

---

1

2        $\mu/\pi$  separation using

3       Convolutional Neural Networks

4       for the MicroBooNE

5       Charged Current Inclusive Cross Section

6       Measurement

---

7

8                   Jessica Nicole Esquivel

9                   Bachelor of Science in Electrical Engineering and Applied Physics

10                  St. Mary's University

11                  San Antonio, TX, USA 2011

12                   DISSERTATION

13                   Submitted in partial fulfillment  
14                  of the requirements for the degree

15                  *Doctor of Philosophy in Physics*

16                   - \* - DRAFT January 30, 2018 - \* -

17                   December, 2017  
18                  Syracuse University  
19                  Syracuse, New York

20  
21  
22  
23

Copyright 2017  
Jessica Nicole Esquivel  
All Rights Reserved  
Syracuse University

24



$\mu$ BooNE







26

## Abstract

27

The purpose of this thesis was to use Convolutional Neural Networks  
28 (CNN) to separate  $\mu'$ s and  $\pi'$ s for use in increasing the acceptance rate  
29 of  $\mu'$ 's below the implemented 75cm track length cut in the Charged  
30 Current Inclusive (CC-Inclusive) event selection for the CC-Inclusive  
31 Cross-Section Measurement. In doing this, we increase acceptance  
32 rate for CC-Inclusive events below a specific momentum range.



33

## Dedication

34

I dedicate this dissertation to the two important women in my life; My  
35 wife and my mom. Both have been there cheering me on giving me strength  
36 and love as I worked towards the hardest accomplishment I've ever done.

37

Jessica Nicole Esquivel



38

## Acknowledgements

39        Of the many people who deserve thanks, some are particularly prominent, such as  
40        my supervisor...



# **Contents**

41	<b>List of figures</b>	xv
43	<b>List of tables</b>	xxv
44	<b>1 Introduction</b>	1
45	<b>2 Neutrinos</b>	3
46	2.1 What are Neutrinos . . . . .	3
47	2.2 History of Neutrinos . . . . .	4
48	2.2.1 Solar Oscillations and the Solar Neutrino Problem . . . . .	4
49	2.2.2 Atmospheric Oscillations and the Atmospheric Neutrino Anomaly	7
50	2.3 Neutrino Oscillations . . . . .	9
51	2.3.1 Two Flavor Neutrino Oscillation Formulation . . . . .	9
52	2.3.2 Three Flavor Neutrino Oscillation Formulation . . . . .	12
53	2.3.3 Reactor Oscillation . . . . .	13
54	<b>3 The MicroBooNE Experiment</b>	15
55	3.1 Liquid argon time projection chambers . . . . .	15
56	3.2 The MicroBooNE Time Projection Chamber . . . . .	16
57	3.3 MicroBooNE’s Physics Goals . . . . .	18
58	3.3.1 The low-energy excess . . . . .	18
59	3.3.2 Cross sections . . . . .	19
60	3.3.3 Liquid argon detector development . . . . .	20
61	3.4 The Booster Neutrino Beam . . . . .	20
62	3.4.1 Creating the Booster Neutrino Beam . . . . .	21
63	3.5 Event Reconstruction . . . . .	22

64	<b>4 Neutrino Identification: Finding MicroBooNE's first Neutrinos</b>	25
65	4.1 Flash Finding . . . . .	25
66	4.1.1 Flash Reconstruction . . . . .	26
67	4.1.2 Beam Timing . . . . .	27
68	4.1.3 Event Rates . . . . .	28
69	4.2 TPC Topology Selection . . . . .	28
70	4.2.1 Cosmic Tagging . . . . .	29
71	4.2.2 2D Cluster Selection . . . . .	29
72	4.2.3 3D Tracks and vertices Selection . . . . .	31
73	4.2.4 TPC Updates . . . . .	33
74	4.3 Conclusion . . . . .	34
75	<b>5 CC-Inclusive Cross Section Selection Filter</b>	37
76	5.1 Data and MC Processing Chain . . . . .	38
77	5.2 Normalization of data and MC . . . . .	39
78	5.3 Optical Software Trigger and Reconstruction . . . . .	40
79	5.3.1 Software Trigger . . . . .	40
80	5.3.2 Flash Reconstruction . . . . .	41
81	5.3.3 Beam Window . . . . .	42
82	5.4 TPC Reconstruction . . . . .	44
83	5.4.1 Hit Reconstruction . . . . .	45
84	5.4.2 Clustering . . . . .	45
85	5.4.3 Pandora . . . . .	46
86	5.4.4 Trackkalmanhit . . . . .	46
87	5.4.5 Cosmic Hit Removal . . . . .	46
88	5.4.6 Projection Matching Algorithm . . . . .	47
89	5.5 Event Selection . . . . .	47
90	5.5.1 Expected Bakgrounds . . . . .	48
91	5.5.2 Truth Distributions . . . . .	51
92	<b>6 Background on Convolutional Neural Networks</b>	59
93	6.1 Image Classification . . . . .	59
94	6.2 CNN Structure . . . . .	60
95	6.2.1 Backpropagation . . . . .	63
96	6.3 Choosing Hyperparameters . . . . .	64

97	<b>7 Training process of Convolutional Neural Networks</b>	<b>67</b>
98	7.1 Hardware Configurations for Convolutional Neural Network Training	67
99	7.2 Creating images using LArTPC data for training/validation of CNNS	68
100	7.3 Convolutional Neural Network Training . . . . .	70
101	7.3.1 Training CNN1075 . . . . .	70
102	7.3.2 Training CNN10000 . . . . .	73
103	7.3.3 Training CNN100000 . . . . .	76
104	<b>8 Using Convolutional Neural Networks for <math>\nu_\mu</math> CC event classification</b>	<b>87</b>
105	8.1 Classification using CNN10000 . . . . .	87
106	8.1.1 Classification of MC data using Selection I CC-Inclusive Filter .	87
107	8.1.2 Conclusions of CNN10000 classification of MC data . . . . .	95
108	8.2 Classification using CNN100000 . . . . .	96
109	8.2.1 Classification of MC data using Selection I CC-Inclusive Filter .	96
110	8.2.2 Classification of MicroBooNE data using Selection I CC-Inclusive	
111	Filter . . . . .	96
112	8.2.3 Comparing two CC-Inclusive Cross Section Selection Filters . .	96
113	<b>9 Conclusion</b>	<b>107</b>
114	<b>Bibliography</b>	<b>109</b>



# <sup>115</sup> List of figures

<sup>116</sup>	2.1	The Standard Solar Model	5
<sup>117</sup>	2.2	Solar Neutrino Experiments	7
<sup>118</sup>	2.2a	Ray Davis's Homestake Experiment	7
<sup>119</sup>	2.2b	Kamiokande Experiment	7
<sup>120</sup>	2.2c	SNO Experiment	7
<sup>121</sup>	2.3	Cosmic Ray Shower	8
<sup>122</sup>	2.4	Measurements of the double ratio for various atmospheric neutrino experiments	8
<sup>124</sup>	2.5	The flavor eigenstates are rotated by an angle $\theta$ with respect to the mass eigenstates	10
<sup>126</sup>	3.1	Low Energy excess seen in MiniBooNE	19
<sup>127</sup>	3.2	Arial view of the Main Injector and the Booster Neutrino Beam at Fermilab	21
<sup>128</sup>	3.3	Arial view of the Main Injector and the Booster Neutrino Beam at Fermilab	23
<sup>129</sup>	3.4	Energy spectrum of the Booster Neutrino Beam at Fermi National Laboratories	24
<sup>131</sup>	4.1	Efficiency for selecting beam events as a function of minimum total PE cut for all PE cuts as well as zoomed into interesting region.	26
<sup>133</sup>	4.2a	Predicted distribution of flash times with respect to trigger time for 1 day of data taking at nominal rate and intensity	31

135	4.2b	Measured distribution of flash times with a 50 PE threshold cut, with respect to trigger time. Shown as a ratio to the expected cosmic rate from off-beam data. A clear excess from neutrinos is visible between 3- 5 $\mu$ s after the trigger time. . . . .	31
139	4.3	Percent of good clusters remaining for neutrinos and cosmics after the primary cuts were applied. This is relative to total number of initial clusters. . . . .	32
142	4.4	Percent matched cluster pairs remaining for neutrinos and cosmics after secondary cuts applied. This is relative to the number of events that contain clusters which pass the primary cuts. . . . .	32
145	4.5	First Neutrino Interaction Candidate Events from MicroBooNE . . . . .	35
146	4.6	First Neutrino Interaction Candidate Events from MicroBooNE . . . . .	36
147	5.1	5.1a Track range distribution for selection I. The track range is defined as the 3D distance between the start and end of the muon candidate track. No data is shown below 75 cm due to the track length cut described previously. 5.1b Efficiency of the selected events by process quasi-elastic (QE), resonant (RES), and deep-inelastic (DIS). Statistical uncertainty is shown in the bands and the distributions are a function of true muon momentum. The rise of the efficiency between 0 GeV and 0.5 GeV is due to the minimum track length cut and the decreasing efficiency for higher momentum tracks is caused by the containment requirement. . . . .	40
156	5.1a	Track range distribution of selection I . . . . .	40
157	5.1b	Selection efficiency as a function of the true muon momentum . . . . .	40
158	5.2	Time distribution of reconstructed optical flashes with a PE value of 50 or more for a sample of BNB unbiased triggered events. . . . .	42
160	5.3	Flash time distribution for all flashes (left plot) and flashes > 20PE (right plot). The different curves are as follows: on-beam data (black), off-beam data (red), CORSIKA inTime MC (light blue), BNB only MC (green), and BNB+Cosmic MC (purple). The dashed vertical lines mark the time window that was chosen for each sample . . . . .	43

165	5.4 Reconstruction chain run on both data and MC. The red stars mean that the algorithm returns reconstructed 3D vertices. . . . .	44
167	5.5 Event selection diagram for selection I and selection II. This analysis focused on optimizing selection I. Boxes with the same color across the two selections symbolize similar cuts. . . . .	49
170	5.6 MC momentum distributions of the muon originating from a $\nu_\mu$ CC interaction. Upper left is the momentum distributions of events with a vertex within the FV (green) and the events with a fully contained track (blue) before the selection. The upper right side is the momentum distribution after the selection (red). The lower plot is the selection efficiencies. . . . .	52
176	5.7 MC $\cos(\theta)$ distributions of the muon originating from a $\nu_\mu$ CC interac- tion. Upper left is the $\cos(\theta)$ distributions of events with a vertex within the FV (green) and the events with a fully contained track (blue) before the selection. The upper right side is the $\cos(\theta)$ distribution after the selection (red). The lower plot is the selection efficiencies. . . . .	53
181	5.8 MC $\phi$ distributions of the muon originating from a $\nu_\mu$ CC interaction. Upper left is the $\phi$ distributions of events with a vertex within the FV (green) and the events with a fully contained track (blue) before the selection. The upper right side is the $\phi$ distribution after the selection (red). The lower plot is the selection efficiencies. . . . .	54
186	5.9 MC momentum distributions of the muon originating from a $\nu_\mu$ CC interaction. Upper left is the momentum distributions of events with a vertex within the FV split up into CCQE (red), CCRES (yellow), and CCDIS (green) before selection. The upper right side is the momentum distribution after the selection with the same color schemes. The lower plot is the selection efficiencies for all three interaction types. The definition of QE, RES, and DIS is based on GENIE. . . . .	55

193	5.10 MC $\cos(\theta)$ distributions of the muon originating from a $\nu_\mu$ CC interaction. Upper left is the $\cos(\theta)$ distributions of events with a vertex within the FV split up into CCQE (red), CCRES (yellow), and CCDIS (green) before selection. The upper right side is the $\cos(\theta)$ distribution after the selection with the same color schemes. The lower plot is the selection efficiencies for all three interaction types. The definition of QE, RES, and DIS is based on GENIE. . . . .	56
200	5.11 MC $\phi$ distributions of the muon originating from a $\nu_\mu$ CC interaction. Upper left is the $\phi$ distributions of events with a vertex within the FV split up into CCQE (red), CCRES (yellow), and CCDIS (green) before selection. The upper right side is the $\phi$ distribution after the selection with the same color schemes. The lower plot is the selection efficiencies for all three interaction types. The definition of QE, RES, and DIS is based on GENIE. . . . .	57
207	6.1 Pixel representation and visualization of a curve detector filter. As you can see, in the pixel representation, the weights of this filter are greater along a curve we are trying to find in the input image . . . . .	61
210	6.2 Visualization of filters found in first layer of a CNN. . . . .	61
211	6.3 Applying a feature mask over a set of fashion items to extract necessary information for auto-encoding. Unnecessary information for example color or brand emblems are not saved. This feature map is an edge detection mask that leaves only shape information which helps to distinguish between different types of clothes. . . . .	62
216	6.4 Pictorial Representation of Convolutional Neural Networks as well as a visual representation on CNN's complexity of layer feature extraction	62
218	6.5 Pictorial representation of the AlexNet model. The AlexNet model consists of 5 convolution layers and 3 fully connected layers. . . . .	65

220	6.6	Pictoral representation of the GoogleNet model. The GoogleNet model 221 consists of 22 layers. The model implements 9 Inception modules which 222 performs covolution and pooling in parallel and strays away from 223 the basis that CNN layers need to be stacked up sequentially. The 224 GoogleNet model also doesn't use fully connected layers, instead it 225 uses average pooling which greatly reduces the amount of parameters. 226 GoogleNet has 12x fewer parameters than AlexNet. . . . .	65
227	7.1	Accuracy vs. Loss of AlexNet 2-output $\mu/\pi$ sample consisting of 2,150 228 images each. . . . .	71
229	7.2	Description of confusion matrix variables: False pion rate = $false\pi/total\pi$ 230 True pion rate = $true\pi/total\pi$ Accuracy = $(true\pi rate + true\mu rate)/2$ 231 Pion prediction value = $true\pi/(true\pi + false\pi)$ Muon prediction value 232 = $true\mu/(true\mu + false\mu)$ . . . . .	72
233	7.2a	Confusion Matrix showing Accuracy of CNN1075 using training 234 MC data . . . . .	72
235	7.2b	Confusion Matrix showing Accuracy of CNN1075 using testing 236 MC data . . . . .	72
237	7.3	Accuracy vs. Loss of AlexNet 2-output $\mu/\pi$ sample consisting of 10,000 238 images each. . . . .	74
239	7.4	Description of confusion matrix variables: False pion rate = $false\pi/total\pi$ ; 240 True pion rate = $true\pi/total\pi$ ; Accuracy = $(true\pi rate + true\mu rate)/2$ ; 241 Pion prediction value = $true\pi/(true\pi + false\pi)$ ; Muon prediction value 242 = $true\mu/(true\mu + false\mu)$ . . . . .	75
243	7.4a	Confusion Matrix showing Accuracy of CNN10000 using training 244 MC data . . . . .	75
245	7.4b	Confusion Matrix showing Accuracy of CNN10000 using testing 246 MC data . . . . .	75
247	7.5	Probability plot of muons and pions from testing set. Images surrounding 248 histogram are a random event from lowest bin and highest bin for 249 each particle. . . . .	76

250	7.6 Training and testing accuracy of CNN trained on 100,000 images of $\mu/\pi/p/\gamma/e$ with 20,000 images of each particle. Each image was a size of 576x576 and the images per particle were split 90% use for training and 10% used for testing the network . . . . .	77
254	7.7 Training and testing loss of CNN trained on 100,000 images of $\mu/\pi/p/\gamma/e$ . . . . .	77
255	7.8 Confusion Matrix of all five particles . . . . .	78
256	7.9 t-SNE of CNN . . . . .	79
257	7.10 Probabilities of different particle classes as well as their contamination from other classes . . . . .	80
259	7.10a Muon Prob . . . . .	80
260	7.10b Pion Prob . . . . .	80
261	7.10c Proton Prob . . . . .	80
262	7.10d Electron Prob . . . . .	80
263	7.10e Gamma Prob . . . . .	80
264	7.11 Muon probability of true muons (blue) versus pions (red), protons (cyan), gammas (green) and electrons (magenta). . . . .	81
266	7.12 Kinematic distributions versus muon probability for true muons and true pions. . . . .	82
268	7.12a Track range versus muon probability for true muons (blue) and true pions (red). . . . .	82
270	7.12b Track range $\leq 75$ cm versus muon probability for true muons (blue) and true pions (red). . . . .	82
272	7.12c Momentum versus muon probability for true muons (blue) and true pions (red). . . . .	82
274	7.13 Kinematic distributions versus muon probability for true muons and true protons. . . . .	83
276	7.13a Track range versus muon probability for true muons (blue) and true protons (cyan). . . . .	83

278	7.13b Track range $\leq 75$ cm versus muon probability for true muons (blue) and true protons (cyan). . . . .	83
279		
280	7.13c Momentum versus muon probability for true muons (blue) and true protons (cyan). . . . .	83
281		
282	7.14 Kinematic distributions versus muon probability for true muons and true electrons. . . . .	84
283		
284	7.14a Track range versus muon probability for true muons (blue) and true electrons (magenta). . . . .	84
285		
286	7.14b Track range $\leq 75$ cm versus muon probability for true muons (blue) and true electrons (magenta). . . . .	84
287		
288	7.14c Momentum versus muon probability for true muons (blue) and true electrons (magenta). . . . .	84
289		
290	7.15 Kinematic distributions versus muon probability for true muons and true gammas. . . . .	85
291		
292	7.15a Track range versus muon probability for true muons (blue) and true gammas (green). . . . .	85
293		
294	7.15b Track range $\leq 75$ cm versus muon probability for true muons (blue) and true gammas (green). . . . .	85
295		
296	7.15c Momentum versus muon probability for true muons (blue) and true gammas (green). . . . .	85
297		
298	8.1 Confusion matrix and probability plot of events passing selection I cc-inclusive cuts right before 75cm track length cut . . . . .	88
299		
300	8.1a Confusion Matrix for CNN10000 classified events from selection I	88
301	8.1b Probability plot for CNN10000 classified events from selection I	88
302	8.2 CNN10000 distributions of track candidate images output from Selection I Modified cc-inclusive filter . . . . .	89
303		
304	8.2a Track range distribution of events from Selection I Modified passing CNN with 70% accuracy . . . . .	89
305		

306	8.2b	Stacked signal and background track range distributions from Selection I Modified passing CNN with 70% accuracy . . . . .	89
307			
308	8.2c	Stacked signal and background track range distributions from Selection I Modified passing 75 cm track length cut . . . . .	89
309			
310	8.2d	Stacked signal muons and background muons/pions of track range distributions from Selection I Modified passing CNN with 70% accuracy . . . . .	89
311			
312			
313	8.2e	Stacked signal muons and background muons/pions of track range distributions from Selection I Modified passing 75 cm track length cut . . . . .	89
314			
315			
316	8.3	CNN10000 stacked signal/background track range distributions of track candidate images output from Selection I Modified cc-inclusive filter	90
317			
318	8.3a	Stacked signal muons and background muons/pions of track range distributions from Selection I Modified passing CNN with 75% accuracy . . . . .	90
319			
320			
321	8.3b	Stacked signal muons and background muons/pions of track range distributions from Selection I Modified passing CNN with 80% accuracy . . . . .	90
322			
323			
324	8.3c	Stacked signal muons and background muons/pions of track range distributions from Selection I Modified passing CNN with 85% accuracy . . . . .	90
325			
326			
327	8.3d	Stacked signal muons and background muons/pions of track range distributions from Selection I Modified passing CNN with 90% accuracy . . . . .	90
328			
329			
330	8.4	CNN10000 momentum distributions of track candidate images output from Selection I Modified cc-inclusive filter . . . . .	91
331			
332	8.4a	Momentum distribution of events from Selection I Modified passing CNN with 70% accuracy . . . . .	91
333			
334	8.4b	Stacked signal and background momentum distributions from Selection I Modified passing CNN with 70% accuracy . . . . .	91
335			

