1.00          | <b>1.88</b>     | 1.08            | 0.89            | <b>1.25</b>   |
|                     | LEM>0.65                    | <b>1.23</b>   | <b>1.82</b>     | <b>1.31</b>     | <b>1.89</b>     | <b>1.58</b>   |

Table 7.12: FD MRCC sideband numbers for different cut levels broken down by MINOS runs. The first set of lines shows the predictions for the run subsets, the second set of lines shows the corresponding data numbers, and the final set of numbers shows the corresponding data/prediction ratios. The high ratios are highlighted in bold.

data was above the prediction for this data set too.

## 7.4 MRE Sideband

The final sideband to be opened, the MRE sideband, was used to evaluate whether there were any problems with signal selection efficiency in the FD. The prediction procedure was similar to the MRCC sideband prediction, although, since the data/MC differences in the ND samples are smaller for the MRE sample, the correction applied to the FD MRE MC was small. The MRE sideband was found to agree well with the predictions and the results are shown in Table 7.13 and Figure 7.14.

| Cut | Data | Prediction | Total<br>Error | Difference<br>in $\sigma$ |
|-----|------|------------|----------------|---------------------------|
| ANN | 159  | 151.4      | 12.3           | 0.6                       |
| LEM | 158  | 144.0      | 12.0           | 1.2                       |

Table 7.13: FD MRE sideband results for the two PIDs. All numbers are scaled to the full FD data set of  $3.14 \times 10^{20}$  POT.

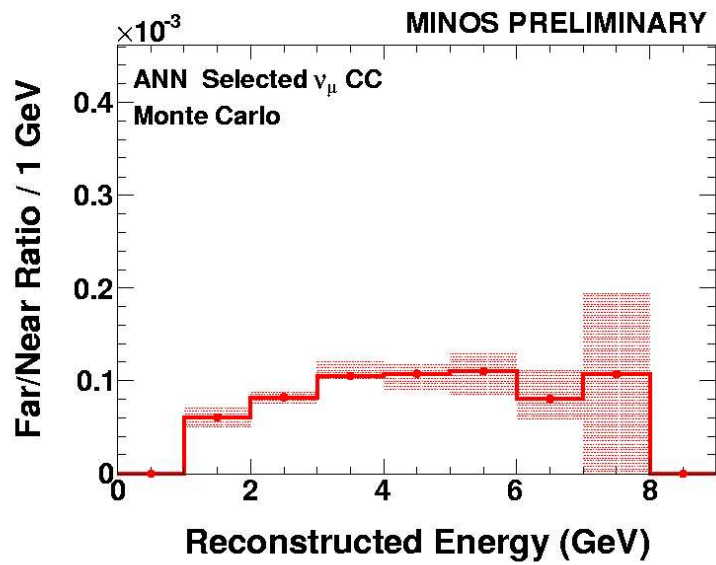
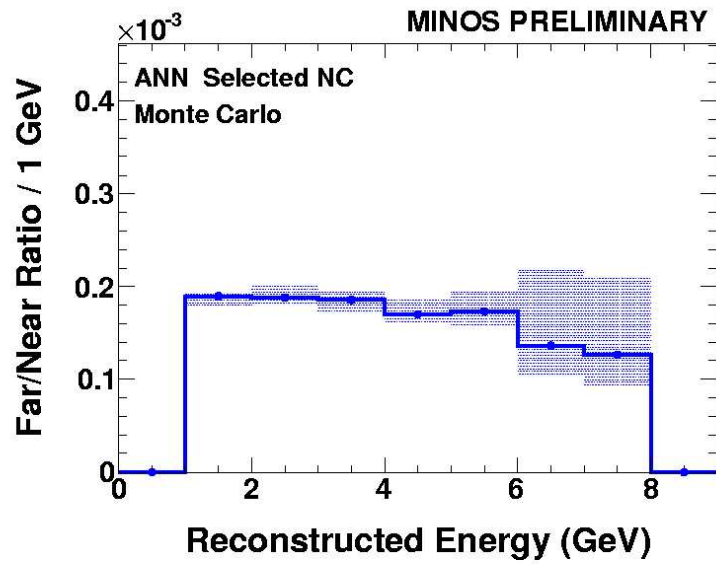


Figure 7.1: Far/Near ratios with systematic error bands for the NC and  $\nu_\mu$ -CC backgrounds for the ANN PID selection.

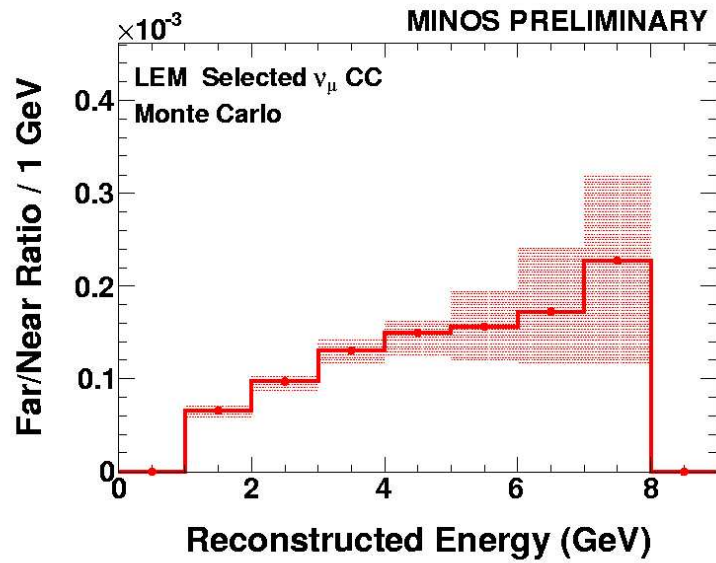
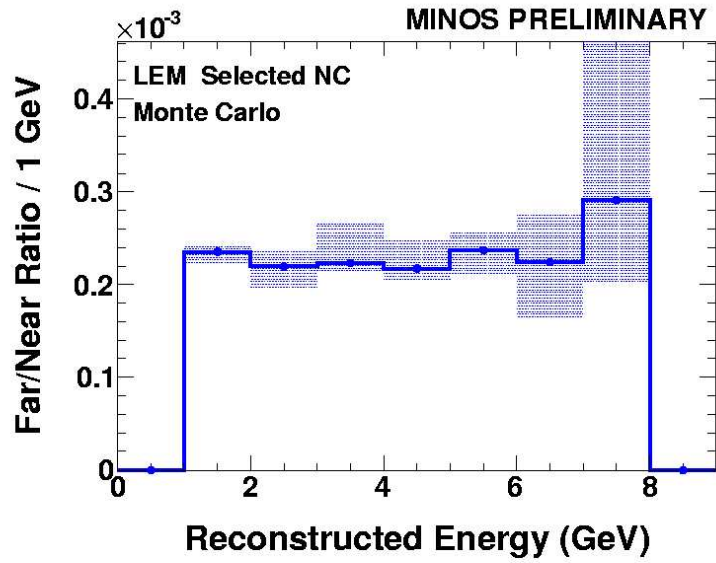


Figure 7.2: Far/Near ratios with systematic error bands for the NC and  $\nu_\mu$ -CC backgrounds for the LEM PID selection.

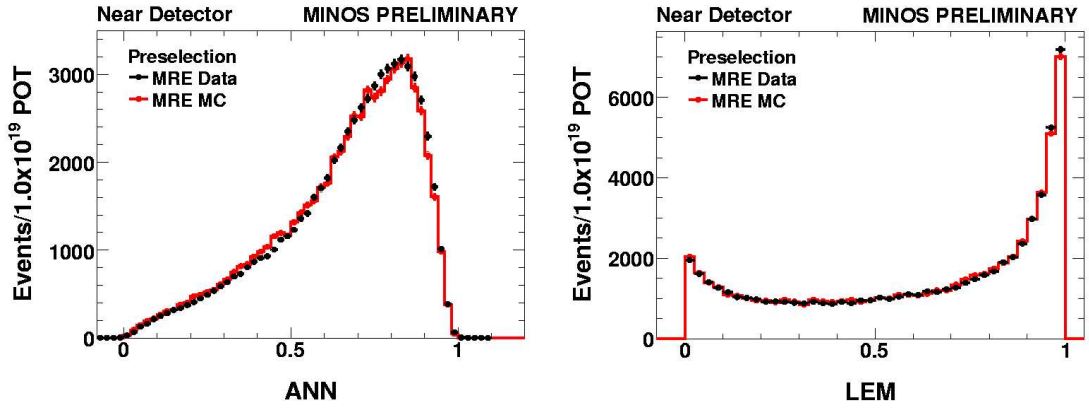


Figure 7.3: ND MRE distributions for the ANN PID on the left and the LEM PID on the right. The MRE data is shown in black, and the MRE MC is shown in red. The  $\nu_e$  preselection is applied and the distributions are normalised by POT.

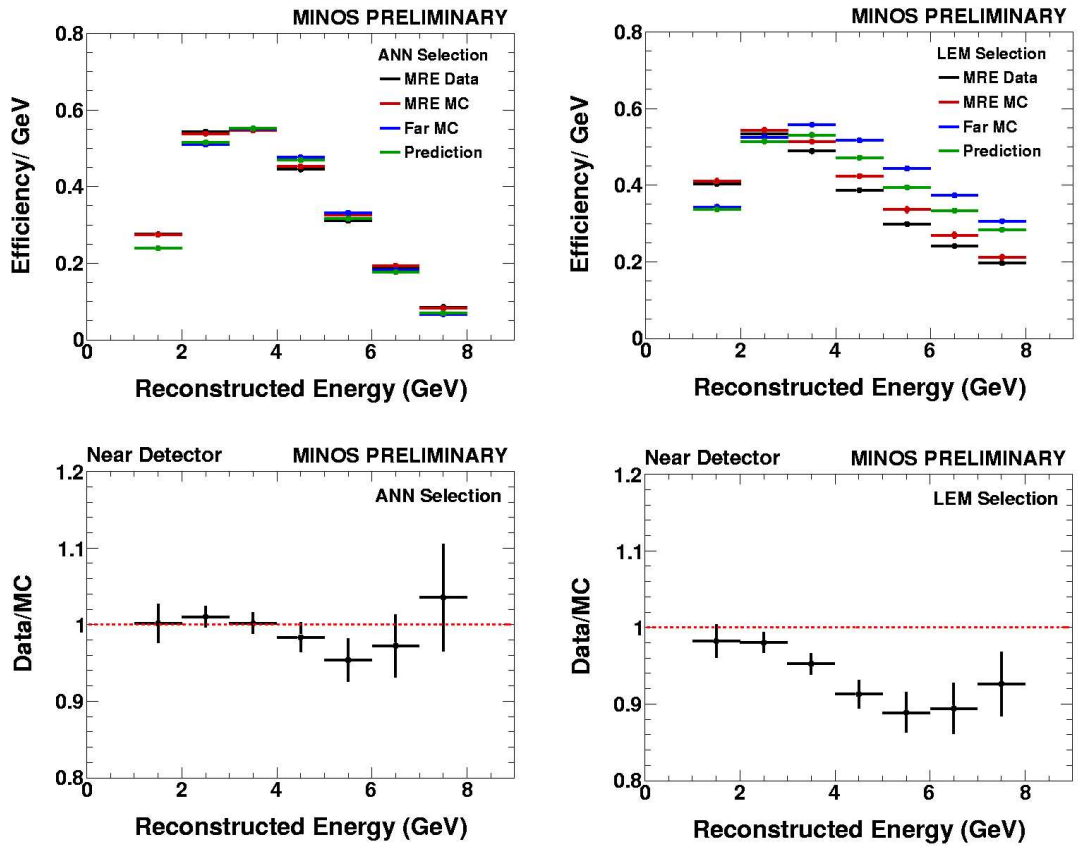


Figure 7.4: The top plots show the ND MRE and standard FD efficiencies after ANN PID selection on the left and LEM PID selection on the right. The ND MRE data is shown in black, the ND MRE MC is shown in red, the standard Far MC signal efficiency is shown in blue, and the prediction, i.e. the corrected FD signal efficiency is shown in green. The bottom plots show the ANN (left) and LEM (right) ND MRE data/MC efficiency ratios.



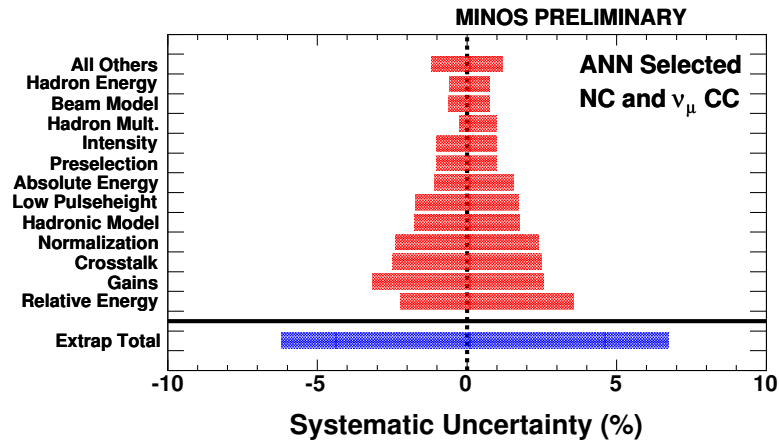


Figure 7.5: Important systematic error bands for the NC and  $\nu_\mu$ -CC backgrounds for the ANN PID selection with the horn on-off method (systematics for the LEM and MRCC method are very similar).

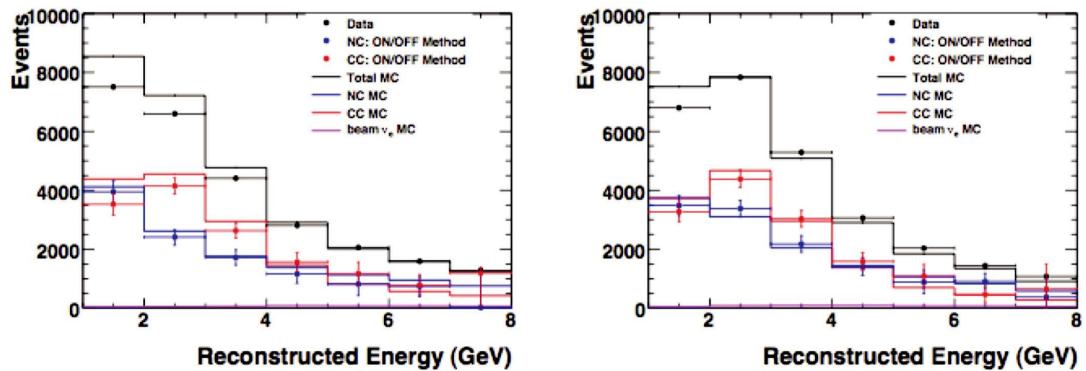


Figure 7.6: This Figure shows the ND background decompositions for the ANN anti-PID on the left plot and the LEM anti-PID on the right using the HOO decomposition method. The solid lines are the default MC and the points are the data decomposed into the ND backgrounds. The data and the uncorrected MC agree relatively well.

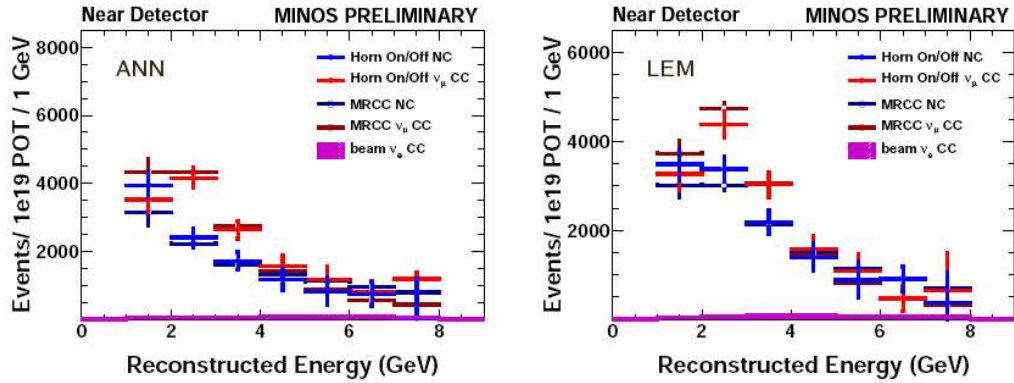


Figure 7.7: This Figure shows the comparison of the ND background decomposition methods for the anti-PID sidebands. The data decomposition for the ANN anti-PID is shown on the left plot, and the decomposition for the LEM anti-PID is shown on the right. The blue points correspond to the NC background, and the red points to the  $\nu_\mu$ -CC background. The HOO decompositions are shown in the lighter color shades, and the MRCC decompositions are shown in the darker color shades. The intrinsic beam  $\nu_e$ -CC background is shown in pink. It can be seen that both background decomposition methods agree with each other within errors.

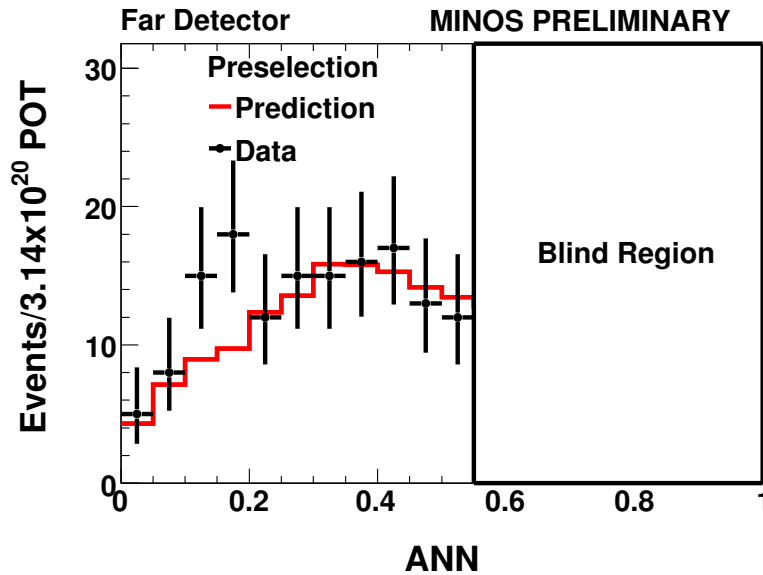


Figure 7.8: This Figure shows the distribution of the ANN PID after preselection cuts, where the signal region above a cut of 0.55 was blinded. The red line is the prediction, and the black points are the data. It can be seen that data and MC agree within statistics. The background decomposition method used for this Figure is the HOO method.

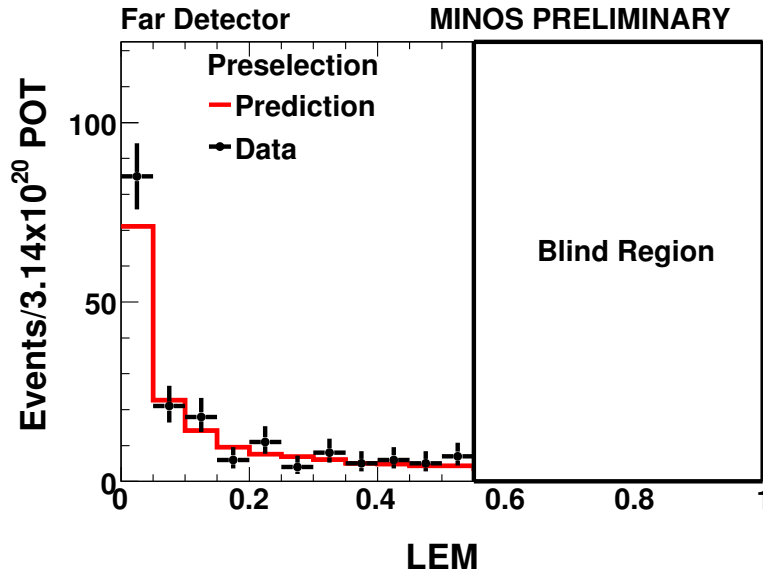


Figure 7.9: This Figure shows the distribution of the LEM PID after preselection cuts, where the signal region above a cut of 0.55 was blinded. The red line is the prediction, and the black points are the data. It can be seen that data and MC agree within statistics. The background decomposition method used for this Figure is the HOO method.

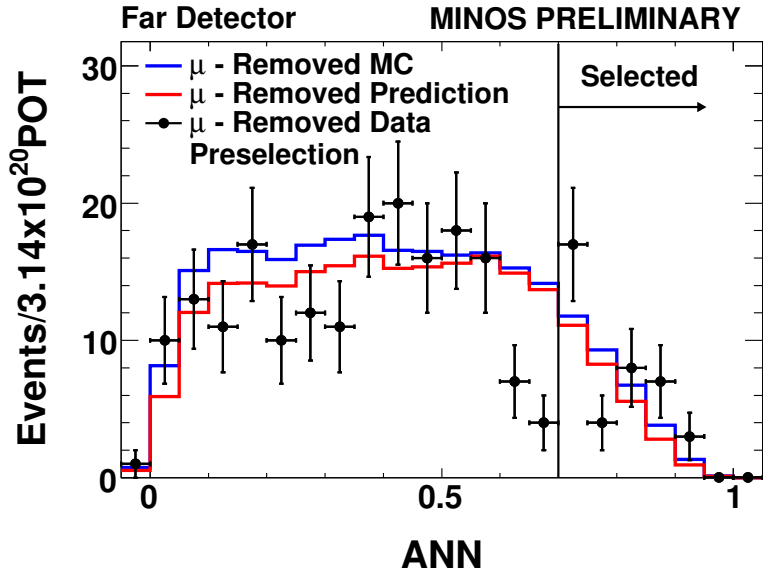


Figure 7.10: ANN PID distributions for the FD MRCC preselected event sample. The black points are the data, the red line is the prediction obtained by scaling the default MRCC MC in blue by the ND MRCC data/MC ratios. The histograms are normalised to the data exposure of  $3.14 \times 10^{20}$  POT.

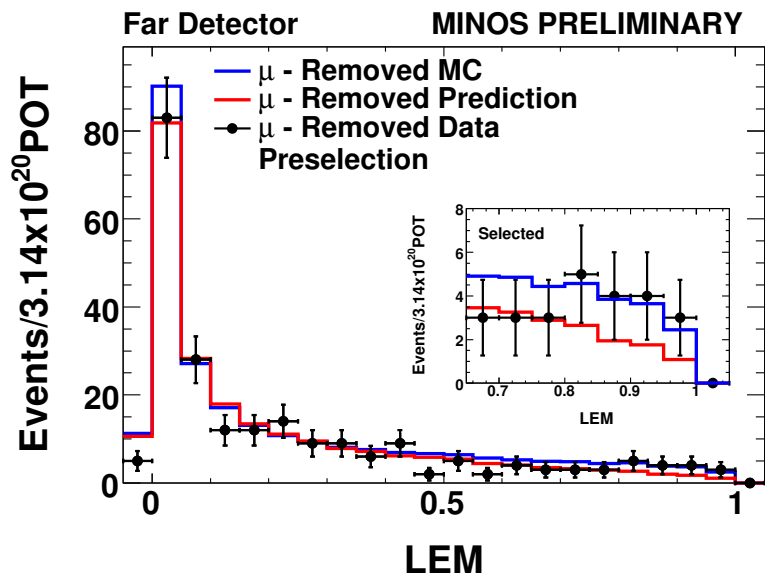


Figure 7.11: LEM PID distributions for the FD MRCC preselected event sample. The black points are the data, the red line is the prediction obtained by scaling the default MRCC MC in blue by the ND MRCC data/MC ratios. The histograms are normalised to the data exposure of  $3.14 \times 10^{20}$  POT.

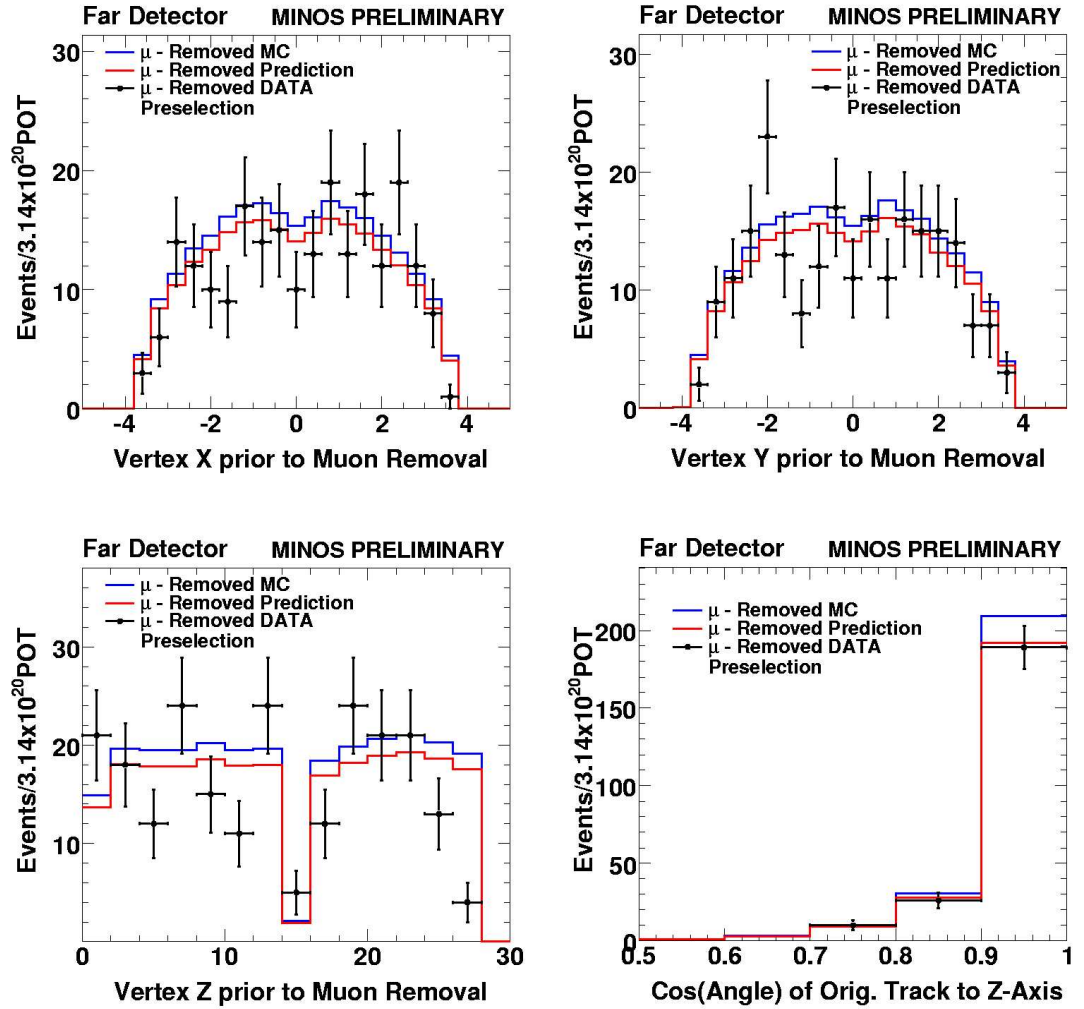


Figure 7.12: This Figure shows data and prediction spatial distributions for the FD MRCC  $\nu_e$  preselected events. The black points are the data, and the red line is the prediction obtained by scaling the default MRCC MC (blue) by the ND MRCC data/MC ratios. The histograms are normalised to the data exposure of  $3.14 \times 10^{20}$  POT. The top left plot shows the original vertex X in (m) (prior to muon removal), the top right plot shows the original vertex Y in (m), the bottom left plot shows the original vertex Z in (m) - note that the low point at 15m is caused by the gap between the two FD super modules. The bottom right plot shows the cosin of the angle of the original muon track to the Z-axis of the detector.

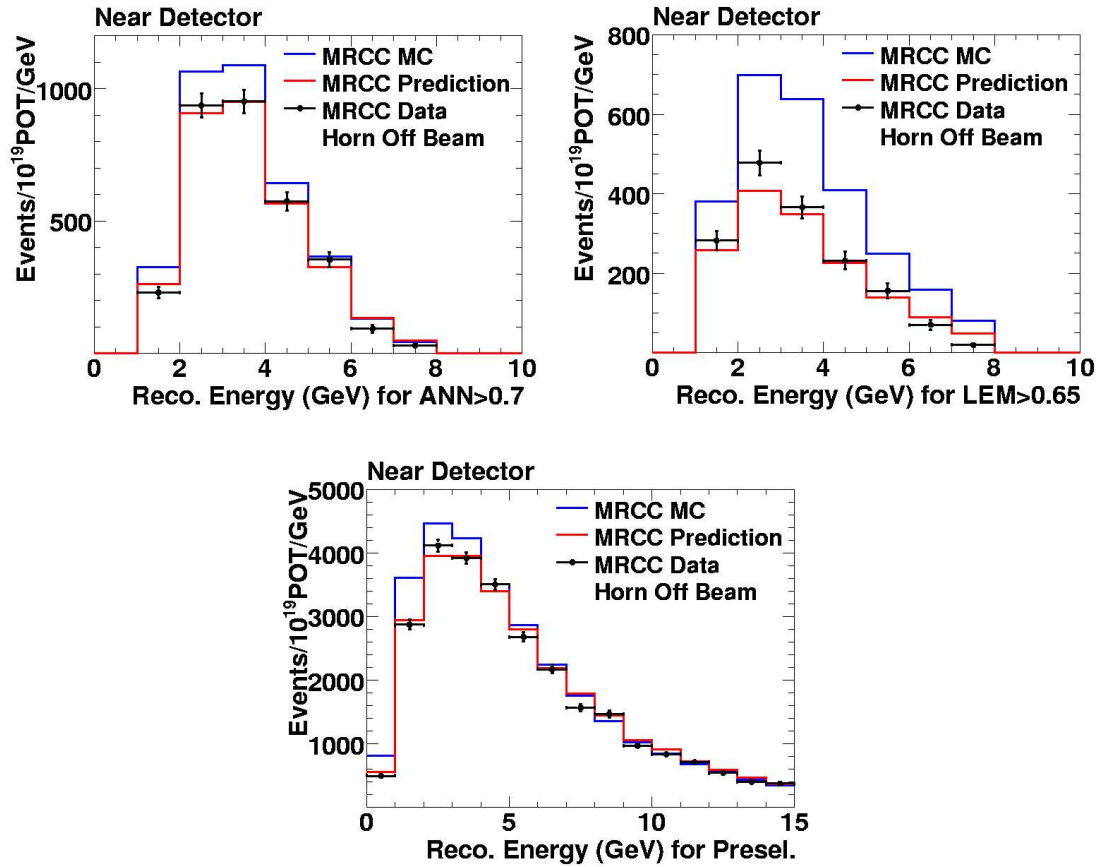


Figure 7.13: This Figure shows a cross-check of the MRCC sideband extrapolation method within one detector using the MRCC horn off sample in place of the MRCC FD Data and MC. The top left plot shows the results for the ANN PID with a cut at 0.7, the top right plot shows the results for the LEM PID with a cut at 0.65, and the bottom plot shows the results for the  $\nu_e$  preselection. It can be seen that the prediction and the data agree well.

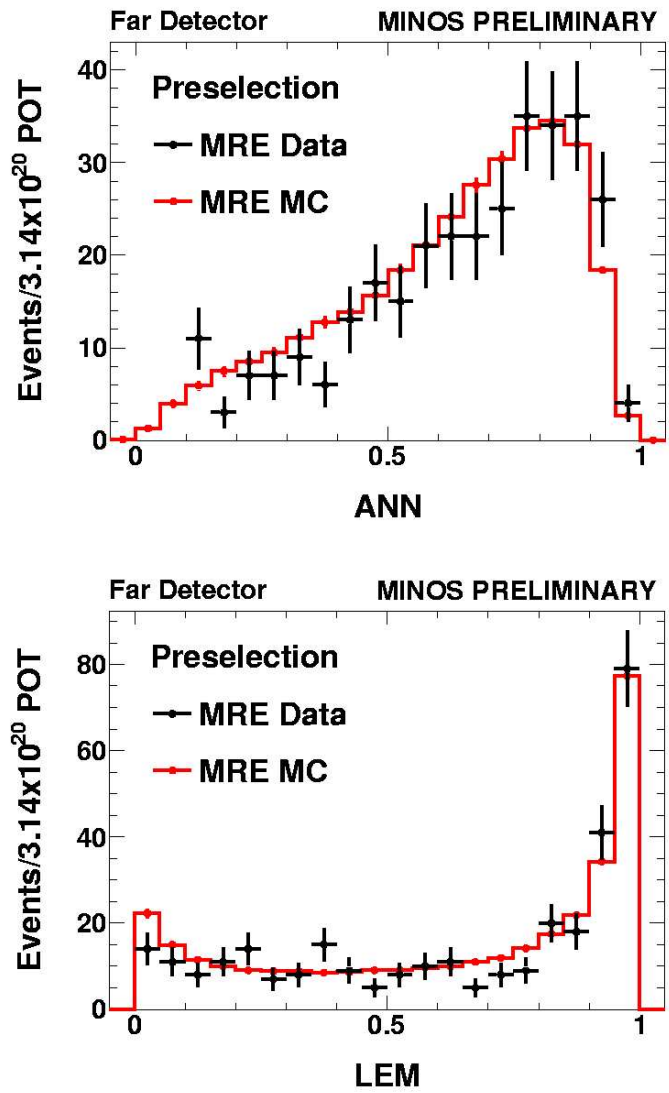


Figure 7.14: PID distributions for the FD MRE preselected event sample. The black points are the data, and the red line is the prediction. The histograms are normalised to the data exposure of  $3.14 \times 10^{20}$  POT. The top plot shows the distributions for the ANN PID, and the bottom plot shows the distributions for the LEM PID.

## Chapter 8

# Final $\nu_e$ -Appearance Analysis

## Result

The  $\nu_e$ -appearance analysis was now prepared to unblind the final analysis box since the following steps had been carried out:

- two PIDs had been decided upon to be carried through the analysis - the ANN PID as a primary PID, and the LEM PID as a cross-check method
- two ND background decomposition methods had been used to calculate the relative contributions of the three different background components in the ND data
- the FD data predictions had been obtained by extrapolating the ND backgrounds to the FD
- systematic errors had been evaluated for both the ND decompositions and the FD extrapolations
- three different sidebands had been opened so as to identify any potential problems prior to box opening. The LEM PID cut was changed as a result.
- a procedure was established to draw the final result contours (this will be explained further in this chapter)

The last step prior to opening the box consisted of looking at area normalized distributions in the FD. No issues were found in those distributions and they all looked consistent within statistics. Two such distributions can be seen in Figure 8.1.



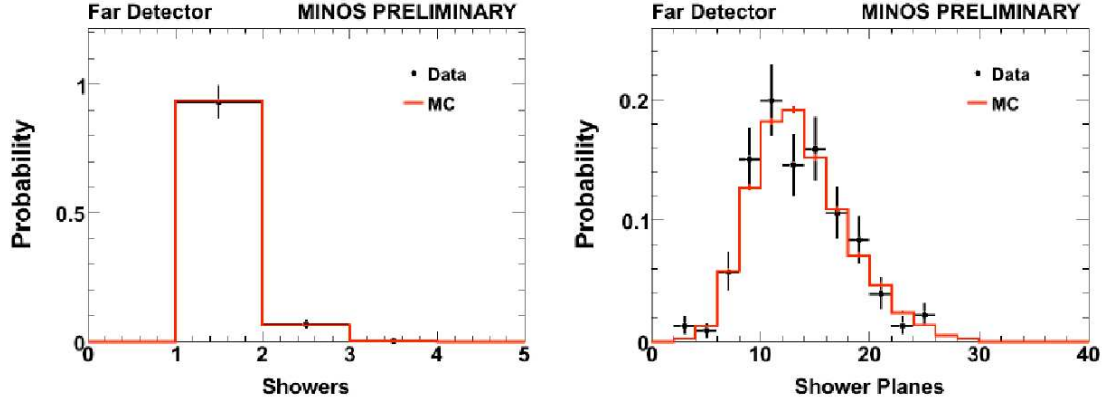


Figure 8.1: This Figure shows two of the area normalized distributions that were looked at prior to opening the final analysis box. The data is in black, and the MC is in red. The two distributions shown are the number of showers and the number of shower planes at preselection level.