336	8.4c	Stacked signal and background momentum distributions from Selection I Modified passing 75 cm track length cut . . . . .	91
337			
338	8.5	Track distribution comparisons of true CC muons plotted vs true CC muons and pions plotted . . . . .	92
339			
340	8.5a	Stacked signal $\mu$ /background $\mu$ and $\pi$ track range distribution of CNN @ 70% . . . . .	92
341			
342	8.5b	Stacked signal $\mu\&\pi$ /background $\mu\&\pi$ track range distribution of CNN @ 70% . . . . .	92
343			
344	8.6	Images of true CC events where the pion was the tagged track candidate . . . . .	93
345	8.6a	Pion reconstructed track range is less than 75 cm and longer than muon track due to dead wires . . . . .	93
346			
347	8.6b	Pion reconstructed track range is less than 75 cm and larger than muon reconstructed track . . . . .	93
348			
349	8.6c	Pion reconstructed track range is greater than 75 cm and larger than muon reconstructed track . . . . .	93
350			
351	8.7	CNN performance of classified muons and pions compared to the already implemented 75 cm track length cut . . . . .	95
352			
353	8.8	$\text{Cos}(\theta)$ distribution at $\text{CNN10000} \geq 85\%$ . . . . .	97
354			
355	8.9	$\text{Cos}(\theta)$ distribution at $\text{CNN10000} \geq 85\%$ . . . . .	98
356			
357	8.10	$\text{Cos}(\theta)$ distribution at $\text{CNN10000} \geq 85\%$ . . . . .	99
358			
359	8.11	$\text{Cos}(\theta)$ distribution at $\text{CNN10000} \geq 85\%$ . . . . .	100
360			
361	8.12	$\text{Cos}(\theta)$ distribution at $\text{CNN10000} \geq 85\%$ . . . . .	101
362			
363	8.13	$\text{Cos}(\theta)$ distribution at $\text{CNN10000} \geq 85\%$ . . . . .	102
364			
365	8.14	$\text{Cos}(\theta)$ distribution at $\text{CNN10000} \geq 85\%$ . . . . .	103
366			
367	8.15	$\text{Cos}(\theta)$ distribution at $\text{CNN10000} \geq 85\%$ . . . . .	104
368			
369	8.16	$\text{Cos}(\theta)$ distribution at $\text{CNN10000} \geq 85\%$ . . . . .	105
370			
371	8.17	$\text{Cos}(\theta)$ distribution at $\text{CNN10000} \geq 85\%$ . . . . .	106
372			



# <sup>363</sup> List of tables

<sup>364</sup> 4.1	Passing rates for 2D cluster cuts for neutrino on MC set and a cosmic only MC set. First column shows event rates with no cuts applied to both sets. Columns two and three show event rates after primary and secondary cuts are applied. Line three shows the second line scaled with the flash finding factor of 0.008. All events are normalized to per day assuming we are running at 5 Hz. . . . .	31
<sup>370</sup> 5.1	Passing rates of Selection I. Numbers are absolute event counts and cosmic background is not scaled. The BNB+Cosmic sample contains all events, not just $\nu_\mu$ CC inclusive. The numbers in brackets give the passing rate wrt the step before (first percentage) and wrt total generated events (second percentage). The BNB+Cosmic MC-Truth column shows how many true $\nu_\mu$ CC inclusive events are left in the sample. This number includes mis-identifications where a cosmic track is picked by the selection instead of the neutrino interaction in the same event. The cosmic only sample is used just to illustrate the cut efficiency. The last column Signal:Cosmic only gives an estimate of the $\nu_\mu$ CC events wrt the cosmic only background at each step. For this number, the cosmic background has been scaled as described in section 5.2. . . . .	50
<sup>382</sup> 5.2	Passing rates for Selection I selection applied to on-beam and off-beam data. The numbers in brackets give the passing rate wrt the step before (first percentage) and wrt the generated events (second percentage). Off-beam data has been scaled with a factor 1.23 to normalize to the on-beam data stream. . . . .	50

387	8.1 Comparing passing rates of CNN at different probabilities versus 75 cm track length cut: Numbers are absolute event counts and Cosmic background is not scaled appropriately. The BNB+Cosmic sample contains all events. The numbers in brackets give the passing rate wrt the step before (first percentage) and wrt the generated events (second percentage). In the BNB+Cosmic MC Truth column shows how many true $\nu_\mu$ CC-inclusive events (in FV) are left in the sample. This number includes possible mis-identifications where a cosmic track is picked by the selection instead of the neutrino interaction in the same event. The CNN MC True generated events were scaled wrt the MC True generated events for the 75 cm cut passing rates due to only running over 188,880 generated events versus the 191362 generated events. The last column Signal:Cosmic only gives an estimate of the $\nu_\mu$ CC events wrt the cosmic only background at each step. For this number, the cosmic background has been scaled as described in [?]. Note that these numbers are not a purity, since other backgrounds can't be determined at this step. . .	93
403	8.2 Signal and background event numbers of selection I and selection I with CNN cut estimated from a BNB+Cosmic sample and Cosmic only sample normalized to $5 * 10^{19}$ PoT. The last column gives the fraction of this signal or background type to the total selected events per CNN probability. . . . .	94

*"If they don't give you a seat at the table,  
bring a folding chair."*

— Shirley Chisholm



<sup>410</sup> **Chapter 1**

<sup>411</sup> **Introduction**

<sup>412</sup> This thesis will be a description of work done to further increase efficiency and purity  
<sup>413</sup> of the charged current inclusive cross section measurement using the MicroBooNE  
<sup>414</sup> detector. It will also describe the MicroBooNE detector, what neutrinos are, the  
<sup>415</sup> charged current inclusive cross section measurement and its importance as well as  
<sup>416</sup> convolutional neural networks and how they can be used in  $\mu/\pi$  separation. Chapter  
<sup>417</sup> **2** will talk about the background of neutrinos and the people and detectors that  
<sup>418</sup> discovered neutrinos as well as an in depth history of neutrino oscillation and the  
<sup>419</sup> discovery that neutrinos have mass.

<sup>420</sup> Chapter **3** will discuss the MicroBooNE experiment, specifically, how Liquid  
<sup>421</sup> Argon Time Projection Chambers work, the Light Collection System and the Electronic  
<sup>422</sup> and Readout Trigger systems. This chapter will also describe the Booster Neutrino  
<sup>423</sup> Beam sationed at Fermilab.

<sup>424</sup> Chapter **4** will discuss the work that was done to detect the first neutrinos seen in  
<sup>425</sup> the MicroBooNE detector and the software reconstruction efforts required to create an  
<sup>426</sup> automated neutrino ID filter that was used to find the first neutrinos and then was  
<sup>427</sup> later expanded on to create the charged current inclusive filter that will be discussed  
<sup>428</sup> in chapter **5**

<sup>429</sup> Chapter **6** will give a brief description of what Convolutional Neural Networks are  
<sup>430</sup> and how it will be used for  $\mu/\pi$  separation in this selection. Chapter **7** will discuss  
<sup>431</sup> the hardware frameworks and training methods used to train multiple Convolutional  
<sup>432</sup> Neural Networks for use in the charged current inclusive cross section measurement.  
<sup>433</sup> Chapters **8** and **??** will discuss the results of using Convolutional Neural Networks on  
<sup>434</sup> monte-carlo and data to sift out charged current inclusive neutrino events.



<sup>435</sup> **Chapter 2**

<sup>436</sup> **Neutrinos**

<sup>437</sup> **2.1 What are Neutrinos**

<sup>438</sup> Neutrinos are fundamental particles which help make up the universe. They are also  
<sup>439</sup> one of the least understood. Neutrinos are not affected by the electromagnetic force  
<sup>440</sup> because they do not have electric charge. Neutrinos are affected by a weak sub-atomic  
<sup>441</sup> force of much shorter range than electromagnetism, and are therefore able to pass  
<sup>442</sup> through great distances in matter without much possibility of being affected by it.  
<sup>443</sup> Until the late 1990's, neutrinos were thought to have no mass. Neutrinos are created  
<sup>444</sup> by radioactive decay such as the ones that happen in the sun, in nuclear reactors or  
<sup>445</sup> when cosmic rays hit atoms. There are three types of neutrinos,  $\nu_e$ ,  $\nu_\mu$  and  $\nu_\tau$  which  
<sup>446</sup> correspond to their charged lepton pairs.

<sup>447</sup> As previously stated, neutrinos are very weakly interacting; in fact, neutrinos can  
<sup>448</sup> pass unscathed through a wall of lead several hundred light-years thick. Because  
<sup>449</sup> neutrinos interact so rarely, studying neutrinos requires a massive detector and a  
<sup>450</sup> powerful neutrino source. With that being said, we can only "see" a neutrino when  
<sup>451</sup> they interact in a detector. In a collision, distinct charged particles are produced with  
<sup>452</sup> each type of neutrino because of the weak force. An electron neutrino will create an  
<sup>453</sup> electron, a muon neutrino will create a muon, and a tau neutrino will create a tau. The  
<sup>454</sup> charged lepton track the particle leaves in the detector is how one figures out what  
<sup>455</sup> type of neutrino interaction was "seen". Liquid Argon Time Projection Chambers are  
<sup>456</sup> being used to study neutrinos due to their excellent imaging and particle identification  
<sup>457</sup> capabilities.

## <sup>458</sup> 2.2 History of Neutrinos

<sup>459</sup> The neutrino was first postulated by Wolfgang Pauli in 1930 to explain how to resolve  
<sup>460</sup> the conservation of energy, momentum and angular momentum problem in beta  
<sup>461</sup> decay. Pauli suggested that this missing energy might be carried off, unseen, by a  
<sup>462</sup> neutral particle (he called neutron) which was escaping detection. James Chadwick  
<sup>463</sup> discovered a much heavier nuclear particle in 1932 that he also named neutron, leaving  
<sup>464</sup> two particles with the same name. Enrico Fermi was the first person to coin the term  
<sup>465</sup> neutrino (which means little neutral one in italian) in 1933 to fix this confusion.

<sup>466</sup> Fermi's paper, which was published in 1934, unified Pauli's neutrino with Paul  
<sup>467</sup> Dirac's positron and Werner Heisenberg's neutron-proton model and his theory ac-  
<sup>468</sup> curately explained many experimentally observed results. Wang Ganchang first  
<sup>469</sup> proposed the use of beta capture to experimentally detect neutrinos and in 1956 Clyde  
<sup>470</sup> Cowan and Frederick Reines published their work stating that they had detected the  
<sup>471</sup> neutrino. The experiment called for antineutrinos created in a nuclear reactor by beta  
<sup>472</sup> decay that reacted with protons producing neutrons and positrons:  $\nu_e + p^+ \rightarrow n^0 + e^+$ .  
<sup>473</sup> Once this happens, the positron finds an electron and they annihilate each other and  
<sup>474</sup> the resulting gamma rays are detectable. The neutron is detected by neutron capture  
<sup>475</sup> and the releasing of another gamma ray.

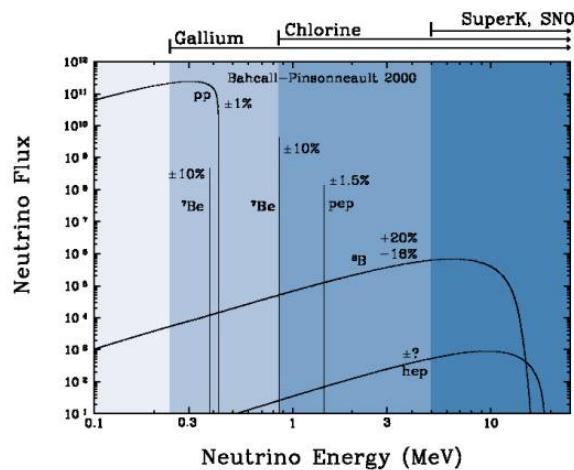
<sup>476</sup> In 1962 Leon M. Lenderman, Melvin Schwartz and Jack Steinberger were the first  
<sup>477</sup> to detect interactions of the muon neutrino [?]. The trio received the 1988 Nobel Prize  
<sup>478</sup> in Physics for their discovery of the muon neutrino. The experiment used a beam  
<sup>479</sup> of energetic protons from Brookhaven's Alternating Gradient Synchrotron (AGS) to  
<sup>480</sup> produce a shower of pions. These pions would then travel 70 ft. toward a 5,000  
<sup>481</sup> ton steel wall. The pions then decayed into muons and neutrinos, the neutrinos  
<sup>482</sup> being the only particle making it through. These neutrinos would enter a neon-filled  
<sup>483</sup> detector producing muon spark trails that were detected and photographed, proving  
<sup>484</sup> the existence of muon neutrinos. The experiment's use of the first ever neutrino beam  
<sup>485</sup> was pioneering work that scientists around the world still use today.

<sup>486</sup> The first detection of the tau neutrino was announced in the summer of 2000 by  
<sup>487</sup> the DONUT collaboration at Fermilab. The scientists used the Tevatron to produce an  
<sup>488</sup> intense neutrino beam. The neutrinos then passed through layers of nuclear emulsion.  
<sup>489</sup> When a neutrino interaction would occur, charged particles would leave visible tracks  
<sup>490</sup> in the emulsion.

### 491 2.2.1 Solar Oscillations and the Solar Neutrino Problem

492 In the late 1960s, many experiments found that the number of electron neutrinos arriving  
493 from the sun was around 1/3 to 1/2 the number predicted by the Standard Solar  
494 Model. This became known as the solar neutrino problem and remained unresolved  
495 for around thirty years. This problem was resolved by the discovery of neutrino  
496 oscillation and mass. [1]

497 The solar neutrino flux derived from Bahcall's Standard Solar Model is shown  
498 in figure 2.1. Nuclear fusion and decay processes produce an abundant amount of  
499 neutrinos. The standard solar model predicts that these reactions produce several  
500 groups of neutrinos, each with differing fluxes and energy spectra. The figure also  
501 shows the ranges of detection of existing solar neutrino experiments in different shades  
502 of blue to illustrate that they sample different portions of the solar neutrino energy  
503 spectrum. Three of these experiments, plus a new one, are discussed below.



**Figure 2.1:** The Standard Solar Model

504 Since neutrinos rarely interact with matter, they pass through the sun and the earth  
505 undetected. About 65-billion neutrinos from the sun stream through every square  
506 centimeter on the Earth every second, yet we are oblivious to their passage in our  
507 every-day lives. [3]

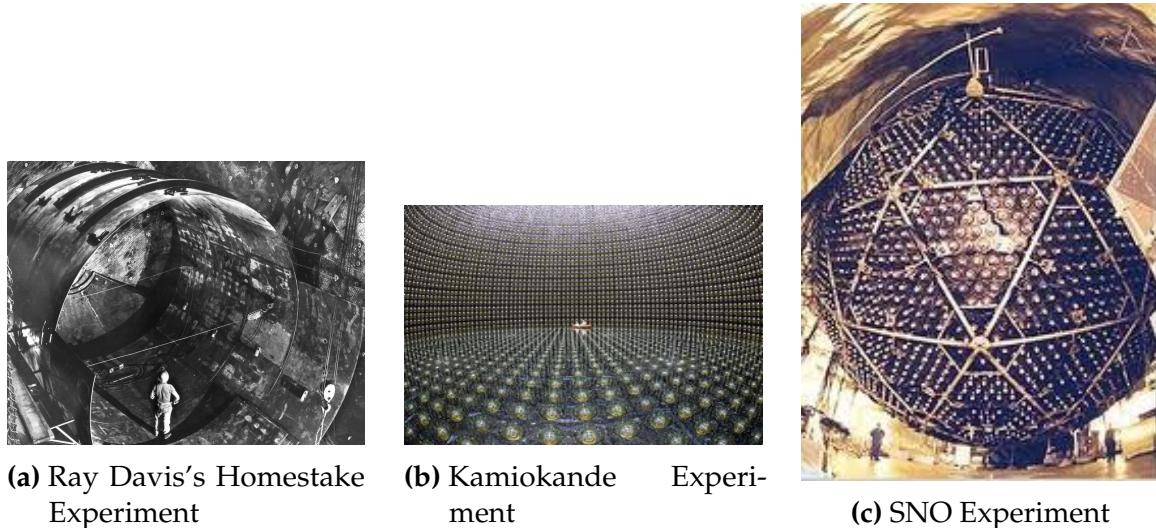
508 The first experiment to detect the effects of neutrino oscillation was the Ray Davis's  
509 Homestake Experiment. The detector was stationed in the Homestake Gold Mine in  
510 Lead, South Dakota. It was 1,478 meters underground and was  $380\text{ m}^3$ . The detector  
511 was filled with perchloroethylene. Perchloroethylene was chosen because of it's high

concentrations of chlorine. When an  $\nu_e$  interacted with chlorine-37 atom, the atom would transform to argon-37 which was then extracted and counted. The neutrino capture reaction is shown in equation 2.1. Davis observed a deficit of about 1/3 the flux of solar neutrinos that was predicted by Bahcall's Standard Solar Model. The unexplained difference between the measured solar neutrino flux and model predictions lead to the Solar Neutrino Problem. [4]



While it is now known that the Homestake Experiment detected neutrinos, some physicist were weary of the results. Conclusive evidence of the Solar Neutrino Problem was provided by the Kamiokande-II experiment, a water Cherenkov detector with a low enough energy threshold to detect neutrinos through neutrino-electron elastic scattering. In the elastic scattering interaction the electrons coming out of the point of reaction strongly point in the direction that the neutrino was traveling, away from the sun. While the neutrinos observed in Kamiokande-II were clearly from the sun, there was still a discrepancy between Kamiokande-II and Homestake; The Kamiokande-II experiment measured about 1/2 the predicted flux, rather than the 1/3 that the Homestake Experiment saw.

The solution to the solar neutrino problem was finally experimentally determined by the Sudbury Neutrino Observatory(SNO). The Ray Davis's Homestake Experiment was only sensitive to electron neutrinos, and the Kamiokande-II Experiment was dominated by the electron neutrino signal. The SNO experiment had the capability to see all three neutrino flavors. Because of this, it was possible to measure the electron neutrinos and total neutrino flux. The experiment demonstrated that the deficit was due to the MSW effect, the conversion of electron neutrinos from their pure flavor state into the second neutrino mass eigenstate as they passed through a resonance due to the changing density of the sun. The resonance is energy dependent, and is visible near 2 MeV. The water cherenkov detectors only detect neutrinos above about 5 MeV, while the radiochemical experiments were sensitive to lower energy (0.8 MeV for chlorine, 0.2 MeV for gallium), and this turned out to be the source of the difference in the observed neutrino rates at the two types of experiments. Figure 2.2 shows Homestake, Kamiokande-II and SNO experiments.