## 8.1 Far Detector Final Box Opening

In the final selected data sets, the ANN PID selected 35 events above a PID cut of 0.7. The LEM PID selected 28 events above the looser cut of 0.65 (for the original cut of 0.8, the LEM PID selected 17 events with a prediction of 13 background events). The PID distributions after preselection are shown in Figures 8.2 and 8.3.

The FD result in reconstructed energy is shown in Figures 8.4 and 8.5 for all four methods (ANN and LEM, for HOO and MRCC) respectively.

The final FD results are summarized in Table 8.1.

| PID/Method | Prediction | Stat. Error | Syst. Error | DATA | Stat. Error | DATA Excess | $\sigma$ Excess |
|------------|------------|-------------|-------------|------|-------------|-------------|-----------------|
| ANN MRCC   | 28         | 5           | 2           | 35   | 6           | 7           | 1.3             |
| ANN HOO    | 27         | 5           | 2           | 35   | 6           | 8           | 1.5             |
| LEM MRCC   | 22         | 5           | 2           | 28   | 5           | 6           | 1.1             |
| LEM HOO    | 22         | 5           | 3           | 28   | 5           | 6           | 1.0             |

Table 8.1: Final FD results using the two different data decomposition methods and the two different PIDs. Normalised to an exposure of  $3.14 \times 10^{20}$  POT.

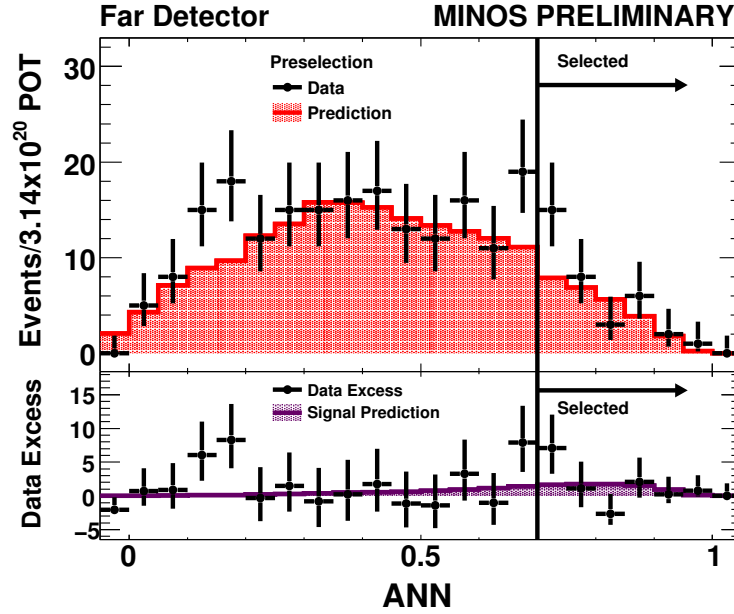


Figure 8.2: The ANN PID distributions for a data exposure of  $3.14 \times 10^{20}$  POT after preselection. The FD data is compared to the prediction. The top plot shows the data in black and the prediction in red, and the bottom plot shows the signal prediction in purple and the data minus background prediction in black. The selected PID region is shown by the line and arrow.

## 8.2 Cross-Checks carried out on the Final Data Set

In order to verify that there were no unforeseen problems with the FD data, various checks were carried out. Data and MC selection efficiencies were investigated. Various distributions in the FD were also looked at, and all the FD events were scanned by eye to check that they were all events with a reasonably looking EM-like shower. The FD data was also split into MINOS runs to see if the excess was in a particular set of the data. Those different checks are described in detail in this section.

### 8.2.1 Far Detector Selection Numbers

FD event numbers and efficiencies as a function of cut level were checked for the FD data and oscillated FD MC to see if they were behaving approximately the same way. To do this, the FD MC was oscillated at the Chooz limit ( $\sin^2 2\theta_{13}=0.15$ ). It is somewhat hard to interpret the resulting selection efficiencies because of statistical limitations and also because the data may contain a different amount of signal than the MC. Despite

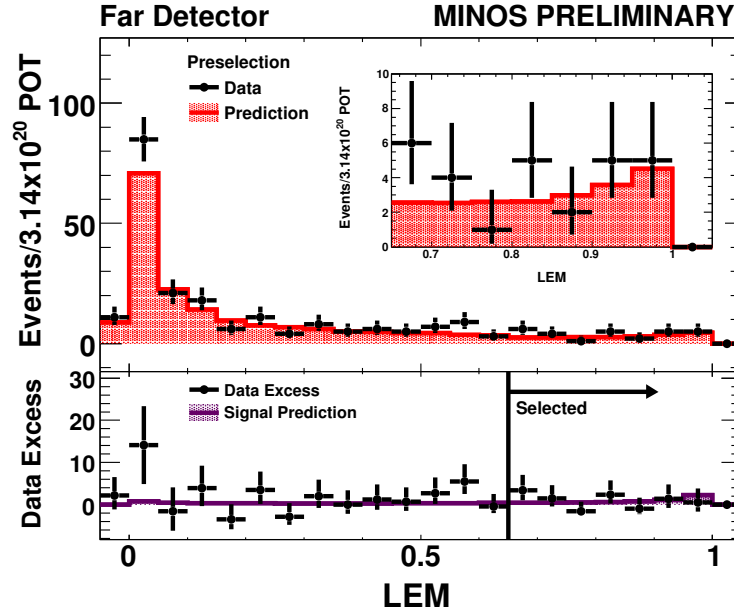


Figure 8.3: The LEM PID distributions for a data exposure of  $3.14 \times 10^{20}$  POT after preselection. The FD data is compared to the prediction. The top plot shows the data in black and the prediction in red, and the bottom plot shows the signal prediction in purple and the data minus background prediction in black. The selected PID region is shown by the line and arrow.

these limitations, the data and MC selection numbers are consistent with each other - especially for the ANN PID selection, as in Table 8.2. One interesting fact to note is that if no preselection (except fiducial volume cuts) is applied to the FD data, the ANN PID selects 37 events instead of 35 and the LEM PID selects 32 events instead of 28. This is because the PIDs are tuned into 2-3GeV EM-shower-like events, so essentially act as a preselection themselves.

## 8.2.2 Far Detector Distributions

Upon opening the final analysis box, many FD distributions were checked. Most of those distributions are included in the Appendix D but several are shown in Figures 8.6, 8.7, and 8.8. All distributions looked reasonable within statistics and no problems were found.

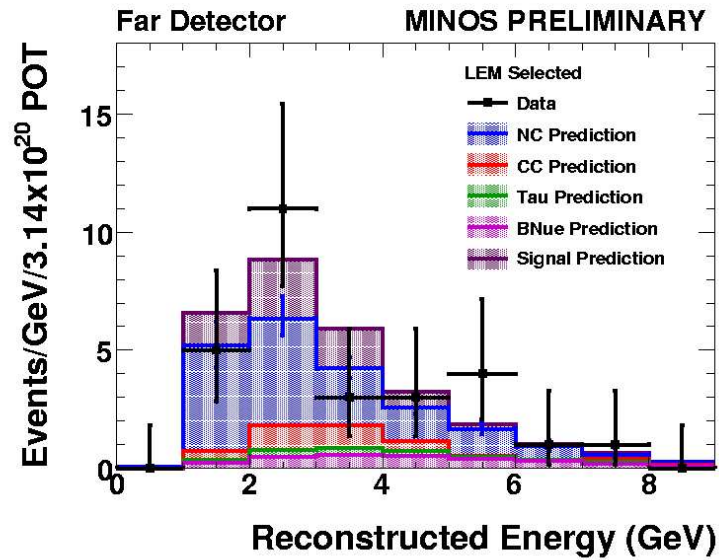
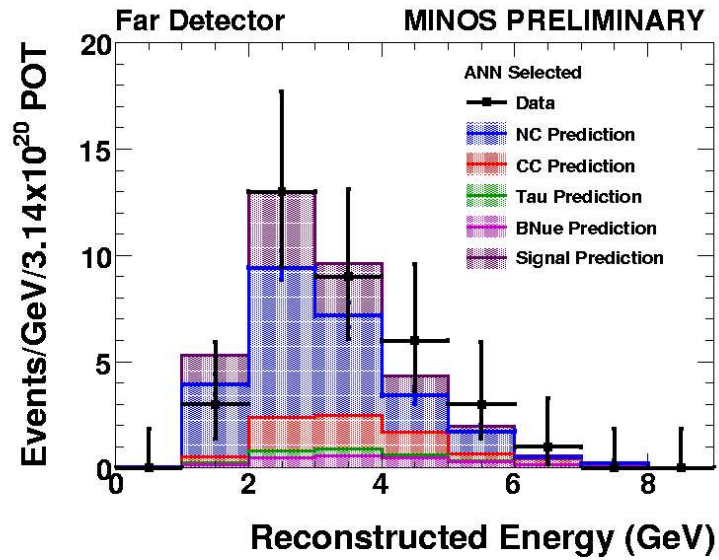


Figure 8.4: This Figure shows the FD data versus prediction reconstructed energy distribution for a data exposure of  $3.14 \times 10^{20}$  POT after ANN or LEM PID selection for the HOO decomposition method. The data is shown in black, and the individual predicted backgrounds are shown as shaded regions. The purple shaded region is the signal.

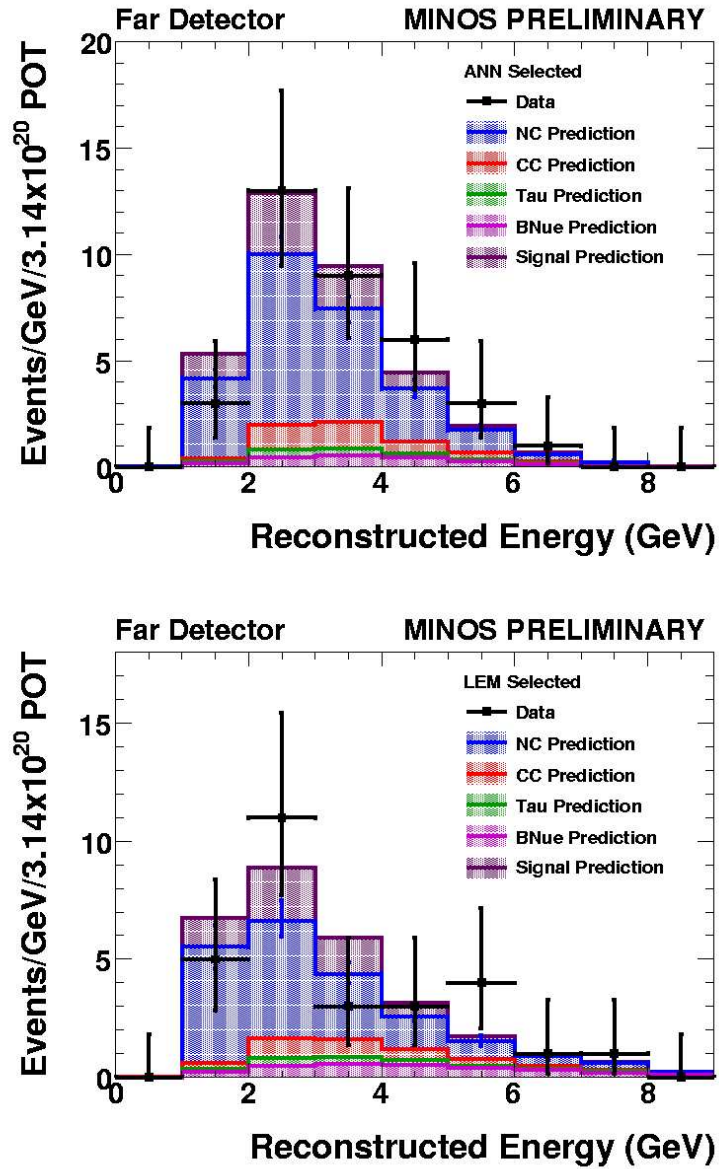


Figure 8.5: This Figure shows the FD data versus prediction reconstructed energy distribution for a data exposure of  $3.14 \times 10^{20}$  POT after ANN or LEM PID selection for the MRCC decomposition method. The data is shown in black, and the individual predicted backgrounds are shown as shaded regions. The purple shaded region is the signal.

| Cut       | Data<br>$\pm$ Stat. | MC All | All<br>Backg. | NC    | $\nu_\mu$ -CC | $\nu_\tau$ -CC | Beam<br>$\nu_e$ -CC | Signal<br>$\nu_e$ -CC |
|-----------|---------------------|--------|---------------|-------|---------------|----------------|---------------------|-----------------------|
| Fiducial  | 1188 $\pm$ 34.5     | 1221.4 | 1196.4        | 320.2 | 852.4         | 7.9            | 15.9                | 25.0                  |
| TrackP.   | 444 $\pm$ 21.1      | 445.6  | 420.9         | 288.8 | 112.4         | 6.3            | 13.4                | 24.7                  |
| TrkLikeP. | 410 $\pm$ 20.2      | 410.3  | 385.9         | 280.8 | 85.8          | 5.9            | 13.3                | 24.4                  |
| Shw.>0    | 406 $\pm$ 20.2      | 406.4  | 382.0         | 277.8 | 84.9          | 5.9            | 13.3                | 24.4                  |
| Cont.P.   | 286 $\pm$ 16.9      | 298.5  | 277.0         | 181.2 | 77.6          | 5.3            | 12.9                | 21.5                  |
| E>1.0     | 271 $\pm$ 16.5      | 282.8  | 261.6         | 168.2 | 75.4          | 5.2            | 12.9                | 21.2                  |
| Presel.   | 227 $\pm$ 15.1      | 229.3  | 208.9         | 136.3 | 63.1          | 4.3            | 5.3                 | 20.4                  |
| ANN       | 35 $\pm$ 5.9        | 42.9   | 32.6          | 24.6  | 5.1           | 1.0            | 1.9                 | 10.3                  |
| LEM       | 28 $\pm$ 5.3        | 46.7   | 35.2          | 25.9  | 6.0           | 1.0            | 2.4                 | 11.5                  |
| Effic.    | %                   | %      | %             | %     | %             | %              | %                   | %                     |
| TrackP.   | 37.4 $\pm$ 1.8      | 36.5   | 35.2          | 90.2  | 13.2          | 79.4           | 84.5                | 98.8                  |
| TrkLikeP. | 34.5 $\pm$ 1.7      | 33.6   | 32.3          | 87.7  | 10.1          | 75.0           | 83.7                | 97.7                  |
| Shw.>0    | 34.2 $\pm$ 1.7      | 33.3   | 31.9          | 86.8  | 10.0          | 74.8           | 83.6                | 97.6                  |
| Cont.P.   | 24.1 $\pm$ 1.4      | 24.4   | 23.2          | 56.6  | 9.1           | 67.2           | 81.3                | 86.0                  |
| E>1.0     | 22.8 $\pm$ 1.4      | 23.2   | 21.9          | 52.5  | 8.8           | 65.7           | 80.7                | 84.7                  |
| Presel.   | 19.1 $\pm$ 1.3      | 18.8   | 17.5          | 42.6  | 7.4           | 54.7           | 33.2                | 81.5                  |
| ANN       | 2.9 $\pm$ 0.5       | 3.5    | 2.7           | 7.7   | 0.6           | 12.4           | 11.9                | 41.2                  |
| LEM       | 2.4 $\pm$ 0.5       | 3.8    | 2.9           | 8.1   | 0.7           | 12.9           | 14.9                | 45.9                  |

Table 8.2: FD event numbers and selection efficiencies with respect to fiducial volume selection. Data versus oscillated standard MC (not FD predictions) for various selection levels. Normalised to an exposure of  $3.14 \times 10^{20}$  POT.

### 8.2.3 Event Scanning

After opening the box, all selected events were scanned by eye to verify that all the events are EM-shower-like and consistent either being  $\nu_e$ -CC like. No problems were found. One golden event with high PID value is shown in Figure 8.9. More are shown in Appendix D.

### 8.2.4 Breakdown of Events by MINOS Data Run

Recall that for the MRCC sideband, for the ANN PID, an excess was observed for Run IIa. In order to verify that the same does not occur for the standard FD data, the selected events were also broken down by MINOS run period. The results are shown in Table 8.3.

Much of the standard FD data excess is in Run IIb, whereas much of the MRCC sideband excess is in Run IIa for the ANN PID selection. This supports a statistical fluctuation explanation for the latter.

| Cut        | Run I                 | Run IIa               | Run IIb               | Run I     | Run IIa   | Run IIb   |
|------------|-----------------------|-----------------------|-----------------------|-----------|-----------|-----------|
| POT        | $1.21 \times 10^{20}$ | $1.22 \times 10^{20}$ | $0.71 \times 10^{20}$ | $10^{20}$ | $10^{20}$ | $10^{20}$ |
| DATA       |                       |                       |                       |           |           |           |
| Presel.    | 77                    | 85                    | 65                    | 63.64     | 69.67     | 91.55     |
| ANN PID    | 10                    | 13                    | 12                    | 8.26      | 10.66     | 16.90     |
| LEM PID    | 10                    | 7                     | 11                    | 8.26      | 5.74      | 15.49     |
| Prediction |                       |                       |                       |           |           |           |
| Presel.    | 75.4                  | 76.4                  | 44.2                  | 62.42     | 62.42     | 62.42     |
| ANN PID    | 10.2                  | 10.3                  | 6.0                   | 8.44      | 8.44      | 8.44      |
| LEM PID    | 8.3                   | 8.4                   | 4.9                   | 6.85      | 6.85      | 6.85      |
| Data/Pred. |                       |                       |                       |           |           |           |
| Presel.    | 1.02                  | 1.11                  | <b>1.47</b>           |           |           |           |
| ANN PID    | 0.98                  | 1.26                  | <b>2.0</b>            |           |           |           |
| LEM PID    | 1.20                  | 0.83                  | <b>2.24</b>           |           |           |           |

Table 8.3: Final FD results broken down by MINOS run period. The right hand set of numbers has been normalised to  $10^{20}$  POT. The first set of numbers are for the FD data. The second set of numbers are for the background prediction, and finally, the last three rows show the data/prediction ratios. It can be seen that much of the excess is in Run IIb and not Run IIa like in the MRCC sideband.

### 8.3 Final Contours for the $3.14 \times 10^{20}$ POT $\nu_e$ -Appearance Analysis

In order to calculate the final contours, an analytical Feldmann-Cousins method was used which was specifically modified for low statistics results [73, 78, 79]. The general idea of this method is that for each point in oscillation space, the fraction of possible observations more likely than the actual observation is calculated and this number becomes the confidence level for that particular point in oscillation space.

The Far/Near extrapolation method ultimately predicts a total number of background events  $b \pm \sigma_b$  for a set of oscillation parameters  $\mu$ . The predicted number of background events is however not necessarily the true observed number of background events  $\beta$ . Hence  $b$  is an estimator of  $\beta$ . Similarly, we have the expected signal  $s$  for a given set of oscillation parameters and the uncertainty of the signal measurement  $\sigma_s$ , where  $s \rightarrow s(\mu)k$  and  $\sigma_s \rightarrow s(\mu)\sigma_k$ . The parameter  $k \pm \sigma_k$  is present to account for the uncertainty on the actual signal measurement (as opposed to just the underlying oscillation parameters) and is a normalization constant on the number of signal events.  $k$  estimates the true normalization scale  $\kappa$ . The true number of background events and

the true normalization are nuisance parameters which can be eliminated by replacing them with their most likely observation values.

The rank  $R(n, b, k)$  is defined as:

$$R(n, b, k) = \frac{(s(\mu)\hat{k} + \hat{\beta})^n e^{-(s(\mu)\hat{k} + \hat{\beta})} e^{-(b-\hat{\beta})^2/2\sigma_b^2} e^{-(k-\hat{k})^2/2\sigma_k^2}}{(s(\mu)\hat{k} + \hat{\beta})^n e^{-(s(\mu)\hat{k} + \hat{\beta})} e^{-(b-\hat{\beta})^2/2\sigma_b^2} e^{-(k-\hat{k})^2/2\sigma_k^2}} \quad (8.1)$$

where  $\hat{k}$  and  $\hat{\beta}$  are the values of  $\kappa$  and  $\beta$  that maximise the numerator, and  $\hat{k}$  and  $\hat{\beta}$  are the values of  $\kappa$  and  $\beta$  that maximise the denominator. Solutions to maximise the numerator and denominator can be calculated for Equation 8.1. In order to obtain a confidence belt  $\alpha$ , all possible measurements are included in decreasing rank order until the sum of those probabilities is equal or greater than  $\alpha$ . So a value of  $\mu$  is included in the confidence interval  $\alpha$  if  $\Omega(\mu) < \alpha$ . The total probability  $\Omega$  is the total sum of the probabilities of all  $n$ ,  $b$ , and  $k$  that have a higher rank than the actual measurement  $n_0$ ,  $b_0$  and  $k_0$ , and is given by:

$$\Omega(\mu) = \frac{1}{2\pi\sigma_b\sigma_k} \sum_{n \neq n_0} \frac{(s(\mu)\hat{k}_0 + \hat{\beta}_0)^n e^{-(s(\mu)\hat{k}_0 + \hat{\beta}_0)}}{n!} \times \int \int_{R_n^+} db dk e^{-(\hat{\beta}_0 - b)^2/2\sigma_b^2} e^{-(\hat{k}_0 - k)^2/2\sigma_k^2} \quad (8.2)$$

where  $R_n^+$  is an area located near the coordinates which satisfy  $n = sk + b$ . A detailed derivation of the above can be found in [73]. The minimum  $\Omega$  represents automatically the best fit, and the lines drawn for a particular  $\alpha$  will represent the confidence limits contours. The final fit is done in  $\sin^2 2\theta_{13}$  and  $\delta_{CP}$  parameter space only, for both the normal and inverted mass hierarchy. All other oscillation parameters are assumed to be fixed, so  $\sin^2 2\theta_{23} = 1$  and  $\Delta m_{32}^2 = 2.43 \times 10^{-3} \text{eV}$ . It has to be noted, that strictly speaking, the quantity that is measured is not  $\sin^2 2\theta_{13}$  but  $\sin^2 2\theta_{23} \sin^2 2\theta_{13}$ . This may be important when comparing the results to reactor experiments, which measure  $\sin^2 2\theta_{13}$  directly.

The best fit contours are shown for all 4 methods in Figures 8.10 and 8.11. The final results look similar for the HOO and the MRCC methods since the FD predictions were very similar. There are however differences arising from the fact that the MRCC method predicts more background events hence the observed signal is smaller.



As has already been shown in Table 8.1, both the MRCC and the HOO methods see a data excess for both PIDs. The best fit points for the  $\sin^2 2\theta_{13}$  parameter are given for a CP-violation phase  $\delta = 0$  in Table 8.4.

| PID/Method    | Best Fit | 90%C.L.     |
|---------------|----------|-------------|
| Normal M.H.   |          |             |
| ANN MRCC      | 0.091    | <0.265      |
| ANN HOO       | 0.112    | 0.001-0.287 |
| LEM MRCC      | 0.076    | <0.230      |
| LEM HOO       | 0.068    | <0.220      |
| Inverted M.H. |          |             |
| ANN MRCC      | 0.154    | <0.395      |
| ANN HOO       | 0.183    | 0.008-0.422 |
| LEM MRCC      | 0.132    | <0.349      |
| LEM HOO       | 0.083    | <0.276      |

Table 8.4: Final FD fit results for  $\sin^2 2\theta_{13}$  at a 90%C.L. using the two different data decomposition methods and the two different PIDs. The CP-violation phase is set to  $\delta = 0$  and results are shown for both mass hierarchies [73].

Lists of events can be found for the FD data in Appendix E and they show that there is a 19 event overlap for the two PIDs, which corresponds to 54% of the ANN PID sample and 68% of the LEM PID sample. This is quite a large overlap, as by nature those two samples will be correlated. As the LEM PID selects fewer events, the overlap is larger for this PID than for the ANN PID. The overlap certainly indicates that the PIDs may possibly be refined further in the future as some events that one PID sees as background are seen as signal by the other.

## 8.4 Comments on Far Detector Data

Even though a small excess is seen for both the MRCC and HOO data decomposition methods, and for both PIDs, it is not yet a conclusive  $\nu_e$ -appearance signal. Ignoring the results of the MRCC sidebands, the anti-PID sideband shows a data excess of approximately the same order of magnitude as the excess in the final sample data. It is possible that there might be some mismodeling of showers or other effect that affects the Far and Near data differently and thus the analysis underestimated the number of background events. Better MC and reconstruction may help to resolve this question in the future.

## 8.5 Future MINOS $\nu_e$ -Appearance Analysis

Currently, there are  $\nu_e$ -appearance analysis improvements underway and the new data set with a total exposure of approximately  $7 \times 10^{20}$  POT will be opened soon. There will be some significant PID improvements which would allow a bigger background reduction. The analysis improvements also include the possibility of evaluating all the ND decompositions in two dimensions of PID and reconstructed energy. Possibly, instead of just fitting the final count of events selected by the PID, the fitting will be carved out in several PID and reconstructed energy bins that will include the full preselected PID region. This should help to discern whether a real signal is seen or whether any excess of events is due to unknown Far/Near differences. The pathway to the new result will remain however very similar to the analysis presented in this document. Projections of potential limits in the absence of a signal are shown in Figures 8.12 and 8.13. The projections are for an analysis identical to the one presented in this document and do not include the aforementioned improvements that are being worked on. There are clear improvements in the 90% confidence limits for a higher exposure of  $10 \times 10^{20}$  POT.

The MINOS experiment has now switched to reverse horn current running which produces predominantly anti-neutrinos focused at low energies (in the 2-3GeV energy peak). It is not clear yet how the  $\nu_e$ -appearance analysis will proceed for this new data set, but it is very likely that MINOS will be switched back to normal (forward) horn current running relatively quickly because of considerations to do with new experiments in the ND Hall that will be taking advantage of the NuMI beam and which primarily require a neutrino beam. If/when the beam is switched back to neutrino mode, it should be possible to accumulate more statistics for the standard  $\nu_e$ -appearance analysis. Presently, the precise run plan is still to be decided.

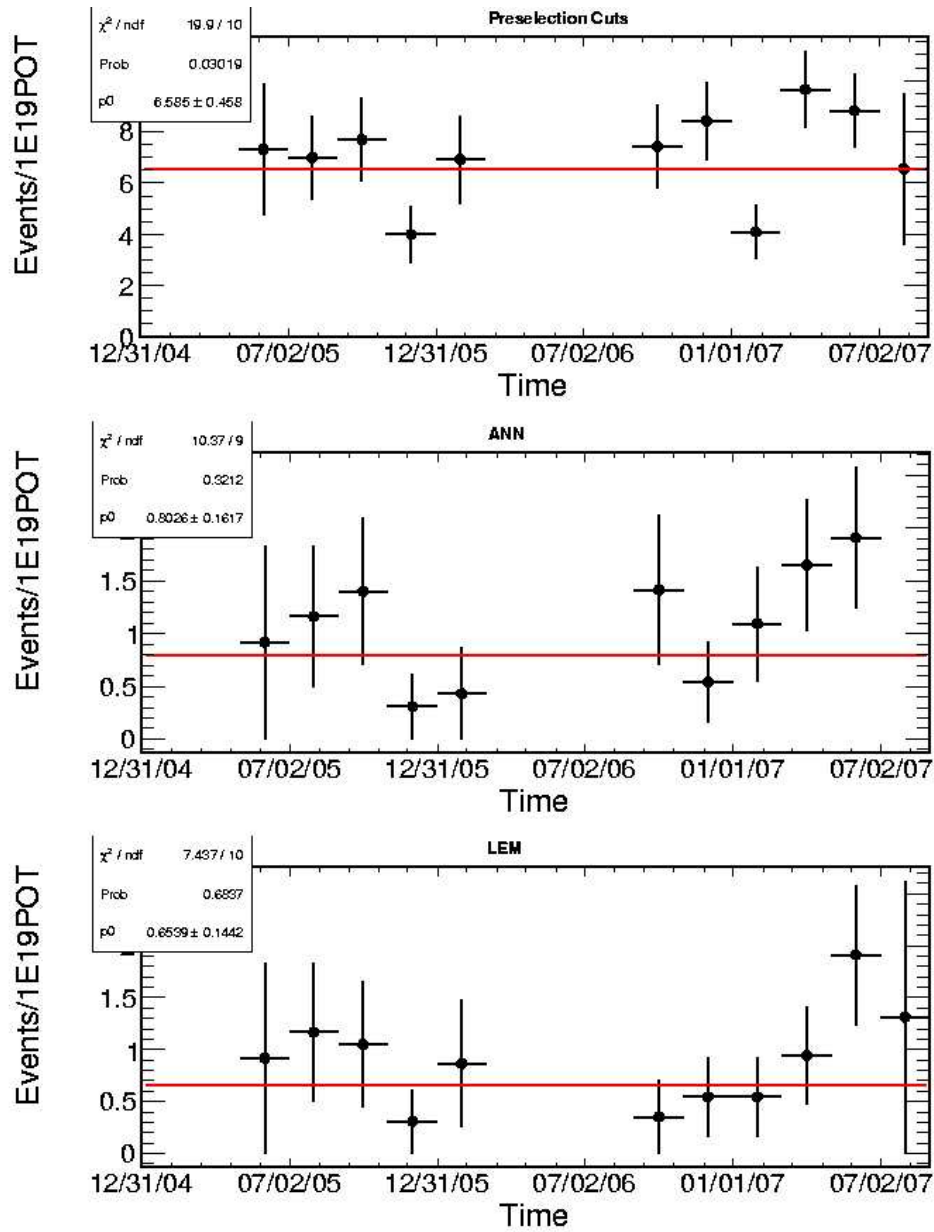


Figure 8.6: FD data events as a function of time for preselection, ANN PID selection, and LEM PID selection.

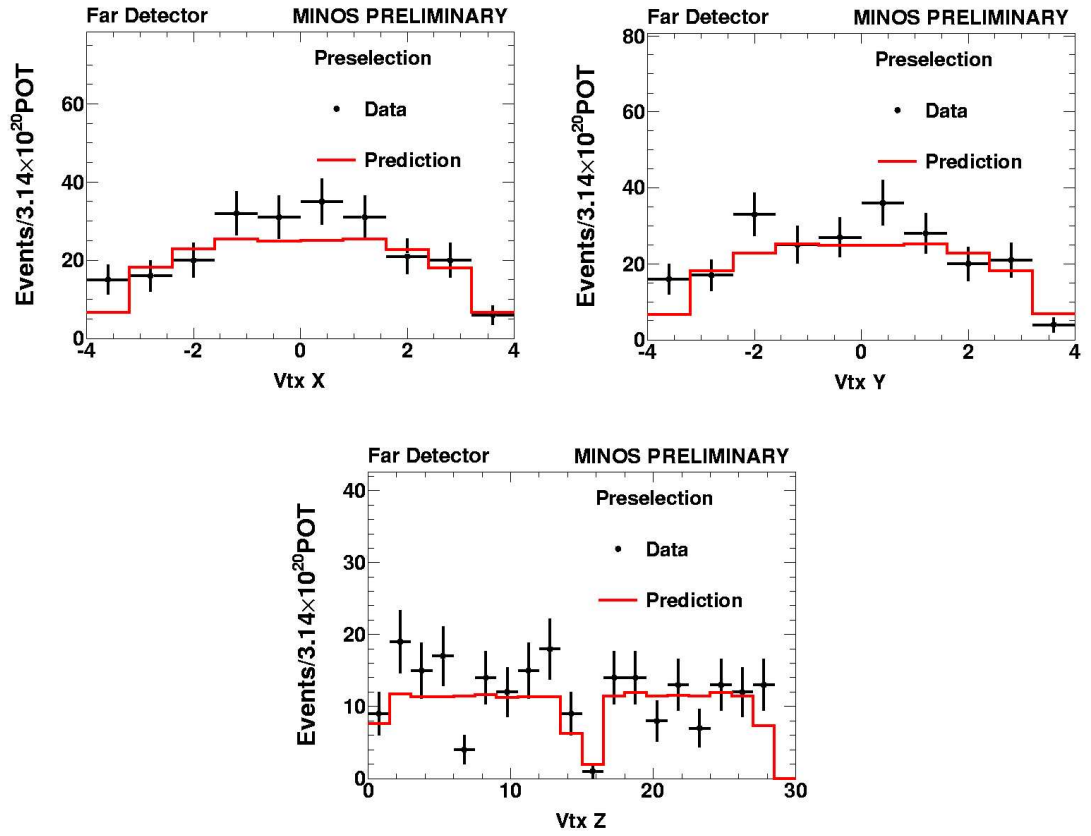


Figure 8.7: FD data event vertices in the transverse and longitudinal directions after preselection. The prediction is shown in red, and the data in black.

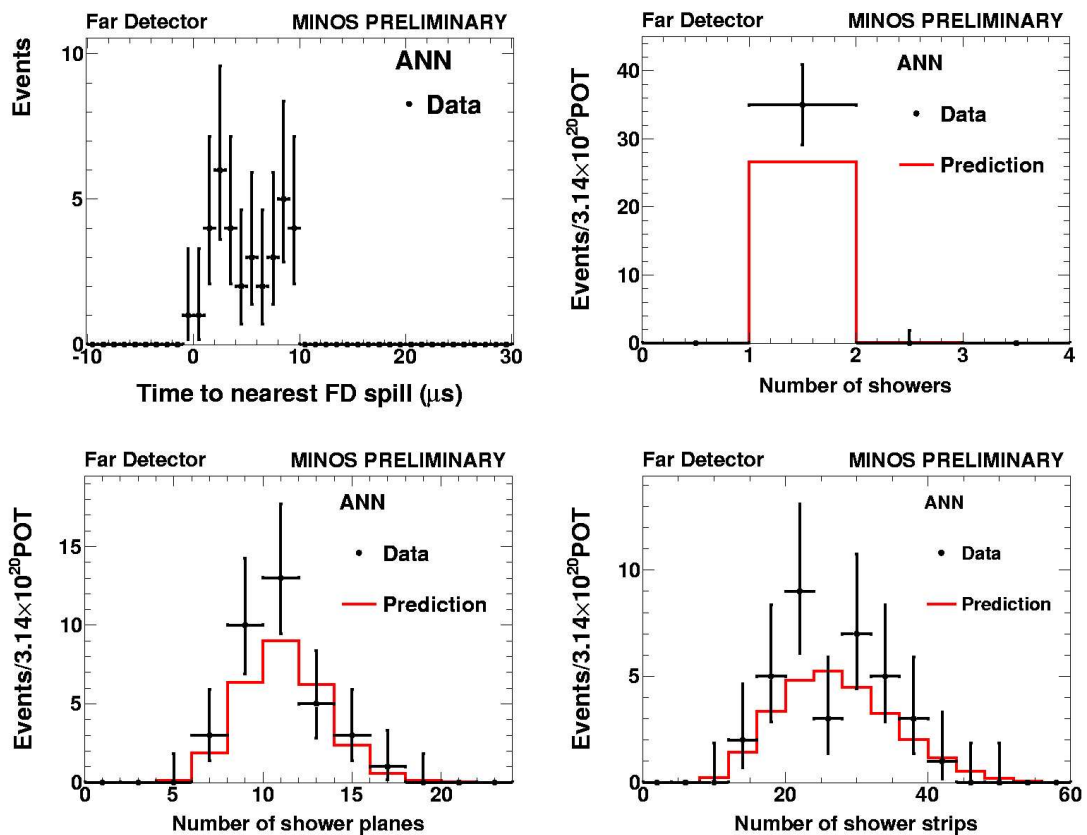


Figure 8.8: Data distributions in the FD after ANN PID selection. The prediction is shown in red, and the data in black. The top left plot shows the spill time window of the selected events. The top right plot show the number of showers. The bottom left shows the number of shower planes and the bottom right shows the number of shower hit strips.

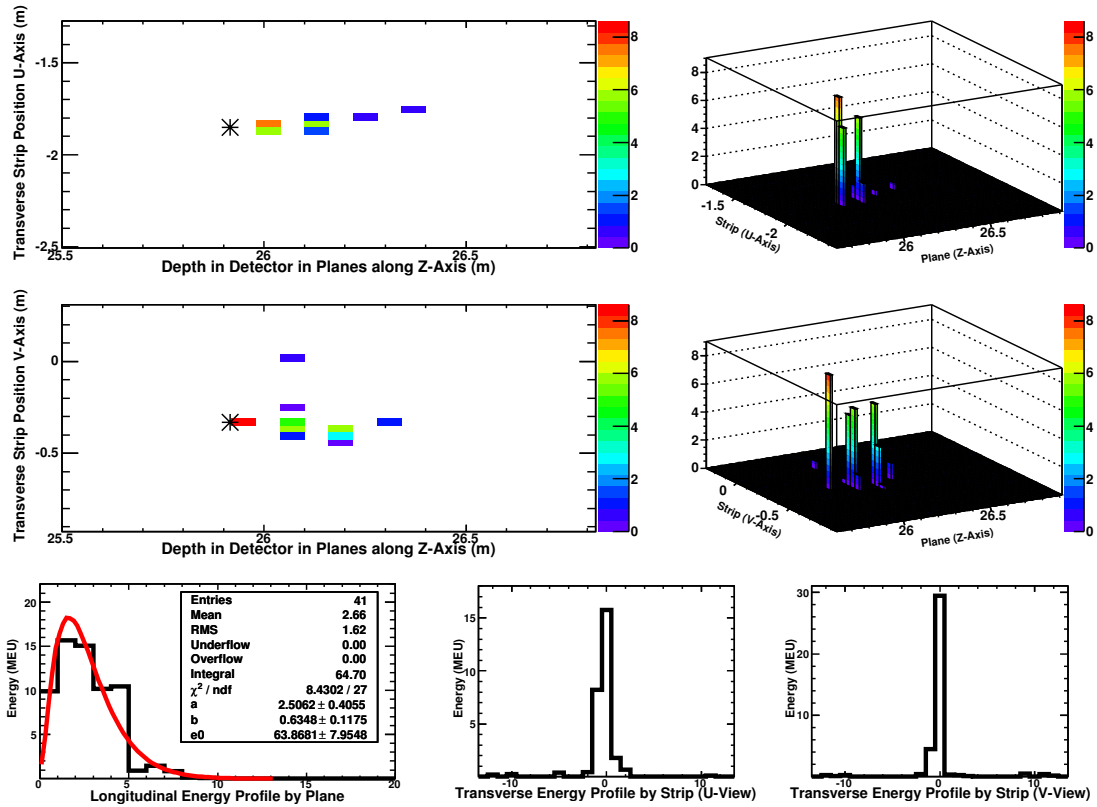


Figure 8.9: Golden  $\nu_e$ -CC event with ANN PID value of 0.92 and LEM PID value of 0.91 and reconstructed energy of 2.72GeV, found in FD subrun 38304. The two top left plots show reconstructed event hit strips as a function of longitudinal direction in the detector and transverse U or V position. The stars represent the reconstructed event vertex. The top right plots show the same information, but in a type of 3-D view. The bottom middle and right plots show the event transverse energy profiles in U and V, and the bottom left plot shows the longitudinal energy profile in black and the fitted EM-distribution in red.

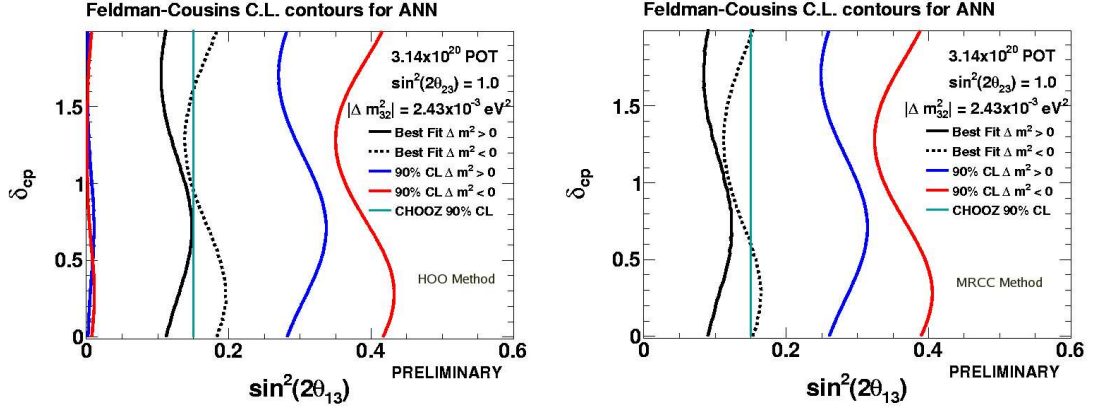


Figure 8.10: This Figure shows the final limit contours for a data exposure of  $3.14 \times 10^{20}$  POT after ANN PID selection. The results are shown for the HOO decomposition method on the left, and the MRCC method on the right. The black lines are the best fit lines for normal (solid) and inverted (dashed) mass hierarchy. The blue lines are the 90% confidence limits for the normal mass hierarchy, and the red lines are the 90% confidence limits for the inverted mass hierarchy. The cyan line is the limit set by the Chooz experiment.

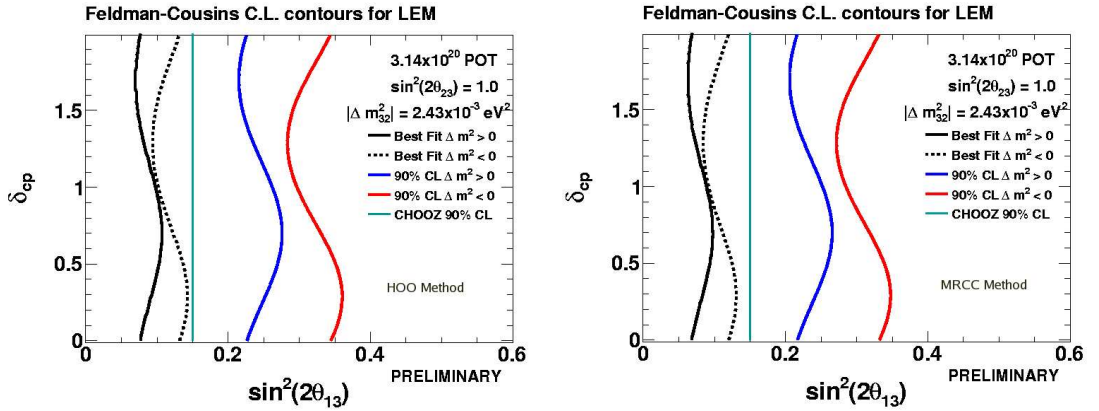


Figure 8.11: This Figure shows the final limit contours for a data exposure of  $3.14 \times 10^{20}$  POT after LEM PID selection. The results are shown for the HOO decomposition method on the left, and the MRCC method on the right. The black lines are the best fit lines for normal (solid) and inverted (dashed) mass hierarchy. The blue lines are the 90% confidence limits for the normal mass hierarchy, and the red lines are the 90% confidence limits for the inverted mass hierarchy. The cyan line is the limit set by the Chooz experiment.

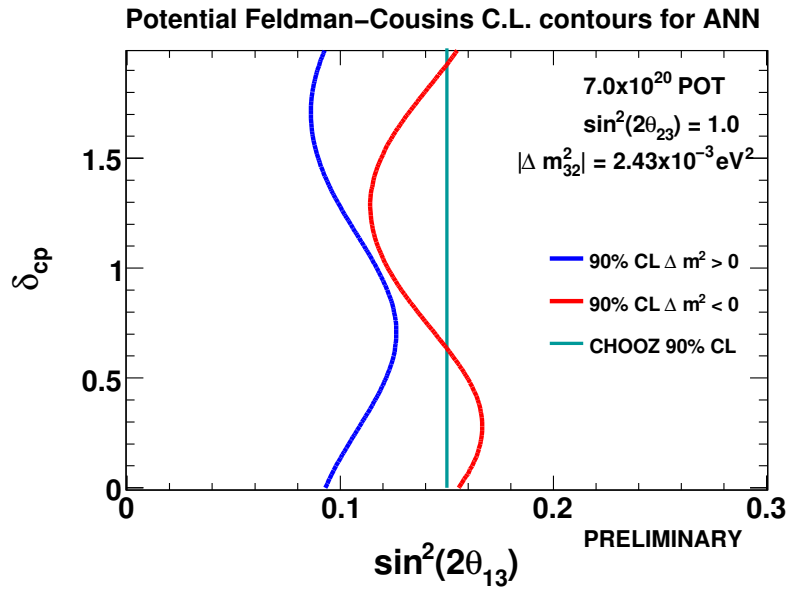


Figure 8.12: This Figure shows the projected 90% confidence limits in the absence of a  $\nu_e$ -appearance signal for a data exposure of  $7.0 \times 10^{20}$  POT. The two possible mass hierarchies have different limits and are compared to the CHOOZ limit of  $\sin^2(2\theta_{13}) = 0.15$ .

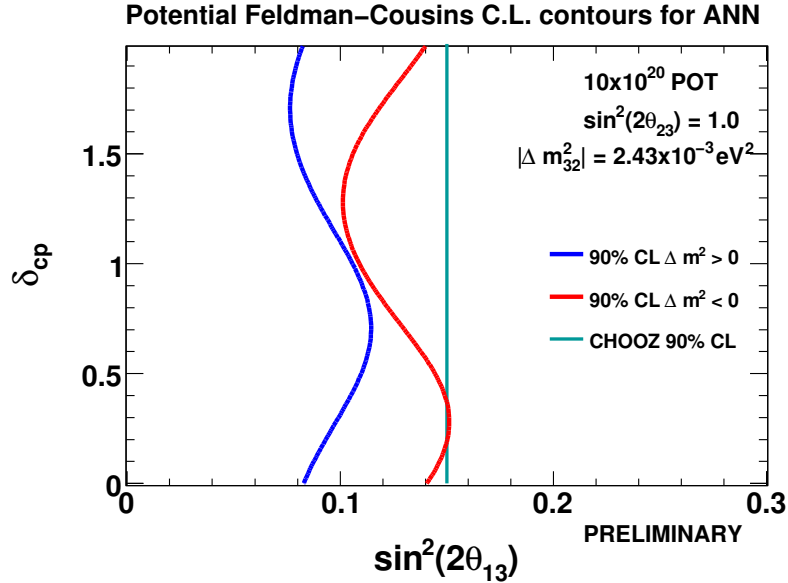


Figure 8.13: This Figure shows the projected 90% confidence limits in the absence of a  $\nu_e$ -appearance signal for a data exposure of  $10.0 \times 10^{20}$  POT. The two possible mass hierarchies have different limits and are compared to the CHOOZ limit of  $\sin^2(2\theta_{13}) = 0.15$ .



## Chapter 9

# Final Summary and Conclusions

The  $\nu_e$ -appearance analysis tries to hone in on a small and very hard-to-identify signal in a large background. As MINOS is a two-detector experiment, it is however possible to use the ND data to obtain a prediction for the FD expected backgrounds. The ND MC disagrees somewhat with the ND data and this is thought to be due to short-comings in hadronic shower modelling. A data driven background decomposition method was thus developed and it relies on muon removed charge current events - MRCC. For this method, the long muon tracks are removed from  $\nu_\mu$ -CC events for both data and MC, thus leaving hadronic showers which can be re-reconstructed as pseudo-NC events. These independent samples can then be used to apply a data-driven correction to the standard MC NC events and to obtain a final ND background decomposition. Using the MRCC method, the backgrounds in the ND are decomposed into three components. For the primary selection PID, those ND backgrounds, after separate extrapolation to the FD, translate into a total predicted FD background of  $28 \pm 5(\text{stat.}) \pm 2(\text{Syst.})$  events for the final FD data set with an exposure of  $3.14 \times 10^{20}$  POT. The other background decomposition method used in the  $\nu_e$ -appearance analysis, the HOO method, predicts  $27 \pm 5(\text{stat.}) \pm 2(\text{Syst.})$  FD data events, in excellent agreement with the MRCC method.

The MRCC method can also be used to provide an important FD analysis sideband prior to opening the final blinded data sample. ND MRCC data and MC ratios can be applied to FD MRCC MC to predict FD MRCC data. This sideband is unique as it provides the only means of probing whether there may be any problems with background events in the selected signal region prior to opening the final data sample. The MRCC FD sideband method predicted  $28.6 \pm 5.8$  events in the region selected by the primary

PID and observed 39 MRCC data events. Being within  $1.78\sigma$  of the prediction, and after exhaustive checks, the excess was judged to be most likely of a statistical nature.

Finally, the blinded FD data sample was opened and a small excess was seen for both selection PIDs and both background decomposition methods. In particular the primary PID observed 35 FD data events giving a small excess of 7 events above the prediction obtained from the MRCC background decomposition method. This may be interpreted as a  $\nu_e$ -appearance signal, however, one should be cautious when doing so considering the results of the sideband data. Nevertheless, taken at face value, and depending on selection PID and the ND decomposition method, the observed excess is between 1.0 and  $1.5\sigma$  above background. If a similar excess is seen in a future analysis, and with analysis improvements, MINOS may even reach a  $3\sigma$  discovery level for the value of the  $\sin^2 2\theta_{13}$  parameter. However, if the excess vanishes in the new analysis, MINOS will be able to set a better limit.

Analysis improvements for future  $\nu_e$ -appearance analysis include improvements in hadronic modelling and in event reconstruction. In addition, increased statistics may provide the possibility of carrying out the FD prediction in multiple bins of energy and / or PID instead of just as a counting experiment. Currently, an exposure of  $7.0 \times 10^{20}$  POT on target has been achieved, thus more than doubling the available data set. More data is being taken in reverse horn current mode at present and this data set should achieve an exposure of  $1.5\text{-}2.0 \times 10^{19}$  POT. This anti-neutrino mode data may be incorporated into a future analysis as well. Once this data set is taken, MINOS will reverse to neutrino mode again which should provide a final total exposure of  $9\text{-}10 \times 10^{19}$  POT in this mode, thus potentially more than tripling the current statistics and possibly allowing to carry out a measurement of  $\sin^2 2\theta_{13}$  for the first time.

# Bibliography

- [1] W. Pauli's letter for example reproduced in: I.J.R. Aitchison and A.J.G. Hey, Gauge Theories in Particle Physics, Vol.II, (2004), p.277
- [2] E. Fermi 1933, Atti Soc. It. Prog. Scient. **2**, (1933), 7-14 55 21-28
- [3] F. Reines and C.L. Cowan, Phys. Rev. **92**, (1953), p.830
- [4] F. Reines, C.L. Cowan, et al., Science **124**, **3212**, **1956**, p.103
- [5] C.S. Wu et al., Phys. Rev. **105**, (1957), p.1413
- [6] M. Goldhaber, L. Grodzins, and A. W. Sunyar, Phys. Rev. **109**, (1958), p. 1015 - 1017
- [7] B. Pontecorvo, Sov. Phys. JETP, (1957), [Zh. Eksp. Teor. Fiz, **33**, p. 549]
- [8] B. Pontecorvo, Sov. Phys. JETP, (1958), [Zh. Eksp. Teor. Fiz, **34**, p. 247]
- [9] B. Pontecorvo, Sov. Phys. JETP, (1967), [Zh. Eksp. Teor. Fiz, **53**, p. 1717]
- [10] Z. Maki, M. Nakagawa, S. Sakata, Progress of Theoretical Physics, Vol. **28**, **5**, (1962)
- [11] G. Danby et al., Phys. Rev. Lett. **9**, **1**, (1962), p.36
- [12] J.N. Bahcall et al., Astrophys. Journal, **137**, **1**, (1963), p. 344
- [13] R. Davis, D. Harmer, K. Hoffman, Phys. Rev. Lett. **20**, **21**, (1968), p.1205
- [14] Y. Fukuda et al., Phys. Rev. Lett. **81**, **6**, (1998)
- [15] W. Hampel et al., Phys. Lett. **447**, **1-2**, (1999)
- [16] J. Abdurashitov et al., Phys. Rev. C **80**, (2009)

- [17] Q.R. Ahmad et al., Phys. Rev. Lett. **87**, **7**, (2001)
- [18] R. Beckerszendy et al., Phys. Rev. D **46**, **9**, (1992)
- [19] K. Hirata et al., Phys. Lett. B, **205**, **2-3**, (1988)
- [20] Y. Fukuda et al., Phys. Rev. Lett. **81**, **8**, (1998)
- [21] R.M. Bionta et al., Phys. Rev. Lett. **58**, **14**, (1987), p. 1494
- [22] K. Hirata et al., Phys. Rev. Lett. **58**, **14**, (1987), p. 1490
- [23] Q.R. Ahmad et al., Phys. Rev. Lett. **89**, **1**, 011301, (2002)
- [24] S. Glashow, Nuclear Physics **22**, **4**, (1961)
- [25] S. Weinberg, Phys. Rev. Lett. **19**, (1967)
- [26] A. Salam, Proc. of the 8th Nobel Symposium, (1968)
- [27] The ALEPH Collaboration, the DELPHI Collaboration, the L3 Collaboration, the OPAL Collaboration, the SLD Collaboration, the LEP Electroweak Working Group and the SLD Electroweak and Heavy Flavour Groups, Physics Reports **427**, **5-6**, (2006), Precision Electroweak Measurements on the Z Resonance
- [28] L. Wolfenstein, Phys. Rev. D **17**, (1978), p.2369
- [29] S.P. Mikheev, A.Y. Smirnov, Sov. Jour. Nucl. Phys. **42**, **6**, (1985), p. 913
- [30] M. Apollonio, et al. [Chooz Collaboration], European Phys. Jour. C **27**, **3**, (2003), p.331
- [31] C. Amsler et al. (Particle Data Group), Phys. Lett. B **667**, **1**, (2008)
- [32] P. Adamson et al., Phys. Rev. D **77**, (2008)
- [33] Numi Technical Design Handbook,  
URL: [http://www-numi.fnal.gov/numwork/tdh/tdh\\_index.html](http://www-numi.fnal.gov/numwork/tdh/tdh_index.html)
- [34] MINOS Technical Design Report,  
URL: [http://www-numi.fnal.gov/Minos/info/minos\\_tdr.html](http://www-numi.fnal.gov/Minos/info/minos_tdr.html)
- [35] S. Kopp, Phys. Rep. **439**, **101**, (2007)
- [36] R.Zwaska et al., IEEE Trans. on Nucl. Science, **50**, **4**, 2003

- [37] D.Indurthy et al., AIP Conf. Proc., **732**, (2004)
- [38] P.Adamson et al., Nucl.Instrum.Meth. A **556**, (2006)
- [39] D.G. Michael et al., Nucl.Instrum.Meth. A **596**, (2008)
- [40] R.Nichol, PHD Thesis, University College London, 2003
- [41] P.Adamson, PHD Thesis, University of Sussex, 2001
- [42] C.B.Smith, PHD Thesis, University College London, 2002
- [43] A.E.Cabrera Serra, PHD Thesis, University of Oxford, The Queen's College, 2005
- [44] P.L.Vahle, PHD Thesis, The University of Texas at Austin, 2004
- [45] M.Kordosky, PHD Thesis, The University of Texas at Austin, 2004
- [46] M.A.Barker, PHD Thesis, University of Oxford, Brasenose College, 2003
- [47] P.Adamson et al., Phys. Rev. Lett. **101**, (2008)
- [48] P.Adamson et al., Phys. Rev. Lett. **97**, (2006)
- [49] P.Adamson et al., Phys. Rev. Lett. **101**, (2008)
- [50] P.Adamson et al., Phys. Rev. D **81**, (2010)
- [51] R. Brun et. al., CERN Program Library W5013 (1984).
- [52] A. Fasso et. al., CERN-2005-10, INFN/TC05/11, SLAC-R-773 (2005).
- [53] H. Gallagher. ArXiv:0806.2119 (2008).
- [54] T. Vahle, NuMI Note 4060, MINOS Document Database,  $\nu_e$  B-Field Considerations
- [55] S. Kopp, Proc. 2005 IEEE Particle Accel. Conference, **1-4**, (2005), ArXiv:0508001
- [56] A. Holin et al., MINOS DocDB 3707,  $\nu_e$  Group Official Cuts and Files
- [57] S.Cavanaugh, NuMI Note 4145, MINOS Document Database, Cosmic Background in NueAna
- [58] S.Cavanaugh, NuMI Note 5382, MINOS Document Database, Cosmic Background in NueAna FireBird

- [59] G.Pawloski, T.Yang, NuMI Note 5319, MINOS Document Database, Cross-Talk Tuning
- [60] P.Ochoa, NuMI Document 4177, MINOS Document Database, Selection Optimisation
- [61] T.Yang, NuMI Document 3393, MINOS Document Database, A Neural Network Technique for  $\nu_e$  Identification
- [62] P.Ochoa, NuMI Document 5031, MINOS Document Database, The Monte Carlo Nearest Neighbours (MCNN) Selection - LEM
- [63] H.Gallagher, T.Yang, J.Boehm, NuMI Document 5392, MINOS Document Database, Hadronization Model Uncertainties for the Electron Neutrino Appearance Analysis
- [64] G.Pawloski, L. Whitehead, NuMI Document 3935, MINOS Document Database, Estimation of the Near Detector NC and CC Spectra for the  $\nu_\mu \rightarrow \nu_e$  Analysis with Horn-On and Horn-Off Data
- [65] H.Gallagher, NuMI Document 5413, MINOS Document Database, The Horn On-Off Method and Physics Simulations Uncertainties
- [66] C. Smith and C. Howcroft, NuMI Document 2415, MINOS Document Database, Measuring  $\nu_e$ -CC Selection Efficiency from  $\nu_\mu$ -CC Data with Muon Removal and MC Electron Addition
- [67] A.Holin, NuMI Document 4141, MINOS Document Database, Muon Removal Background Estimation Method
- [68] H.R. Gallagher, NuMI Document 3541, Modeling  $\nu_e$ -like showers in NC and CC events
- [69] C. Andreopoulos et al., NuMI Document 2989, MINOS Document Database
- [70] J.Boehm, NuMI Document 5355, MINOS Document Database,  $\nu_e$  Oscillation Note
- [71] E.K.Akhmedov et al., J. High Energy Phys. 04, **78**, (2004)
- [72] J.Boehm, NuMI Document 5405, MINOS Document Database, Far/Near Extrapolation Method in the  $\nu_e$  Analysis

- [73] J.Boehm, PHD Thesis, Harvard University, 2009
- [74] J.Boehm, NuMI Document 5385, MINOS Document Database, Measuring  $\nu_e$ -CC Selection Efficiency and Systematic Error using the Muon Removal with MC Electron Addition Process
- [75] J.Boehm, NuMI Document 5387, MINOS Document Database, Evaluation of Systematic Uncertainties on the Measurement of the  $\nu_e$ -CC Appearance
- [76] R.B. Toner, NuMI Document 5257, MINOS Document Database,  $\nu_e$  Preselection Systematics
- [77] G.Pawloski, NuMI Document 5455, MINOS Document Database,  $\nu_e$  Anti-PID Sideband
- [78] G.Feldman, NuMI Document 5650, MINOS Document Database, Calculation of Unified Approach Confidence Limits for the  $\nu_e$  Analysis
- [79] G.Feldman, R.Cousins, Phys. Rev. D **57**, **7**, (1998)

## Appendix A

# Additional ND Distributions

This appendix contains additional MRCC and standard sample ND comparisons. The first set of plots is for the kinematic truth variables as discussed in section 6.1.1. The second set of comparisons includes additional data/MC distributions as discussed in section 6.2.



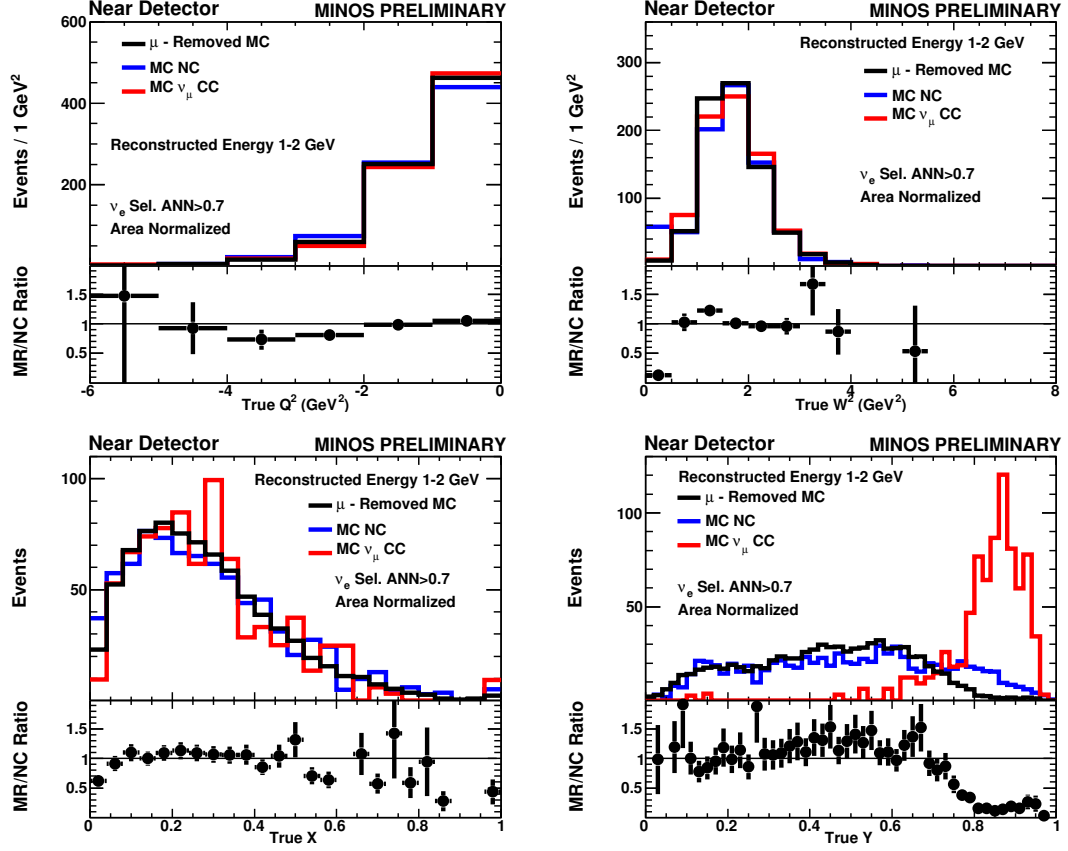


Figure A.1: Kinematic truth distributions after  $\nu_e$  selection (ANN PID) for ND MRCC and standard MC. All distributions are for an energy slice between 1 and 2 GeV. In the four sets of plots, the top plots show the distributions of MRCC events in black, standard MC NC in blue, and standard MC  $\nu_\mu$ -CC in red. The top plots are area normalised to reveal shape differences. The bottom plots show the ratios of the MRCC to the standard MC NC distributions. The top left set of plots shows true  $Q^2$  - the four-momentum squared transferred by the neutrino to the struck nucleon. The top right set of plots shows true  $W^2$  - the mass squared of the system recoiling against the scattered neutrino. The bottom left set of plots shows true  $x$  - the fraction of the nucleon's momentum carried by the struck quark. Finally, the bottom right set of plots shows true  $y$  - the fraction of the neutrino's energy lost in the nucleon rest frame.

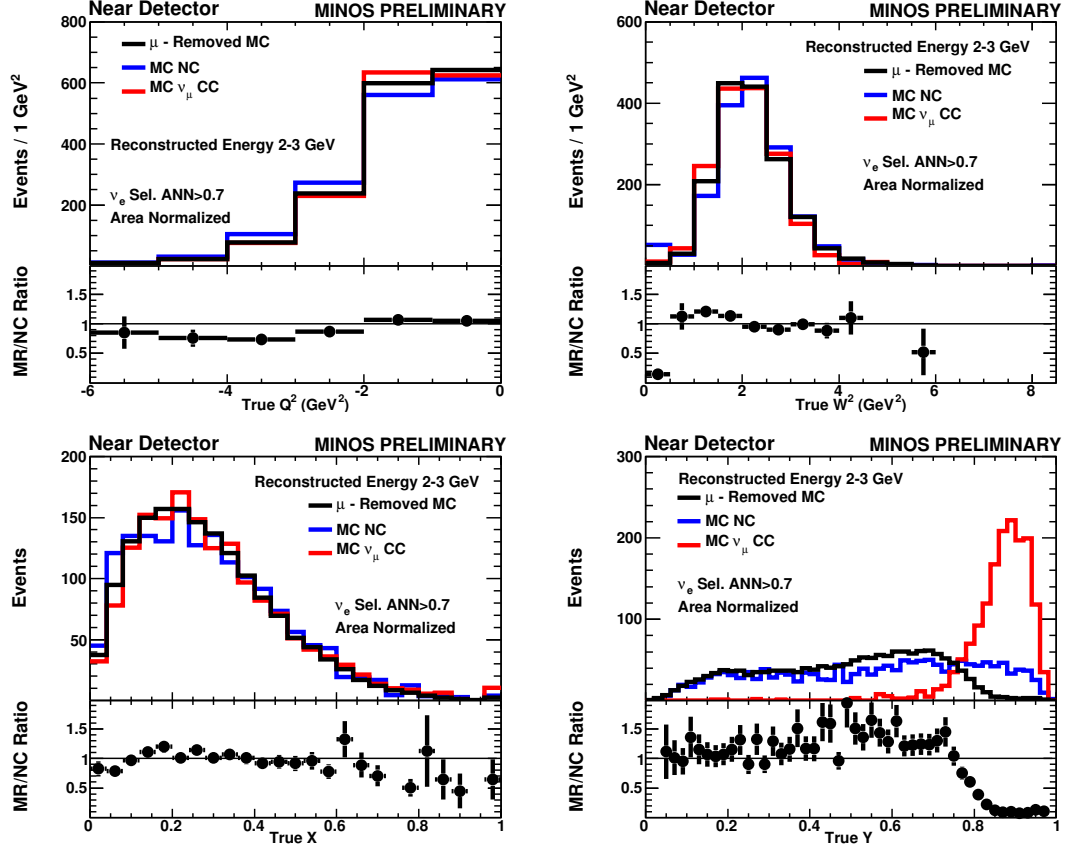


Figure A.2: Kinematic truth distributions after  $\nu_e$  selection (ANN PID) for ND MRCC and standard MC. All distributions are for an energy slice between 2 and 3 GeV. In the four sets of plots, the top plots show the distributions of MRCC events in black, standard MC NC in blue, and standard MC  $\nu_\mu$ -CC in red. The top plots are area normalised to reveal shape differences. The bottom plots show the ratios of the MRCC to the standard MC NC distributions. The top left set of plots shows true  $Q^2$  - the four-momentum squared transferred by the neutrino to the struck nucleon. The top right set of plots shows true  $W^2$  - the mass squared of the system recoiling against the scattered neutrino. The bottom left set of plots shows true  $x$  - the fraction of the nucleon's momentum carried by the struck quark. Finally, the bottom right set of plots shows true  $y$  - the fraction of the neutrino's energy lost in the nucleon rest frame.

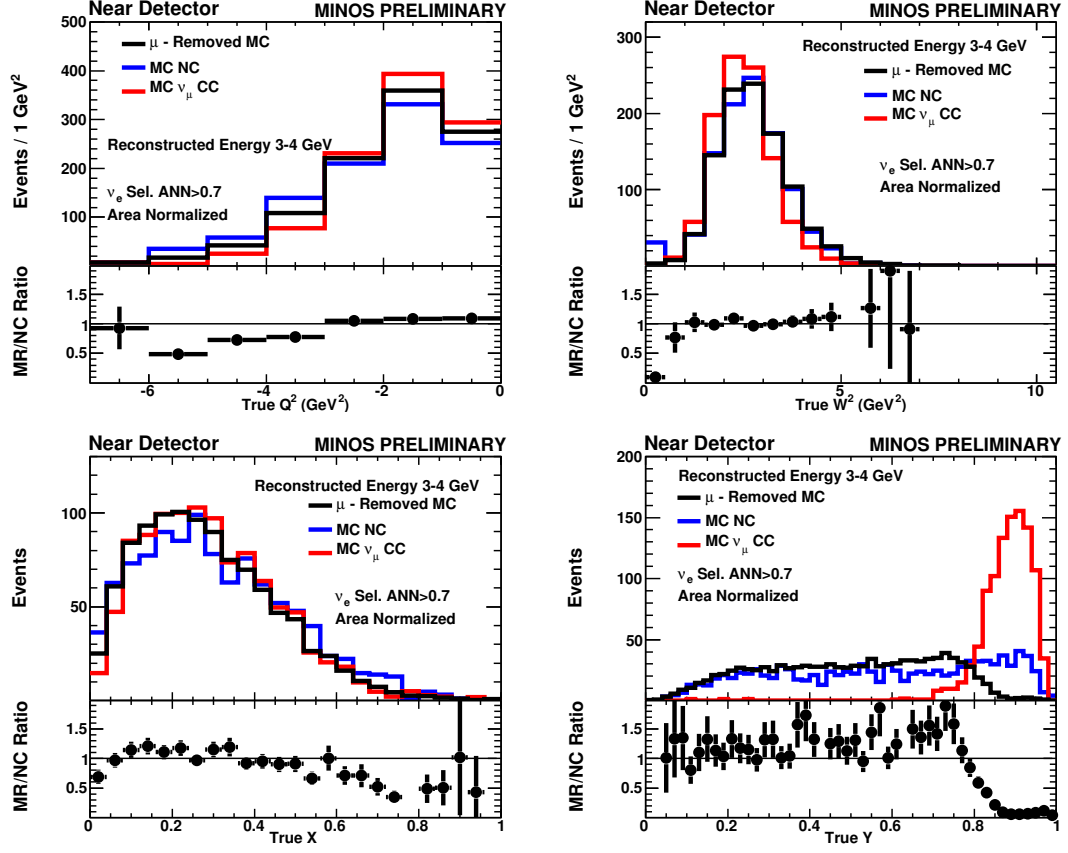


Figure A.3: Kinematic truth distributions after  $\nu_e$  selection (ANN PID) for ND MRCC and standard MC. All distributions are for an energy slice between 3 and 4 GeV. In the four sets of plots, the top plots show the distributions of MRCC events in black, standard MC NC in blue, and standard MC  $\nu_\mu$ -CC in red. The top plots are area normalised to reveal shape differences. The bottom plots show the ratios of the MRCC to the standard MC NC distributions. The top left set of plots shows true  $Q^2$  - the four-momentum squared transferred by the neutrino to the struck nucleon. The top right set of plots shows true  $W^2$  - the mass squared of the system recoiling against the scattered neutrino. The bottom left set of plots shows true  $x$  - the fraction of the nucleon's momentum carried by the struck quark. Finally, the bottom right set of plots shows true  $y$  - the fraction of the neutrino's energy lost in the nucleon rest frame.