**Figure 2.2:** Solar Neutrino Experiments

## 542 MSW Effect

543 The Mikheyev-Smirnov-Wolfenstein effect is a process which acts to modify neutrino oscillations in matter. The presence of electrons in matter changes the energy  
 544 levels of the mass eigenstates of neutrinos due to charged current coherent forward  
 545 scattering of the electron neutrinos. This coherent forward scattering is similar to  
 546 the electromagnetic process with respect to the refractive index of light in a medium.  
 547 Because of this MSW Effect, neutrinos in vacuum have a different effective mass than  
 548 neutrinos in matter and because neutrino oscillations depend on the squared mass  
 549 difference of the neutrinos, the neutrino oscillations are different in matter than in  
 550 vacuum. This effect is important at the sun where electron neutrinos are produced.  
 551 The neutrinos of high energy leaving the sun are in a vacuum propagation eigenstate  
 552  $\nu_2$  that has a very small overlap with the electron neutrino  $\nu_e = \nu_1 \cos(\theta) + \nu_2 \sin(\theta)$   
 553 seen by the charged current reactions in Kamiokande-II and SNO. The discrepancy of  
 554 the deficit between SNO, Kamiokande-II and Homestake is due to the energy of the  
 555 solar neutrinos. The MSW effect "turns on" at about 2 MeV and at lower energies, this  
 556 MSW effect is negligible. [5]

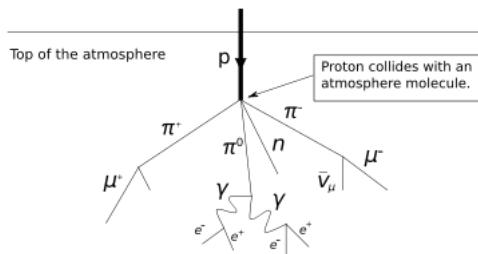
558 **2.2.2 Atmospheric Oscillations and the Atmospheric Neutrino  
559 Anomaly**

560 Atmospheric neutrinos are neutrinos that stem from the decay hadrons coming from  
561 primary cosmic rays. The dominant part of the decay chain is shown in equations 2.2  
562 and 2.3

$$\pi^+ \rightarrow \mu^+ \nu_\mu \mu^+ \rightarrow e^+ \nu_e \bar{\nu}_\mu \quad (2.2)$$

563

$$\pi^- \rightarrow \mu^- \bar{\nu}_\mu \mu^- \rightarrow e^- \bar{\nu}_e \nu_\mu \quad (2.3)$$



**Figure 2.3:** Cosmic Ray Shower

564 Figure 2.3 shows the cosmic ray shower. In general, these neutrinos have energies  
565 from 1 GeV to 100s of GeV and the ratio of  $\nu_\mu$ s to  $\nu_e$ s equals to 2 (see equation 2.4)

$$R = \frac{(\nu_\mu + \bar{\nu}_\mu)}{(\nu_e + \bar{\nu}_e)} \quad (2.4)$$

566 There have been two types of detectors used to study atmospheric neutrinos: Water  
567 Cherenkov detectors and tracking calorimeters. Super-Kamiokande is the detector we  
568 will focus on. These atmospheric detector experiments measure the ratio of  $\nu_\mu$  to  $\nu_e$ .  
569 They also measure the zenith angle distribution of the neutrinos. These experiments  
570 report a double ratio (shown in equation 2.5). This double ratio is the ratio measured  
571 in the detector to the ratio that's expected which is 2. If the double ratio equals to 1, the  
572 data agrees with the prediction. Various measurements from multiple experiments  
573 are shown in figure 2.4. Except for Frejus, all R measurements are less than 1. This  
574 discrepancy between the predicted R and the measured R became known as the

575 Atmospheric Neutrino Anomaly.

$$R = \frac{(N_\mu/N_e)_{DATA}}{(N_\mu/N_e)_{SIM}} \quad (2.5)$$

Experiment	Type of experiment	R
Super-Kamiokande	Water Cerenkov	$0.675 \pm 0.085$
Soudan2	Iron Tracking Calorimeter	$0.69 \pm 0.13$
IMB	Water Cerenkov	$0.54 \pm 0.12$
Kamiokande	Water Cerenkov	$0.60 \pm 0.07$
Frejus	Iron Tracking Calorimeter	$1.0 \pm 0.15$

**Figure 2.4:** Measurements of the double ratio for various atmospheric neutrino experiments

576 Kamiokande-II has the the capability of measuring the direction of the incoming  
577 neutrinos. The expectation of atmospheric neutrino detection is that the flux be  
578 isotropic due to the fact that atmospheric neutrinos can reach the detector from all  
579 directions. Kamiokande-II noticed that muon-like data did not agree well with this  
580 expectation. At low energies approximately half of the  $\nu_\mu$  are missing over the full  
581 range of zenith angles. At high energies the number of  $\nu_\mu$  coming down from above  
582 the detector seems to agree with expectation, but half of the same  $\nu_\mu$  coming up from  
583 below the detector are missing. This anomaly can be easily explained by neutrino  
584 flavor oscillations. Due to the fact that the neutrino travels less distance coming  
585 straight down into the detector (about 15km) than coming up from the bottom of the  
586 detector(13000km) changes the probability of oscillation. The probability of oscillation  
587 for the muon neutrinos coming down into the detector is roughly zero, whereas  
588 for neutrinos coming up, the oscillation probability is  $\sin^2(2\theta)$ . Both the solar and  
589 atmospheric neutrino problems can be explained by neutrino oscillation so its fitting  
590 to derive this phenomenon mathematically. In the next two sections, two flavor and  
591 three flavor neutrino oscillation derivations will be explained.

## 592 2.3 Neutrino Oscillations

593 Neutrino oscillation was first predicted by Bruno Pontecorvo. It describes the phe-  
594 nomenon of a neutrino created with a specific lepton flavor (electron, muon or tau)  
595 that is later measured to have a different flavor. Neutrino oscillation is important  
596 theoretically and experimentally due to the fact that this observation implies that the

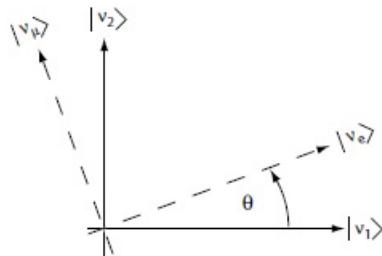
<sup>597</sup> neutrino has a non-zero mass, which is not part of the original Standard Model of  
<sup>598</sup> particle physics. [2]

### <sup>599</sup> 2.3.1 Two Flavor Neutrino Oscillation Formulation

<sup>600</sup> The flavor eigenstates can oscillate between each other because they are composed  
<sup>601</sup> of an add-mixture of mass eigenstates( $\nu_1, \nu_2$ ). Figure 2.5 shows the mass and flavor  
<sup>602</sup> eigenstates rotated by an angle  $\theta$  which is the mixing angle.

<sup>603</sup> In matrix form the wavefunctions are:

$$\begin{pmatrix} \nu_\mu \\ \nu_e \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} * \begin{pmatrix} \nu_1 \\ \nu_2 \end{pmatrix} \quad (2.6)$$



**Figure 2.5:** The flavor eigenstates are rotated by an angle  $\theta$  with respect to the mass eigenstates

<sup>604</sup> Applying the time evolution operator to  $\nu_\mu$ :

$$|\nu_\mu(t)\rangle = -\sin\theta|\nu_1\rangle e^{-i\frac{E_1 t}{\hbar}} + \cos\theta|\nu_2\rangle e^{-i\frac{E_2 t}{\hbar}} \quad (2.7)$$

<sup>605</sup> where  $E_1 = \sqrt{p^2 c^2 + m_1^2 c^4}$  and  $E_2 = \sqrt{p^2 c^2 + m_2^2 c^4}$  and  $p_1 = p_2$ . For the time  
<sup>606</sup> being, let us assume  $\hbar = c = 1$ . With this assumption:  $E_1 = \sqrt{p^2 + m_1^2}$  and  $E_2 =$   
<sup>607</sup>  $\sqrt{p^2 + m_2^2}$ . The next modifications is to assume neutrinos are relativistic:

$$\gamma = \frac{E}{m_o c^2} = \frac{\sqrt{p^2 c^2 + m_o^2 c^4}}{m_o c^2} \gg 1 \quad (2.8)$$

<sub>608</sub> because of this,

$$p \gg m_o \quad (2.9)$$

<sub>609</sub>

$$E = \sqrt{p^2 + m_o^2} = p\sqrt{1 + m_o^2/p^2} \simeq p + \frac{1}{2}\frac{m_o^2}{p} \quad (2.10)$$

<sub>610</sub> where the binomial expansion is used. Now  $E_1$  and  $E_2$  can be written as:

$$E_1 \simeq p + \frac{1}{2}\frac{m_1^2}{p} \text{ and } E_2 \simeq p + \frac{1}{2}\frac{m_2^2}{p} \quad (2.11)$$

<sub>611</sub> Now applying all these assumptions back into equation 2.7 gives us:

$$|\nu_\mu(t)\rangle = -\sin\theta|\nu_1\rangle e^{-i(p+\frac{1}{2}\frac{m_1^2}{p})t} + \cos\theta|\nu_2\rangle e^{-i(p+\frac{1}{2}\frac{m_2^2}{p})t} \quad (2.12)$$

<sub>612</sub>

$$|\nu_\mu(t)\rangle = e^{-i(p+\frac{1}{2}\frac{m_1^2-m_2^2}{p})t} (-\sin\theta|\nu_1\rangle + \cos\theta|\nu_2\rangle) \quad (2.13)$$

<sub>613</sub> Substituting  $\Delta m^2 = m_1^2 - m_2^2$  and  $t = \frac{x}{c} = x$  and  $e^{-iz} = e^{-i(p+\frac{1}{2}\frac{m_1^2}{p})t}$  gives us:

$$|\nu_\mu(t)\rangle = e^{-iz} \left( -\sin\theta|\nu_1\rangle + \cos\theta|\nu_2\rangle e^{+ix(\frac{1}{2}\frac{\Delta m^2}{p})} \right) \quad (2.14)$$

<sub>614</sub> Finding the Probability for a  $\nu_\mu \rightarrow \nu_e$ :

$$P(\nu_\mu \rightarrow \nu_e) = |\langle \nu_e | \nu_\mu(t) \rangle|^2 \quad (2.15)$$

<sub>615</sub> Remembering that  $\langle \nu_i | \nu_j \rangle = \delta_{ij}$

$$\langle \nu_e | \nu_\mu(t) \rangle = e^{-iz} \left( -\sin\theta \cos\theta + \sin\theta \cos\theta e^{\frac{i\Delta m^2 x}{p}} \right) \quad (2.16)$$

616 Taking the absolute value squared gives us:

$$P(\nu_\mu \rightarrow \nu_e) = |<\nu_e|\nu_\mu(t)>|^2 = e^{+iz}e^{-iz}\sin^2\theta\cos^2\theta \left(-1 + e^{\frac{i\Delta m^2 x}{p}}\right) \left(-1 + e^{\frac{i\Delta m^2 x}{p}}\right) \quad (2.17)$$

617 Since the neutrino is relativistic we can set  $p = E_\nu$  and change  $x = L$ . Also 618 recognizing the trigonometric relation  $(1 - \cos 2\theta)/2 = \sin^2\theta$  the above equation 619 becomes:

$$P(\nu_\mu \rightarrow \nu_e) = \sin^2 2\theta \sin^2 \left( \frac{\Delta m^2 L}{4E_\nu} \right) \quad (2.18)$$

620 All that's left to do now is re-introduce  $\hbar$  and  $c$  doing this we get:

$$P_{\nu_\mu \rightarrow \nu_e}(L, E) = \sin^2 2\theta \sin^2 \left( 1.27 \Delta m^2 \frac{L}{E_\nu} \right) \quad (2.19)$$

621 This equations has three important variables.

- 622 • The angle  $\theta$ : This angle, as mentioned before, is called the mixing angle. It defines 623 the difference between the flavor and the mass eigenstates. When  $\theta = 0$  the mass 624 and flavor eigenstates are identical and no oscillations occur.
- 625 • The mass squared difference,  $\Delta m^2$ : Again  $\Delta m^2 = m_1^2 - m_2^2$ . The reason this is an 626 important variable is because it implies that for neutrinos to oscillate, neutrinos 627 must have mass. Furthermore, the mass squared difference also tells us that the 628 neutrino mass eigenstates must be different.
- 629 • L/E: This is the variable that is of most interest to experimental physicists due to 630 the fact that it is the variable that we set. L is the distance between the source and 631 detector and E is the energy of the neutrino. For a given  $\Delta m^2$ , the probability of 632 oscillation changes with respect to L/E.

### 633 2.3.2 Three Flavor Neutrino Oscillation Formulation

634 Seeing the quantum mechanics involved in deriving the probability of a two flavor 635 neutrino oscillation, it is now possible to formulate the three flavor neutrino oscillation.

<sup>636</sup> The three flavor neutrino oscillation formulation begins similarly to the two flavor,  
<sup>637</sup> but there is the Pontecorvo-Maki-Nakagawa-Sakata matrix (PMNS) instead of the 2X2  
<sup>638</sup> matrix in the previous section. The PMNS matrix is show below:

$$\begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix} * \begin{pmatrix} e^{i\alpha_1/2} & 0 & 0 \\ 0 & e^{i\alpha_2/2} & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (2.20)$$

<sup>639</sup> where  $c_{ij} = \cos\theta_{ij}$  and  $s_{ij} = \sin\theta_{ij}$

<sup>640</sup> Following the same steps as before we get:

$$P_{\alpha \rightarrow \beta} = \delta_{\alpha\beta} - 4\sum Re(U_{\alpha i}^* U_{\beta i} U_{\alpha j} U_{\beta j}^*) \sin^2 \left( \frac{\Delta m_{ij}^2 L}{4E} \right) - 2\sum Im(U_{\alpha i}^* U_{\beta i} U_{\alpha j} U_{\beta j}^*) \sin \left( \frac{\Delta m_{ij}^2 L}{2E} \right) \quad (2.21)$$

<sup>641</sup> The main things to notice here are  $\delta_{ij}$  which is the CP violating term and has not  
<sup>642</sup> been measured yet, and  $\theta_{13}$  which has just been measured. CP violation is a violation  
<sup>643</sup> of the postulated CP-symmetry. CP-symmetry states that the laws of physics should  
<sup>644</sup> be the same if a particle were to be exchanged with its antiparticle and then if the left  
<sup>645</sup> hand side of a decay were switched with the right hand side.

### <sup>646</sup> 2.3.3 Reactor Oscillation

<sup>647</sup> Many experiments have searched for oscillation of electron anti-neutrinos produced at  
<sup>648</sup> nuclear reactors. Such oscillations give the value of the parameter  $\theta_{13}$ . The KamLAND  
<sup>649</sup> experiment, started in 2002, has made a high precision observation of reactor neutrino  
<sup>650</sup> oscillation. Neutrinos produced in nuclear reactors have energies similar to solar  
<sup>651</sup> neutrinos, a few MeV. The baselines of these experiments have ranged from tens  
<sup>652</sup> of meters to over 100 km. On 8 March 2012, the Daya Bay team announced a  $5.2\sigma$   
<sup>653</sup> discovery that  $\theta_{13} \neq 0$ .



654 **Chapter 3**

655 **The MicroBooNE Experiment**

656 The purpose of this chapter is to discuss and understand the details of the MicroBooNE  
657 detector. A thorough understanding of MicroBooNE and the technology behind liquid  
658 argon time projection chambers is important for understanding results as well as  
659 understanding how images were made for use in deep learning efforts that will be  
660 outlined in later chapters.

661 **3.1 Liquid argon time projection chambers**

662 Liquid Argon Time Projection Chambers (LArTPCs) are an exciting detector technol-  
663 ogy that provide excellent imaging and particle identification, and are now being  
664 used to study neutrinos. The Time Projection Chamber (TPC) was first invented by  
665 Nygren in 1974 [?] and the proposal for a LArTPC for neutrino physics was made  
666 by Rubbia [?] in 1977 with the ICARUS collaboration implementing this concept [?].  
667 A LArTPC is a three-dimensional imaging detector that uses planes of wires at the  
668 edge of an active volume to read out an interaction. When a neutrino interacts with an  
669 argon atom, the charged particles that are produced ionize the LAr as they travel away  
670 from the interaction. By placing a uniform electric field throughout the LAr volume,  
671 the ionization is made to drift towards a set of anode planes, which consist of wires  
672 spaced very closely together collecting the ionized charge, which is subsequently read  
673 out by electronics connected to the anode wires. The collected ionization creates a  
674 spatial image of what happened in the detector on each anode plane. The position  
675 resolution of the interaction along the beam direction (perpendicular to drift direction)  
676 relies on the wire pitch, while the resolution in drift direction is dependent on the

timing resolution of the electronics used and the longitudinal diffusion in the volume.  
 The drift time of the ionization relative to the time of the original signal allows the  
 signal to be projected back along the drift coordinate, hence the name LArTPC. Having  
 very small distances between each wire within an anode plane allows for very fine  
 granularity and detail to be captured, and having multiple wire planes at different  
 angles provides independent two-dimensional views that can be combined into a  
 three dimensional picture of the interaction. Once the charge signal is created on the  
 anode planes, software analysis packages identify particles in the detector by using  
 deposited energy on the wires along their track length. The 30 year development of the  
 ICARUS detector has led to LArTPCs being used as cosmic ray [?], solar neutrino [?]  
 and accelerator neutrino [?] detectors. The ArgoNeuT experiment at Fermilab was  
 the first United States based liquid argon neutrino program that has since produced  
 short-baseline  $\nu - Ar$  cross-section measurements in the NUMI beamline [?]. The  
 MicroBooNE experiment is the second experiment in the US based LArTPC neutrino  
 program and will be discussed thoroughly in the next sections. The next phases of  
 the liquid argon neutrino program are under way and are the Fermilab Short Base-  
 line Neutrino (SBN) program [?] and the Deep Underground Neutrino Experiment  
 (DUNE) [6]. The SBN program will include three LArTPC detectors, including the  
 MicroBooNE detector, on the Booster Neutrino Beam (BNB) to do multiple-baseline  
 oscillation measurements. The detector closest to the beam will be the 40 ton Short  
 Baseline Neutrino Detector (SBND) [?] at 150 m and the detector furthest is the 600 ton  
 ICARUS T600 [?] detector positioned at 600 m. The DUNE collaboration will deliver  
 a 30 GeV neutrino beam 1300 km from Fermilab to a 34 kiloton LArTPC detector  
 at Homestake, SD. DUNE will study the leptonic CP phase,  $\delta_{cp}$ , as well as measure  
 neutrino and antineutrino oscillations.

## 702 3.2 The MicroBooNE Time Projection Chamber

703 MicroBooNE (Micro Booster Neutrino Experiment) is a 89 ton active volume (180 ton  
 704 total mass) LArTPC which is then inserted into a cylindrical cryostat on axis of the  
 705 Booster Neutrino Beam (BNB) stationed at Fermilab in Batavia, Illinois. Understanding  
 706 LArTPC technology and detector physics is necessary to build a LArTPC the size of  
 707 DUNE, and MicroBooNE has made many advances in developing this technology [7]  
 708 [8].