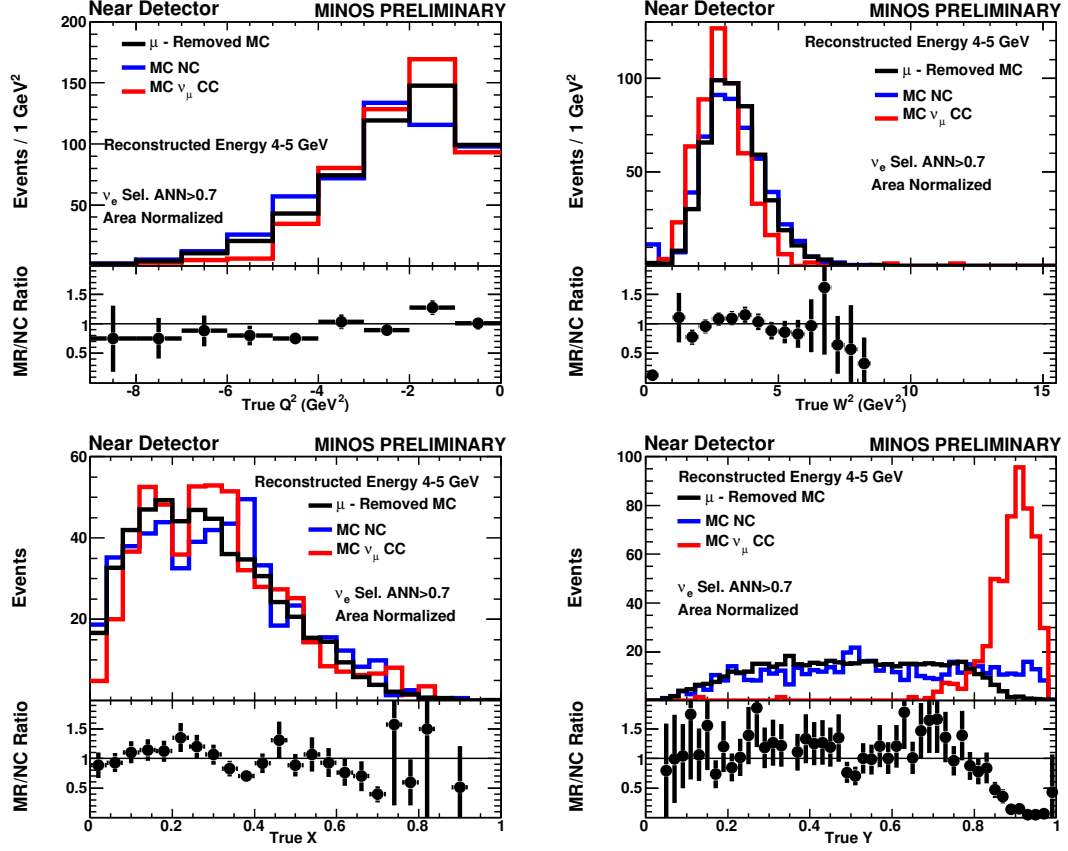


Figure A.4: Kinematic truth distributions after  $\nu_e$  selection (ANN PID) for ND MRCC and standard MC. All distributions are for an energy slice between 4 and 5 GeV. In the four sets of plots, the top plots show the distributions of MRCC events in black, standard MC NC in blue, and standard MC  $\nu_\mu$ -CC in red. The top plots are area normalised to reveal shape differences. The bottom plots show the ratios of the MRCC to the standard MC NC distributions. The top left set of plots shows true  $Q^2$  - the four-momentum squared transferred by the neutrino to the struck nucleon. The top right set of plots shows true  $W^2$  - the mass squared of the system recoiling against the scattered neutrino. The bottom left set of plots shows true  $x$  - the fraction of the nucleon's momentum carried by the struck quark. Finally, the bottom right set of plots shows true  $y$  - the fraction of the neutrino's energy lost in the nucleon rest frame.

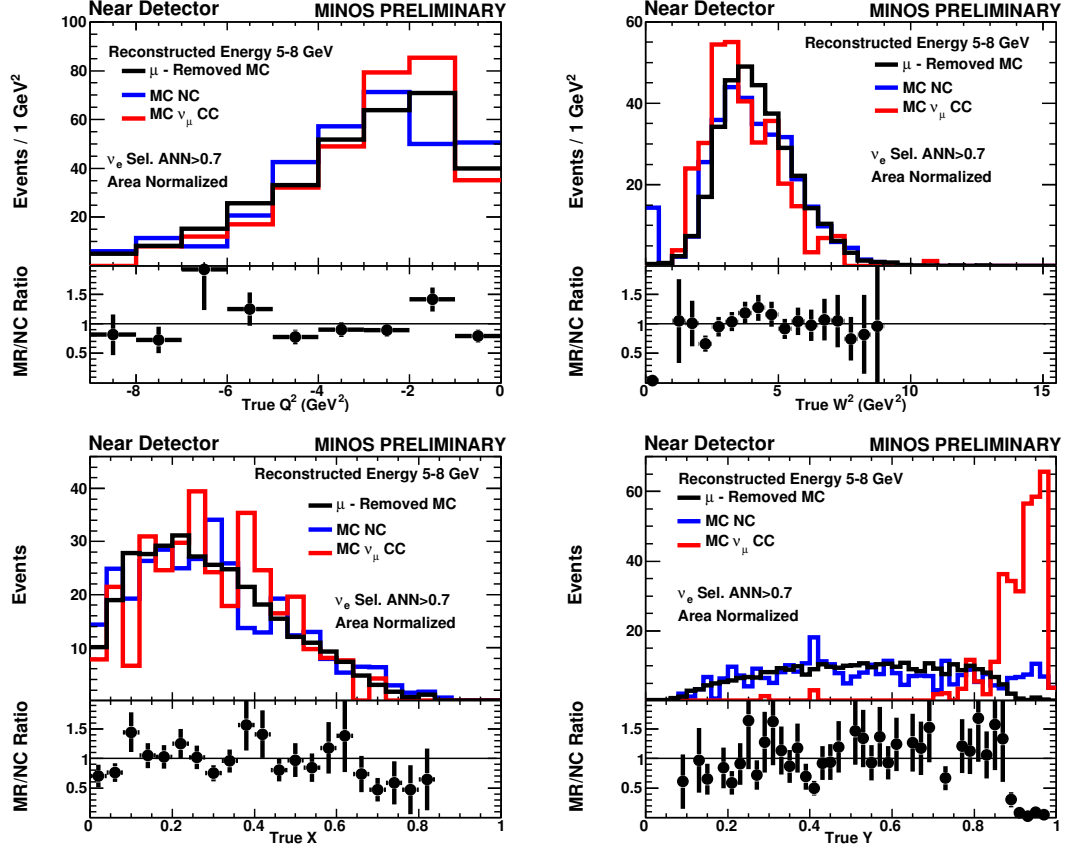


Figure A.5: Kinematic truth distributions after  $\nu_e$  selection (ANN PID) for ND MRCC and standard MC. All distributions are for an energy slice between 5 and 8 GeV. In the four sets of plots, the top plots show the distributions of MRCC events in black, standard MC NC in blue, and standard MC  $\nu_\mu$ -CC in red. The top plots are area normalised to reveal shape differences. The bottom plots show the ratios of the MRCC to the standard MC NC distributions. The top left set of plots shows true  $Q^2$  - the four-momentum squared transferred by the neutrino to the struck nucleon. The top right set of plots shows true  $W^2$  - the mass squared of the system recoiling against the scattered neutrino. The bottom left set of plots shows true  $x$  - the fraction of the nucleon's momentum carried by the struck quark. Finally, the bottom right set of plots shows true  $y$  - the fraction of the neutrino's energy lost in the nucleon rest frame.

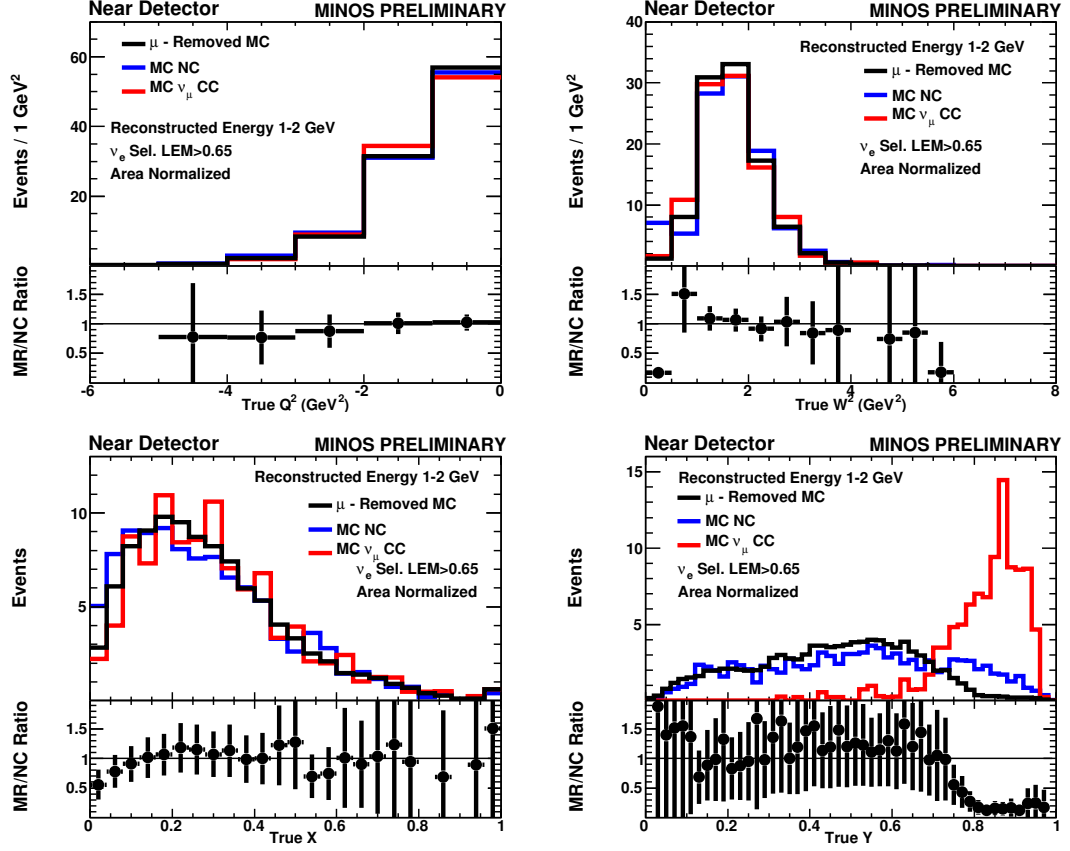


Figure A.6: Kinematic truth distributions after  $\nu_e$  selection (LEM PID) for ND MRCC and standard MC. All distributions are for an energy slice between 1 and 2 GeV. In the four sets of plots, the top plots show the distributions of MRCC events in black, standard MC NC in blue, and standard MC  $\nu_\mu$ -CC in red. The top plots are area normalized to reveal shape differences. The bottom plots show the ratios of the MRCC to the standard MC NC distributions. The top left set of plots shows true  $Q^2$  - the four-momentum squared transferred by the neutrino to the struck nucleon. The top right set of plots shows true  $W^2$  - the mass squared of the system recoiling against the scattered neutrino. The bottom left set of plots shows true  $x$  - the fraction of the nucleon's momentum carried by the struck quark. Finally, the bottom right set of plots shows true  $y$  - the fraction of the neutrino's energy lost in the nucleon rest frame.

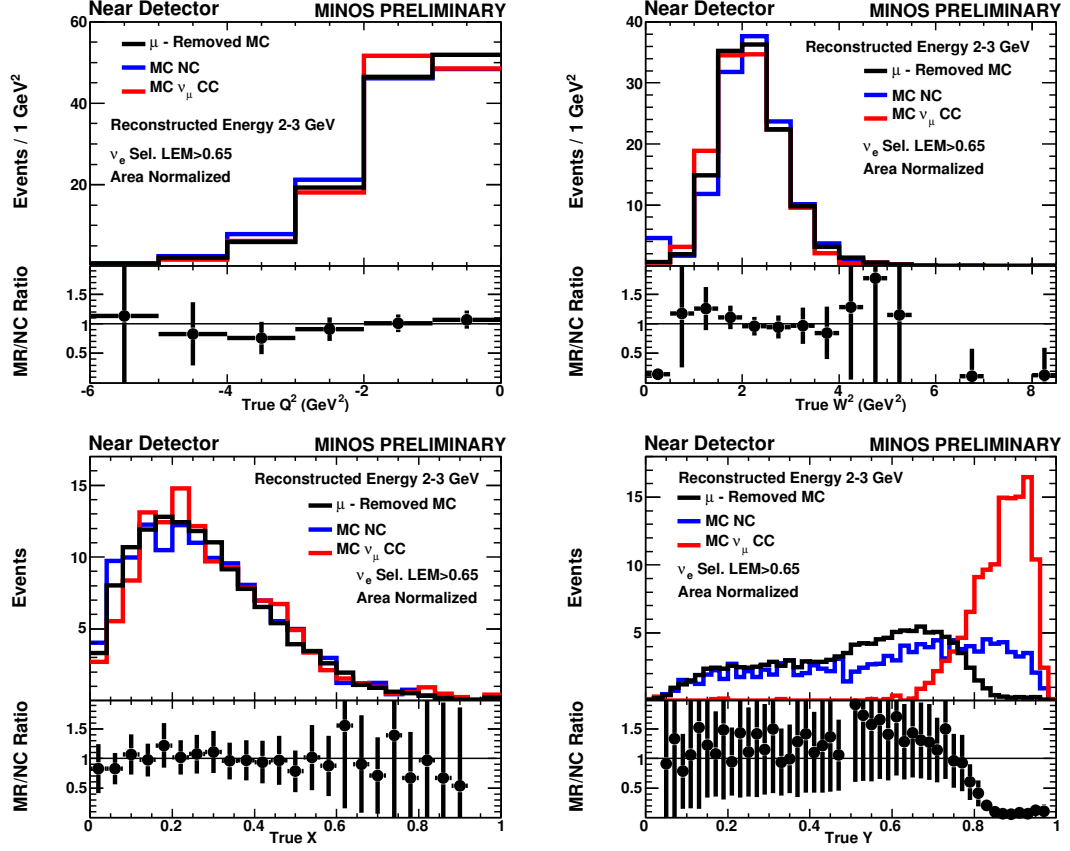


Figure A.7: Kinematic truth distributions after  $\nu_e$  selection (LEM PID) for ND MRCC and standard MC. All distributions are for an energy slice between 2 and 3 GeV. In the four sets of plots, the top plots show the distributions of MRCC events in black, standard MC NC in blue, and standard MC  $\nu_\mu$ -CC in red. The top plots are area normalised to reveal shape differences. The bottom plots show the ratios of the MRCC to the standard MC NC distributions. The top left set of plots shows true  $Q^2$  - the four-momentum squared transferred by the neutrino to the struck nucleon. The top right set of plots shows true  $W^2$  - the mass squared of the system recoiling against the scattered neutrino. The bottom left set of plots shows true  $x$  - the fraction of the nucleon's momentum carried by the struck quark. Finally, the bottom right set of plots shows true  $y$  - the fraction of the neutrino's energy lost in the nucleon rest frame.

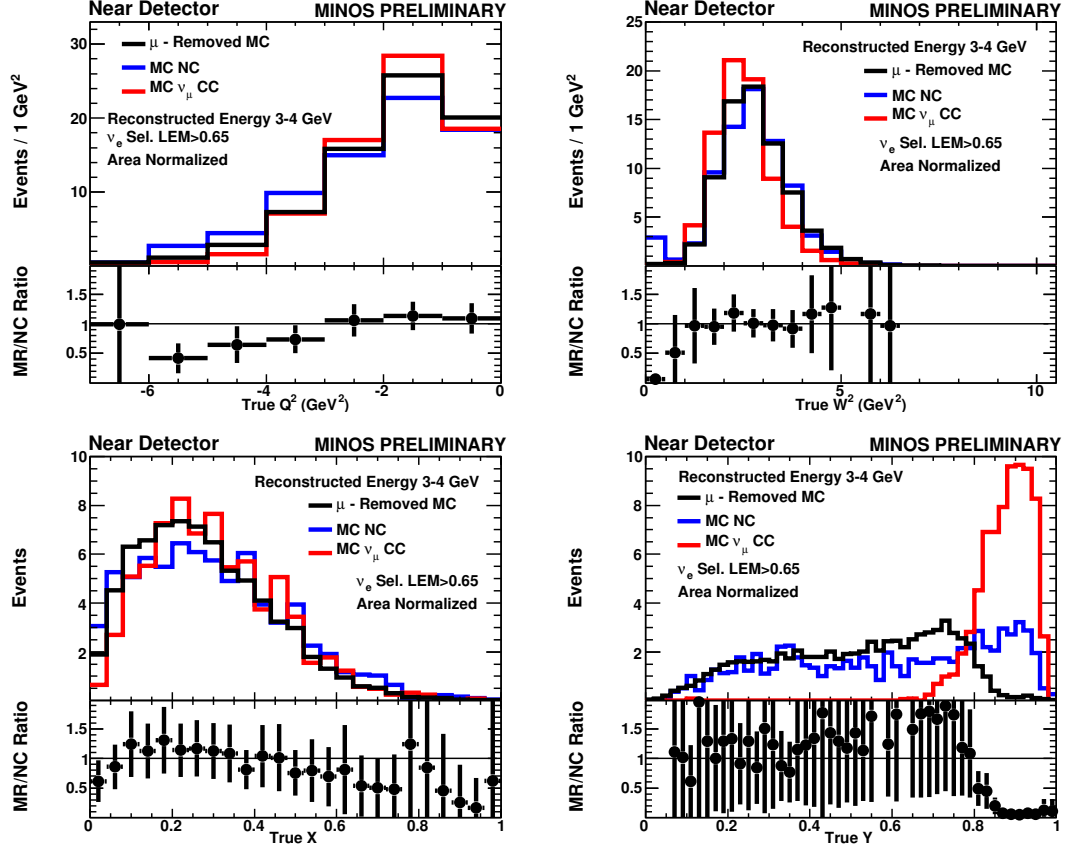


Figure A.8: Kinematic truth distributions after  $\nu_e$  selection (LEM PID) for ND MRCC and standard MC. All distributions are for an energy slice between 3 and 4 GeV. In the four sets of plots, the top plots show the distributions of MRCC events in black, standard MC NC in blue, and standard MC  $\nu_\mu$ -CC in red. The top plots are area normalized to reveal shape differences. The bottom plots show the ratios of the MRCC to the standard MC NC distributions. The top left set of plots shows true  $Q^2$  - the four-momentum squared transferred by the neutrino to the struck nucleon. The top right set of plots shows true  $W^2$  - the mass squared of the system recoiling against the scattered neutrino. The bottom left set of plots shows true  $x$  - the fraction of the nucleon's momentum carried by the struck quark. Finally, the bottom right set of plots shows true  $y$  - the fraction of the neutrino's energy lost in the nucleon rest frame.



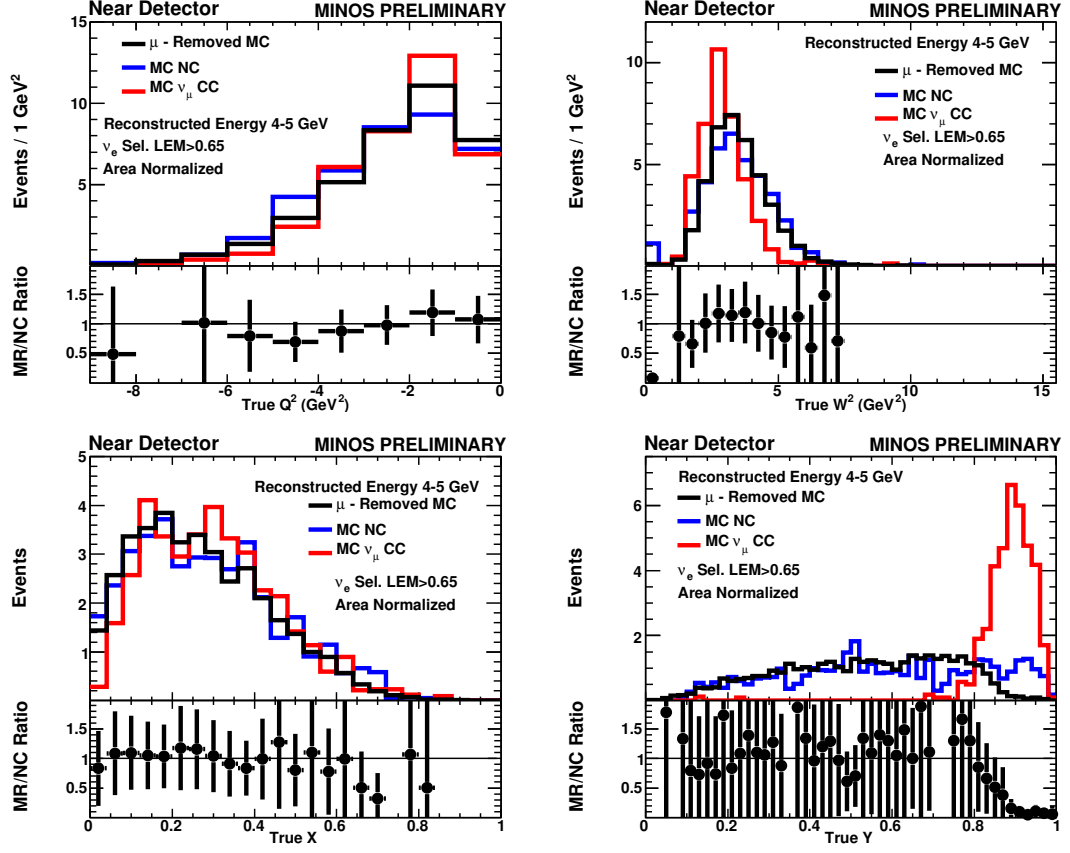


Figure A.9: Kinematic truth distributions after  $\nu_e$  selection (LEM PID) for ND MRCC and standard MC. All distributions are for an energy slice between 4 and 5 GeV. In the four sets of plots, the top plots show the distributions of MRCC events in black, standard MC NC in blue, and standard MC  $\nu_\mu$ -CC in red. The top plots are area normalised to reveal shape differences. The bottom plots show the ratios of the MRCC to the standard MC NC distributions. The top left set of plots shows true  $Q^2$  - the four-momentum squared transferred by the neutrino to the struck nucleon. The top right set of plots shows true  $W^2$  - the mass squared of the system recoiling against the scattered neutrino. The bottom left set of plots shows true  $x$  - the fraction of the nucleon's momentum carried by the struck quark. Finally, the bottom right set of plots shows true  $y$  - the fraction of the neutrino's energy lost in the nucleon rest frame.

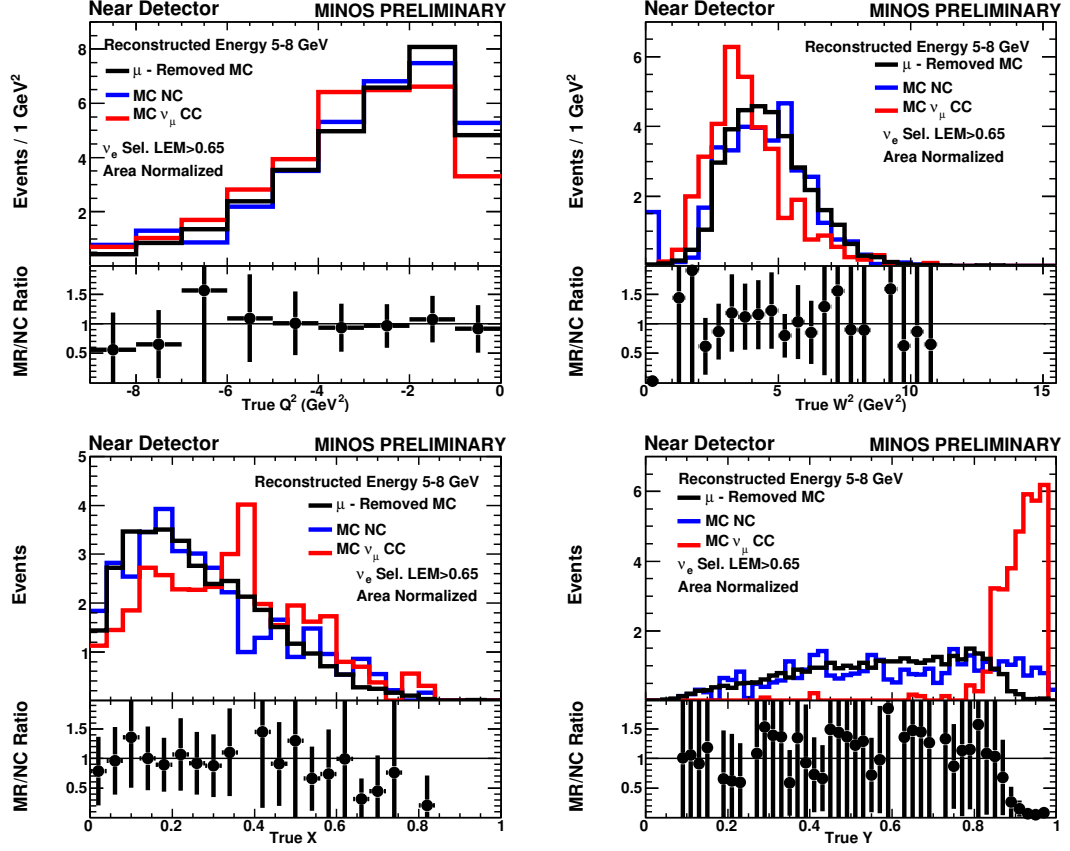


Figure A.10: Kinematic truth distributions after  $\nu_e$  selection (LEM PID) for ND MRCC and standard MC. All distributions are for an energy slice between 5 and 8 GeV. In the four sets of plots, the top plots show the distributions of MRCC events in black, standard MC NC in blue, and standard MC  $\nu_\mu$ -CC in red. The top plots are area normalized to reveal shape differences. The bottom plots show the ratios of the MRCC to the standard MC NC distributions. The top left set of plots shows true  $Q^2$  - the four-momentum squared transferred by the neutrino to the struck nucleon. The top right set of plots shows true  $W^2$  - the mass squared of the system recoiling against the scattered neutrino. The bottom left set of plots shows true  $x$  - the fraction of the nucleon's momentum carried by the struck quark. Finally, the bottom right set of plots shows true  $y$  - the fraction of the neutrino's energy lost in the nucleon rest frame.

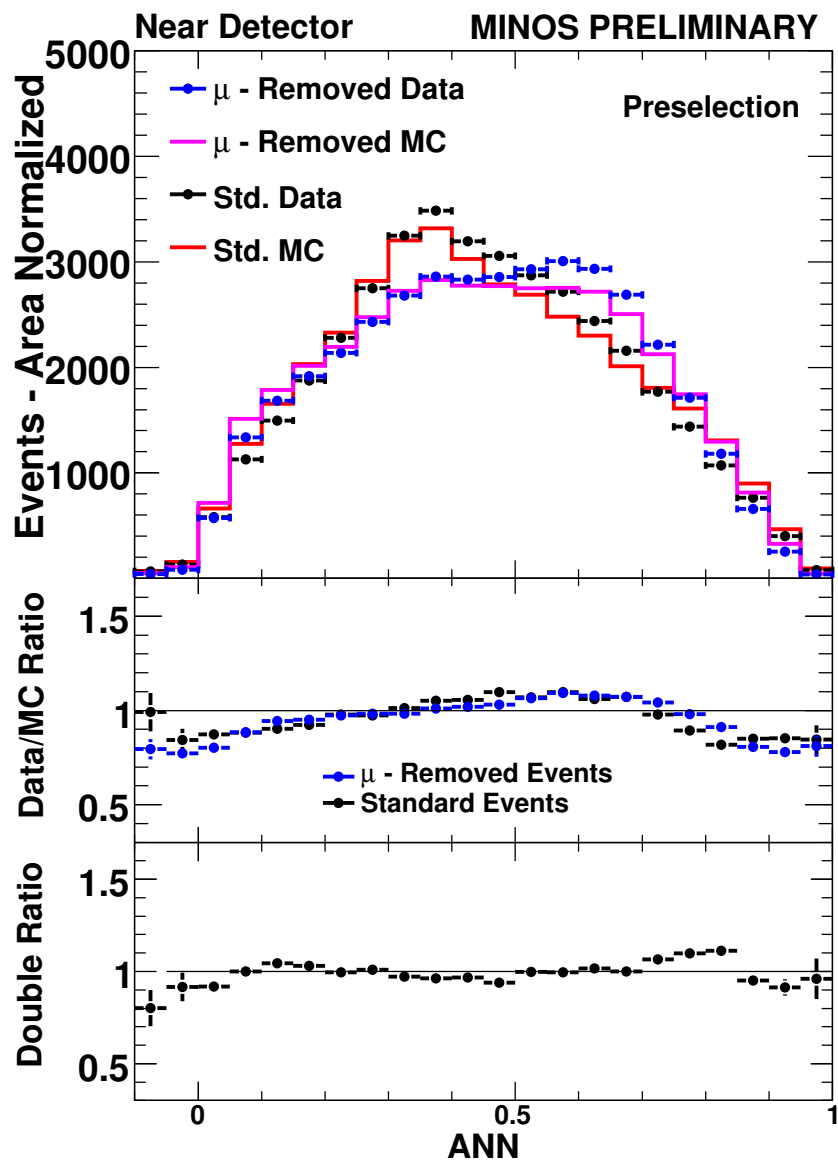


Figure A.11: Near Detector Data/MC MRCC/Standard comparisons after  $\nu_e$  preselection cuts. Distributions are normalised to the  $10^{19}$  POT standard data area to show shape differences. The variable shown here is the ANN PID used to select CC- $\nu_e$  events.

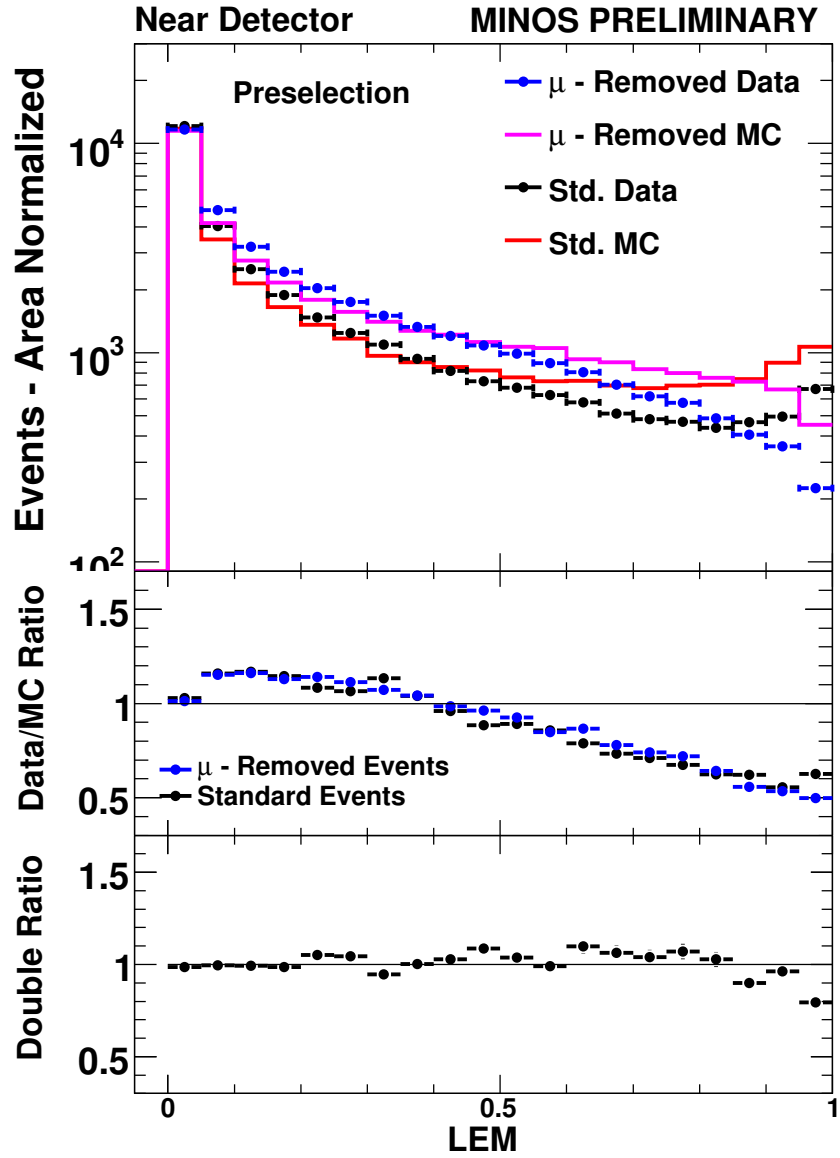


Figure A.12: Near Detector Data/MC MRCC/Standard comparisons after  $\nu_e$  preselection cuts. Distributions are normalised to the  $10^{19}$  POT standard data area to show shape differences. The variable shown here is the LEM PID used to select CC- $\nu_e$  events.

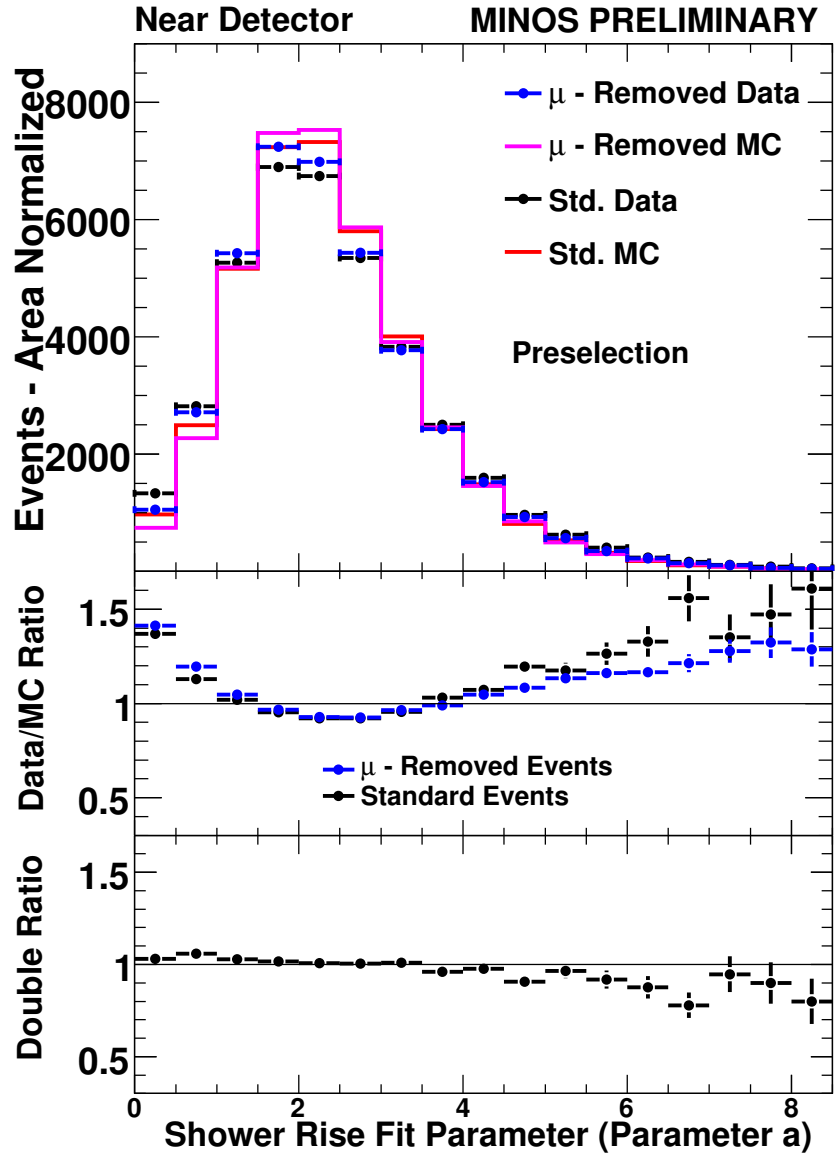


Figure A.13: Near Detector Data/MC MRCC/Standard comparisons after  $\nu_e$  preselection cuts. Distributions are normalised to the  $10^{19}$  POT standard data area to show shape differences. The variable shown here is the parameter associated with the rise of the shower.