MicroBooNE's Time Projection Chamber (TPC) is 10.3 m long (beamline direction), 2.3 m high and 2.5 m wide (which corresponds to the drift distance). The TPC is shown in figure ?? . MicroBooNE is the largest LArTPC currently running in the world [9]. This LArTPC has 3 wire planes: 1 plane that collects the ionization in the wires and is  $0^\circ$  to the vertical with 3456 wires spaced 3 mm apart, and 2 planes where the ionization drifts passed and induces a signal at  $\pm 60^\circ$  to the vertical each with 2400 wires also spaced 3 mm apart. Each plane has a spacing also of 3 mm from eachother. The first two planes are the induction planes and the last is the collection. The 270 V/cm electric field of the TPC is created using 64 stainless steel tubes shaped into rectangles around the TPC and held in place by G10 to form a field cage. The cathode is charged at a high voltage of -70 kV and this voltage is stepped down across the field cage tubes using a voltage divider chain with an equivalent resistance of  $240 \text{ M}\Omega$  between the tubes. The field cage tubes are separated by 4 cm from center to center. The electron drift distance is 2.5 m in the x direction with a drift time of 2.3 ms. Maintaining high charge yield is done by continuously recirculating and purifying the argon. The purity is monitored using MicroBooNE's light collection system. Another use of the light collection system is initial timing and drift coordinate of the interaction.

MicroBooNE's light collection system is a crucial part for 3D reconstruction of all particle interactions in the LArTPC. The initial interaction time,  $t_0$ , and initial drift coordinate,  $x_0$ , are not known from the TPC alone. For beam events, the accelerator clock is used to determine  $t_0$  of the interaction and the  $x_0$  can be inferred using drift time. Non-beam events, however, do not have this capability, which is why scintillation light from an interaction is used. The  $\nu - Ar$  interaction produces scintillation light which is collected by photomultiplier tubes (PMTs) which allows the exact time,  $t_0$  of the neutrino interaction to be determined. The scintillation light created propagates within nanoseconds to the light collection system compared to the milliseconds it takes the ionized electrons from the interaction to reach the anode wire planes. Therefore we can precisely know where along the drift direction the particle interaction first took place. The scintillation light is also localized, so combining the PMT information with the wire plane information allows for cosmic background rejection happening outside the beam timing window.

The light collection system is made up of 32 Hamamatsu R5912-02mod cryogenic PMTs with a diameter of 8-inches. The PMTs are located behind the 3 wire anode planes and provides 0.85% photocathode coverage. Each PMT has an acrylic plate mounted in front of it that is coated with a wave-length shifting material called TPB.

<sup>744</sup> The acrylic plates take in the scintillation light, at 128 nm, and re-emits it visible  
<sup>745</sup> wavelengths visible to the PMTs, with a peak at 425 nm.

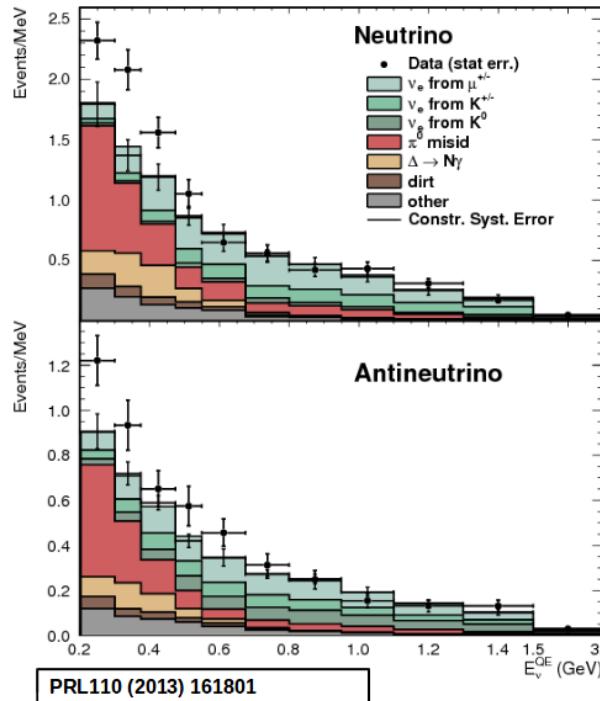
<sup>746</sup> Both the light collection system and the TPC create analog signal that is read out and  
<sup>747</sup> digitized by the electronics system. The process requires amplification and shaping of  
<sup>748</sup> the signal which then goes to the data acquisition (DAQ) software for writing of the  
<sup>749</sup> digitized data to disk. The anode plane wires are connected to detector specific circuit  
<sup>750</sup> boards (ASICS) that are submerged and operate inside the liquid argon volume. These  
<sup>751</sup> ASICS send amplified signal to 11 feed-throughs where further amplification of the  
<sup>752</sup> signal happens outside the cryostat. The signal is received by custom LArTPC readout  
<sup>753</sup> modules distributed over nine readout crates which do the digitization. The TPC wires  
<sup>754</sup> are digitized at 16 MHz then downsampled to 2 MHz. The TPC system reads out 4  
<sup>755</sup> frames of wire signal data per event, 1 frame before a trigger and 2 frames after the  
<sup>756</sup> triggered frame. The four frames allows for identification of a neutrino interaction as  
<sup>757</sup> well as cosmic background rejection. The process of digitization is similar for the light  
<sup>758</sup> collection system. Each PMT signal undergoes a shaping with a 60 ns peaking time  
<sup>759</sup> for digitization of multiple samples. The digitization occurs at 64 MHz but are not  
<sup>760</sup> read out continuously during the TPC readout time. Only shaped PMT signal samples  
<sup>761</sup> above a small threshold are read out and saved. Both the TPC and PMT readouts are  
<sup>762</sup> initiated via triggers on a separate trigger board located in a warm electronics crate.  
<sup>763</sup> The timing trigger is created by a timing signal from the BNB accelerator which is  
<sup>764</sup> shaped and sent to the trigger board. The PMT trigger is generated when the PMT  
<sup>765</sup> signal multiplicity is greater than 1 and the summed PMT pulse-height is more than 2  
<sup>766</sup> photo-electrons summed up over all PMT channels. When the trigger board gets both  
<sup>767</sup> a timing trigger and a PMT trigger in coincidence, at BNB trigger is then generated by  
<sup>768</sup> the board. This signal is then passed to all readout crates initiating the readout of data.  
<sup>769</sup> The data is then sent to the DAQ software which then saves the data to disk into one  
<sup>770</sup> event memory.

### <sup>771</sup> 3.3 MicroBooNE's Physics Goals

#### <sup>772</sup> 3.3.1 The low-energy excess

<sup>773</sup> The primary goal of the MicroBooNE experiment is to study and investigate the low-  
<sup>774</sup> energy excess seen in MiniBooNE 3.1. MicroBooNE has the capability of confirming or

denying this excess as electrons or photons due to the detector being in the same beam, having a similar baseline, and lastly the detector being able to clearly distinguish between electrons and photons. LArTPCs use the topology of events as well as energy loss near the vertex to differentiate between single  $e^-$  tracks and photon-induced induced pair production  $\gamma \rightarrow e^+ e^-$ , which wasn't possible in MiniBooNE, a Cherenkov detector. This technique has been shown in the ArogoNeuT detector [?] and a side by side comparison of both event types in a LArTPC can be seen in figure ???. An excess in electrons would point towards new oscillation physics beyond the standard model, while photons would be within the standard model. MicroBooNE will observe a  $4-5\sigma$  signal.



**Figure 3.1:** Low Energy excess seen in MiniBooNE

### 3.3.2 Cross sections

MicroBooNE's neutrino cross-section program will be the first  $\nu - Ar$  cross-section in the 1 GeV energy range and one of only a few cross-section measurements of  $\nu - Ar$  in the world. MicroBooNE is also the first liquid argon detector to collect the highest statistics sample of neutrino interactions. Investigating final-state-interactions in the 1GeV energy range provides information about short range nuclear correlations that affect the interpretations of neutrino oscillation experiment data.

One of the cross-section measurements MicroBooNE can make is an inclusive charged-current cross-section measurement (referred to as CC-inclusive). CC-inclusive events consist of a neutrino exchanging a  $W^\pm$  boson with an argon atom, producing a charged lepton and any number of other final state particles. In MicroBooNE's case, a CC-inclusive event will mostly have a defining muon track coming out of the vertex due to our neutrinos being predominately  $\nu_\mu$ s. A cross-section measurement is the energy dependent probability of  $\nu - Ar$  interaction in the detector. Cross-sections however are independent of the intensity or focus of the particle beam so they can be compared among different experiments. A background for a CC-inclusive cross-section measurement are the neutral-current events that contain a pion. It is possible to have a neutral current interaction with a  $\pi + p$  event signature that looks like a charged current  $\mu + p$  event. Reconstruction tools implemented to date don't efficiently separate muons from pions. A common way to separate these two particles species is to implement a track length cut. On average, muons tend to have longer track lengths in LArTPCs so by requiring that the hypothesized lepton be above a threshold track length, it is possible to increase signal to background.

### 3.3.3 Liquid argon detector development

The last physics goal for the MicroBooNE collaboration is to provide important information regarding LArTPC technology. Being the first in large scare LArTPCs in the US, MicroBooNE will be able to provide improvements to High Voltage (HV) distribution, Noise Characterization [?], and Michel Electron Reconstruction [8].

## 3.4 The Booster Neutrino Beam

The MicroBooNE detector is stationed at Fermi National Accelerator Laboratory (FNAL) where it receives neutrinos from both the Booster Neutrino Beam (BNB) and Neutrinos from the Main Injector (NuMI) beams. MicroBooNE is on-axis for the BNB and off-axis by 135 mrad for NuMI. For the purpose of this analysis, only data from the BNB was used. This section will discuss how neutrinos are created using the BNB. How these neutrinos are produced as well as their flux through the MicroBooNE detector is necessary for any analysis because of the systematic uncertainties the beam

<sup>821</sup> introduces to a measurement. An aerial view of fermilab as well as the BNB is shown  
<sup>822</sup> in figure 3.2

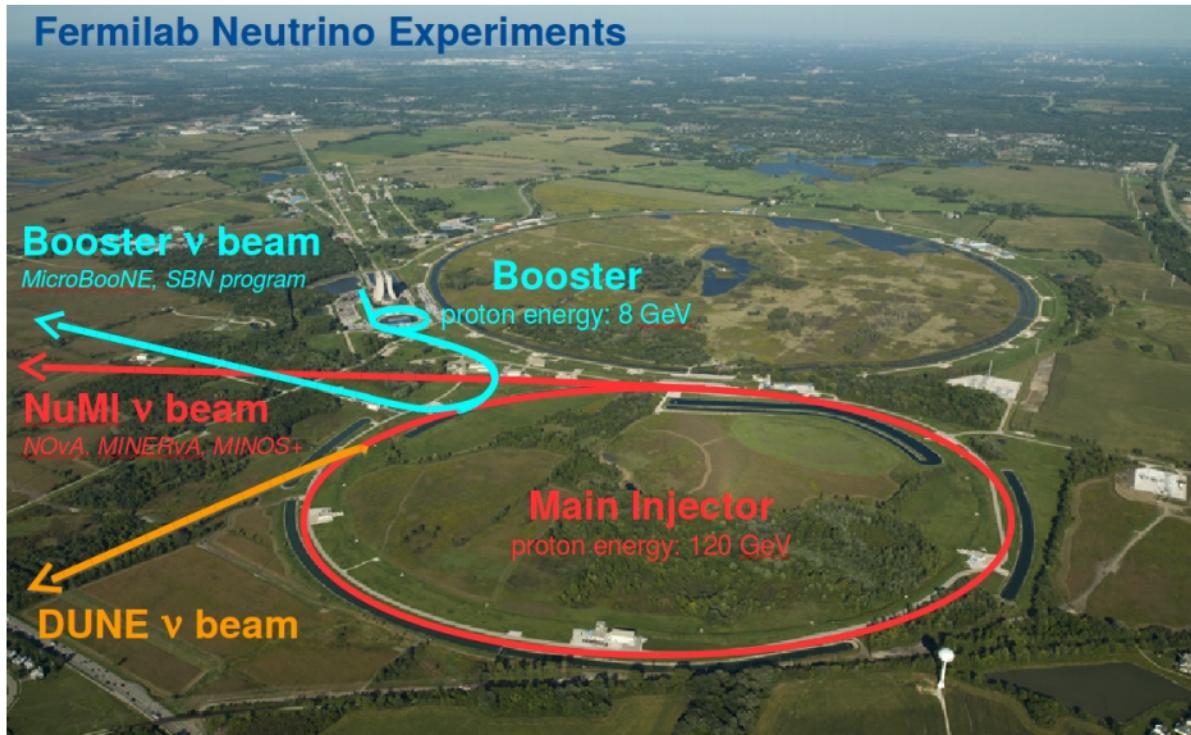


Figure 3.2: Arial view of the Main Injector and the Booster Neutrino Beam at Fermilab

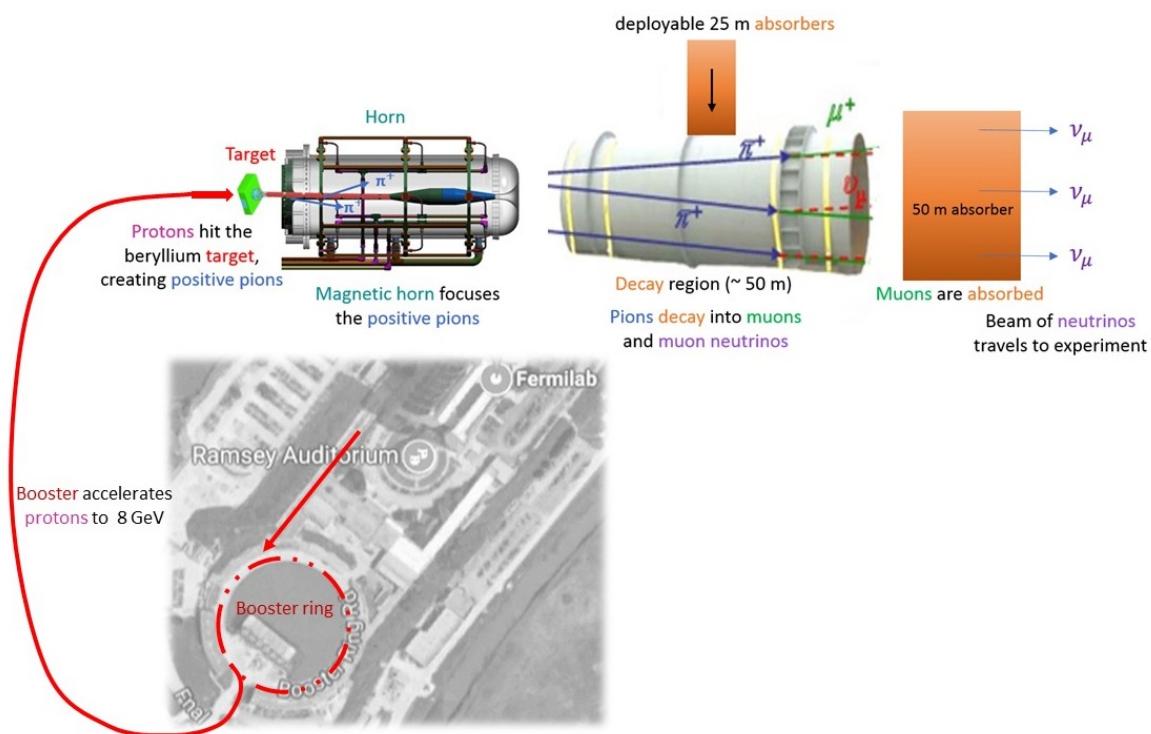
### <sup>823</sup> 3.4.1 Creating the Booster Neutrino Beam

<sup>824</sup> The BNB is a very pure  $\nu_\mu$  beam, with only 0.6% contamination from  $\nu_e$ s. The energy  
<sup>825</sup> also peaks around 700 MeV which is desired based on the probability of oscillation  
<sup>826</sup> equation which depends on the the value of  $L/E$ , where  $L$  is the distance of the  
<sup>827</sup> detector from the neutrino beam and  $E$  is the energy of the neutrino beam.  $L/E$  was  
<sup>828</sup> chosen to increase the probability of seeing neutrino oscillations in the MiniBooNE  
<sup>829</sup> Low Energy Excess (LEE) range based on the probability of oscillation equation, which  
<sup>830</sup> is  $P_{\nu_\mu \rightarrow \nu_e}(L, E) = \sin^2 2\theta \sin^2 \left( 1.27 \Delta m^2 \frac{L}{E} \right)$ . The BNB collides 8.9 GeV/c momentum  
<sup>831</sup> protons from the FNAL booster synchrotron into a beryllium target which produces a  
<sup>832</sup> high flux of neutrinos. The protons originate from  $H^2$  gas molecules that are turned  
<sup>833</sup> into  $H^-$  ions by a Cockcroft-Walton generator shown in figure ???. The  $H^-$  initially are  
<sup>834</sup> accelerated to 1MeV kinetic energy and are then passed to a linear accelerator using  
<sup>835</sup> alternating electromagnetic fields to increase their energy to 400MeV. The ions are  
<sup>836</sup> stripped of electrons by passing them through a carbon foil. The protons are bunched

837 into beam spills which contain  $4 * 10^{12}$  protons in a  $1.6 \mu\text{s}$  time window per spill. It's  
838 at this point that the protons are directed towards the beryllium target. The amount  
839 of protons directed towards the target (POT) is measured by two toroids upstream of  
840 the target with an error of 2%. Beam intensity, timing, width, position, and direction  
841 are monitored by beam position monitors, multi-wire chamber and resistive monitors.  
842 The beryllium target is 71.1 cm long, 1.7 proton interaction lengths, and is 0.51 cm in  
843 radius. The target is located inside a larger focusing electromagnet called the horn.  
844 The horn is an aluminum alloy pulsed toroidal electromagnet. The pulsed current  
845 peaks at 170 kA with a time-width of  $143 \mu\text{s}$  which coincides with the protons arriving  
846 on the target. The current flows from the inner conductor to the outer conductor  
847 with a maximum magnetic field of 1.5 Tesla. The magnetic field focuses the charged  
848 secondary particles produced by the p-Be interactions. The direction of current can be  
849 switched to change the polarity of the secondary particles being focused creating a  
850 beam of either primarily neutrinos, with positively charged secondary particles, or  
851 antineutrinos.

852 Further down the beamline is a concrete collimator which absorbs particles not  
853 necessary to the neutrino flux. The collimator is 214 cm long and 30 cm in radius.  
854 After the collimator comes a 45 meter long, 1 meter radius, air-filled cylindrical decay  
855 region which then ends in a beam-stop made of steel and concrete. The beam-stop  
856 contains an array of gas proportional counters to detect muons. The BNB is shown in  
857 figure 3.3.

858 **3.5 Event Reconstruction**



**Figure 3.3:** Aerial view of the Main Injector and the Booster Neutrino Beam at Fermilab

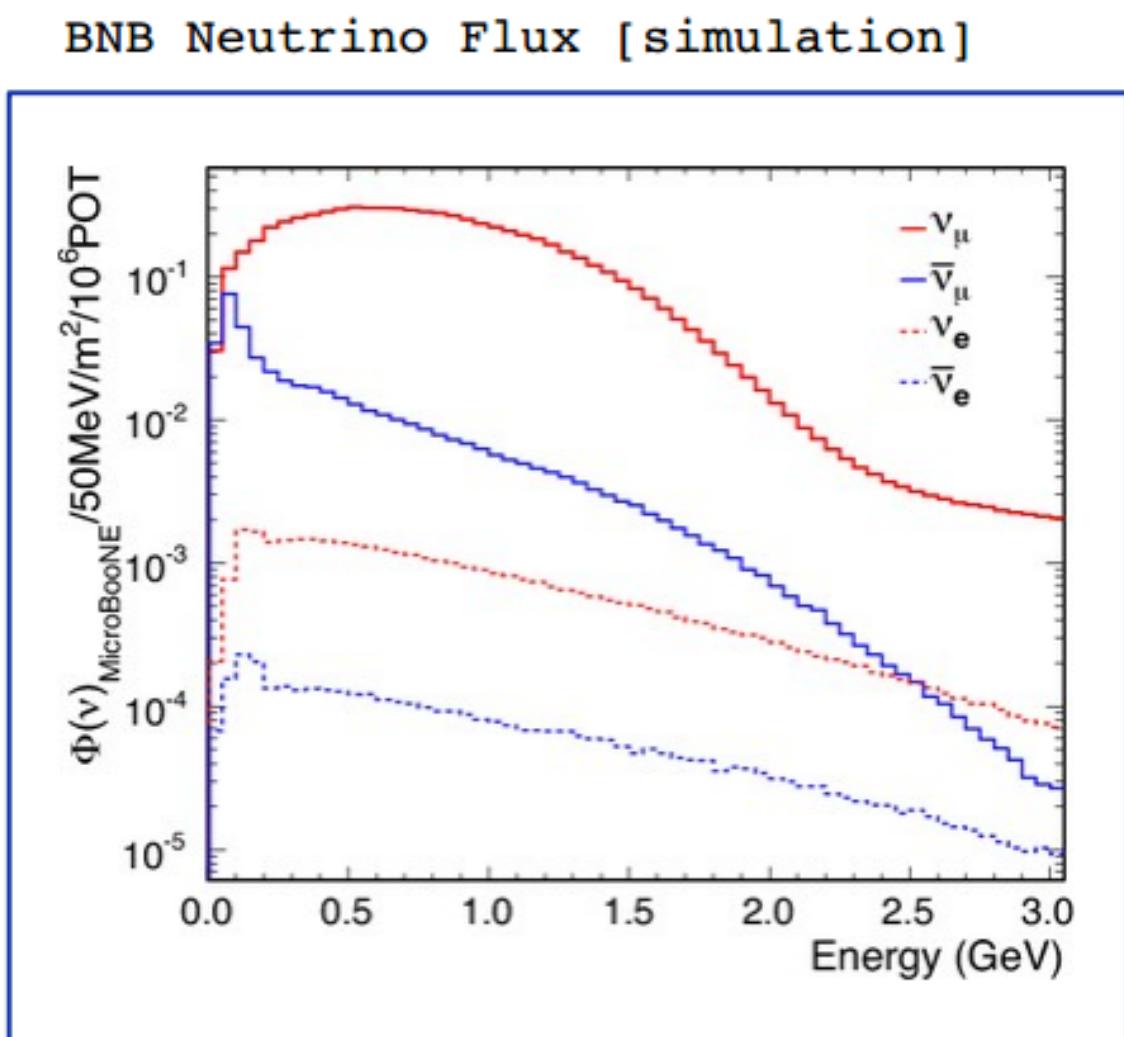


Figure 3.4: Energy spectrum of the Booster Neutrino Beam at Fermi National Laboratories

859 **Chapter 4**

860 **Neutrino Identification: Finding  
861 MicroBooNE's first Neutrinos**

862 The goal of the Neutrino Identification analysis was to positively identify BNB neutrino  
863 interactions in the MicroBooNE detector collected during the first days of running.  
864 Neutrino event candidates were identified in part by using a cut on detected flash of  
865 scintillation light during the  $1.6 \mu\text{s}$  beam-spill length of the BNB as well as identifying  
866 reconstructed object from the TPC that are neutrino like. After this selection, 2D  
867 and 3D event displays were used for verification of the selection performance. This  
868 selection was targeted to reduce the ratio of neutrino events to cosmic-only events from  
869 the initial 1 neutrino to 675 cosmics to a ratio of 1 to 0.5 or better which is equivalent to  
870 a background reduction by a factor of 1000 or more. These selected events were used  
871 for MicroBooNE's public displays of neutrino interactions. A clearly visible neutrino  
872 interaction with an identifiable vertex and at least 2 tracks originating from the vertex  
873 was what the analysis focused on. This analysis wasn't optimized for high purity  
874 or efficiency, but rather for very distinguishable neutrino interactions that could be  
875 identified by the public.

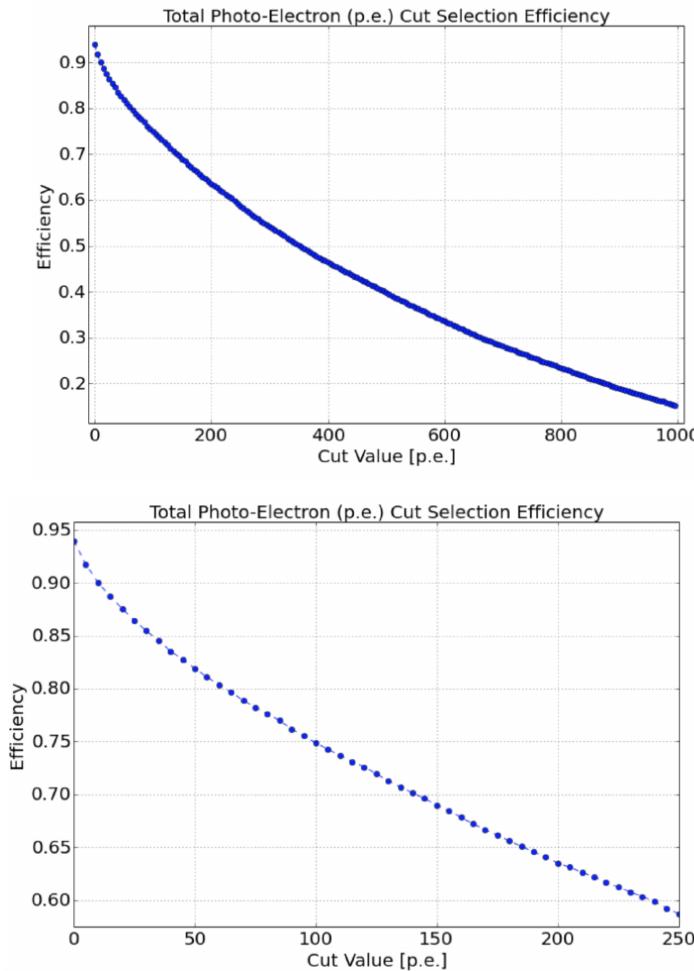
876 **4.1 Flash Finding**

877 Flash finding is the first step used in finding neutrino interactions. This section will  
878 detail how optical information is reconstructed as well as analysis scripts and event  
879 filters were used.

---

880

### 4.1.1 Flash Reconstruction

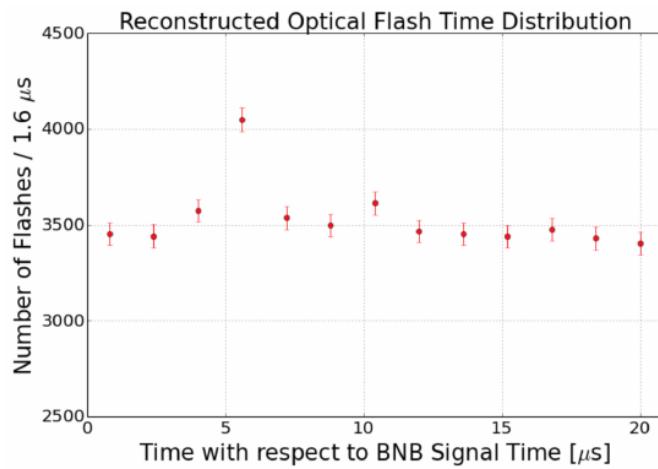


**Figure 4.1:** Efficiency for selecting beam events as a function of minimum total PE cut for all PE cuts as well as zoomed into interesting region.

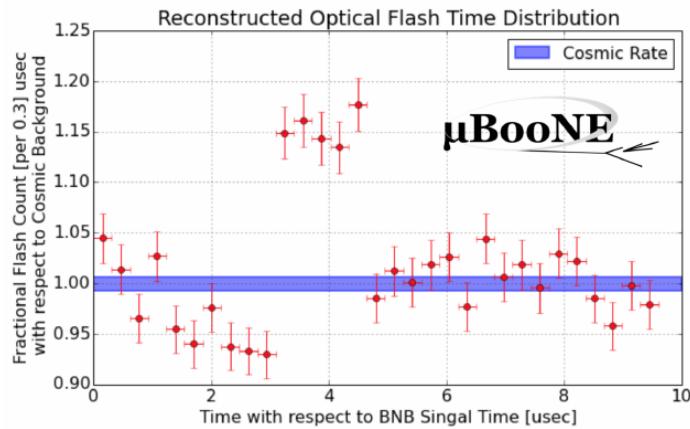
881        A flash is described as a collection of light seen at the same time within the detector.  
 882        They are then reconstructed by identifying signal from the PMTs above a specific  
 883        photoelectron (PE) threshold. These signals are called optical hits. Optical hits from  
 884        all the PMTs are then accumulated into  $1\text{ }\mu\text{s}$  bins of time. If a specific bin is above a  
 885        set PE threshold, then the optical hits that overlap in time are the labeled as the hits  
 886        from the flash. All flash reconstructed properties like average time and x/y positions  
 887        are then found via the flash labeled optical hits. The total size of the flash is found by  
 888        summing up the total number of photoelectrons from all PMTs. Neutrino interactions  
 889        and cosmic muons will have a larger flash size compared to noise and other low-energy  
 890        backgrounds, therefore a total PE cut is used to reject these backgrounds. A total PE

<sup>891</sup> cut of 50 PE was deemed sufficient for this analysis. Figure 4.1 show the total PE  
<sup>892</sup> versus the selection efficency of selecting neutrino beam events.

### <sup>893</sup> 4.1.2 Beam Timing



(a) Predicted distribution of flash times with respect to trigger time for 1 day of data taking at nominal rate and intensity



(b) Measured distribution of flash times with a 50 PE threshold cut, with respect to trigger time. Shown as a ratio to the expected cosmic rate from off-beam data. A clear excess from neutrinos is visible between 3- 5  $\mu$ s after the trigger time.

<sup>894</sup> It is necessary to get the specific time from flashes if one uses flashes to filter out  
<sup>895</sup> neutrino interactions coincident with the neutrino beam spill period and background.  
<sup>896</sup> Before a filter can be applied, an understanding of the timing of the trigger and PMT

897 readout with respect to the arrival of neutrinos from the BNB. To do this, a  $1.6 \mu\text{s}$   
898 window near the expected beamtime was created and verified by finding that the  
899 number of flashes was significantly above the cosmic-ray background flashes. Beam  
900 data during the first week of running, October 16th 2016 through October 22nd 2016  
901 and were used for a timing measurement. The total POT uses corresponds to roughly  
902 24 hours of data taking at nominal intensity ( $4 \times 10^{12} \text{ ppp}$ ) and a 5 Hz repetition rate.  
903 Figure 4.2a shows size of the expected neutrino signal in time using Monte Carlo  
904 predictions and figure 4.2b shows the neutrino signal in data. The intensity in data is  
905 lower, however there can still be seen a significant excess above data.

906 **4.1.3 Event Rates**

907 Applying a 50 PE threshold cut inside a  $1.6 \mu\text{s}$  window reduces the cosmic-ray passing  
908 rate to 0.8%. With a 5 Hz beam rate, this corresponds to 135 cosmics passing per  
909 hour. The neutrino passing rate for this filter is about 22 events per hour. To further  
910 increase the neutrino to cosmic ratio, TPC topology cuts were implemented and will  
911 be discussed in the following section.

912 **4.2 TPC Topology Selection**

913 In order to further reduce the background of cosmic events, two independent selection  
914 streams using TPC wire data reconstruction was implemented. The first using 2D  
915 reconstructed clusters, and the second using 3D reconstructed tracks. Both streams  
916 look for neutrino interactions in the active TPC volume which are identifiable by two  
917 or more tracks originating from the same vertex.

918 Both 2D and 3D channels were optimized using monte carlo simulation which  
919 used a 128 kV cathode voltage. Passing rates were calculated using a 0.008 efficiency  
920 factor for cosmic events passing to simulate the flash finding described in section 4.1.  
921 This efficiency factor was an overestimation and was just used to get a general feel of  
922 what signal and background rates we would actually see in data.

### 923 4.2.1 Cosmic Tagging

924 The first step in TPC selection was based on the geometry of cosmic tracks in an event.  
925 The cosmic ray muon geometry tagger runs on 3D tracks and assigns a score to each  
926 reconstructed track on the likeliness of the track originating from a cosmic. The cosmic  
927 scores are detailed below:

- 928 • 1: The track is tagged as entering or exiting the TPC
- 929 • 0.95: The track is a delta ray associated with a tagged track
- 930 • 0.5: The track is either entering or exiting, but not both
- 931 • 0.4: The track is entering or exiting through the Z boundary
- 932 • 0: The track isn't tagged

933 Clusters are assigned either a 0 or 1, 1 being a cosmic. In simulation, 90% of cosmics  
934 are tagged as cosmics. These tracks are no longer considered when looking for a  
935 neutrino topology. Requiring that the tracks be contained in turn affects the neutrino  
936 efficiency by 20%. The algorithm checks that each track is contained within a boundary  
937 region of 10 cm from all sides of the TPC. This boundary region was optimized via  
938 handscanning of experimental data.

939 As can be expected, cosmic tagging is more efficient in the 3D channel (tracks) than  
940 the 2D channel (clusters) because the reconstructed tracks can use the full 3D position  
941 information of the entering and exiting points while the 2D channel mainly use the  
942 reconstructed x position of the cluster which is associated to timing.

943 Cosmic tagging uses timing information to reject tracks and clusters that are outside  
944 of drift window. The drift window for 128 kV is 1.6  $\mu$ s while for 70 kV, the actual  
945 voltage MicroBooNE is running at, is 2.3  $\mu$ s. Due to this variation between simulation  
946 and data, we expect to see  $2.3/1.6 = 1.44$  times more cosmic induced tracks or clusters  
947 in the drift window.

### 948 4.2.2 2D Cluster Selection

949 This selection was spearheaded by myself and Katherine Woodruff. After looking at  
950 experimental cosmics data, 2D clustering performs well, while 3D track reconstruction  
951 is affected by more variations in simulation, for example noise filters. This was the

952 motivation for having a selection only on 2D clusters in the collection (Y) plane. As  
 953 stated previously, the goal of this analysis was to find identifiable neutrino interactions  
 954 for use in public event displays, in future analyses, the 3D track reconstruction has  
 955 been modified to further increase the tracking efficiency and has more information  
 956 than just the clusters. For this analysis, however, 2D cluster information was sufficient  
 957 enough for neutrino selection.

958 **Primary Cuts**

959 The first cuts were used to select which clusters to consider. First the clusters must  
 960 have at least ten hits on the collection plane and have a cosmic tagging score < 0.4.  
 961 Only events that have at least two clusters that satisfy these primary cuts continue on.

962 After the initial cosmic tagging is applied, the following cuts are used to further  
 963 separate identifiable neutrinos from background cosmics.

964 The next cut was to remove long, vertical clusters. This was applied after seeing  
 965 that most cosmic induced clusters passing were long with high angles, while neutrino  
 966 induced clusters were mainly forward going. We required a good cluster to either  
 967 have a projected start angle less than 30 degrees from the z axis or be less than 200  
 968 wires long. The length cut was added to make sure we don't cut any short high angle  
 969 clusters that can correspond with a proton, or other highly ionizing particle associated  
 970 with a long muon cluster. The 200 wire cut roughly equates to 0.6 m in the z direction,  
 971 with a 3 mm wire pitch. Also, the projected angle is defined by  $\tan \alpha = \Delta T / \Delta W$  where  
 972 T is the time ticks and W is the wires.

973 The last cut requires the clusters to be either 30 time ticks or 30 wires. This cut was  
 974 applied to reduce small delta rays associated with a cosmic without removing proton  
 975 clusters associated with a long muon cluster, which saves ideal neutrino events that  
 976 have both a long minimum ionizing muon like cluster and a short highly ionizing  
 977 proton like cluster.

978 **Secondary Cuts**

979 The secondary cuts look to match long, low-angle clusters with short, high-charge  
 980 clusters. Only clusters that have passed previous cuts are used. First clusters with  
 981 length greater than 100 wires are chosen, which is approximately 0.3 m in the z

Cluster set	No Cuts	Primary Cuts	Secondary Cuts
Neutrinos only	570	303	32
Cosmics only ( no flash)	308,016	291,879	602
Cosmics only (w/ flash)	2464	2335	5
Neutrinos/Cosmics	0.23	0.13	6.4

**Table 4.1:** Passing rates for 2D cluster cuts for neutrino on MC set and a cosmic only MC set. First column shows event rates with no cuts applied to both sets. Columns two and three show event rates after primary and secondary cuts are applied. Line three shows the second line scaled with the flash finding factor of 0.008. All events are normalized to per day assuming we are running at 5 Hz.

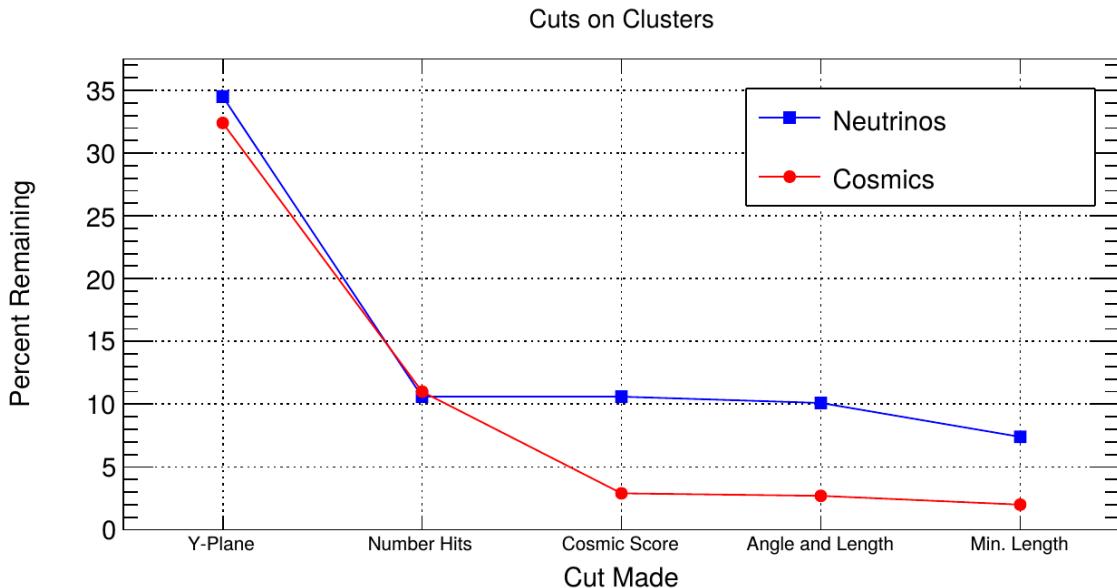
982 direction. Then we search for any cluster that is within approximately 3 cm ( 10 wires  
 983 and 30 time ticks) away from the low-z end of the long cluster. This cluster must also  
 984 be shorter than the first. In our reconstruction, the start and end point of a cluster can  
 985 be swapped so both ends of the short cluster are compared to the long cluster.