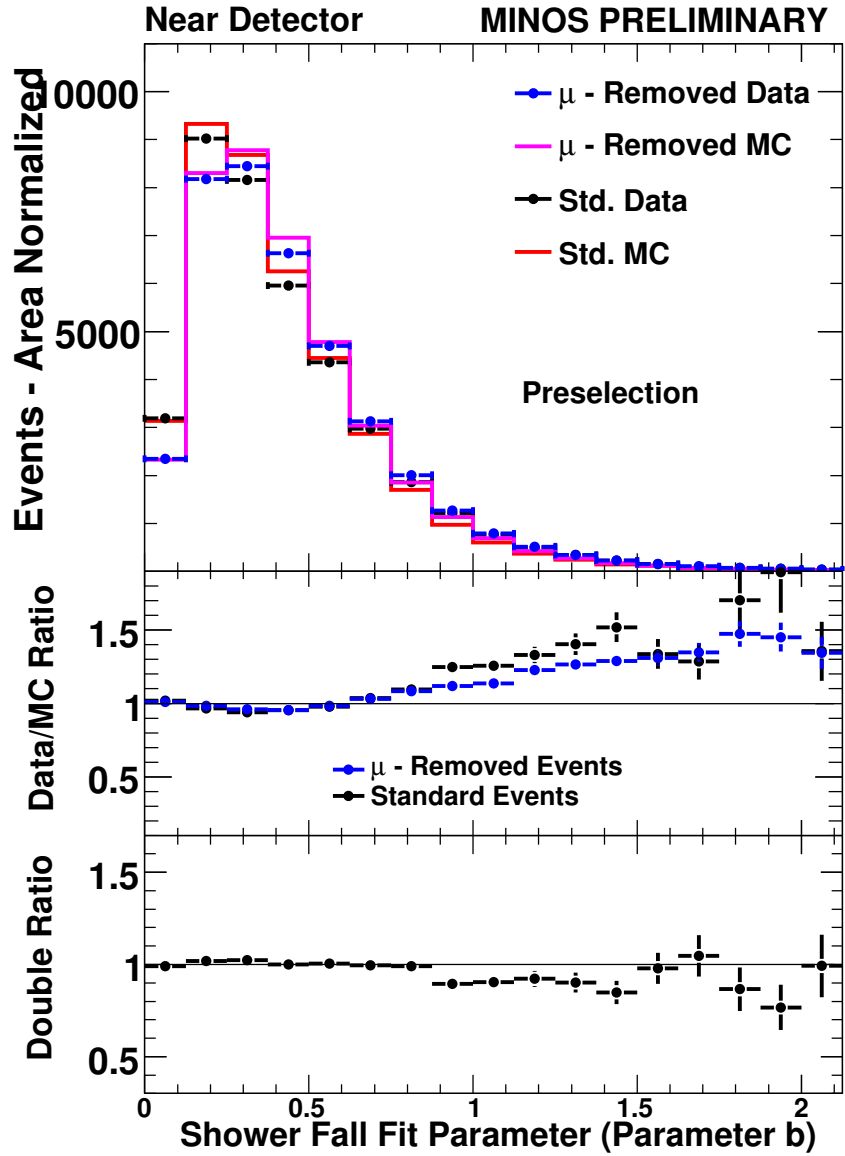


Figure A.14: Near Detector Data/MC MRCC/Standard comparisons after  $\nu_e$  preselection cuts. Distributions are normalised to the  $10^{19}$  POT standard data area to show shape differences. The variable shown here is the parameter associated with the fall of the shower.

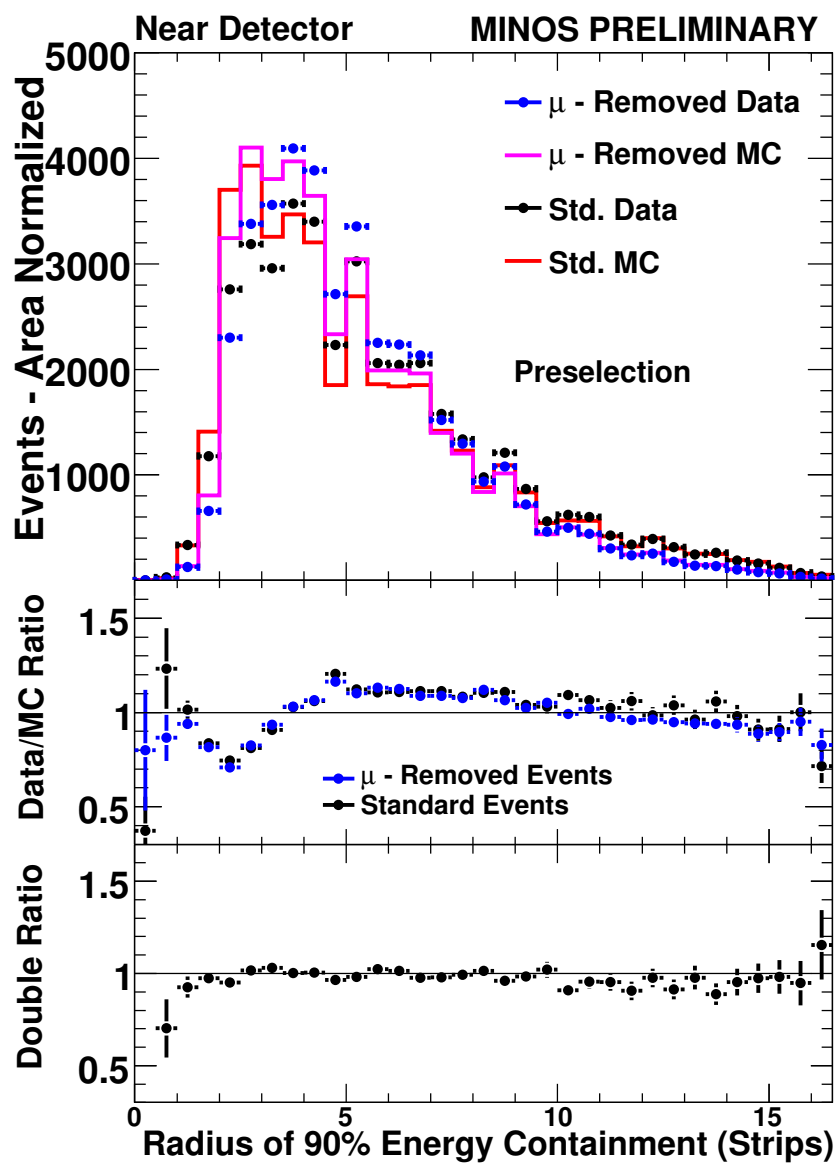


Figure A.15: Near Detector Data/MC MRCC/Standard comparison after  $\nu_e$  preselection cuts. Distributions are normalised to the same area to show shape differences. The variable shown here is the Moliere radius - radius of 90% energy containment - of the shower.

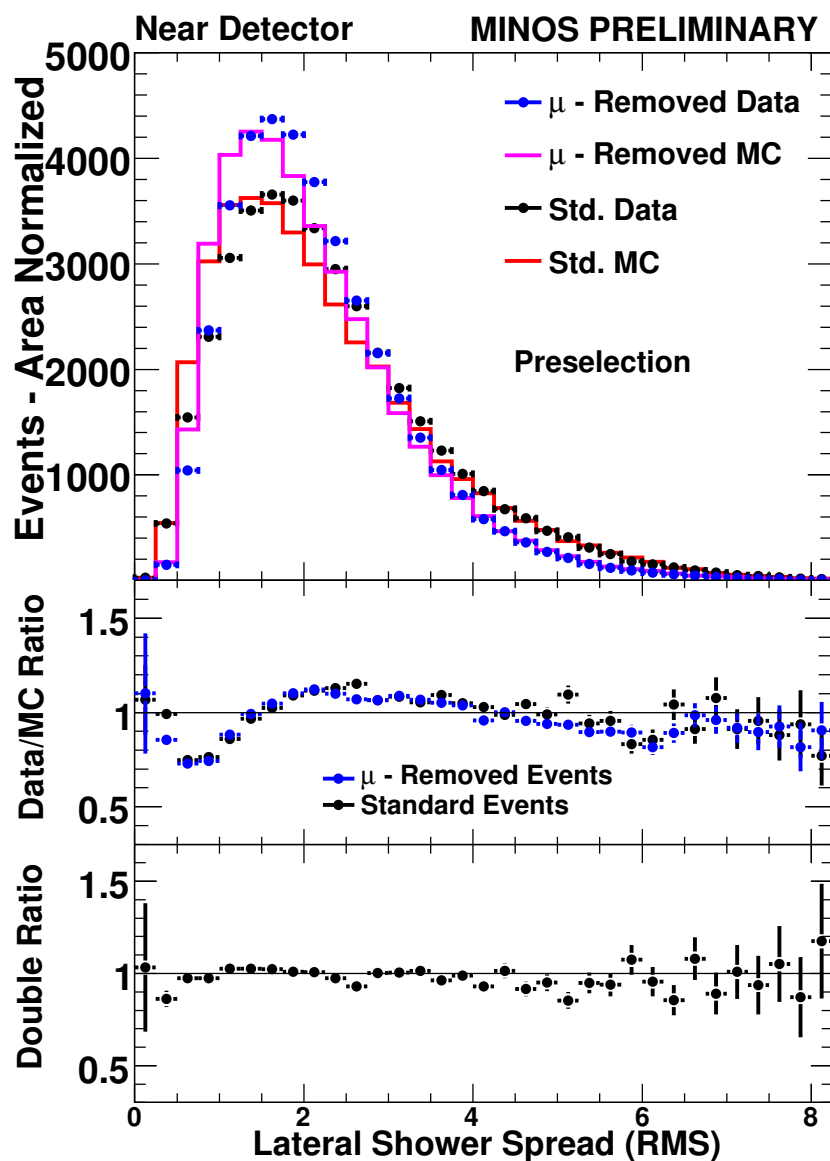


Figure A.16: Near Detector Data/MC MRCC/Standard comparison after  $\nu_e$  preselection cuts. Distributions are normalised to the same area to show shape differences. The variable shown here is the UV RMS of the shower fit, i.e. the lateral spread of the shower.



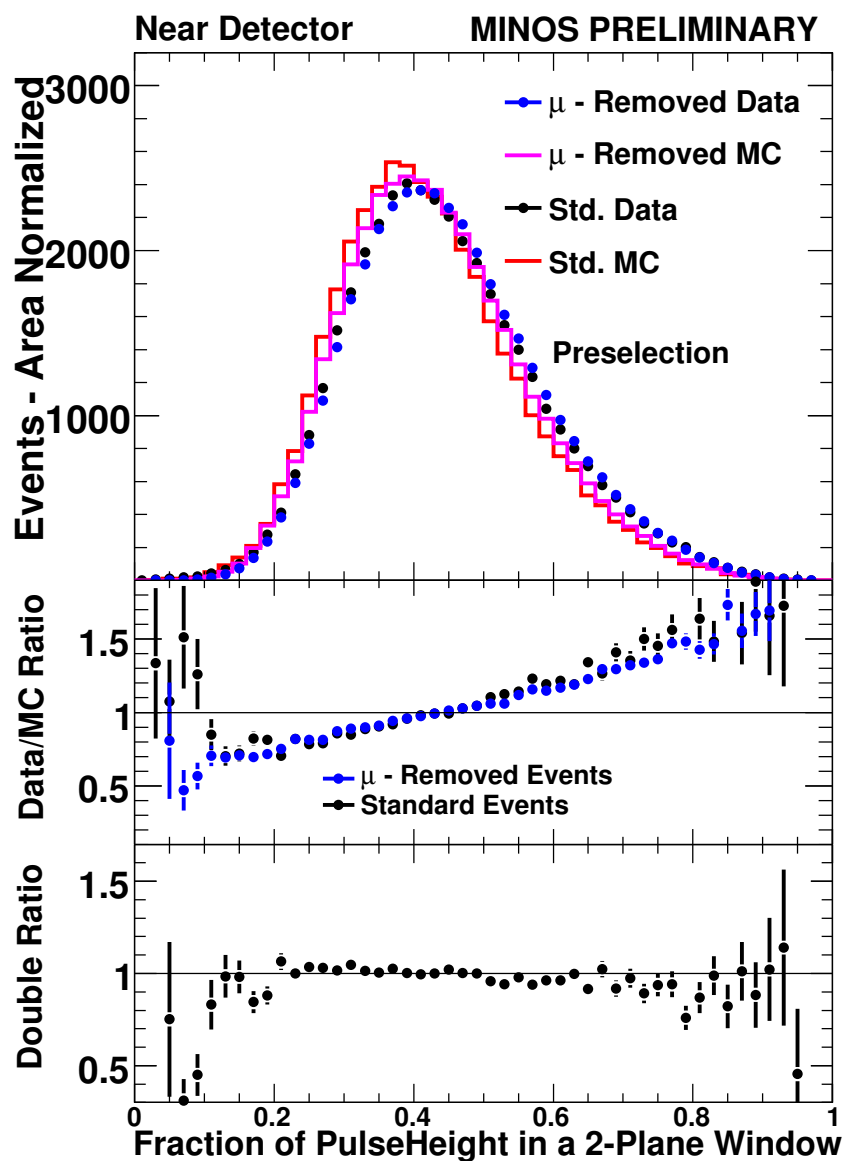


Figure A.17: Near Detector Data/MC MRCC/Standard comparisons after  $\nu_e$  preselection cuts. Distributions are normalised to the  $10^{19}$  POT standard data area to show shape differences. The variable shown here is the fraction of pulseheight in a 2-plane-window of the event.

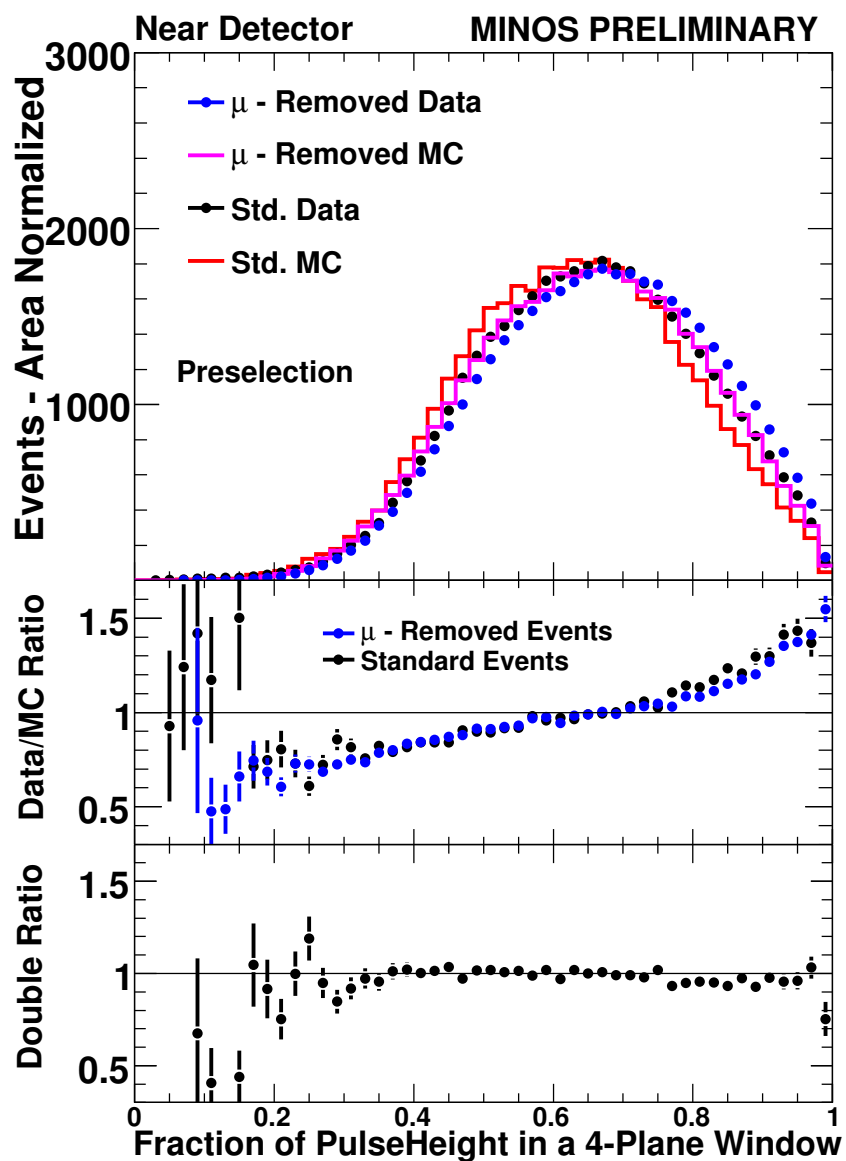


Figure A.18: Near Detector Data/MC MRCC/Standard comparisons after  $\nu_e$  preselection cuts. Distributions are normalised to the  $10^{19}$  POT standard data area to show shape differences. The variable shown here is the fraction of pulseheight in a 4-plane-window of the event.

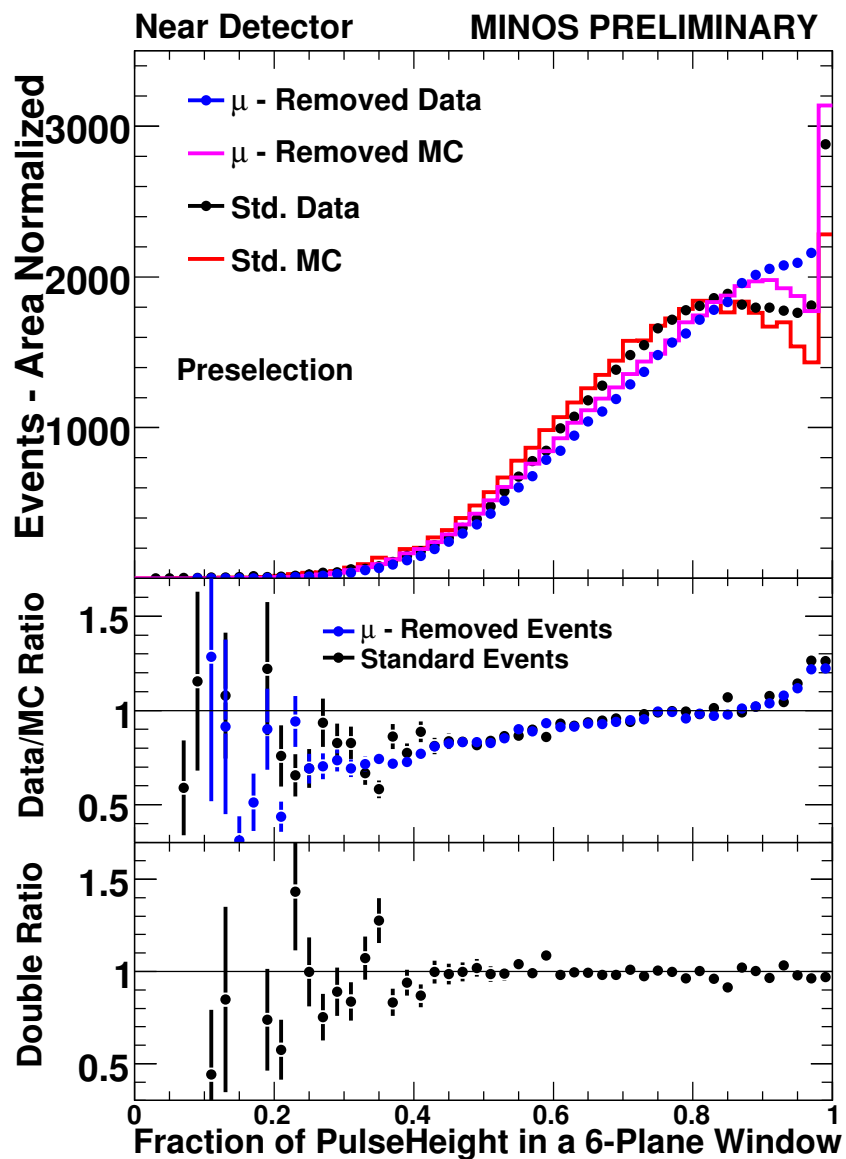


Figure A.19: Near Detector Data/MC MRCC/Standard comparisons after  $\nu_e$  preselection cuts. Distributions are normalised to the  $10^{19}$  POT standard data area to show shape differences. The variable shown here is the fraction of pulseheight in a 6-plane-window of the event.

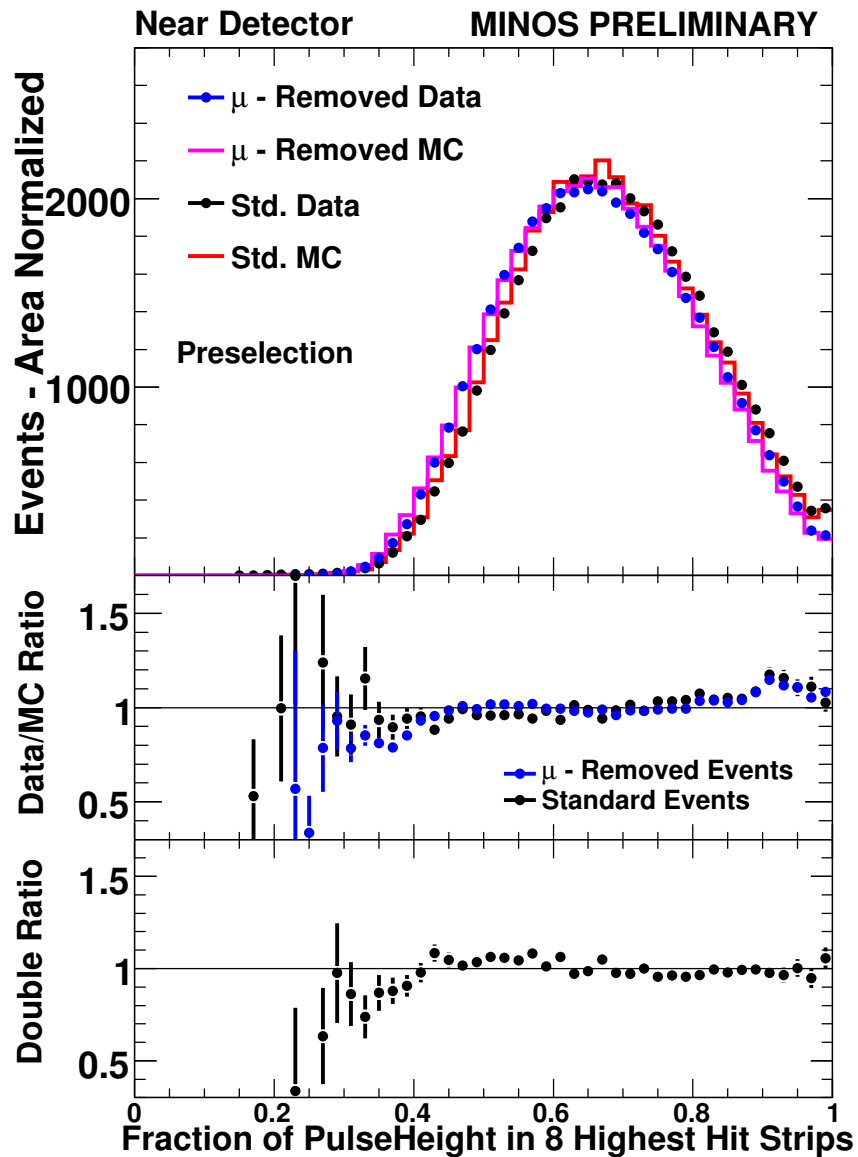


Figure A.20: Near Detector Data/MC MRCC/Standard comparisons after  $\nu_e$  preselection cuts. Distributions are normalised to the  $10^{19}$  POT standard data area to show shape differences. The variable shown here is the fraction of pulseheight in the 8 biggest hit strips in the event.

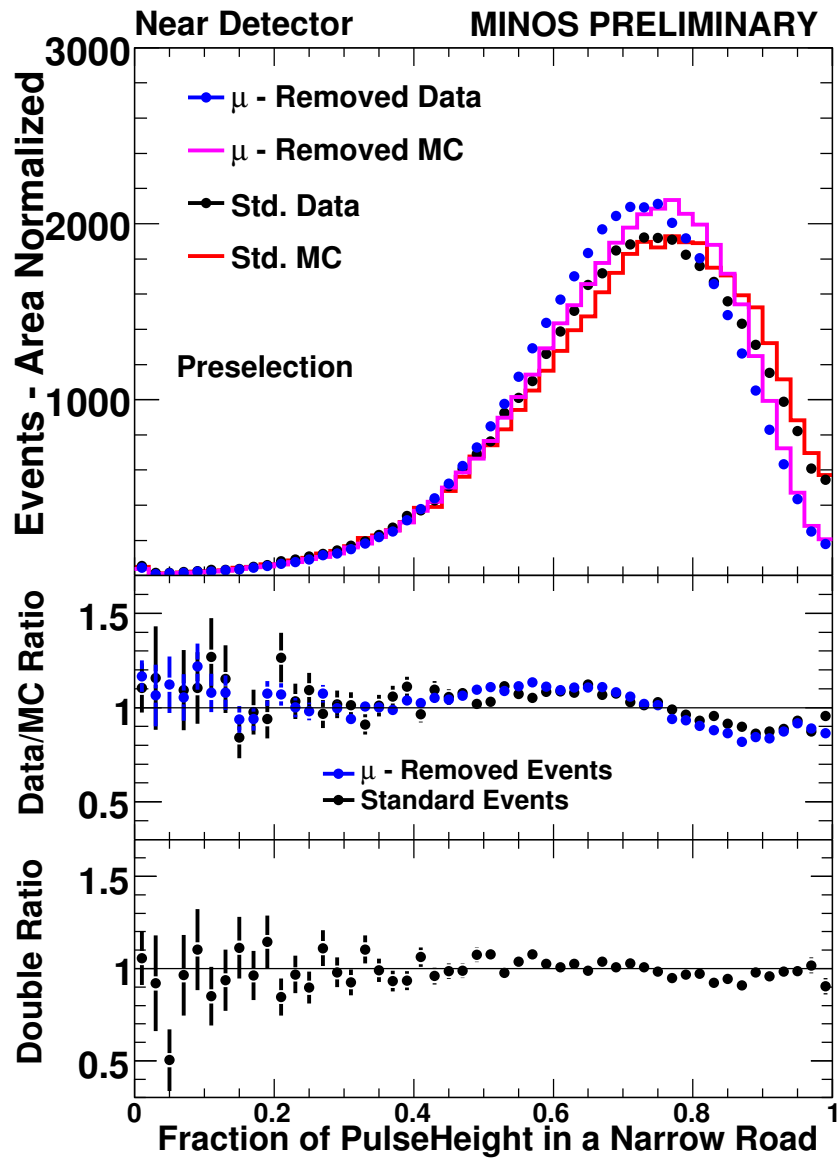


Figure A.21: Near Detector Data/MC MRCC/Standard comparisons after  $\nu_e$  preselection cuts. Distributions are normalised to the  $10^{19}$  POT standard data area to show shape differences. The variable shown here is the fraction of pulseheight in a narrow road along the shower direction.

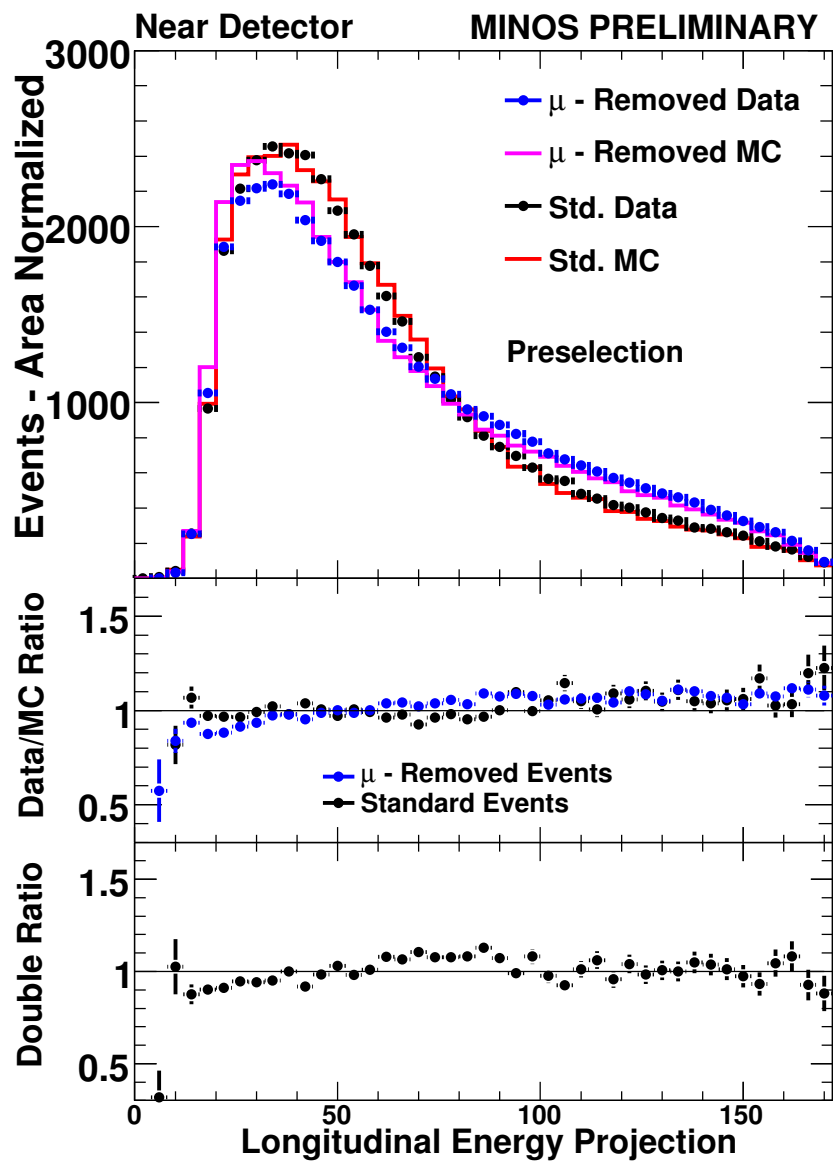


Figure A.22: Near Detector Data/MC MRCC/Standard comparisons after  $\nu_e$  preselection cuts. Distributions are normalised to the  $10^{19}$  POT standard data area to show shape differences. The variable shown here is the energy weighted sum of the distance of each hit to the Z-axis.

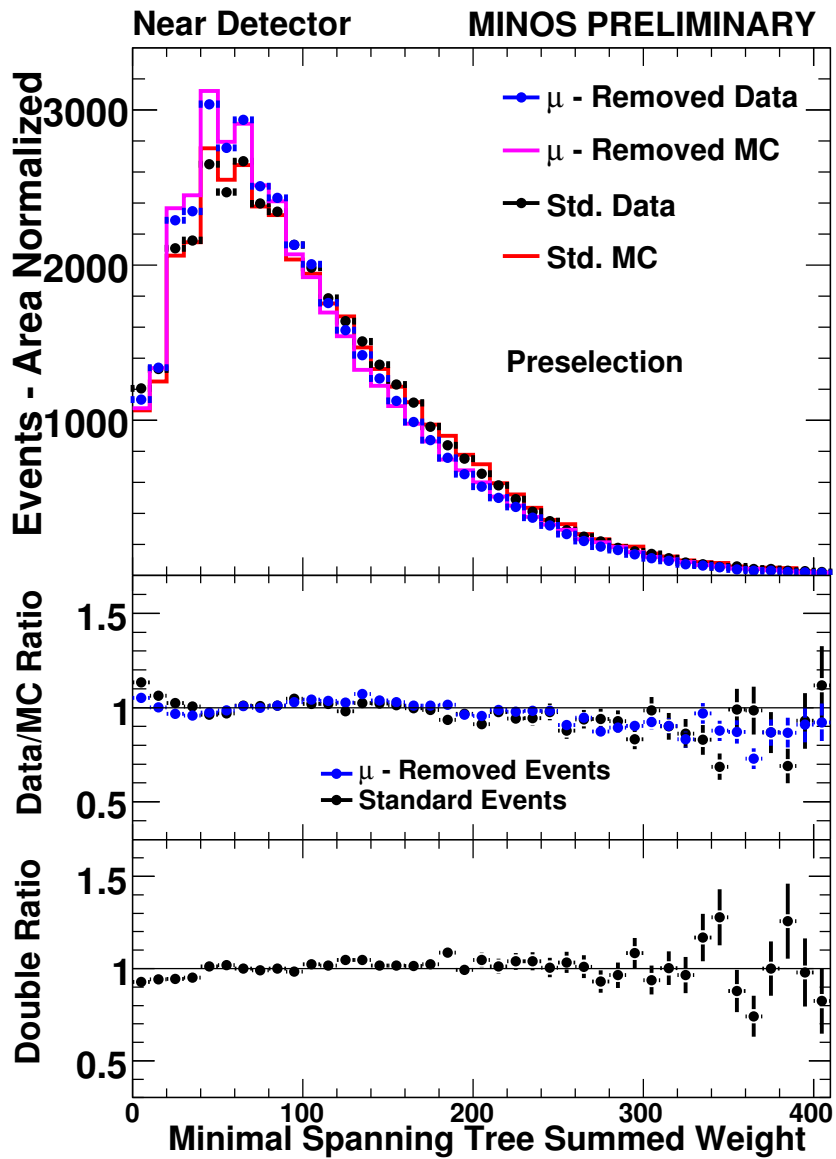


Figure A.23: Near Detector Data/MC MRCC/Standard comparisons after  $\nu_e$  preselection cuts. Distributions are normalised to the  $10^{19}$  POT standard data area to show shape differences. The variable shown here is the sum of the minimum distances between the larger than average hits.

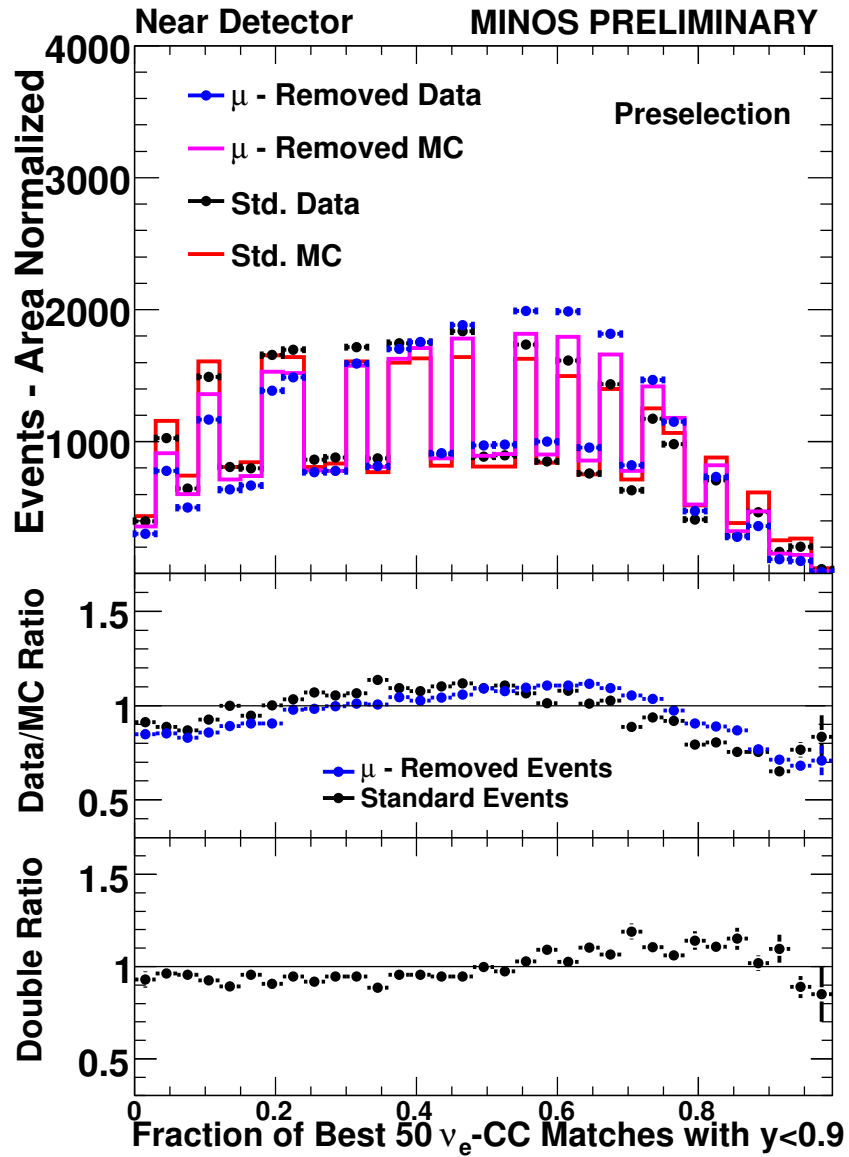


Figure A.24: Near Detector Data/MC MRCC/Standard comparisons after  $\nu_e$  preselection cuts. Distributions are normalised to the  $10^{19}$  POT standard data area to show shape differences. The variable shown here is the fraction of the best 50 LEM CC- $\nu_e$  matches with  $\gamma > 0.9$ .