986 Now that there is a vertex match, cuts based on charge and projected opening angle  
 987 are implemented. We require the short cluster to have a higher start charge than the  
 988 long cluster or the long cluster be longer than 500 wires. Start charge is defined as  
 989 the charge on the first wire in ADC counts. The projected opening angle must also  
 990 be between 11 and 90 degrees. This last cut is intended to remove clusters that are  
 991 entirely overlapping or are part of the same long track. The resulting neutrino/cosmic  
 992 event rate per day is shown in table 4.1. Figures 4.3 and 4.4 shows the percentages of  
 993 clusters that pass each primary and secondary cuts.

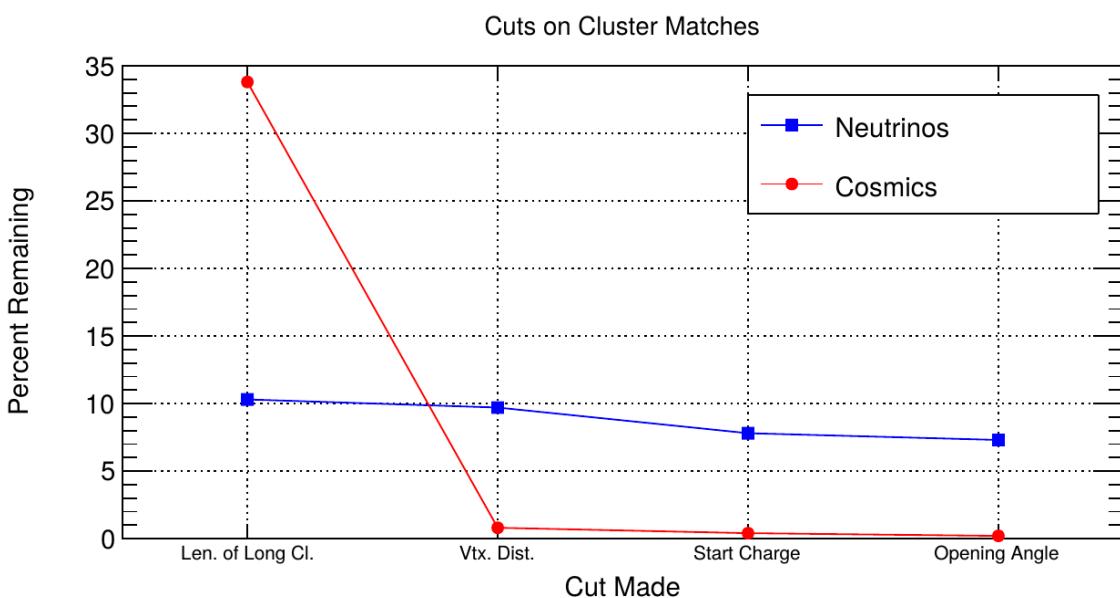
### 994 4.2.3 3D Tracks and vertices Selection

995 The neutrino selection for the 3D channel was based on a reconstructed vertex and  
 996 two tracks. All vertices and tracks were looped over that had a cosmic tag score < 0.4  
 997 and the distances below were calculated:

- 998 •  $d$ : distance between the start points of the two tracks.
- 999 •  $d_1$ : distance between vertex and start of track 1.
- 1000 •  $d_2$ : distance between vertex and start of track 2.



**Figure 4.3:** Percent of good clusters remaining for neutrinos and cosmics after the primary cuts were applied. This is relative to total number of initial clusters.



**Figure 4.4:** Percent matched cluster pairs remaining for neutrinos and cosmics after secondary cuts applied. This is relative to the number of events that contain clusters which pass the primary cuts.

1001 The maximum distance of all three is then selected as the important characteristic per  
1002 trio. The best trio is the one that has the smallest maximum distance. The  $\min(\max_d)$   
1003 for all trios in an event were plotted for BNB neutrino events and for cosmics to  
1004 find the best cut value for each tracking algorithm. The distribution of  $\min(\max_{d,i})$   
1005 is smaller for neutrinos than for cosmics. The cut values for different tracking and  
1006 clustering algorithms are shown below. These cut values were chosen to minimize the  
1007 cosmic background to 20%.

- 1008 • trackkalmanhit with cccluster  $\min(\max_{d,i}) < 3$  cm.  
1009 • trackkalmanhit with pandoraNu  $\min(\max_{d,i}) < 4.5$  cm.  
1010 • pandoraNu with cccluster  $\min(\max_{d,i}) < 5$  cm.

#### 1011 4.2.4 TPC Updates

1012 After doing a visual hand-scanning of the first beam data processed with the filters  
1013 detailed above, the events passing had a larger contamination of background than  
1014 expected. This was mainly in part due to the reconstruction performing better on  
1015 simulation than on data. Due to this, additional cuts on both streams needed to be  
1016 implemented in order to increase signal/background ratio. These cuts were added on  
1017 top of the filters described above and further reduce the event count.

#### 1018 2D Filter Updates

1019 The main background observed in the 2D filter were Michel events, where the muon  
1020 and electron formed two connected clusters. These events were rejected by comparing  
1021 the start and end charge deposition of the long cluster (i.e muon particle). The start  
1022 charge deposition must be less than the end charge deposition. This cut is implemented  
1023 because muons have a higher ionizaiton loss at the end.

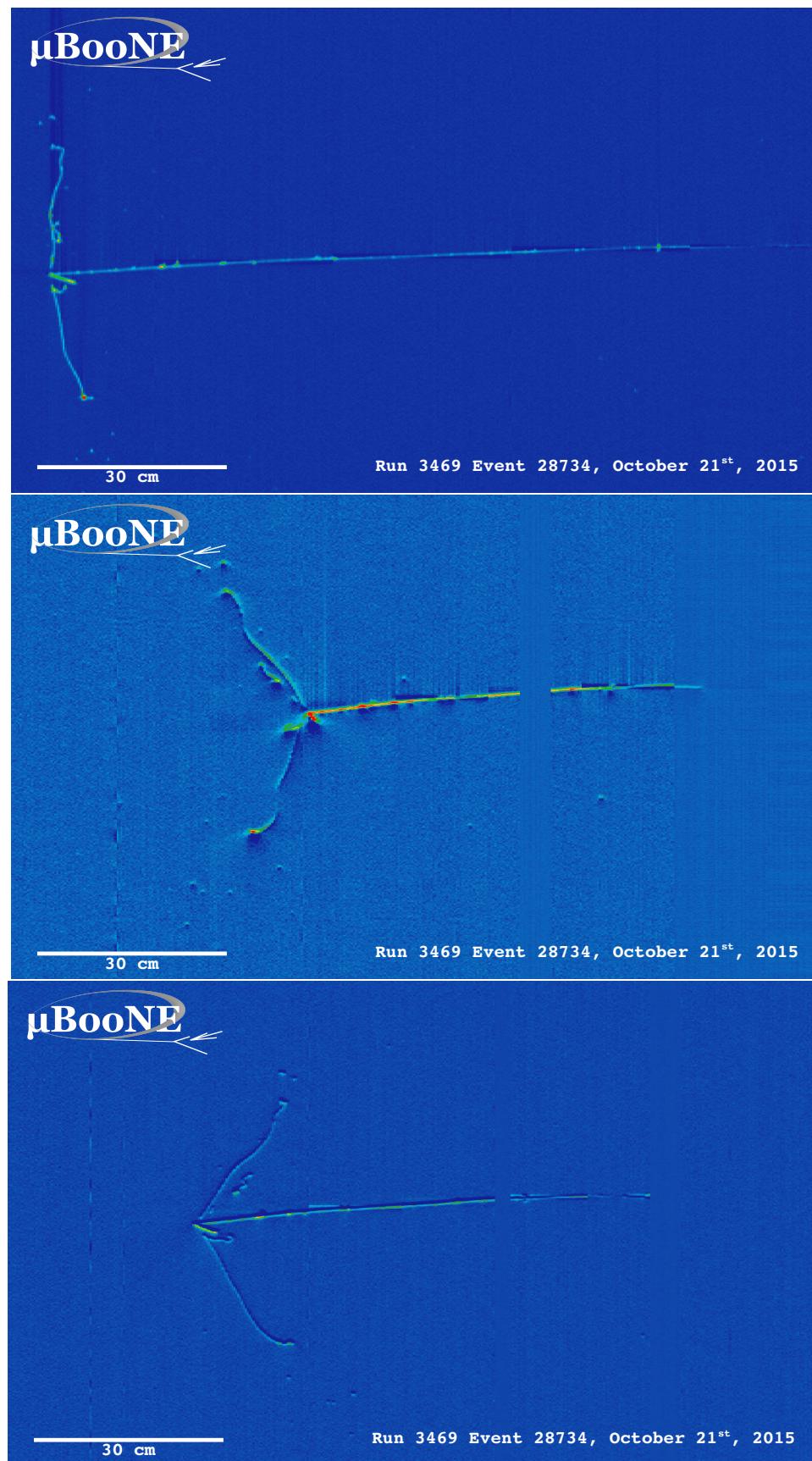
#### 1024 3D Filter Updates

1025 It was seen that cosmic tracks can often originate or end at the same point, therefore  
1026 faking a signal. Cosmic tracks, however, are mostly vertical. By requiring the angle  
1027 of the longer track have a cosine greater than 0.85 with respect to the z-axis as well

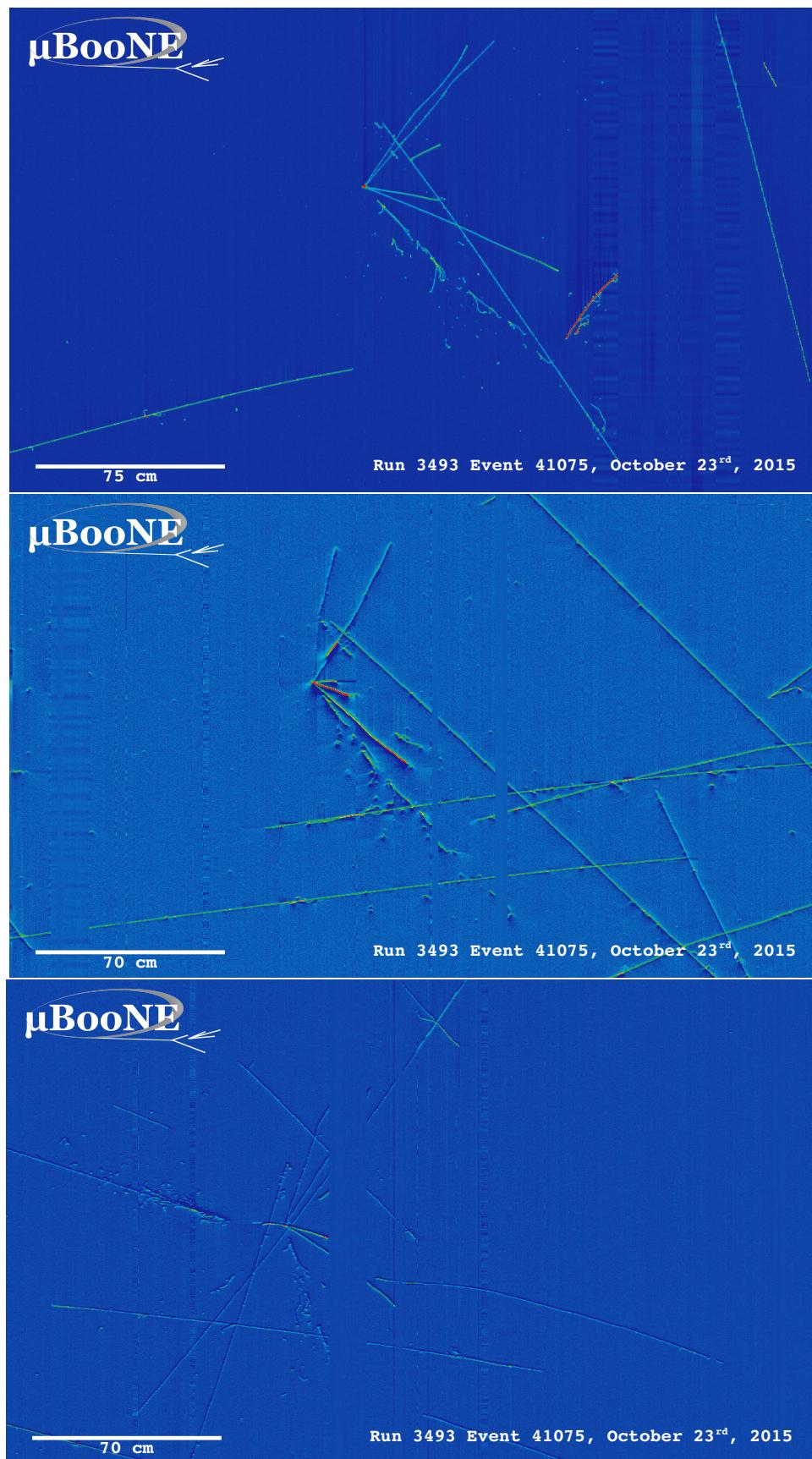
<sub>1028</sub> as requiring the longer track to have a length greater than 10 cm, we can reduce this  
<sub>1029</sub> background.

<sub>1030</sub> **4.3 Conclusion**

<sub>1031</sub> After proccesing these filters in parallel, it was shown that the 3D filter had a higher  
<sub>1032</sub> purity than the 2D filter because of the higher cosmic rejection being used due to 3D  
<sub>1033</sub> reconstruction. The 2D filter is blind to track entering/exiting from the top or bottom  
<sub>1034</sub> of the TPC. Although the 3D filter had a higher purity, the 2D filter was still able to  
<sub>1035</sub> find identifiable events in data that were used as public event displays. A sample of  
<sub>1036</sub> event displays are shown in figures 4.5 and 4.6.



**Figure 4.5:** First Neutrino Interaction Candidate Events from MicroBooNE



**Figure 4.6:** First Neutrino Interaction Candidate Events from MicroBooNE

<sup>1037</sup> **Chapter 5**

<sup>1038</sup> **CC-Inclusive Cross Section Selection**

<sup>1039</sup> **Filter**

<sup>1040</sup> The CC-Inclusive cross-section selection I and selection I modified filters used in this  
<sup>1041</sup> analysis will be described in the following sections below. These filters are an expansion  
<sup>1042</sup> of the Neutrino ID filter. The work done in this thesis was to further improve these  
<sup>1043</sup> selections by increasing both efficiency and purity as well as increasing acceptance  
<sup>1044</sup> without further affecting the kinematic distributions of the selected neutrino events.

<sup>1045</sup> MicroBooNE requires fully automated event reconstruction and selection algorithms for use in the many physics measurements being worked on to date due to  
<sup>1046</sup> the large data rate MicroBooNE receives. Being able to automatically pluck out the  
<sup>1047</sup> neutrino interaction among a sea of cosmics proved to be challenging but was accomplished.  
<sup>1048</sup> MicroBooNE has developed two complementary and preliminary selection  
<sup>1049</sup> algorithms to select charged-current  $\nu_\mu - Ar$  interactions. Both are fully automated  
<sup>1050</sup> and cut based. The results of this thesis will focus on selection I and selection I modified  
<sup>1051</sup> and will focus on further improving these algorithms using Convolutional Neural  
<sup>1052</sup> Network (CNN) implementations. These selections identify the muon from a neutrino  
<sup>1053</sup> interaction without biasing towards track multiplicity. To combat cosmic and neutral  
<sup>1054</sup> current background, the analysis is strongly biased towards forward-going long tracks  
<sup>1055</sup> which are contained. This limits phase space and reduces acceptance.

## 1057 5.1 Data and MC Processing Chain

1058 The data used for this analysis were based on hardware and software triggers. Events  
1059 used came from the *BNB\_INCLUSIVE* and *EXT\_BNB\_INCLUSIVE* streams and were  
1060 used for signal and background. The *BNB\_INCLUSIVE* stream is chosen by requiring  
1061 that the hardware trigger bit is fired and that the event passed an optical software  
1062 trigger within a BNB spill timing window. The *EXT\_BNB\_INCLUSIVE* stream requires  
1063 the EXT hardware trigger to fire as well as pass the same optical software trigger  
1064 within a BNB spill size timing window similar to the *BNB\_INCLUSIVE*.

1065 The two MC samples used in this analysis and for determining selection efficiencies  
1066 and purities were GENIE BNB neutrino interactions with CORSIKA cosmic ray overlay  
1067 within the readout window and inTime CORSIKA cosmic rays. The MC samples  
1068 generated used *uboonecode v04\_36\_00* and are based on the following packages:

- 1069 • larsoft v04\_36\_00
- 1070 • GEANT v04\_09\_06\_p04d
- 1071 • GENIE v02\_08\_06d
- 1072 • GENIE xsec v02\_08\_06a
- 1073 • pandora v02\_03\_0a
- 1074 • CORSIKA v07\_4003

1075 Both data and MC samples were processed using the same reconstruction release,  
1076 *uboonecode v05\_08\_00* and the fcl files used for reconstruction are listed below:

- 1077 • MC fcl files
  - 1078 – reco\_uboone\_mcc7\_driver\_stage1.fcl
  - 1079 – reco\_uboone\_mcc7\_driver\_stage2.fcl
- 1080 • Data fcl files
  - 1081 • reco\_uboone\_data\_Feb2016\_driver\_stage1.fcl
  - 1082 • reco\_uboone\_data\_Feb2016\_driver\_stage2.fcl

1083 On top of the hardware and software triggers, the data also had to pass more  
1084 criteria to be identified as part of the good run list. The criteria is detailed below.

- 1085 • **Detector conditions:** the detector has to be in a good operating condition. The  
 1086 detector conditions are read from the slow monitoring database and are required  
 1087 to be within the alarm thresholds. The variables of interest for events passing  
 1088 the good run list criteria include DAQ, PMT, HV, Drift HV, wire bias, electron  
 1089 lifetime and detector power. These conditions need to be met on a run-by-run  
 1090 basis in order to pass the selection.
- 1091 • **Data quality:** normal and stable behavior for basic reconstruction quantities.  
 1092 These reconstruction variables include average number of tracks, hits, and flashes  
 1093 in each event, the average length of tracks, the average amplitude and area of  
 1094 hits, the average PE and the average spread of each one of these quantities.
- 1095 • **Beam Conditions:** the BNB must be on and stable and the POT per spill needs  
 1096 to above the intensity threshold. Beam quality conditions include checking the  
 1097 fraction of proton beam interacting within the target, the horn current, and the  
 1098 intensity of protons per spill. The final sample is  $5 * 10^{19}$  and a per-spill intensity  
 1099 of  $4 * 10^{12}$
- 1100 • **Run processed:** the full run must be processed completely without missing  
 1101 subruns or crashes in the data processing.

## 1102 5.2 Normalization of data and MC

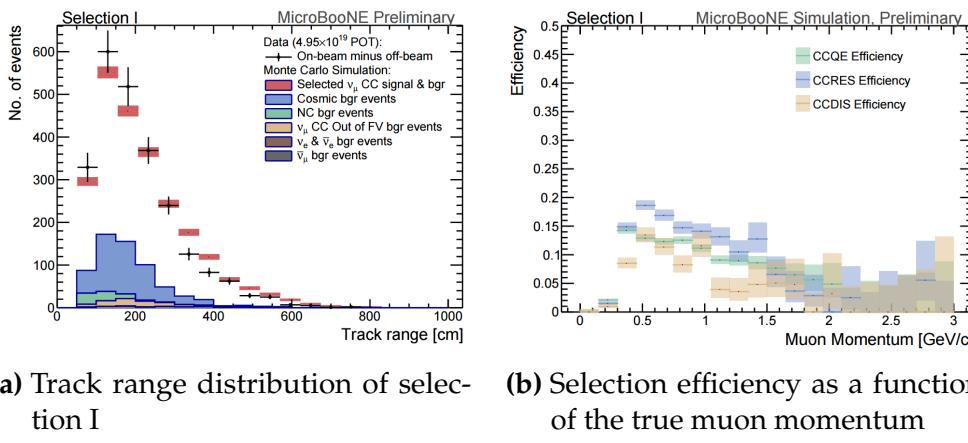
1103 The off-beam sample is used to measure beam unrelated backgrounds. For normalization,  
 1104 one needs the total number of BNB spills ( $N_{BNB}$ ) and the total number of external  
 1105 triggers. The BNB spills used need to pass the beam quality cuts. The normalization  
 1106 factor is then  $N_{BNB}/N_{EXT}$  which is 1.23.

1107 To normalize generated BNB MC events to POT, we used the following:

- 1108 •  $5 * 10^{19} POT = 41524.3$  generated events

1109 where this scaling factor only applies to mcc7 generated events. The inTime cosmic  
 1110 sample is normalized with respect to the open cosmic sample so an understanding  
 1111 of both is necessary. The POT per beam spill for mcc7 BNB samples is  $5 * 10^{12}$ . To  
 1112 calculate how many spills are necessary to produce a specific POT one would multiply  
 1113 the total POT by the average 1/POT per spill. For a total POT of  $5 * 10^{19}$  the amount  
 1114 of spills necessary is  $\frac{5 * 10^{19}}{5 * 10^{12}} = 1 * 10^7$ . This is only one in  $\sim 241$  events therefore each

cosmic event needs to be scaled up by a factor of 240.8 when comparing to BNB MC. For inTime cosmics however, two filters are applied to reduce computing and processing time and only leave cosmics that will interact within the detector. The passing rate after these two filters is 0.02125, therefore the total inTime cosmic scaling factor to compare inTime cosmics to BNB is  $0.02125 * 240.8 = 5.12$ .



**Figure 5.1:** 5.1a Track range distribution for selection I. The track range is defined as the 3D distance between the start and end of the muon candidate track. No data is shown below 75 cm due to the track length cut described previously. 5.1b Efficiency of the selected events by process quasi-elastic (QE), resonant (RES), and deep-inelastic (DIS). Statistical uncertainty is shown in the bands and the distributions are a function of true muon momentum. The rise of the efficiency between 0 GeV and 0.5 GeV is due to the minimum track length cut and the decreasing efficiency for higher momentum tracks is caused by the containment requirement.

## 5.3 Optical Software Trigger and Reconstruction

### 5.3.1 Software Trigger

Most of the BNB spills from the accelerator do not have a neutrino interaction in MicroBooNE. To save computation resources and reduce data-rates, we require a burst of light in the light collection system in coincidence with the 1.6  $\mu\text{s}$  beam spill. Requiring light activity in coincidence with the beam spill eliminates the vast majority of triggers with no neutrino interaction in the detector, however, it doesn't guarantee the activity in the detector is a neutrino interaction since a cosmic ray can interact in coincidence with the beam spill as well.

1129 To implement this, a software trigger was used on the PMT waveforms to decide  
1130 whether or not to keep that event. The software trigger is implemented after the event  
1131 builder combines data from the PMTs and triggers into a single event. The software  
1132 trigger uses the digitized output of the 32 PMT channels in the light collection system.  
1133 Only the waveform region in coincidence with the beam spill is used to search for  
1134 possible triggers. For each PMT, a waveform is found by taking the difference of ADC  
1135 values is calculated between  $t$  and  $t + s$ . This waveform is then scanned for ADC  
1136 values above a threshold  $X_0$ . Once an ADC is above this threshold, a discriminator  
1137 window is opened for a fixed number of time ticks ( $W_0$ ). If the ADC count within this  
1138 window  $W_0$  is greater than a second larger threshold  $X_3$ , a final window of width  $W_3$   
1139 is opened. The max ADC value within this final window is set as the peak amplitude  
1140 for the PMT and then summed across all 32 PMTs and set to the variable PHMAX. The  
1141 software trigger places a final cut on the PHMAX variable to decide whether or not  
1142 to keep the event. The thresholds were found by the Trigger task force using Monte  
1143 Carlo Studies and are as follows:

- 1144 •  $X_0 = 5$  ADC
- 1145 •  $X_3 = 10$  ADC
- 1146 •  $W_0 = 6$  Ticks
- 1147 •  $W_3 = 6$  Ticks
- 1148 • PHMAX cut = 130 ADC

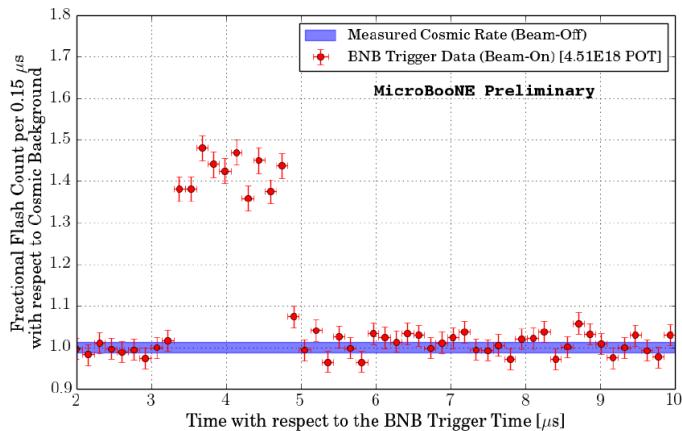
### 1149 5.3.2 Flash Reconstruction

1150 MicroBooNE collects light from each of the 32 PMTs either in a continuous readout  
1151 window of  $23.4 \mu\text{s}$  activated by a beam gate signal on the trigger board, or in discrimi-  
1152 nated pulses of  $\sim 1 \mu\text{s}$  duration activated if the ADC count for any PMT goes above 80  
1153 ADC count. These two formats are saved as output waveforms and put onto an event.  
1154 Additionally, each PMT can provide two output streams, high-gain ( $\sim 20$  ADC/PE)  
1155 and low-gain ( $\sim 2$  ADC/PE) channels. The first step in the reconstruction is to merge  
1156 both these channels into a “saturation corrected waveform” which uses information  
1157 from the low-gain waveform to correct for saturating high-gain pulses.

1158 The saturation corrected waveform in the continuous readout window is used to  
1159 reconstruct optical hits. Each PMT’s waveform is scanned for hits then a threshold

1160 based hit reconstruction algorithm is applied which requires pulses of a minimum  
1161 area in order to be reconstructed. Each reconstructed hit is associated to a PMT, a time  
1162 in  $\mu\text{s}$ , and a PE count.

1163 Once hits are reconstructed for all 32 PMTs, all PMT information is then combined  
1164 into optical flashes which represent optical information seen by the PMTs from interac-  
1165 tions in the detector. Each flash has information on total light seen per interaction, the  
1166 distribution of the light across all 32 PMTs, the flash time with respect to the trigger  
1167 time of the flash, and lastly, the spacial information of the flash in Y-Z plane of the  
1168 detector. These flashes are reconstructed by requiring that there is a  $\sim 1 \mu\text{s}$  coincidence  
1169 between the reconstructed hits in all 32 PMTs. The total PE is summed up among  
1170 all coincident hits across the PMTs and if the total PE is greater than 2 PE, a flash is  
1171 reconstructed. There are also safe guards in place to take care of late scintillation light.

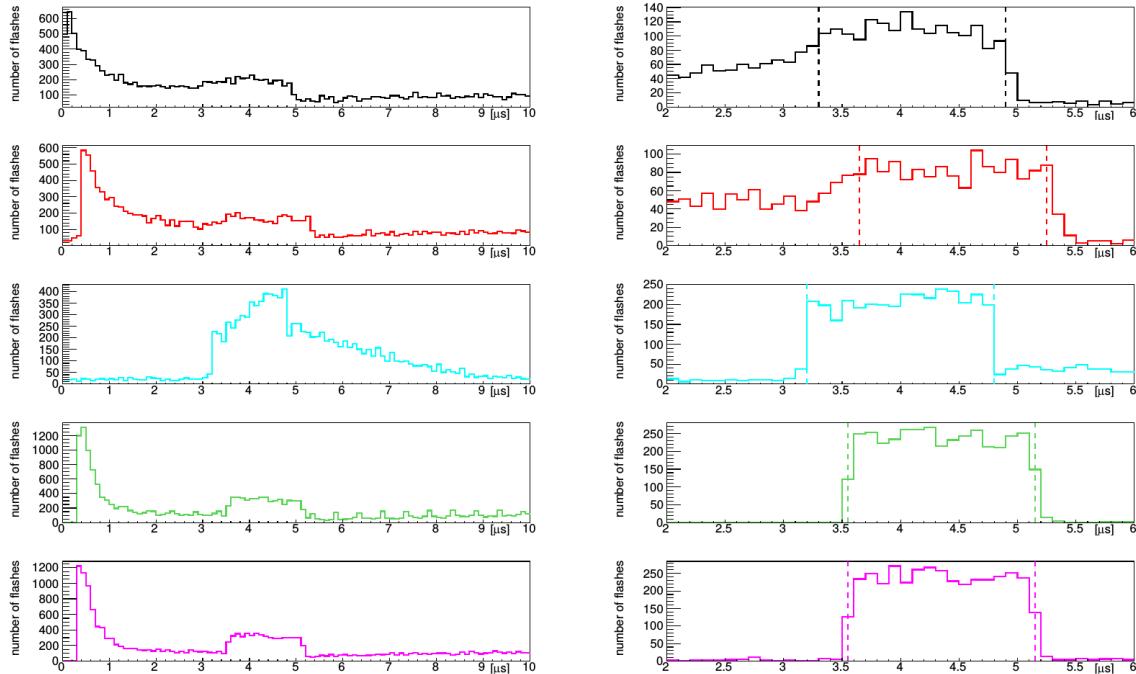


**Figure 5.2:** Time distribution of reconstructed optical flashes with a PE value of 50 or more for a sample of BNB unbiased triggered events.

1172 Figure 5.2 shows the time distribution of reconstructed optical flashes using the  
1173 BNB continuous stream. You can see a clear excess in coincidence with the expected  
1174 arrival time of neutrinos. The same flash reconstruction that was used in the cc-  
1175 inclusive filter detailed here was used to create this plot in data.

### 1176 5.3.3 Beam Window

1177 Figure 5.3 shows the distribution of flashes for on-beam, off-beam and various MC  
1178 samples. The software trigger has been applied to these samples. The pile-up seen just  
1179 after 0  $\mu\text{s}$  is a feature of the flash finding algorithm and consists of low PE flashes and



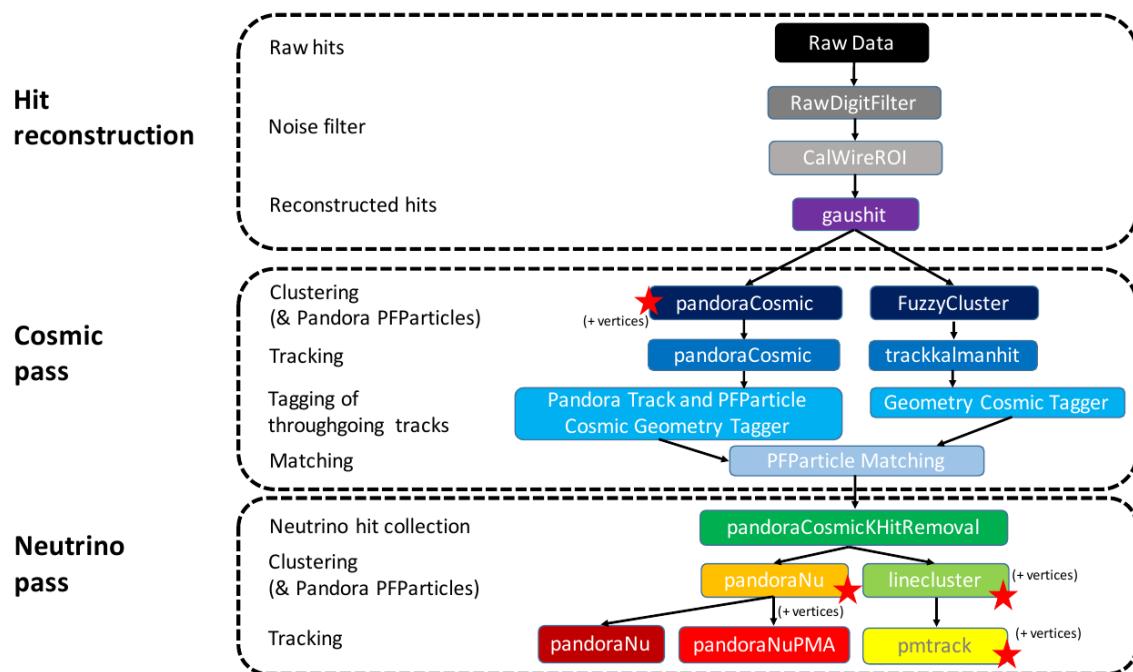
**Figure 5.3:** Flash time distribution for all flashes (left plot) and flashes  $> 20\text{PE}$  (right plot). The different curves are as follows: on-beam data (black), off-beam data (red), CORSIKA inTime MC (light blue), BNB only MC (green), and BNB+Cosmic MC (purple). The dashed vertical lines mark the time window that was chosen for each sample

is removed in the second column of distributions with a low 20 PE threshold cut. The plots show that the time window for the distributions are shifted a small amount from each-other. This is caused by different hardware configurations per sample. Using these distributions, the windows chosen per sample are as follows:

- On-Beam: 3.3 to 4.9  $\mu$ s
- Off-Beam: 3.65 to 5.25  $\mu$ s
- CORSIKA inTime: 3.2 to 4.8  $\mu$ s
- BNB only: 3.55 to 5.15  $\mu$ s
- BNB+Cosmic: 3.55 to 5.15  $\mu$ s

Each window has a width of 1.6  $\mu$ s.

## 5.4 TPC Reconstruction



**Figure 5.4:** Reconstruction chain run on both data and MC. The red stars mean that the algorithm returns reconstructed 3D vertices.

1191     Figure 5.4 summarizes the reconstruction chain applied to both MC and data for  
1192     this analysis. After the hit reconstruction, a cosmic pass is applied which removes all  
1193     hits associated to through-going tracks. A description of these TPC reconstruction  
1194     algorithms will be detailed below.

1195     **5.4.1 Hit Reconstruction**

1196     The waveforms used for hit reconstruction consist of charge deposited on the sense  
1197     wire in drift time. The first step in hit reconstruction is to pass the waveforms through  
1198     a filtering algorithm to filter out the noise introduced from the electronics. The input  
1199     waveforms are also truncated from 9600 time ticks to 6400 time ticks in this first step  
1200     to reduce the data footprint of these waveforms.

1201     Once noise filtering is complete, a deconvolution algorithm is applied to the wave-  
1202     forms to remove the drift field and electronics response, therefore leaving only the  
1203     ionized electrons kicked off the argon atoms by an incident track. During this process,  
1204     Region of Interests (ROI) are identified and cut out of the waveforms to further reduce  
1205     the data volume.

1206     The hit finding algorithm then finds candidate peaks in these ROI's and fits the  
1207     peaks to Gaussian curves. These Gaussian shaped peaks are now called hits and  
1208     represent the charge deposition on a wire by the incoming track. These hit objects  
1209     have a peak time and width and are the basic object input to further algorithms down  
1210     the reconstruction chain.

1211     **5.4.2 Clustering**

1212     There are multiple clustering algorithms used in this analysis. The main purpose of all  
1213     the clustering algorithms is to associate hits together in 2D space to create objects like  
1214     tracks, vertices and showers. For the fuzzy cluster algorithm, three steps are used to  
1215     achieve this. The first step is to associate hits to each-other using a fuzzy clustering  
1216     algorithm which gives each hit a degree of belonging to the cluster. Second, a Hough  
1217     transform is used to find hits associated to candidate tracks and showers within each  
1218     of the clusters found in the first step. The last step merges smaller candidate tracks  
1219     and showers into large clusters. The last step also associates unclustered hits into

1220 nearby objects which helps shower reconstruction. The result is a set of clusters made  
1221 up of associate hits that represent tracks or showers per plane.

1222 The pandora algorithm utilizes it's own clustering algorithm and will be detailed  
1223 in the next section. The last clustering algorithm is called linecluster. The linecluster  
1224 algorithm reconstructs 2D linear clusters per plane by fitting a line onto nearby hits  
1225 which is then extrapolated to neighboring wires. 2D vertices are found per plane by  
1226 using the intersection points of the ends of nearby clusters. These 2D vertices are then  
1227 matched in time across all three planes to get a 3D vertex in space.

### 1228 5.4.3 Pandora

### 1229 5.4.4 Trackkalmanhit

1230 The trackkalmanhit algorithm takes 2D clusters returned from the fuzzy cluster algo-  
1231 rithm and outputs track objects. There are no hierarchy structure as it is in pandora,  
1232 each track is independent. There also is no vertex reconstruction with this algorithm  
1233 as well.

### 1234 5.4.5 Cosmic Hit Removal

1235 The Pandora algorithm is applied to the events twice, the first to remove downward  
1236 going tracks primarily from cosmic ray muon like particles. The second pass only runs  
1237 on a subset of hits that aren't associated with cosmic ray muon tracks.