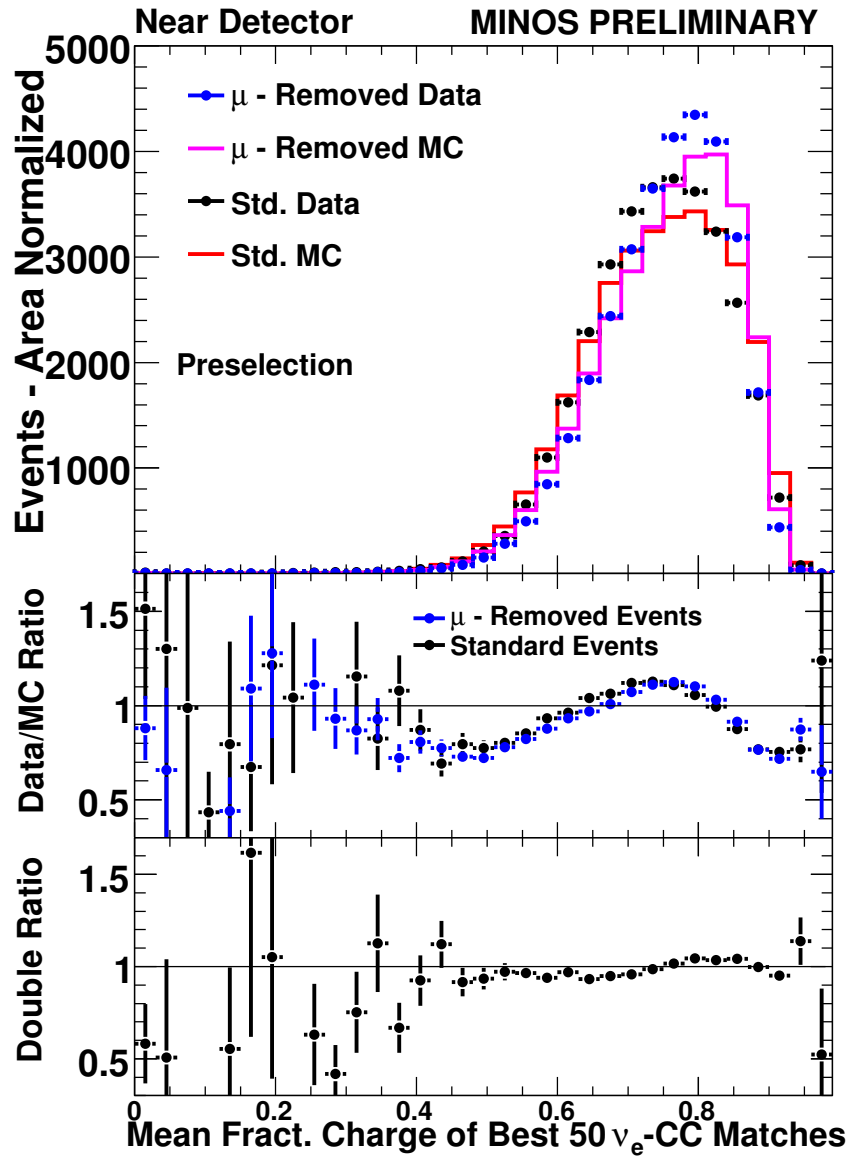


Figure A.25: Near Detector Data/MC MRCC/Standard comparisons after  $\nu_e$  preselection cuts. Distributions are normalised to the  $10^{19}$  POT standard data area to show shape differences. The variable shown here is the mean fractional charge of the best 50 LEM CC- $\nu_e$  matches with  $y > 0.9$ .

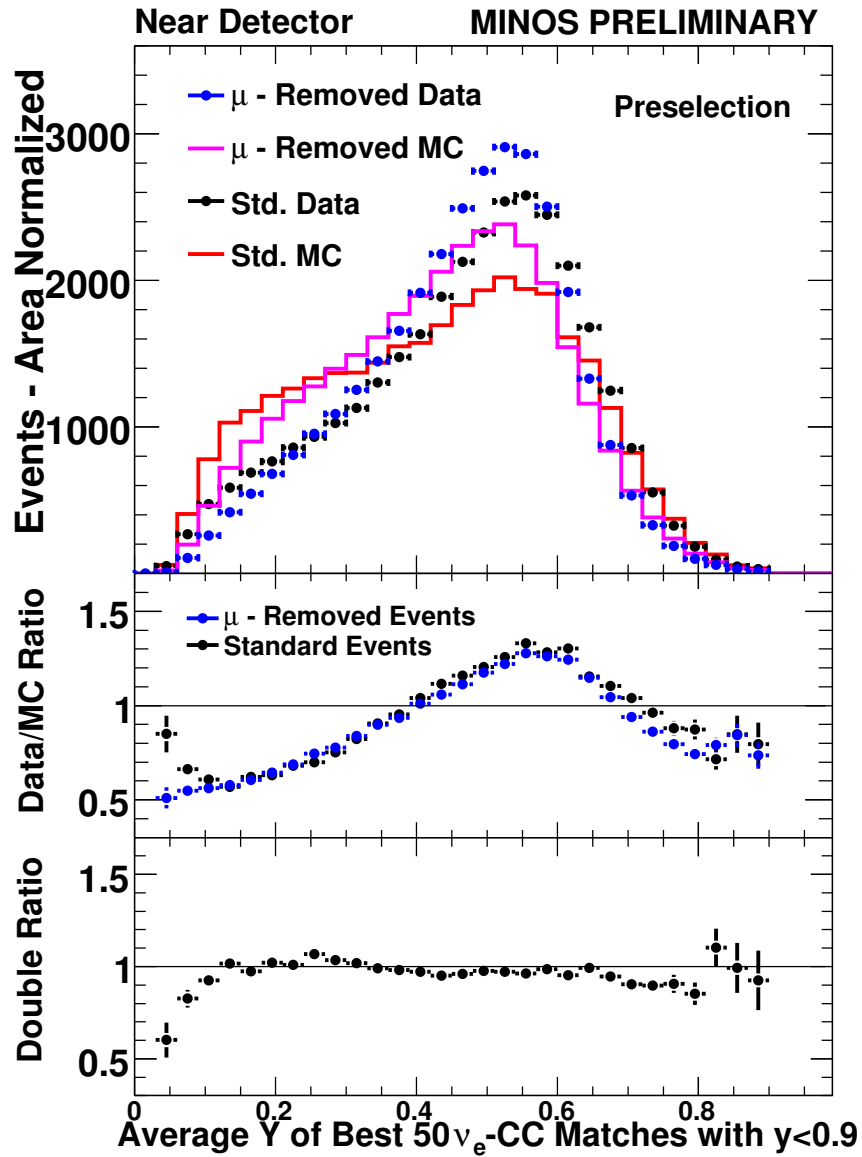


Figure A.26: Near Detector Data/MC MRCC/Standard comparisons after  $\nu_e$  preselection cuts. Distributions are normalised to the  $10^{19}$  POT standard data area to show shape differences. The variable shown here is the average  $y$  of the best 50 LEM CC- $\nu_e$  matches with  $y > 0.9$ .

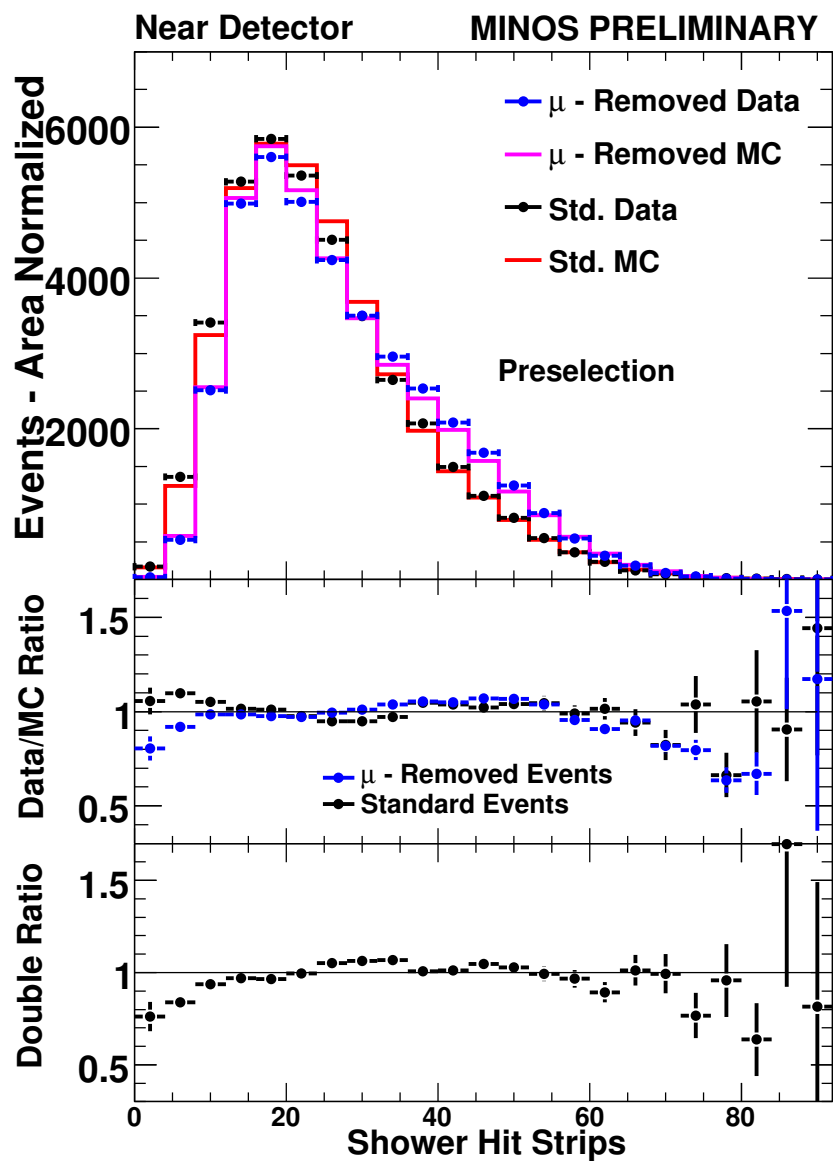


Figure A.27: Near Detector Data/MC MRCC/Standard comparisons after  $\nu_e$  preselection cuts. Distributions are normalised to the  $10^{19}$  POT standard data area to show shape differences. The variable shown here is the total number of hit strips in the primary shower.

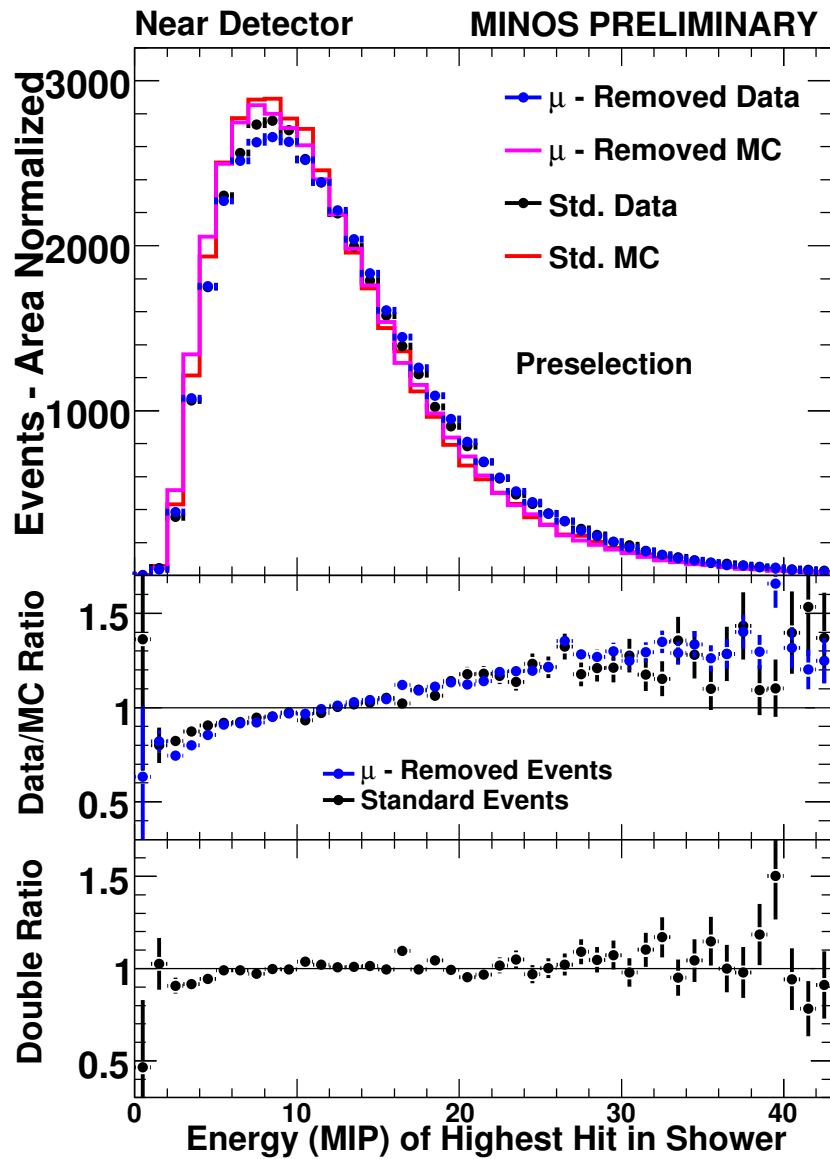


Figure A.28: Near Detector Data/MC MRCC/Standard comparisons after  $\nu_e$  preselection cuts. Distributions are normalised to the  $10^{19}$  POT standard data area to show shape differences. The variable shown here is the energy in MIP of the biggest hit in the primary shower.

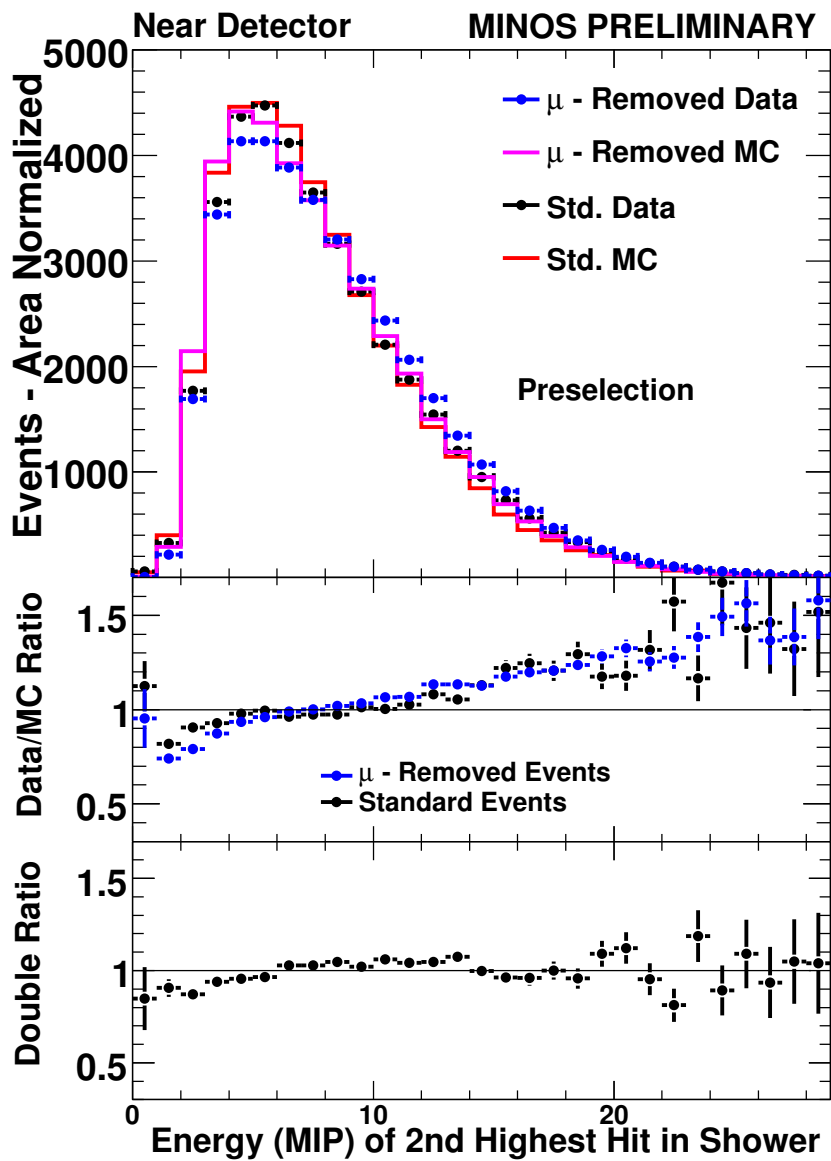


Figure A.29: Near Detector Data/MC MRCC/Standard comparisons after  $\nu_e$  preselection cuts. Distributions are normalised to the  $10^{19}$  POT standard data area to show shape differences. The variable shown here is the energy in MIP of the second biggest hit in the primary shower.

## Appendix B

# Oscillations in the $\nu_e$ Analysis

The oscillation probabilities as used in the  $\nu_e$  analysis and detailed in document [70] are given below. Equations 7.2, 7.3, and 7.5 in section 7.1.1 detail the expressions replaced by  $\alpha$ ,  $\Delta$ , and  $A$ . Furthermore,  $C_{ij}$  is given by  $C_{ij} = \sqrt{\sin^2 2\theta_{ij} + (A - \cos 2\theta_{ij})^2}$ .

$$\begin{aligned}
 P_{\nu_e \rightarrow \nu_e} = & 1 - \frac{\sin^2 2\theta_{13}}{C_{13}^2} \sin^2 C_{13} \Delta + 2\alpha \sin^2 \theta_{12} \frac{\sin^2 2\theta_{13}}{C_{13}^2} \sin C_{13} \Delta \\
 & \times \left[ \Delta \frac{\cos C_{13} \Delta}{C_{13}} (1 - A \cos 2\theta_{13}) - A \frac{\sin C_{13} \Delta \cos 2\theta_{13} - A}{C_{13}} \right] \\
 & + 1 - \frac{\sin^2 2\theta_{12}}{C_{12}^2} \sin^2 \alpha C_{12} \Delta - 1
 \end{aligned} \tag{B.1}$$

$$\begin{aligned}
 P_{\nu_\mu \rightarrow \nu_e} = & \sin^2 \theta_{23} \frac{\sin^2 2\theta_{13}}{C_{13}^2} \sin^2 C_{13} \Delta - 2\alpha \sin^2 \theta_{12} \sin^2 \theta_{23} \frac{\sin^2 2\theta_{13}}{C_{13}^2} \sin C_{13} \Delta \\
 & \times \left[ \Delta \frac{\cos C_{13} \Delta}{C_{13}} (1 - A \cos 2\theta_{13}) - A \frac{\sin C_{13} \Delta \cos 2\theta_{13} - A}{C_{13}} \right] \\
 & + \alpha \sin \theta_{13} \sin 2\theta_{12} \sin 2\theta_{23} \frac{\sin C_{13} \Delta}{AC_{13}^2} \left\{ \cos \delta [C_{13} \sin(1 + A)\Delta \right. \\
 & \left. - (1 - A \cos 2\theta_{13})_{13} \Delta] - C_{13} \sin \delta [{}_{13} \Delta - \cos(1 + A)\Delta] \right\} \\
 & + \cos^2 \theta_{23} \frac{\sin^2 2\theta_{12}}{C_{12}^2} \sin^2 \alpha C_{12} \Delta \\
 & - \sin \theta_{13} \frac{\sin 2\theta_{12}}{C_{12}} \sin 2\theta_{23} \frac{(1 - \alpha) \sin \alpha C_{12} \Delta}{1 + A - \alpha + A\alpha \cos^2 \theta_{12}} \left\{ \sin \delta [\cos \alpha C_{12} \Delta - \cos(A + \alpha - 2)\Delta] \right. \\
 & \left. + \cos \delta \left[ \sin(A + \alpha - 2)\Delta - \sin \alpha C_{12} \Delta \left( \frac{\cos 2\theta_{12} - (A/\alpha)}{C_{12}} - \frac{\alpha AC_{12} \sin^2 2\theta_{12}}{2(1 - \alpha) C_{12}^2} \right) \right] \right\} \\
 & - 2\alpha \sin \theta_{13} \sin 2\theta_{12} \sin 2\theta_{23} \cos(\Delta + \delta) \frac{\sin A \Delta \sin(A - 1)\Delta}{A(A - 1)}
 \end{aligned} \tag{B.2}$$

$$\begin{aligned}
P_{\nu_\mu \rightarrow \nu_\tau} = & \frac{1}{2} \sin^2 2\theta_{23} \left[ \left( 1 - \frac{\cos 2\theta_{13} - A}{C_{13}} \right) \sin^2 \frac{1}{2} (1 + A - C_{13}) \Delta \right. \\
& + \left. \left( 1 + \frac{\cos 2\theta_{13} - A}{C_{13}} \right) \sin^2 \frac{1}{2} (1 + A - C_{13}) \Delta - \frac{1}{2} \frac{\sin^2 2\theta_{13}}{C_{13}^2} \sin^2 C_{13} \Delta \right] \\
& - \frac{\alpha}{2} \sin^2 2\theta_{23} \Delta \left\{ 2 \left[ \cos^2 \theta_{12} - \sin^2 \theta_{12} \sin^2 \theta_{13} \frac{(1 + 2 \sin^2 \theta_{13} A + A^2)}{C_{13}^2} \right] \right. \\
& \times \cos C_{13} \Delta \sin(1 + A) \Delta + 2 [\cos^2 \theta_{12} \cos^2 \theta_{13} - \cos^2 \theta_{12} \sin^2 \theta_{13} \\
& + \sin^2 \theta_{12} \sin^2 \theta_{13} + (\sin^2 \theta_{12} \sin^2 \theta_{13} - \cos^2 \theta_{12}) A] \frac{\sin C_{13} \Delta}{C_{13}} \cos(1 + A) \Delta \\
& + \sin^2 \theta_{12} \frac{\sin^2 2\theta_{13}}{C_{13}^2} \frac{\sin C_{13} \Delta}{C_{13}} \\
& \times \left. \left[ \frac{A}{\Delta} \sin(1 + A) \Delta + \frac{A \cos 2\theta_{13} - A}{\Delta} \frac{\sin C_{13} \Delta}{C_{13}} - (1 - A \cos 2\theta_{13}) \cos C_{13} \Delta \right] \right\} \\
& + \frac{\alpha \sin \theta_{13} \sin 2\theta_{12} \sin 2\theta_{23}}{2A \cos^2 \theta_{13}} \left\{ 2 \cos^2 \theta_{13} \sin \delta \frac{\sin C_{13} \Delta}{C_{13}} [\cos C_{13} \Delta - \cos(1 + A) \Delta] \right. \\
& - \cos 2\theta_{23} \cos \delta (1 + A) [\cos C_{13} \Delta - \cos(1 + A) \Delta]^2 \\
& + \cos 2\theta_{23} \cos \delta \left[ \sin(1 + A) \Delta + \frac{\cos 2\theta_{13} - A}{C_{13}} \sin C_{13} \Delta \right] \\
& \times \left. \left[ (1 + 2 \sin^2 \theta_{13} A + A^2) \frac{13\Delta}{C_{13}} - (1 + A) \sin(1 + A) \Delta \right] \right\} \\
& + \frac{1}{2} \sin^2 2\theta_{23} \left[ 1 - \frac{\sin^2 2\theta_{12}}{2C_{12}^2} \sin^2 \alpha C_{12} \Delta - \cos(\alpha C_{12} + A + \alpha - 2) \Delta \right. \\
& - \left. \left( 1 - \frac{\cos 2\theta_{12} - (A/\alpha)}{C_{12}} \right) \sin \alpha C_{12} \Delta \sin(A + \alpha - 2) \Delta \right] \\
& + \frac{\sin \theta_{13} \sin 2\theta_{12}}{C_{12}} \frac{\sin 2\theta_{23}}{1 - A - \alpha + A\alpha \cos^2 \theta_{12}} \\
& \times \left\{ \frac{\alpha A C_{12}}{2} \cos 2\theta_{23} \cos \delta \left[ (\cos \alpha C_{12} \Delta - \cos(A + \alpha - 2) \Delta) \right]^2 \right. \\
& + \left( \frac{\cos 2\theta_{12} - (A/\alpha)}{C_{12}} \sin \alpha C_{12} \Delta + \sin(A + \alpha - 2) \Delta \right) \\
& \times \left. \left[ \left( \frac{\cos 2\theta_{12} - (A/\alpha)}{C_{12}} + \frac{2(1 - \alpha)}{\alpha A C_{12}} \right) \sin \alpha C_{12} \Delta + \sin(A + \alpha - 2) \Delta \right] \right\} \\
& + \sin \delta (1 - \alpha) [\cos \alpha C_{12} \Delta - \cos(A + \alpha - 2) \Delta] \sin \alpha C_{12} \Delta \left. \right\} \\
& - \sin^2 2\theta_{23} \sin^2 \Delta + \alpha \cos^2 \theta_{12} \sin^2 2\theta_{23} \sin 2\Delta \\
& - 2\alpha \sin \theta_{13} \sin 2\theta_{12} \sin 2\theta_{23} \left[ \sin \delta \sin \Delta \frac{\sin A \Delta}{A} \frac{\sin(A - 1) \Delta}{(A - 1)} \right. \\
& \left. - \frac{1}{A - 1} \cos 2\theta_{23} \cos \delta \sin \Delta \left[ A \sin \Delta - \frac{\sin A \Delta}{A} \cos(A - 1) \Delta \right] \right]
\end{aligned}$$

(B.3)

# Appendix C

## Additional MRCC FD

### Distributions

This Appendix contains FD MRCC sideband distributions, as referred to in section 7.3.2. Each plot shows the FD MRCC MC, the sideband prediction, and the FD MRCC data.

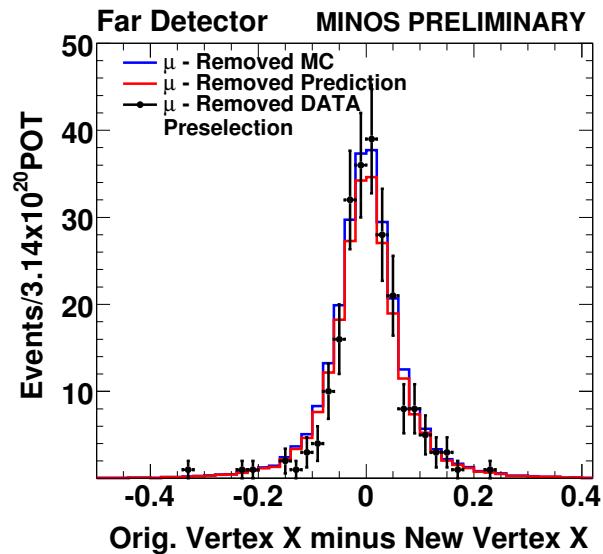


Figure C.1: This figure shows data and prediction distributions of the MRCC minus the original events reconstructed vertex X for the FD MRCC  $\nu_e$  preselected events. The black points are the data, and the red line is the prediction obtained by scaling the default MRCC MC (blue) by the ND MRCC Data/MC ratios. The histograms are normalised to the data exposure of  $3.14 \times 10^{20}$  POT.



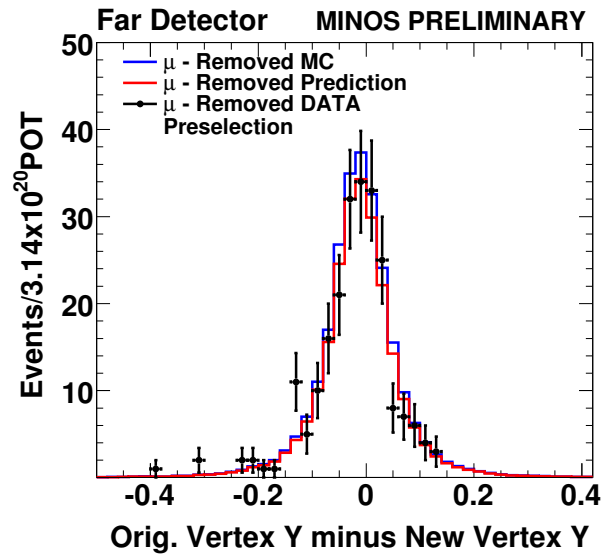


Figure C.2: This figure shows data and prediction distributions of the MRCC minus the original events reconstructed vertex Y for the FD MRCC  $\nu_e$  preselected events. The black points are the data, and the red line is the prediction obtained by scaling the default MRCC MC (blue) by the ND MRCC Data/MC ratios. The histograms are normalised to the data exposure of  $3.14 \times 10^{20}$  POT.

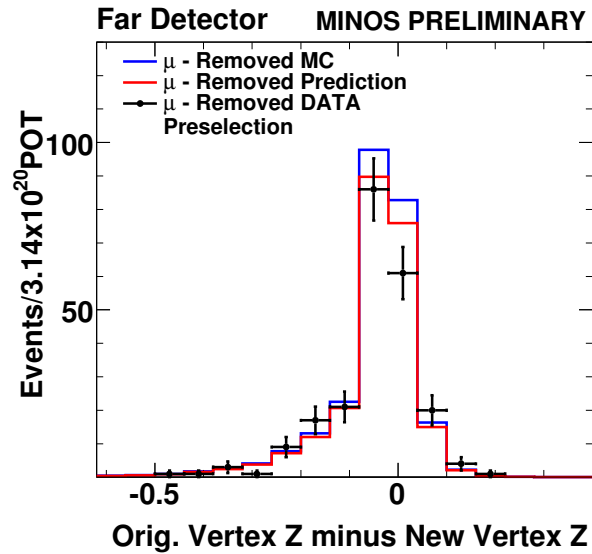


Figure C.3: This figure shows data and prediction distributions of the MRCC minus the original events reconstructed vertex Z for the FD MRCC  $\nu_e$  preselected events. The black points are the data, and the red line is the prediction obtained by scaling the default MRCC MC (blue) by the ND MRCC Data/MC ratios. The histograms are normalised to the data exposure of  $3.14 \times 10^{20}$  POT.

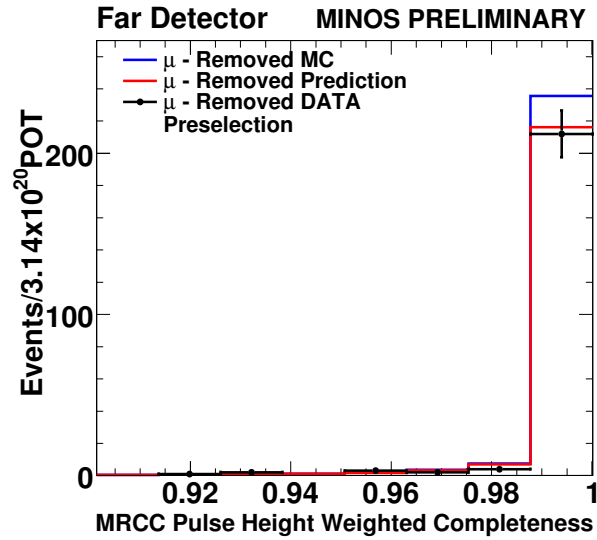


Figure C.4: This figure shows data and prediction distributions of the MRCC pulse height weighted completeness for the FD MRCC  $\nu_e$  preselected events. The black points are the data, and the red line is the prediction obtained by scaling the default MRCC MC (blue) by the ND MRCC Data/MC ratios. The histograms are normalised to the data exposure of  $3.14 \times 10^{20}$  POT.

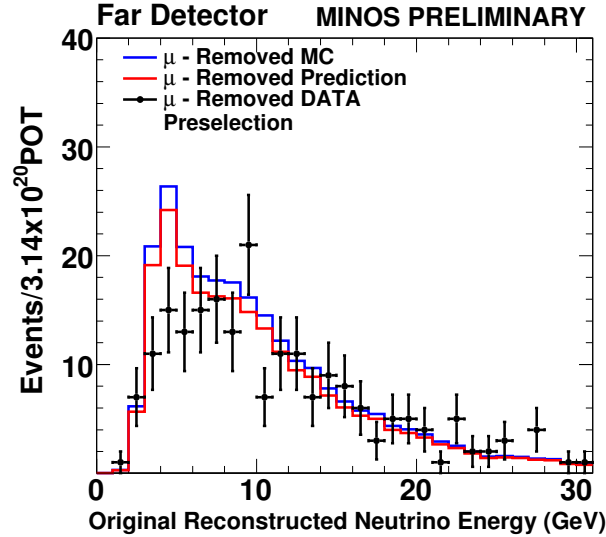


Figure C.5: This figure shows data and prediction distributions of the original reconstructed neutrino energy for the FD MRCC  $\nu_e$  preselected events. The black points are the data, and the red line is the prediction obtained by scaling the default MRCC MC (blue) by the ND MRCC Data/MC ratios. The histograms are normalised to the data exposure of  $3.14 \times 10^{20}$  POT.

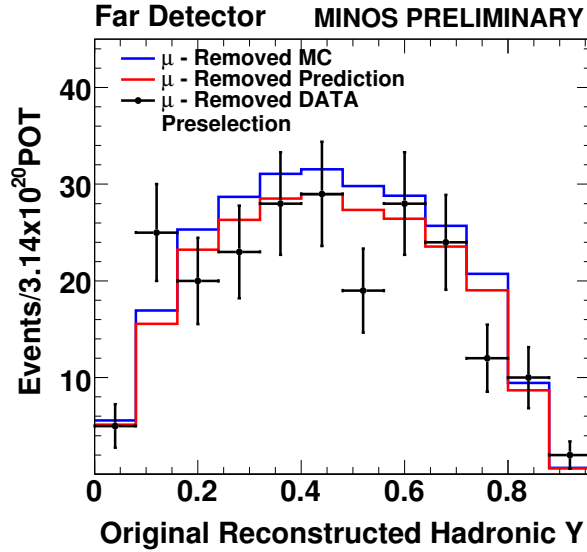


Figure C.6: This figure shows data and prediction distributions of the original reconstructed hadronic  $y$ , the fraction of the neutrino's energy lost in the nucleon rest frame, for the FD MRCC  $\nu_e$  preselected events. The black points are the data, and the red line is the prediction obtained by scaling the default MRCC MC (blue) by the ND MRCC Data/MC ratios. The histograms are normalised to the data exposure of  $3.14 \times 10^{20}$  POT.

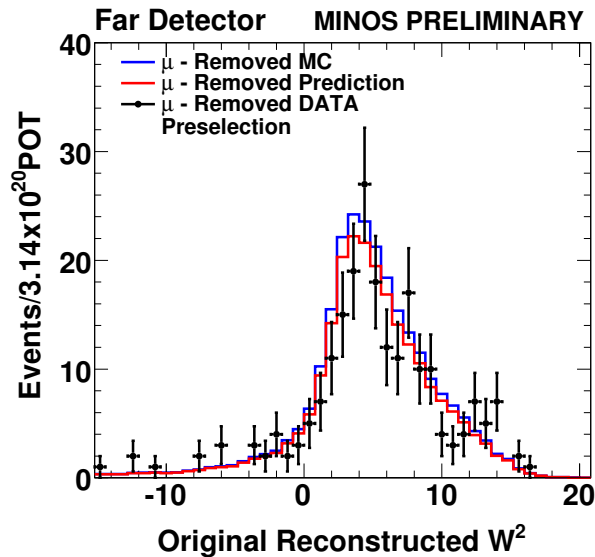


Figure C.7: This figure shows data and prediction distributions of the original reconstructed  $W^2$ , the mass squared of the system recoiling against the scattered neutrino, for the FD MRCC  $\nu_e$  preselected events. The black points are the data, and the red line is the prediction obtained by scaling the default MRCC MC (blue) by the ND MRCC Data/MC ratios. The histograms are normalised to the data exposure of  $3.14 \times 10^{20}$  POT.

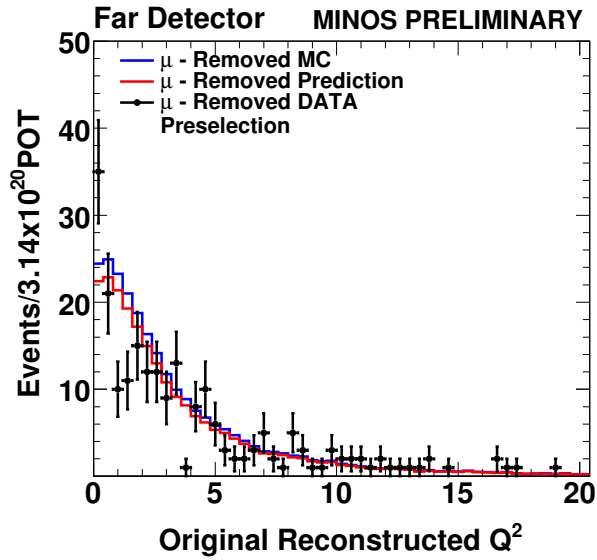


Figure C.8: This figure shows data and prediction distributions of the original reconstructed  $Q^2$ , the four-momentum squared transferred by the neutrino to the struck nucleon, for the FD MRCC  $\nu_e$  preselected events. The black points are the data, and the red line is the prediction obtained by scaling the default MRCC MC (blue) by the ND MRCC Data/MC ratios. The histograms are normalised to the data exposure of  $3.14 \times 10^{20}$  POT.

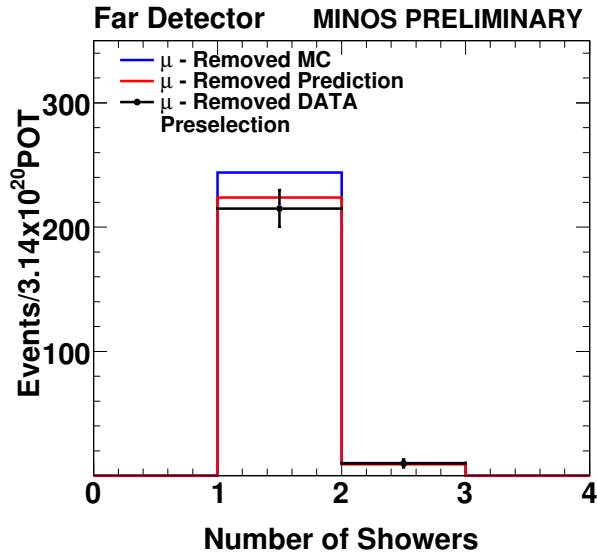


Figure C.9: This figure shows data and prediction distributions of the number of event showers for the FD MRCC  $\nu_e$  preselected events. The black points are the data, and the red line is the prediction obtained by scaling the default MRCC MC (blue) by the ND MRCC Data/MC ratios. The histograms are normalised to the data exposure of  $3.14 \times 10^{20}$  POT.

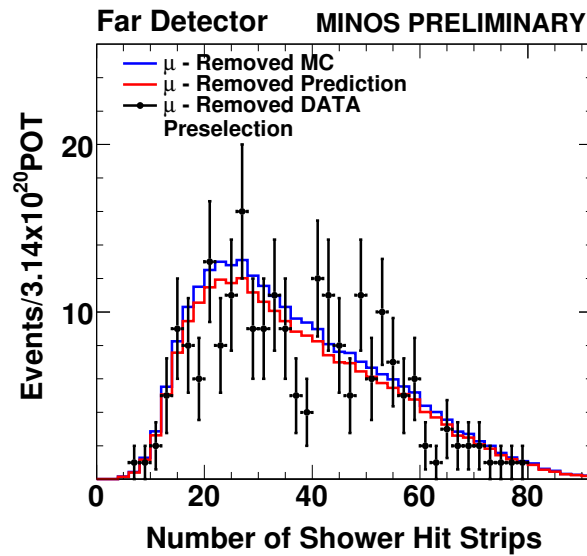


Figure C.10: This figure shows data and prediction distributions of the number of shower hit strips for the FD MRCC  $\nu_e$  preselected events. The black points are the data, and the red line is the prediction obtained by scaling the default MRCC MC (blue) by the ND MRCC Data/MC ratios. The histograms are normalised to the data exposure of  $3.14 \times 10^{20}$  POT.

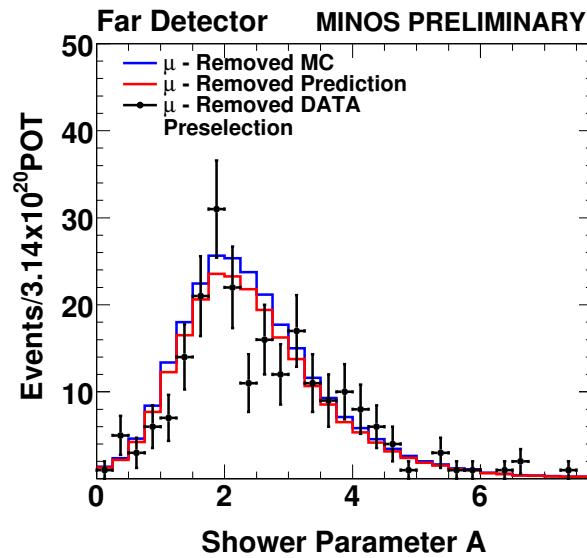


Figure C.11: This figure shows data and prediction distributions of the shower rise parameter  $a$  for the FD MRCC  $\nu_e$  preselected events. The black points are the data, and the red line is the prediction obtained by scaling the default MRCC MC (blue) by the ND MRCC Data/MC ratios. The histograms are normalised to the data exposure of  $3.14 \times 10^{20}$  POT.

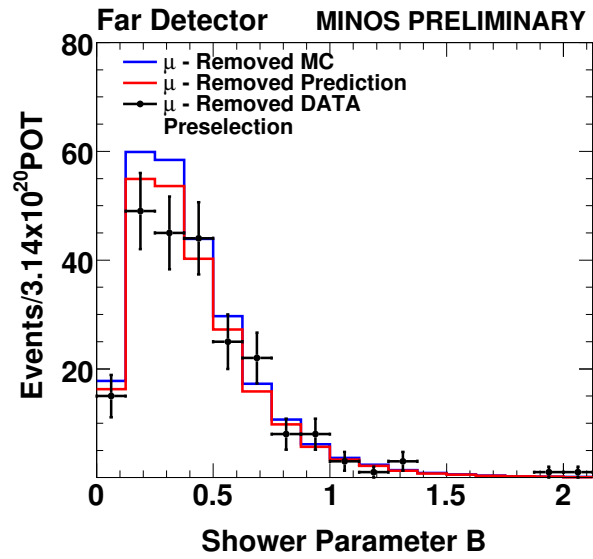


Figure C.12: This figure shows data and prediction distributions of the shower fall parameter  $a$  for the FD MRCC  $\nu_e$  preselected events. The black points are the data, and the red line is the prediction obtained by scaling the default MRCC MC (blue) by the ND MRCC Data/MC ratios. The histograms are normalised to the data exposure of  $3.14 \times 10^{20}$  POT.

## Appendix D

# FD Distributions and Golden Event Displays

This Appendix contains FD distributions as discussed in section 8.2.2. It also contains three additional golden FD data  $\nu_e$ -CC events (see section 8.2.3), where golden means that the events have high PID values and thus a high probability of being true  $\nu_e$ -CC events.

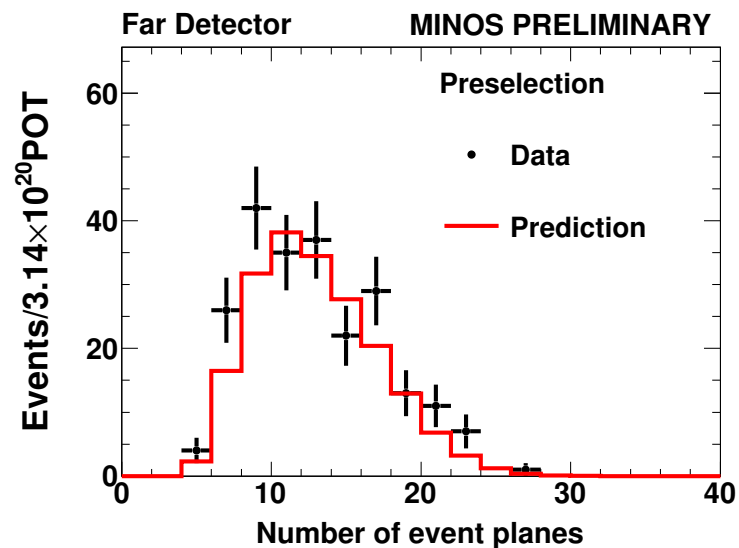


Figure D.1: FD data number of event hit planes after preselection. The prediction is shown in red, and the data in black.

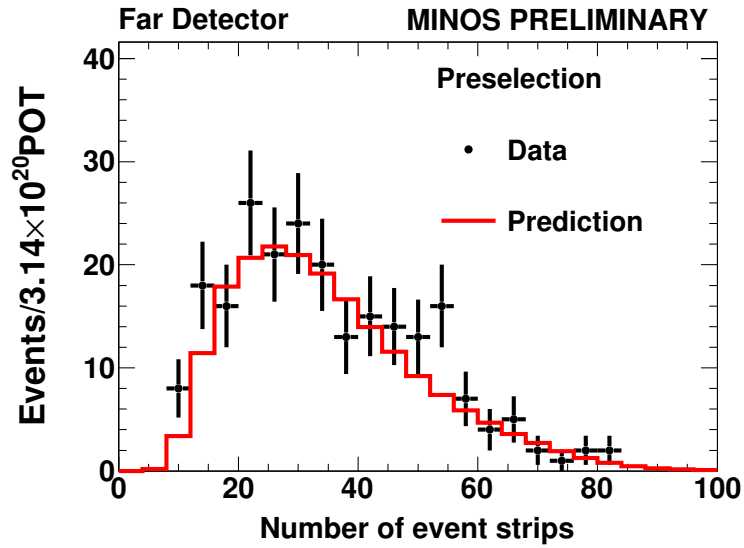


Figure D.2: FD data number of event hit strips after preselection. The prediction is shown in red, and the data in black.

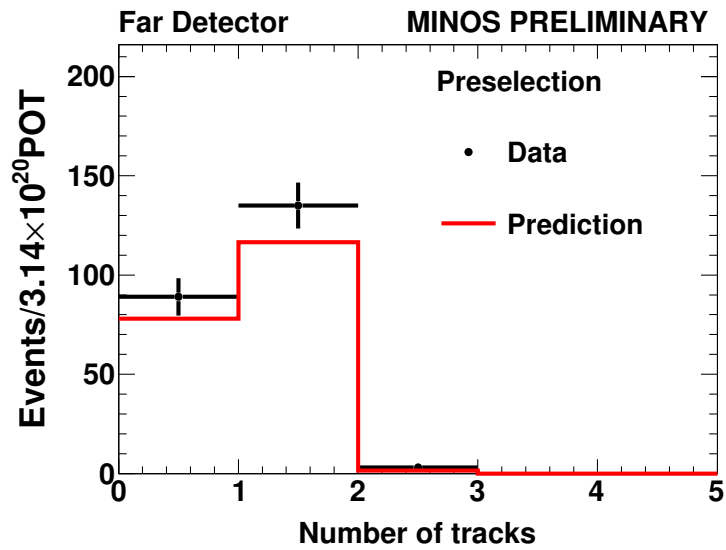


Figure D.3: FD data number of event tracks after preselection. The prediction is shown in red, and the data in black.



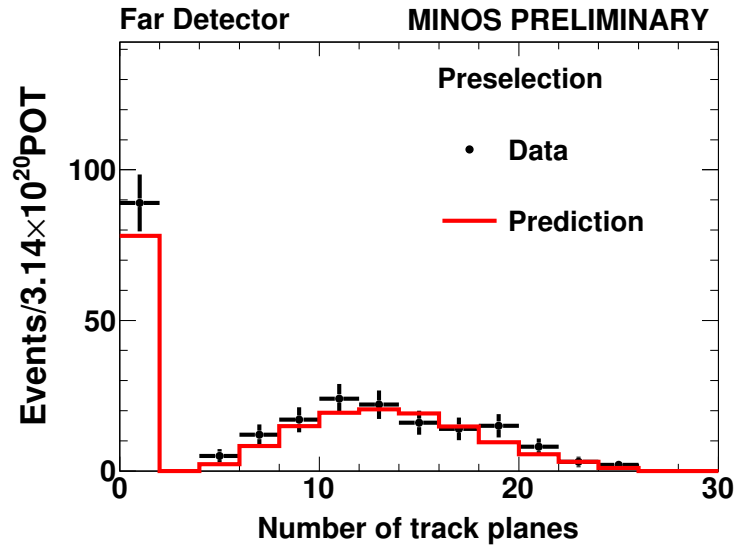


Figure D.4: FD data number of track hit planes after preselection. The prediction is shown in red, and the data in black.

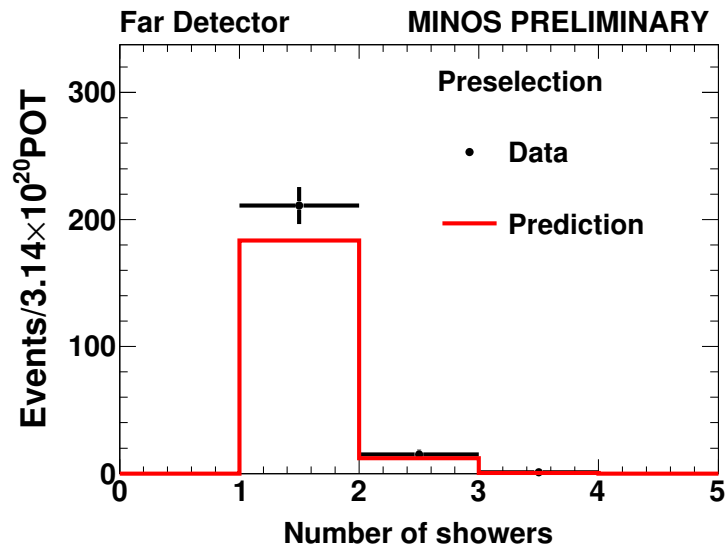


Figure D.5: FD data number of event showers after preselection. The prediction is shown in red, and the data in black.

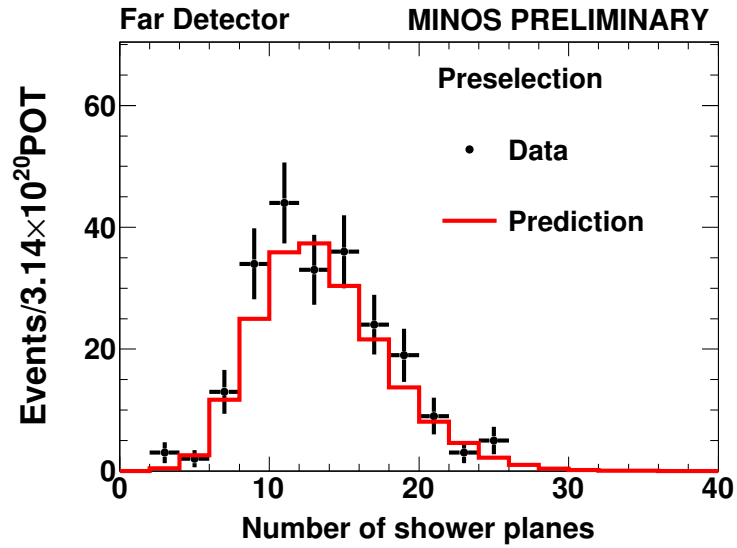


Figure D.6: FD data number of shower hit planes after preselection. The prediction is shown in red, and the data in black.

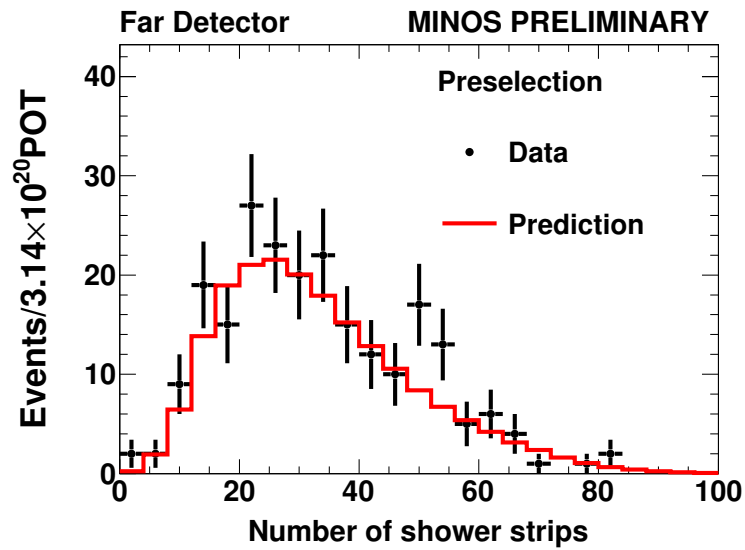


Figure D.7: FD data number of shower hit strips after preselection. The prediction is shown in red, and the data in black.

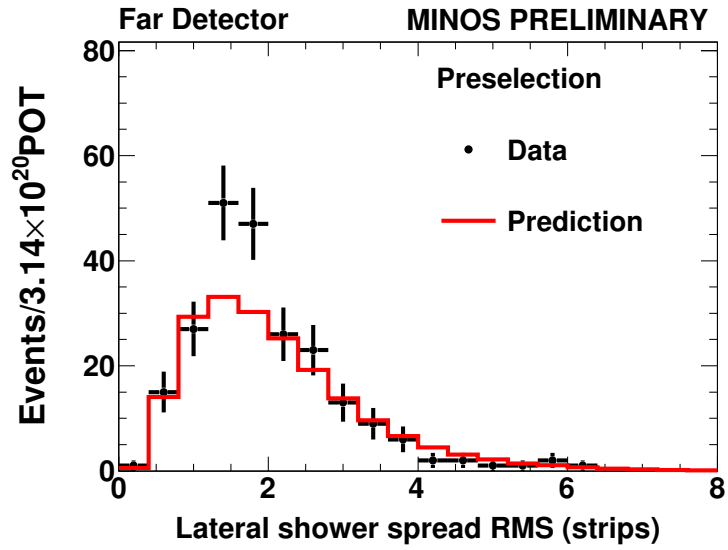


Figure D.8: FD data shower lateral spread after preselection. The prediction is shown in red, and the data in black.

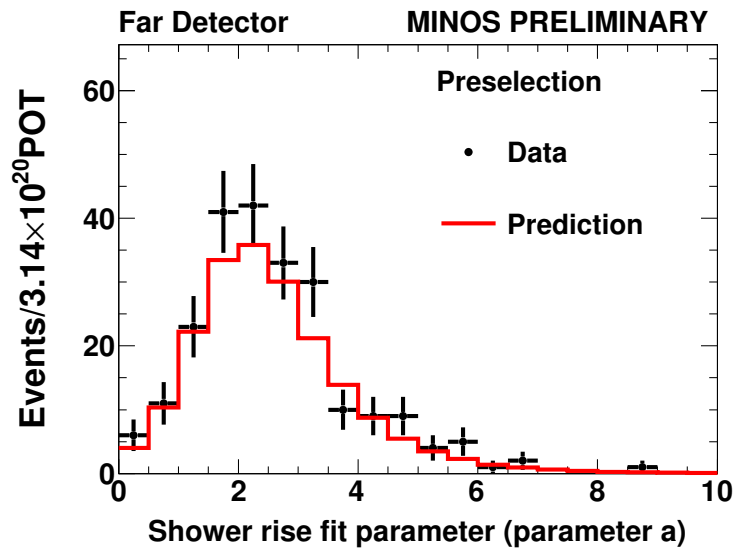


Figure D.9: FD data shower fit parameter a after preselection. The prediction is shown in red, and the data in black.

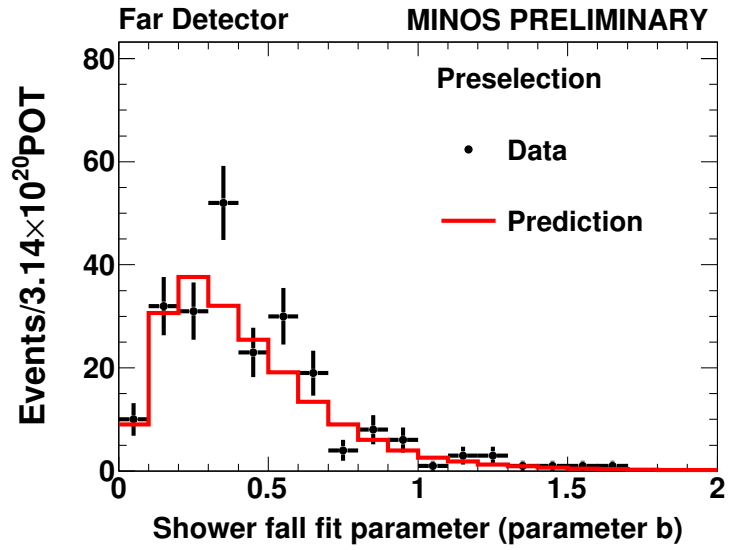


Figure D.10: FD data shower fit parameter b after preselection. The prediction is shown in red, and the data in black.

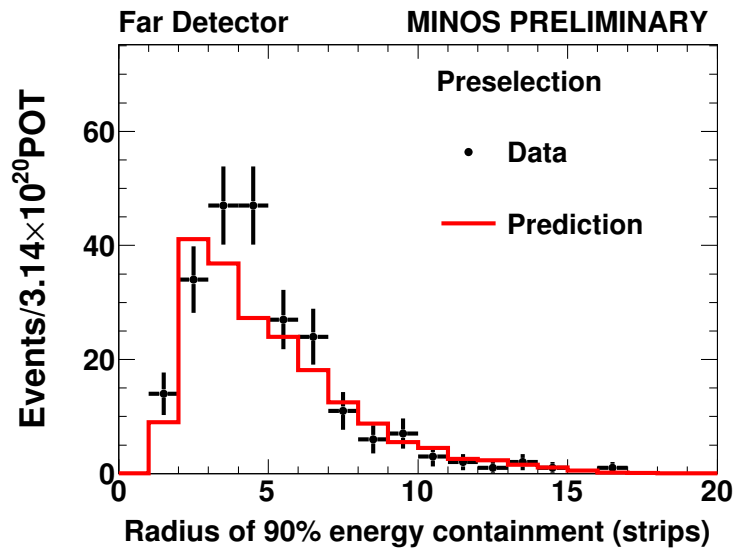


Figure D.11: FD data shower moliere radius - radius of 90% energy containment - after preselection. The prediction is shown in red, and the data in black.

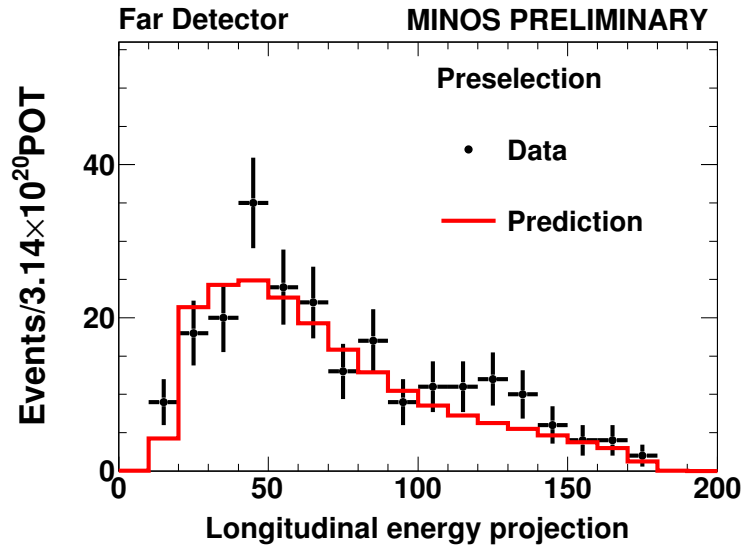


Figure D.12: FD data energy weighted sum of the distance of each hit to the Z-axis after preselection. The prediction is shown in red, and the data in black.

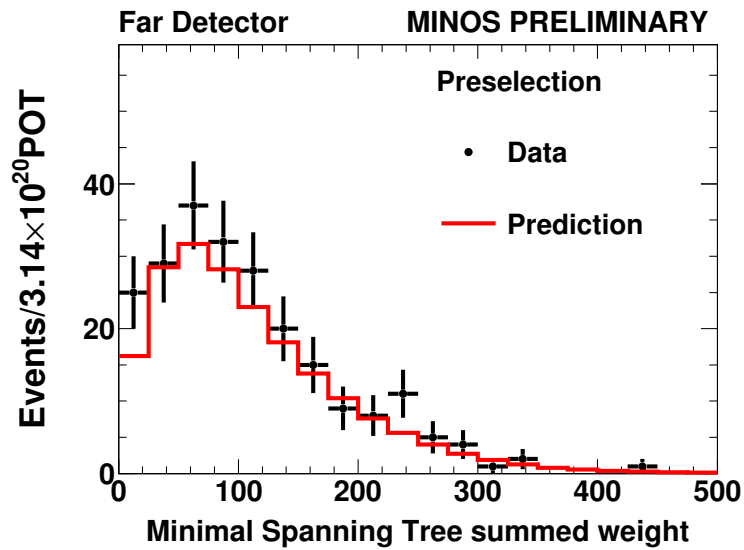


Figure D.13: FD data sum of the minimum distances between the larger than average hits after preselection. The prediction is shown in red, and the data in black.

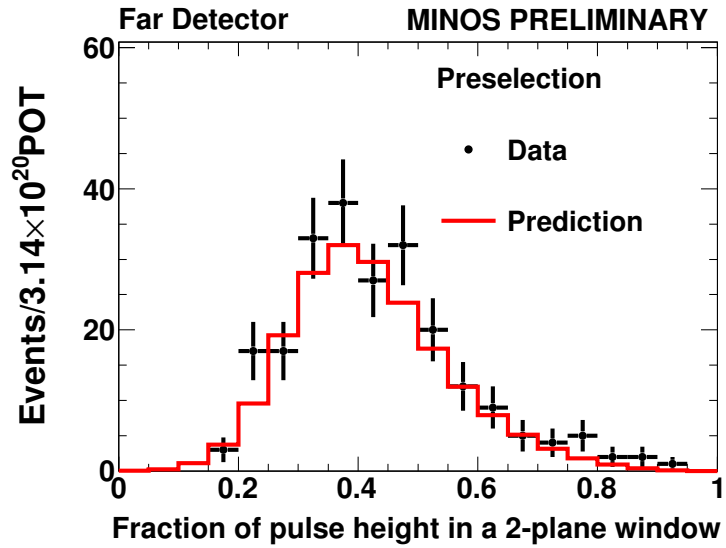


Figure D.14: FD data fraction of pulseheight in a 2-plane-window of the event after preselection. The prediction is shown in red, and the data in black.

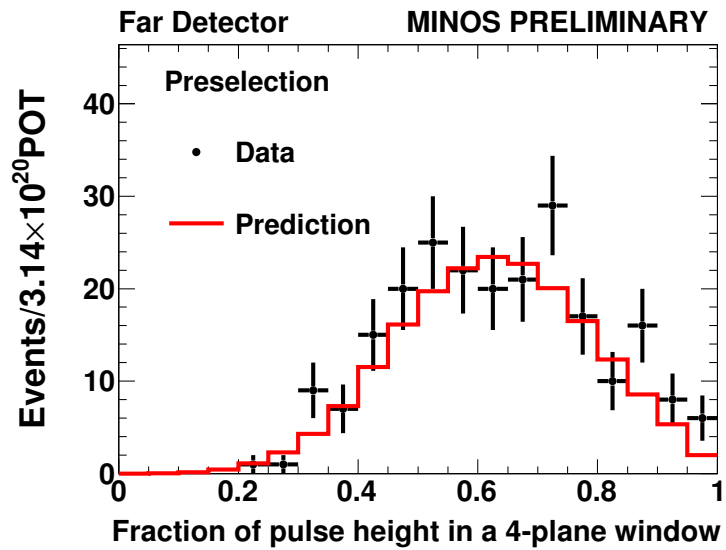


Figure D.15: FD data fraction of pulseheight in a 4-plane-window of the event after preselection. The prediction is shown in red, and the data in black.

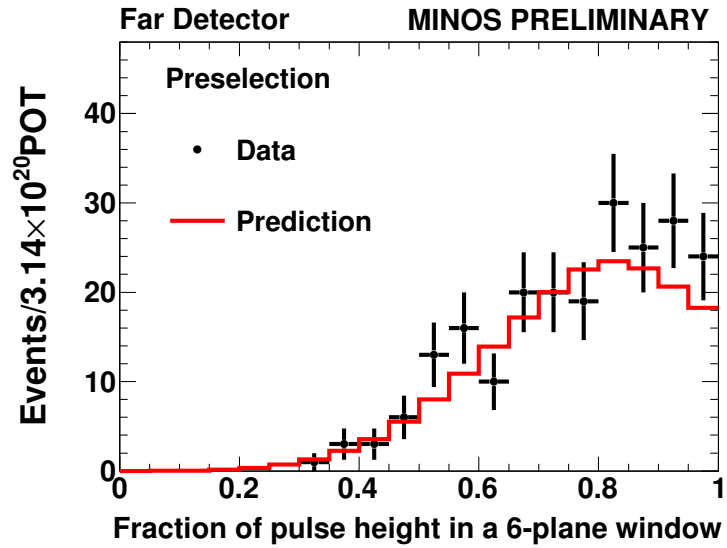


Figure D.16: FD data fraction of pulseheight in a 6-plane-window of the event after preselection. The prediction is shown in red, and the data in black.

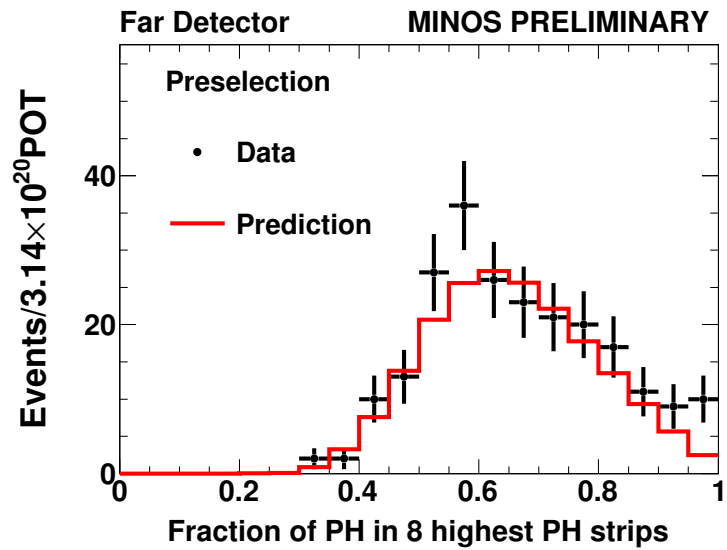


Figure D.17: FD data fraction of pulseheight in the 8 biggest hit strips of the event after preselection. The prediction is shown in red, and the data in black.

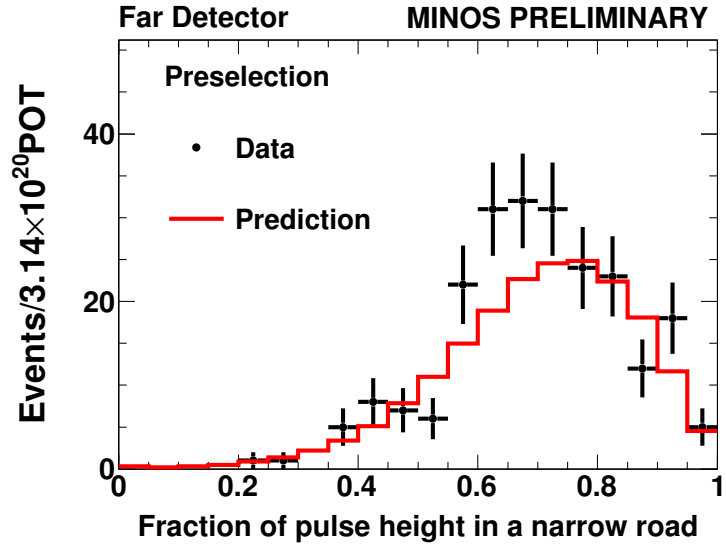


Figure D.18: FD data fraction of pulseheight in a narrow road along the shower direction after preselection. The prediction is shown in red, and the data in black.

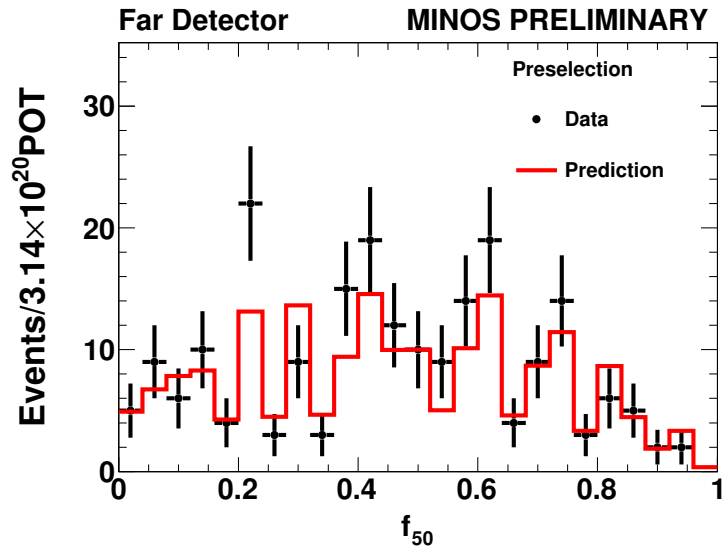


Figure D.19: FD data fraction of the best 50 LEM CC- $\nu_e$  matches with  $y > 0.9$  after preselection. The prediction is shown in red, and the data in black.



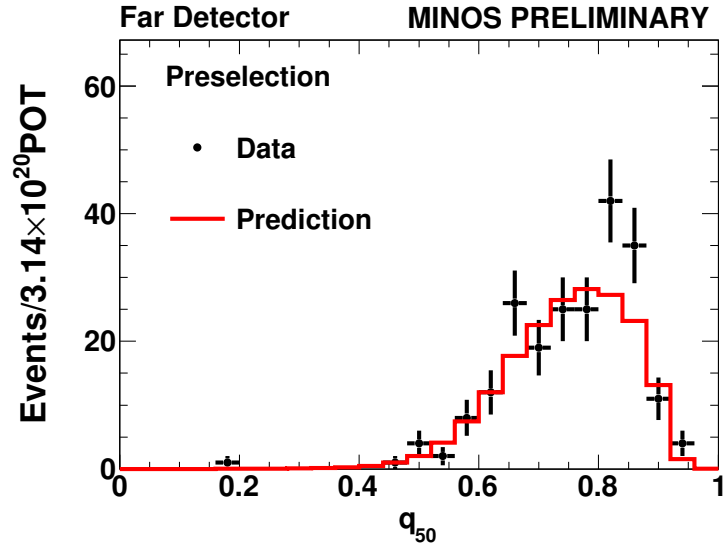


Figure D.20: FD data mean fractional charge of the best 50 LEM CC- $\nu_e$  matches with  $y > 0.9$  after preselection. The prediction is shown in red, and the data in black.

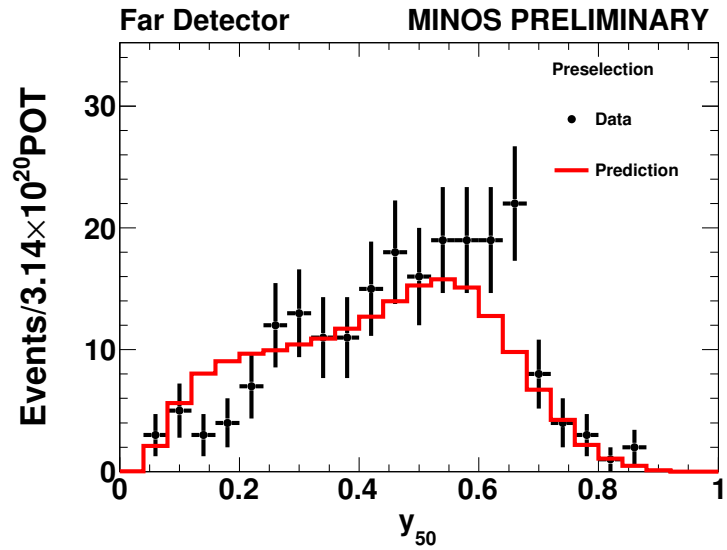


Figure D.21: FD data average  $y$  of the best 50 LEM CC- $\nu_e$  matches with  $y > 0.9$  after preselection. The prediction is shown in red, and the data in black.

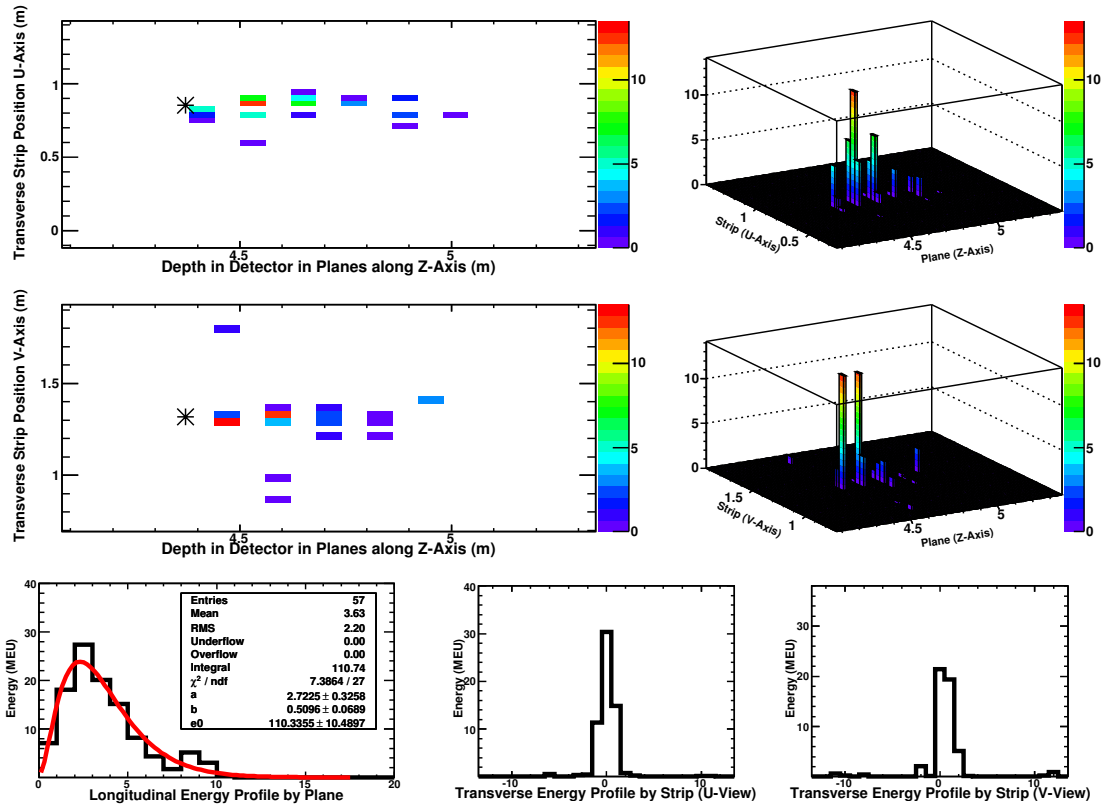


Figure D.22: Golden  $\nu_e$ -CC event with ANN PID value of 0.90 and LEM PID value of 0.97 and reconstructed energy of 4.62GeV, found in FD subrun 32687. The two top left plots show reconstructed event hit strips as a function of longitudinal direction in the detector and transverse U or V position. The top right plots show the same information, but in a type of 3-D view. The bottom middle and right plots show the event transverse energy profiles in U and V, and the left bottom plot shows the longitudinal energy profile in black and the fitted EM-distribution in red.

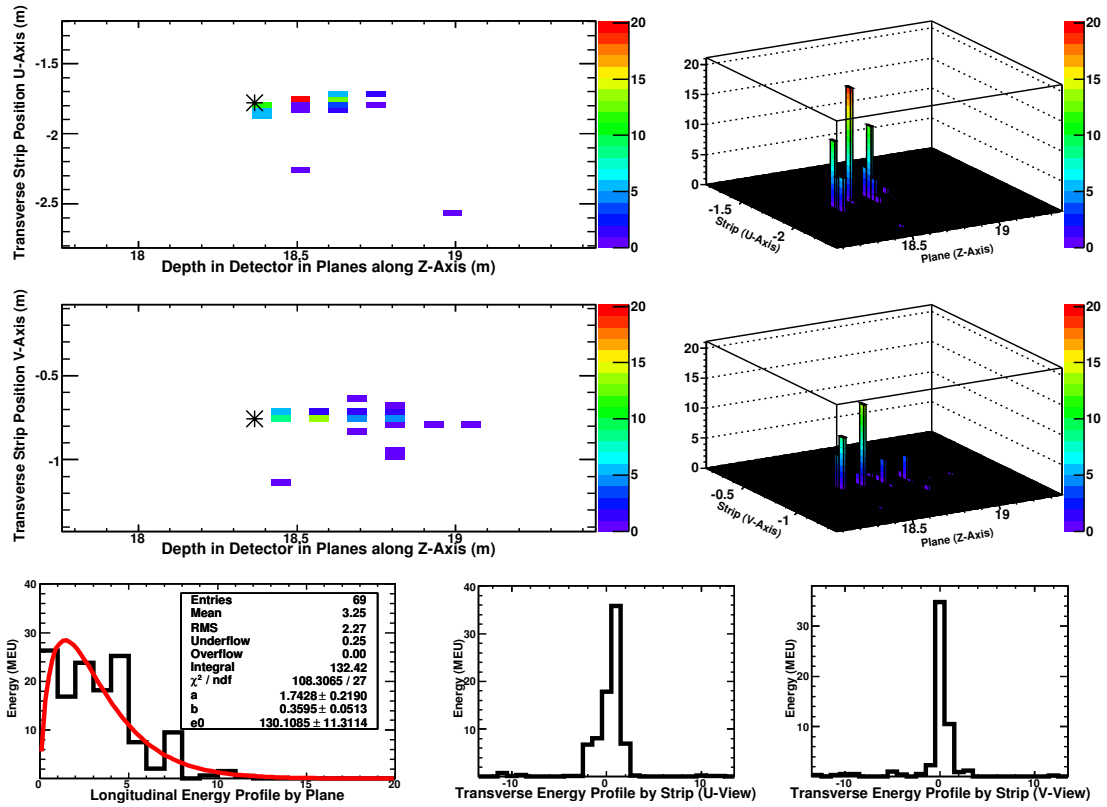


Figure D.23: Golden  $\nu_e$ -CC event with ANN PID value of 0.92 and LEM PID value of 0.94 and reconstructed energy of 5.55GeV, found in FD subrun 37242. The two top left plots show reconstructed event hit strips as a function of longitudinal direction in the detector and transverse U or V position. The top right plots show the same information, but in a type of 3-D view. The bottom middle and right plots show the event transverse energy profiles in U and V, and the left bottom plot shows the longitudinal energy profile in black and the fitted EM-distribution in red.

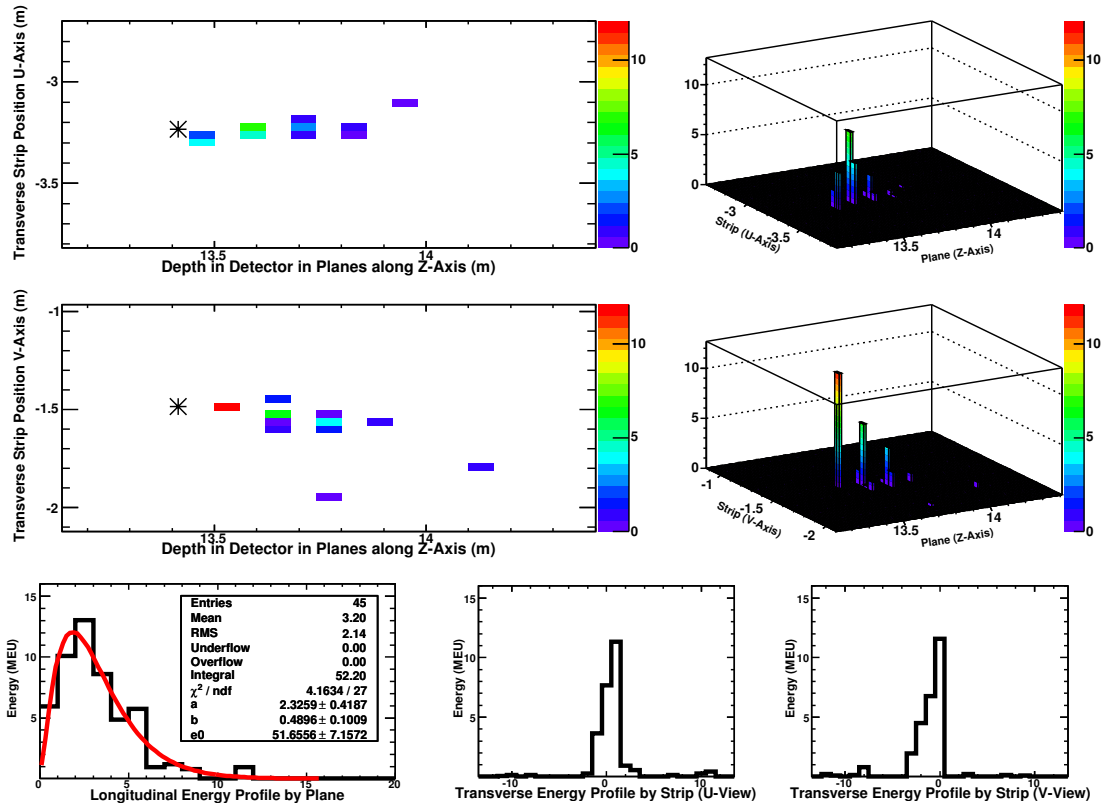


Figure D.24: Golden  $\nu_e$ -CC event with ANN PID value of 0.86 and LEM PID value of 0.93 and reconstructed energy of 2.21GeV, found in FD subrun 38197. The two top left plots show reconstructed event hit strips as a function of longitudinal direction in the detector and transverse U or V position. The top right plots show the same information, but in a type of 3-D view. The bottom middle and right plots show the event transverse energy profiles in U and V, and the left bottom plot shows the longitudinal energy profile in black and the fitted EM-distribution in red.

## Appendix E

### FD Data Event Lists

The following tables contain lists of FD data events selected by the PIDs and, one level lower, by the  $\nu_e$  preselection cuts. Table E.1 contains the events selected by the ANN PID (this is the final  $\nu_e$ -appearance analysis sample). Table E.2 contains events selected by the LEM PID (cross-check PID). Table E.3 contains events selected by both PIDs, and finally, Table E.4 contains events preselected by the  $\nu_e$  analysis cuts, before PID selection has been applied.

| Subrun | Event-ID | ANN PID | LEM PID | Reco. Energy (GeV) |
|--------|----------|---------|---------|--------------------|
| 31745  | 52747    | 0.80    | 0.75    | 3.26               |
| 32569  | 144300   | 0.81    | 0.66    | 3.63               |
| 32617  | 293192   | 0.71    | 0.18    | 4.52               |
| 32623  | 297082   | 0.74    | 0.70    | 2.67               |
| 32684  | 252276   | 0.86    | 0.83    | 3.34               |
| 32687  | 90343    | 0.90    | 0.97    | 4.62               |
| 32805  | 6787     | 0.74    | 0.82    | 2.34               |
| 32976  | 144315   | 0.86    | 0.36    | 3.46               |
| 33193  | 134313   | 0.86    | 0.92    | 6.34               |
| 33483  | 195918   | 0.74    | 0.43    | 4.16               |
| 36576  | 220329   | 0.73    | 0.47    | 2.59               |
| 36592  | 79919    | 0.77    | 0.68    | 2.92               |
| 36614  | 181849   | 0.74    | 0.09    | 5.42               |
| 36730  | 155202   | 0.73    | 0.54    | 2.11               |
| 37096  | 98225    | 0.72    | 0.59    | 3.60               |
| 37112  | 125231   | 0.73    | 0.77    | 1.26               |
| 37204  | 170038   | 0.96    | 0.41    | 3.07               |
| 37207  | 99553    | 0.72    | 0.25    | 2.03               |
| 37242  | 45488    | 0.92    | 0.94    | 5.55               |
| 37401  | 164509   | 0.76    | 0.67    | 4.56               |
| 37731  | 53689    | 0.71    | 0.58    | 1.84               |
| 37746  | 79659    | 0.79    | 0.70    | 2.22               |
| 37761  | 13602    | 0.86    | 0.99    | 2.53               |
| 37841  | 35291    | 0.71    | 0.40    | 2.79               |
| 37853  | 58189    | 0.76    | 0.12    | 3.17               |
| 37904  | 102112   | 0.75    | 0.38    | 3.34               |
| 37929  | 68373    | 0.75    | 0.54    | 4.42               |
| 37980  | 225613   | 0.83    | 1.00    | 5.14               |
| 37986  | 50091    | 0.71    | 0.99    | 4.03               |
| 38011  | 65753    | 0.77    | 0.40    | 3.26               |
| 38163  | 174860   | 0.83    | 0.59    | 2.03               |
| 38188  | 78767    | 0.74    | 0.85    | 2.07               |
| 38197  | 191202   | 0.86    | 0.93    | 2.21               |
| 38221  | 117488   | 0.76    | 0.80    | 1.88               |
| 38304  | 155011   | 0.92    | 0.91    | 2.72               |

Table E.1: FD data event list for an exposure of  $3.14 \times 10^{20}$  POT and selected by the ANN PID (cut value ANN PID > 0.7).

| Subrun | Event-ID | ANN PID | LEM PID | Reco. Energy (GeV) |
|--------|----------|---------|---------|--------------------|
| 31745  | 52747    | 0.80    | 0.75    | 3.26               |
| 32157  | 124035   | 0.45    | 0.93    | 5.14               |
| 32569  | 144300   | 0.81    | 0.66    | 3.63               |
| 32623  | 297082   | 0.74    | 0.70    | 2.67               |
| 32684  | 252276   | 0.86    | 0.83    | 3.34               |
| 32687  | 90343    | 0.90    | 0.97    | 4.62               |
| 32805  | 6787     | 0.74    | 0.82    | 2.34               |
| 33193  | 134313   | 0.86    | 0.92    | 6.34               |
| 33517  | 244910   | 0.60    | 0.68    | 1.07               |
| 33520  | 87625    | 0.63    | 0.74    | 2.86               |
| 36592  | 79919    | 0.77    | 0.68    | 2.92               |
| 37082  | 228034   | 0.58    | 0.65    | 2.80               |
| 37112  | 125231   | 0.73    | 0.77    | 1.26               |
| 37242  | 45488    | 0.92    | 0.94    | 5.55               |
| 37401  | 164509   | 0.76    | 0.67    | 4.56               |
| 37746  | 79659    | 0.79    | 0.70    | 2.22               |
| 37761  | 13602    | 0.86    | 0.99    | 2.53               |
| 37944  | 91945    | 0.68    | 0.84    | 1.81               |
| 37953  | 61714    | 0.55    | 0.82    | 5.48               |
| 37974  | 184541   | 0.44    | 0.66    | 1.65               |
| 37980  | 225613   | 0.83    | 1.00    | 5.14               |
| 37986  | 50091    | 0.71    | 0.99    | 4.03               |
| 38188  | 78767    | 0.74    | 0.85    | 2.07               |
| 38197  | 191202   | 0.86    | 0.93    | 2.21               |
| 38221  | 117488   | 0.76    | 0.80    | 1.88               |
| 38263  | 7555     | 0.70    | 0.85    | 2.07               |
| 38304  | 155011   | 0.92    | 0.91    | 2.72               |
| 38340  | 253126   | 0.59    | 0.98    | 7.91               |

Table E.2: FD data event list for an exposure of  $3.14 \times 10^{20}$  POT and selected by the LEM PID (cut value LEM PID > 0.65).

| Subrun | Event-ID | ANN PID | LEM PID | Reco. Energy (GeV) |
|--------|----------|---------|---------|--------------------|
| 31745  | 52747    | 0.80    | 0.75    | 3.26               |
| 32569  | 144300   | 0.81    | 0.66    | 3.63               |
| 32623  | 297082   | 0.74    | 0.70    | 2.67               |
| 32684  | 252276   | 0.86    | 0.83    | 3.34               |
| 32687  | 90343    | 0.90    | 0.97    | 4.62               |
| 32805  | 6787     | 0.74    | 0.82    | 2.34               |
| 33193  | 134313   | 0.86    | 0.92    | 6.34               |
| 36592  | 79919    | 0.77    | 0.68    | 2.92               |
| 37112  | 125231   | 0.73    | 0.77    | 1.26               |
| 37242  | 45488    | 0.92    | 0.94    | 5.55               |
| 37401  | 164509   | 0.76    | 0.67    | 4.56               |
| 37746  | 79659    | 0.79    | 0.70    | 2.22               |
| 37761  | 13602    | 0.86    | 0.99    | 2.53               |
| 37980  | 225613   | 0.83    | 1.00    | 5.14               |
| 37986  | 50091    | 0.71    | 0.99    | 4.03               |
| 38188  | 78767    | 0.74    | 0.85    | 2.07               |
| 38197  | 191202   | 0.86    | 0.93    | 2.21               |
| 38221  | 117488   | 0.76    | 0.80    | 1.88               |
| 38304  | 155011   | 0.92    | 0.91    | 2.72               |

Table E.3: FD data event list for an exposure of  $3.14 \times 10^{20}$  POT and selected by both the ANN and the LEM PIDs (cut value ANN PID  $> 0.7$  and LEM PID  $> 0.65$ ).



| Subrun | Event-ID | ANN PID | LEM PID | Reco. Energy (GeV) |
|--------|----------|---------|---------|--------------------|
| 31733  | 83627    | 0.42    | 0.16    | 4.40               |
| 31745  | 52747    | 0.80    | 0.75    | 3.26               |
| 31788  | 58801    | 0.68    | 0.14    | 2.04               |
| 31823  | 123634   | 0.47    | 0.07    | 7.89               |
| 31838  | 938      | 0.21    | 0.00    | 2.76               |
| 31891  | 104327   | 0.67    | 0.39    | 2.40               |
| 31979  | 62647    | 0.52    | 0.48    | 1.27               |
| 31990  | 80936    | 0.33    | 0.47    | 1.03               |
| 32120  | 106624   | 0.17    | 0.02    | 1.11               |
| 32153  | 24070    | 0.11    | 0.01    | 6.80               |
| 32157  | 91878    | 0.46    | 0.13    | 5.18               |
| 32157  | 124035   | 0.45    | 0.93    | 5.14               |
| 32163  | 127381   | 0.52    | 0.09    | 2.35               |
| 32454  | 76630    | 0.12    | 0.01    | 3.86               |
| 32506  | 67168    | 0.26    | 0.00    | 3.08               |
| 32537  | 100328   | 0.39    | 0.01    | 3.01               |
| 32566  | 23176    | 0.44    | 0.31    | 1.32               |
| 32566  | 140737   | 0.48    | 0.12    | 6.98               |
| 32569  | 144300   | 0.81    | 0.66    | 3.63               |
| 32575  | 120373   | 0.64    | 0.42    | 3.10               |
| 32578  | 216091   | 0.29    | 0.15    | 2.47               |
| 32605  | 55989    | 0.31    | 0.12    | 1.48               |
| 32617  | 293192   | 0.71    | 0.18    | 4.52               |
| 32620  | 340097   | 0.44    | 0.05    | 2.30               |
| 32623  | 297082   | 0.74    | 0.70    | 2.67               |
| 32632  | 279597   | 0.18    | 0.13    | 7.81               |
| 32663  | 58993    | 0.54    | 0.08    | 4.17               |
| 32672  | 317398   | 0.37    | 0.02    | 5.28               |
| 32684  | 252276   | 0.86    | 0.83    | 3.34               |
| 32687  | 90343    | 0.90    | 0.97    | 4.62               |
| 32690  | 90636    | 0.56    | 0.09    | 1.08               |
| 32713  | 350908   | 0.31    | -1.00   | 6.48               |
| 32728  | 146269   | 0.06    | -1.00   | 7.06               |
| 32731  | 73725    | 0.58    | 0.27    | 1.51               |
| 32731  | 128313   | 0.61    | 0.48    | 2.96               |
| 32737  | 9092     | 0.39    | 0.02    | 4.93               |
| 32740  | 94287    | 0.31    | 0.01    | 2.02               |
| 32750  | 50652    | 0.39    | 0.27    | 7.32               |
| 32788  | 121216   | 0.27    | -1.00   | 2.15               |
| 32805  | 6787     | 0.74    | 0.82    | 2.34               |
| 32805  | 48791    | 0.54    | 0.28    | 5.55               |
| 32808  | 43797    | 0.23    | 0.03    | 1.27               |
| 32811  | 146140   | 0.34    | 0.00    | 3.54               |
| 32901  | 75037    | 0.51    | 0.41    | 5.23               |
| 32973  | 143972   | 0.66    | 0.47    | 1.94               |
| 32976  | 144315   | 0.86    | 0.36    | 3.46               |
| 33014  | 55534    | 0.48    | 0.07    | 2.60               |
| 33017  | 109438   | 0.41    | 0.04    | 3.73               |
| 33102  | 176345   | 0.61    | 0.08    | 2.30               |

| Subrun | Event-ID | ANN PID | LEM PID | Reco. Energy (GeV) |
|--------|----------|---------|---------|--------------------|
| 33102  | 233273   | 0.56    | 0.55    | 2.00               |
| 33129  | 41304    | 0.44    | 0.02    | 2.60               |
| 33132  | 251716   | 0.13    | 0.00    | 5.19               |
| 33138  | 147520   | 0.67    | 0.23    | 2.66               |
| 33181  | 210172   | 0.60    | 0.09    | 3.24               |
| 33184  | 100297   | 0.18    | 0.01    | 3.22               |
| 33193  | 112568   | 0.09    | 0.01    | 7.37               |
| 33193  | 134313   | 0.86    | 0.92    | 6.34               |
| 33259  | 32043    | 0.69    | 0.15    | 2.39               |
| 33271  | 223819   | 0.47    | 0.04    | 1.86               |
| 33280  | 184432   | 0.15    | 0.02    | 5.30               |
| 33458  | 84703    | 0.15    | 0.00    | 4.60               |
| 33483  | 195918   | 0.74    | 0.43    | 4.16               |
| 33505  | 270457   | 0.70    | 0.32    | 5.25               |
| 33508  | 74227    | 0.28    | 0.14    | 1.15               |
| 33508  | 180309   | 0.20    | 0.12    | 6.18               |
| 33517  | 244910   | 0.60    | 0.68    | 1.07               |
| 33520  | 87625    | 0.63    | 0.74    | 2.86               |
| 33545  | 78856    | 0.48    | 0.59    | 2.33               |
| 33545  | 112147   | 0.09    | 0.00    | 4.78               |
| 33606  | 52307    | 0.31    | 0.01    | 5.56               |
| 33619  | 242651   | 0.21    | 0.24    | 3.75               |
| 33619  | 245147   | 0.14    | -1.00   | 5.66               |
| 33622  | 216994   | 0.41    | 0.02    | 4.74               |
| 33646  | 141227   | 0.01    | 0.00    | 2.72               |
| 33646  | 200790   | 0.24    | 0.01    | 5.02               |
| 33677  | 128872   | 0.39    | 0.15    | 6.52               |
| 33785  | 15910    | 0.33    | 0.20    | 3.79               |
| 36576  | 18849    | 0.51    | 0.54    | 1.15               |
| 36576  | 220329   | 0.73    | 0.47    | 2.59               |
| 36579  | 279718   | 0.66    | 0.09    | 1.84               |
| 36589  | 312977   | 0.36    | 0.03    | 2.04               |
| 36592  | 79919    | 0.77    | 0.68    | 2.92               |
| 36605  | 190721   | 0.24    | 0.00    | 5.00               |
| 36614  | 181849   | 0.74    | 0.09    | 5.42               |
| 36614  | 282986   | 0.31    | 0.03    | 4.52               |
| 36643  | 160638   | 0.41    | 0.12    | 2.20               |
| 36649  | 254975   | 0.11    | 0.04    | 6.86               |
| 36652  | 193466   | 0.39    | 0.24    | 4.52               |
| 36652  | 238349   | 0.68    | 0.53    | 1.74               |
| 36662  | 116957   | 0.57    | 0.28    | 1.82               |
| 36668  | 171187   | 0.03    | -1.00   | 7.83               |
| 36718  | 46773    | 0.36    | 0.05    | 1.92               |
| 36724  | 30029    | 0.39    | 0.52    | 4.67               |
| 36727  | 47867    | 0.52    | 0.03    | 1.17               |
| 36730  | 155202   | 0.73    | 0.54    | 2.11               |
| 36747  | 281      | 0.07    | 0.00    | 7.44               |
| 36747  | 34319    | 0.29    | 0.04    | 7.25               |
| 36750  | 182222   | 0.11    | 0.00    | 4.81               |

| Subrun | Event-ID | ANN PID | LEM PID | Reco. Energy (GeV) |
|--------|----------|---------|---------|--------------------|
| 36862  | 117781   | 0.50    | 0.03    | 2.81               |
| 37013  | 173412   | 0.09    | 0.02    | 1.37               |
| 37016  | 9610     | 0.25    | 0.54    | 1.33               |
| 37022  | 96729    | 0.59    | 0.09    | 3.90               |
| 37028  | 11507    | 0.11    | 0.02    | 1.03               |
| 37037  | 20185    | 0.45    | 0.32    | 4.49               |
| 37040  | 73957    | 0.42    | 0.15    | 3.57               |
| 37040  | 210484   | 0.27    | 0.12    | 2.18               |
| 37043  | 180363   | 0.52    | 0.17    | 3.65               |
| 37050  | 17678    | 0.42    | 0.06    | 3.14               |
| 37050  | 19546    | 0.25    | 0.11    | 6.09               |
| 37054  | 19743    | 0.14    | 0.01    | 2.09               |
| 37057  | 41094    | 0.13    | 0.01    | 6.76               |
| 37063  | 251884   | 0.16    | 0.02    | 6.42               |
| 37069  | 95078    | 0.26    | 0.00    | 4.90               |
| 37069  | 147868   | 0.66    | 0.62    | 2.45               |
| 37076  | 29938    | 0.40    | 0.01    | 3.52               |
| 37082  | 228034   | 0.58    | 0.65    | 2.80               |
| 37096  | 98225    | 0.72    | 0.59    | 3.60               |
| 37096  | 158117   | 0.24    | -1.00   | 6.50               |
| 37109  | 176639   | 0.20    | 0.02    | 4.30               |
| 37112  | 125231   | 0.73    | 0.77    | 1.26               |
| 37112  | 181168   | 0.15    | 0.02    | 5.17               |
| 37115  | 83230    | 0.36    | 0.04    | 7.92               |
| 37118  | 127433   | 0.40    | 0.09    | 1.94               |
| 37126  | 71692    | 0.33    | 0.01    | 3.66               |
| 37129  | 5499     | 0.69    | 0.18    | 2.63               |
| 37129  | 45362    | 0.59    | 0.64    | 4.60               |
| 37129  | 150548   | 0.27    | 0.04    | 1.49               |
| 37135  | 62500    | 0.19    | 0.08    | 6.41               |
| 37159  | 51174    | 0.53    | 0.13    | 3.69               |
| 37201  | 247490   | 0.46    | 0.03    | 4.85               |
| 37204  | 170038   | 0.96    | 0.41    | 3.07               |
| 37207  | 21680    | 0.41    | 0.61    | 5.85               |
| 37207  | 99553    | 0.72    | 0.25    | 2.03               |
| 37210  | 69391    | 0.19    | 0.02    | 4.71               |
| 37210  | 141549   | 0.63    | 0.59    | 4.76               |
| 37218  | 111278   | 0.25    | 0.00    | 2.80               |
| 37242  | 45488    | 0.92    | 0.94    | 5.55               |
| 37242  | 108753   | 0.64    | 0.21    | 3.90               |
| 37246  | 116995   | 0.51    | 0.22    | 3.03               |
| 37262  | 209412   | 0.28    | 0.01    | 3.22               |
| 37357  | 118427   | 0.60    | 0.24    | 3.38               |
| 37401  | 164509   | 0.76    | 0.67    | 4.56               |
| 37636  | 22235    | 0.29    | 0.01    | 3.10               |
| 37686  | 14919    | 0.21    | 0.00    | 4.15               |
| 37724  | 7105     | 0.52    | 0.08    | 6.72               |
| 37728  | 210611   | 0.67    | 0.38    | 2.63               |
| 37731  | 53689    | 0.71    | 0.58    | 1.84               |

| Subrun | Event-ID | ANN PID | LEM PID | Reco. Energy (GeV) |
|--------|----------|---------|---------|--------------------|
| 37734  | 90123    | 0.57    | 0.25    | 5.51               |
| 37734  | 148335   | 0.17    | 0.00    | 4.27               |
| 37740  | 30567    | 0.48    | 0.56    | 1.13               |
| 37746  | 79659    | 0.79    | 0.70    | 2.22               |
| 37746  | 82105    | 0.26    | 0.02    | 2.30               |
| 37755  | 28528    | 0.16    | 0.00    | 6.45               |
| 37761  | 13602    | 0.86    | 0.99    | 2.53               |
| 37779  | 177138   | 0.03    | 0.00    | 1.02               |
| 37783  | 185883   | 0.20    | 0.09    | 5.36               |
| 37789  | 83124    | 0.31    | 0.04    | 1.34               |
| 37793  | 17024    | 0.40    | 0.07    | 3.97               |
| 37798  | 65889    | 0.57    | 0.15    | 4.02               |
| 37813  | 33050    | 0.07    | -1.00   | 7.99               |
| 37813  | 208405   | 0.29    | 0.02    | 4.32               |
| 37826  | 18432    | 0.19    | 0.16    | 3.73               |
| 37835  | 75082    | 0.26    | 0.00    | 3.65               |
| 37838  | 77412    | 0.16    | 0.04    | 3.70               |
| 37841  | 35291    | 0.71    | 0.40    | 2.79               |
| 37853  | 58189    | 0.76    | 0.12    | 3.17               |
| 37868  | 85398    | 0.36    | 0.02    | 3.38               |
| 37868  | 137937   | 0.19    | 0.00    | 6.21               |
| 37874  | 43667    | 0.16    | 0.00    | 5.89               |
| 37874  | 51162    | 0.38    | 0.03    | 2.29               |
| 37874  | 54216    | 0.25    | 0.01    | 3.23               |
| 37901  | 7375     | 0.20    | 0.02    | 1.14               |
| 37901  | 214624   | 0.37    | 0.02    | 3.31               |
| 37904  | 102112   | 0.75    | 0.38    | 3.34               |
| 37910  | 69653    | 0.57    | 0.56    | 2.39               |
| 37929  | 5791     | 0.41    | 0.04    | 5.92               |
| 37929  | 68373    | 0.75    | 0.54    | 4.42               |
| 37944  | 91945    | 0.68    | 0.84    | 1.81               |
| 37947  | 92685    | 0.42    | -1.00   | 6.22               |
| 37953  | 61714    | 0.55    | 0.82    | 5.48               |
| 37953  | 200150   | 0.46    | 0.60    | 5.14               |
| 37956  | 234985   | 0.31    | 0.06    | 1.01               |
| 37959  | 109353   | 0.35    | 0.10    | 5.46               |
| 37962  | 66672    | 0.09    | 0.00    | 4.49               |
| 37962  | 95031    | 0.70    | 0.56    | 1.98               |
| 37971  | 17054    | 0.58    | 0.01    | 2.63               |
| 37974  | 98527    | 0.30    | 0.10    | 1.89               |
| 37974  | 127659   | 0.20    | 0.05    | 6.83               |
| 37974  | 184541   | 0.44    | 0.66    | 1.65               |
| 37974  | 239257   | 0.63    | 0.32    | 1.27               |
| 37980  | 225613   | 0.83    | 1.00    | 5.14               |
| 37986  | 50091    | 0.71    | 0.99    | 4.03               |
| 37996  | 21131    | 0.57    | 0.14    | 2.08               |
| 38002  | 45895    | 0.66    | 0.42    | 2.08               |
| 38008  | 236595   | 0.43    | 0.10    | 3.81               |
| 38011  | 65753    | 0.77    | 0.40    | 3.26               |

| Subrun | Event-ID | ANN PID | LEM PID | Reco. Energy (GeV) |
|--------|----------|---------|---------|--------------------|
| 38018  | 92216    | 0.13    | 0.00    | 1.84               |
| 38033  | 36940    | 0.02    | 0.00    | 6.01               |
| 38163  | 48247    | 0.31    | 0.01    | 2.93               |
| 38163  | 174860   | 0.83    | 0.59    | 2.03               |
| 38169  | 158807   | 0.48    | 0.02    | 3.21               |
| 38172  | 23913    | 0.17    | 0.00    | 2.38               |
| 38185  | 206844   | 0.16    | 0.03    | 5.89               |
| 38185  | 298455   | 0.10    | -1.00   | 7.37               |
| 38188  | 78767    | 0.74    | 0.85    | 2.07               |
| 38197  | 191202   | 0.86    | 0.93    | 2.21               |
| 38206  | 66129    | 0.58    | 0.13    | 1.91               |
| 38209  | 43935    | 0.62    | 0.02    | 3.25               |
| 38218  | 260429   | 0.66    | 0.33    | 2.54               |
| 38221  | 117488   | 0.76    | 0.80    | 1.88               |
| 38243  | 130786   | 0.68    | 0.01    | 3.26               |
| 38246  | 153837   | 0.69    | 0.22    | 1.89               |
| 38252  | 158002   | 0.11    | 0.01    | 7.00               |
| 38252  | 188758   | 0.19    | 0.04    | 2.26               |
| 38255  | 127392   | 0.65    | 0.31    | 2.56               |
| 38263  | 7555     | 0.70    | 0.85    | 2.07               |
| 38263  | 171321   | 0.32    | 0.21    | 2.31               |
| 38269  | 71643    | 0.34    | 0.00    | 3.56               |
| 38283  | 54209    | 0.12    | 0.00    | 5.69               |
| 38292  | 33057    | 0.35    | 0.00    | 2.03               |
| 38304  | 155011   | 0.92    | 0.91    | 2.72               |
| 38304  | 205257   | 0.37    | 0.05    | 1.41               |
| 38324  | 167528   | 0.00    | 0.01    | 1.74               |
| 38330  | 24194    | 0.64    | 0.33    | 2.99               |
| 38340  | 253126   | 0.59    | 0.98    | 7.91               |
| 38346  | 135543   | 0.50    | 0.03    | 3.41               |
| 38349  | 4232     | 0.60    | 0.31    | 4.11               |

Table E.4: FD data event list for an exposure of  $3.14 \times 10^{20}$  POT and preselected by the  $\nu_e$  analysis cuts (prior to applying a final PID selection).