1238 After the first pass, the output of PFParticle hierarchy is then passed to a cosmic  
1239 ray tagger to look through all hits to determine start and end points. If the start or  
1240 end point trajectories are consistent with entering or exiting the TPC, then these hits  
1241 are removed from the second pass. Hits are considered entering or exiting the TPC  
1242 if the drift time are outside of the neutrino drift window or outside of the fiducial  
1243 volume of the TPC. The fiducial volume was based on a montecarlo study and is 20  
1244 cm from the top or bottom of the TPC and 10 cm from the TPC ends. Hits associated  
1245 with candidate cosmic ray tracks are removed from the input hit collection and the  
1246 remaining hits are passed to the neutrino optimized pass of Pandora.

### **1247 5.4.6 Projection Matching Algorithm**

1248 The projection matching algorithm (PMA) was inherited from ICARUS and has been  
1249 implemented in LArSoft. PMA differs from traditional LArSoft 3D reconstruction  
1250 algorithms. Most 3D reconstruction attempts to match 2D objects from all three planes  
1251 by drift time, while the PMA algorithm projects a track hypothesis on each plane  
1252 then the distance between this projection and the hits on each plane is minimized  
1253 simultaneously. More information can be found in [?].

## **1254 5.5 Event Selection**

1255 The first requirement for selecting  $\nu_\mu$  CC events is that the event has at least one  
1256 scintillation light flash in the beam trigger window with more than 50 PE on all PMTs  
1257 combined. From the flashes that pass, the most intense is chosen and considered to be  
1258 originating from a neutrino interaction and will be the only flash used in further cuts.

1259 Vertices are then required to have at least one reconstructed track start or endpoint  
1260 within a 5 cm radius. Showers associated with a vertex do not pass this cut. All  
1261 tracks associated with a vertex are then used to calculate a track length weighted  
1262 average of the  $\theta$ -angle. Of all the vertices that do pass, only the vertex with the most  
1263 forward going  $\theta$ -angle average of all associated tracks is considered the neutrino vertex  
1264 candidate. The most forward going  $\theta$ -angle average is chosen by picking the largest  
1265 track range weighted average of  $|\cos(\theta)|$ , seeing as  $\cos(\theta) = 1$  is the beam direction.  
1266 Next, it is required that the reconstructed neutrino vertex candidate be within the  
1267 fiducial volume as well as within the drift time starting at  $t_0$ . The fiducial volume  
1268 boundaries chosen are 10 cm from the edges of the TPC in x and z which is the drift  
1269 direction and beam direction respectively, and 20 cm from the edges of the TPC in y  
1270 which is the vertical direction. For all further cuts, only the longest track associated  
1271 with the neutrino vertex candidate and this track is assumed to be the muon candidate  
1272 of the neutrino event.

1273 The next cut requires the position of the flash in the z-direction and the track z-  
1274 projection to be compared. This basic flash matching algorithm is rudimentary and a  
1275 placeholder for a more sophisticated algorithm. The z-position of the flash needs to be  
1276 within 80 cm to the z-positions of track start or endpoints. If the flash is between the  
1277 track start and endpoint, the distance of the flash to the track is considered to be 0 cm.

1278     Lastly, the track needs to be fully contained within the fiducial volume and have a  
 1279     track range greater than 75 cm. The range is the 3D distance between the track's start  
 1280     and endpoint. The length cut was optimized to remove NC background that contain  
 1281     a pion due to the pion interaction rate to be  $\sim 70$  cm. A track that makes all the cuts  
 1282     is considered to be the muon of a  $\nu_\mu$  CC event. The list of cuts for this selection is  
 1283     described below:

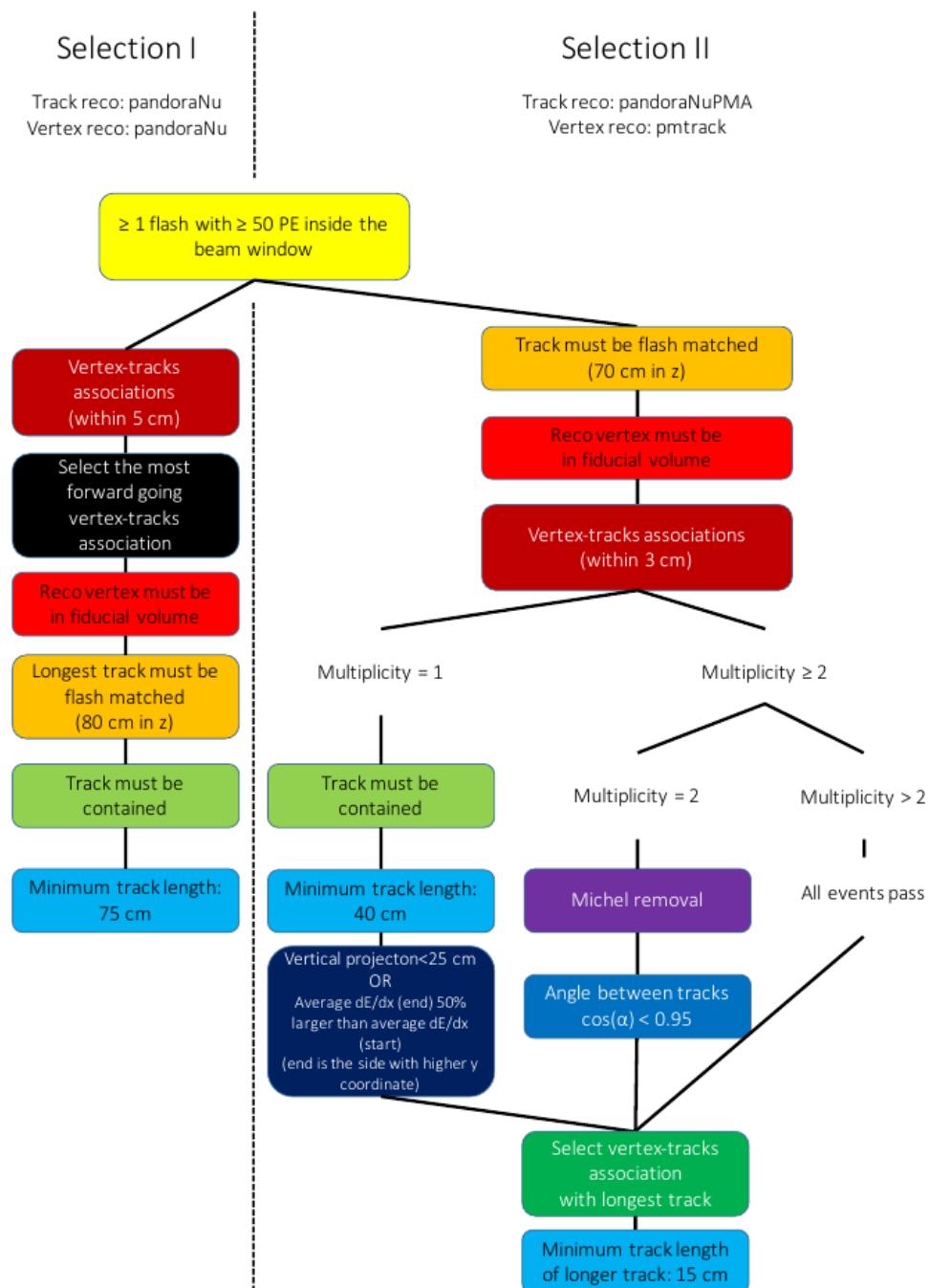
- 1284     1. At least one flash > 50 PE within the beam gate.
- 1285     2. At least one track within 5 cm around a vertex.
- 1286     3. Vertex with flattest tracks is chosen to be vertex candidate.
- 1287     4. Vertex candidate in fiducial volume.
- 1288     5. Longest track associated with vertex candidate is chosen to be track candidate.
- 1289     6. Longest track is within 80 cm (z-axis only) of the flash.
- 1290     7. Longest track is fully contained.
- 1291     8. Longest track is greater than 75 cm.

1292     The event selection scheme can also be seen in figure 5.5. Table 5.1 lists the passing  
 1293     rates for MC events for the selection scheme described above. Table 5.2 lists the passing  
 1294     rates for on-beam and off-beam data for the selection scheme. The normalization  
 1295     factors applied between on-beam and off-beam data are described in section 5.2.

### 1296     5.5.1 Expected Backgrounds

1297     Most of the selected background events will be of cosmic origin. There are two types  
 1298     of cosmic background, one triggered by a cosmic-ray event occurring in the beam  
 1299     gate time window, the other triggered by a beam induced interaction in the cryostat  
 1300     followed by a misidentification of a cosmic event as a neutrino event. The first  
 1301     cosmic background can be subtracted from the selected events using the off-beam  
 1302     BNBEXT sample normalized to the on-beam. The second cosmic background events  
 1303     are modeled by MC by using BNB+Cosmic MC sample.

1304     Other backgrounds originate from neutrino beam contaminants. A major contribu-  
 1305     tion in this sector is by neutral current neutrino events for example a charged pion track  
 1306     misidentified as a muon. Another contribution are  $\nu_e$ -like and anti-muon-neutrino



**Figure 5.5:** Event selection diagram for selection I and selection II. This analysis focused on optimizing selection I. Boxes with the same color across the two selections symbolize similar cuts.

	BNB+Cosmic Selection	BNB+ Cosmic MC-Truth	Cosmic Only	Signal:Cosmic Only
Generated Events	191362	45273	4804	1:22
$\geq 1$ flash with $\geq 50$ PE	136219 (71%/71%)	44002 (97%/97%)	2970 (62%/62%)	1:14
$\geq 1$ track within 5 cm of vertex	135830 (99%/71%)	43974 (99%/97%)	2975 (99%/62%)	1:14
vertex candidate in FV	79112 (58%/41%)	34891 (79%/77%)	1482 (50%/31%)	1:8.9
flash matching of longest track	40267 (51%/21%)	25891 (74%/57%)	340 (23%/7.1%)	1:2.8
track containment	19391 (48%/10%)	11693 (45%/26%)	129 (38%/2.7%)	1:2.3
track $\geq 75$ cm	6920 (36%/3.6%)	5780 (49%/13%)	17 (13%/0.4%)	1:0.6

**Table 5.1:** Passing rates of Selection I. Numbers are absolute event counts and cosmic background is not scaled. The BNB+Cosmic sample contains all events, not just  $\nu_\mu$  CC inclusive. The numbers in brackets give the passing rate wrt the step before (first percentage) and wrt total generated events (second percentage). The BNB+Cosmic MC-Truth column shows how many true  $\nu_\mu$  CC inclusive events are left in the sample. This number includes mis-identifications where a cosmic track is picked by the selection instead of the neutrino interaction in the same event. The cosmic only sample is used just to illustrate the cut efficiency. The last column Signal:Cosmic only gives an estimate of the  $\nu_\mu$  CC events wrt the cosmic only background at each step. For this number, the cosmic background has been scaled as described in section 5.2.

	on-beam	off-beam
Generated Events	546910	477819
$\geq 1$ flash with $\geq 50$ PE	135923 (25%/25%)	96748 (20%/20%)
$\geq 1$ track within 5 cm of vertex	134744 (99%/25%)	95778 (99%/20%)
vertex candidate in FV	74827 (55%/14%)	51468 (54%/11%)
flash matching of longest track	22059 (29%/4.0%)	12234 (24%/2.6%)
track containment	10722 (49%/1.9%)	5283 (43%/1.1%)
track $\geq 75$ cm	3213 (30%/0.6%)	1328 (25%/0.3%)

**Table 5.2:** Passing rates for Selection I selection applied to on-beam and off-beam data. The numbers in brackets give the passing rate wrt the step before (first percentage) and wrt the generated events (second percentage). Off-beam data has been scaled with a factor 1.23 to normalize to the on-beam data stream.

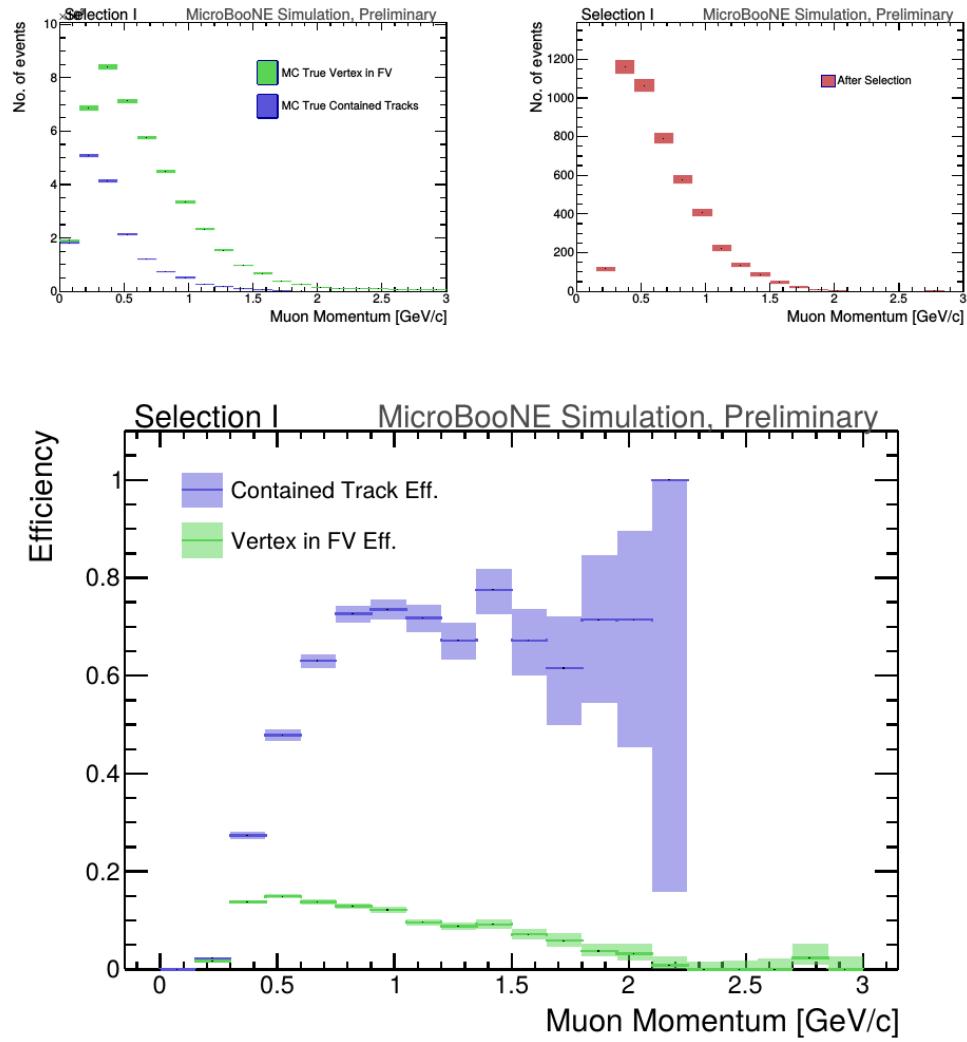
1307 events. These beam related backgrounds are an order of magnitude smaller than the  
1308 cosmic misidentification backgrounds. These backgrounds can not be subtracted and  
1309 are estimated using MC truth.

1310 The efficiency and purity of Selection I are calculated below:

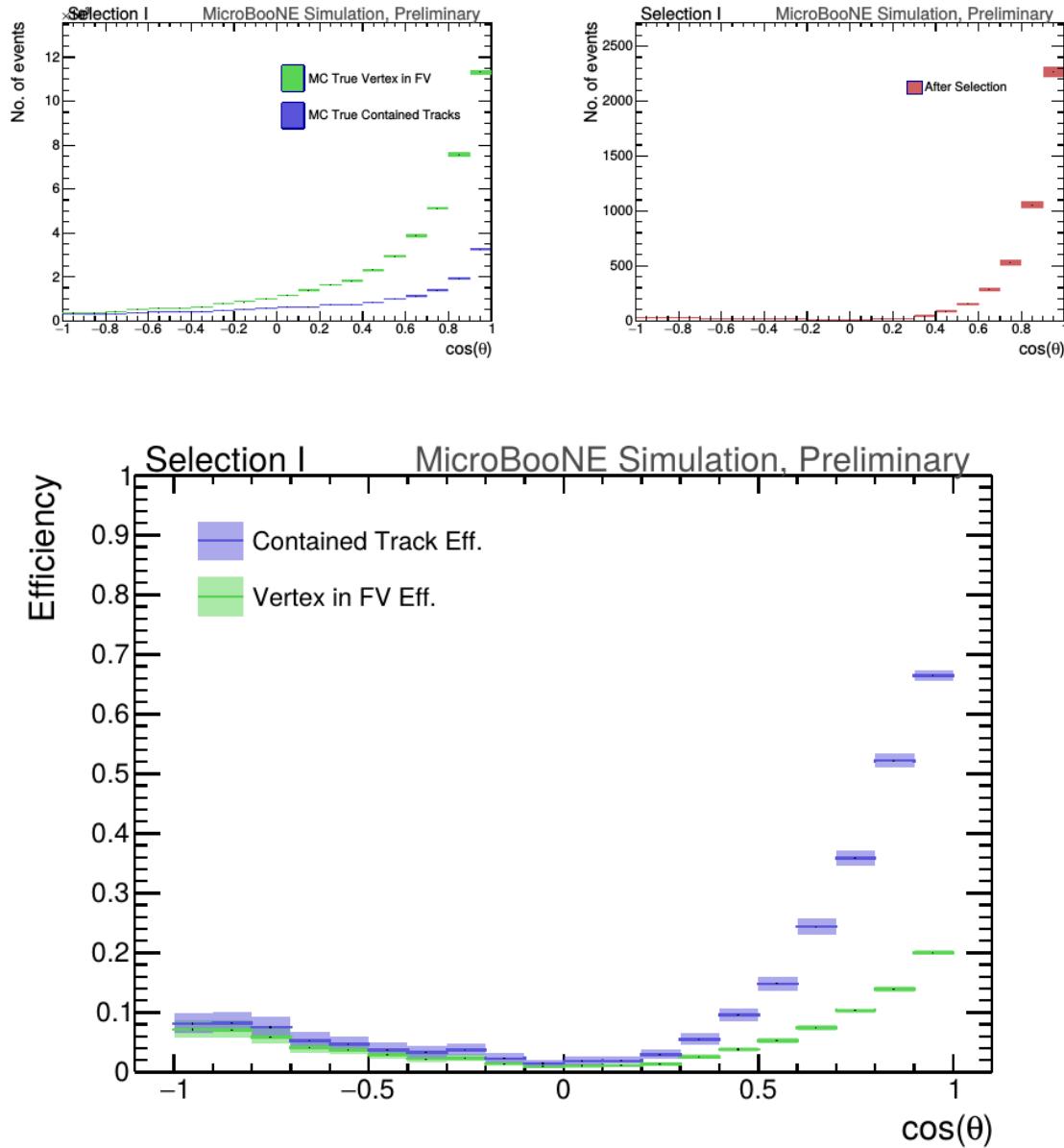
- 1311 • Efficiency: Number of selected true  $\nu_\mu$  CC events divided by the number of  
1312 expected true  $\nu_\mu$  CC events with interaction in the FV.
  - 1313 –  $(12.3 \pm 3.4) \%$
- 1314 • Purity: Number of selected true  $\nu_\mu$  CC events divided by the sum of itself and  
1315 the number of all backgrounds.
  - 1316 –  $(53.8 \pm 4.4) \%$

### 1317 5.5.2 Truth Distributions

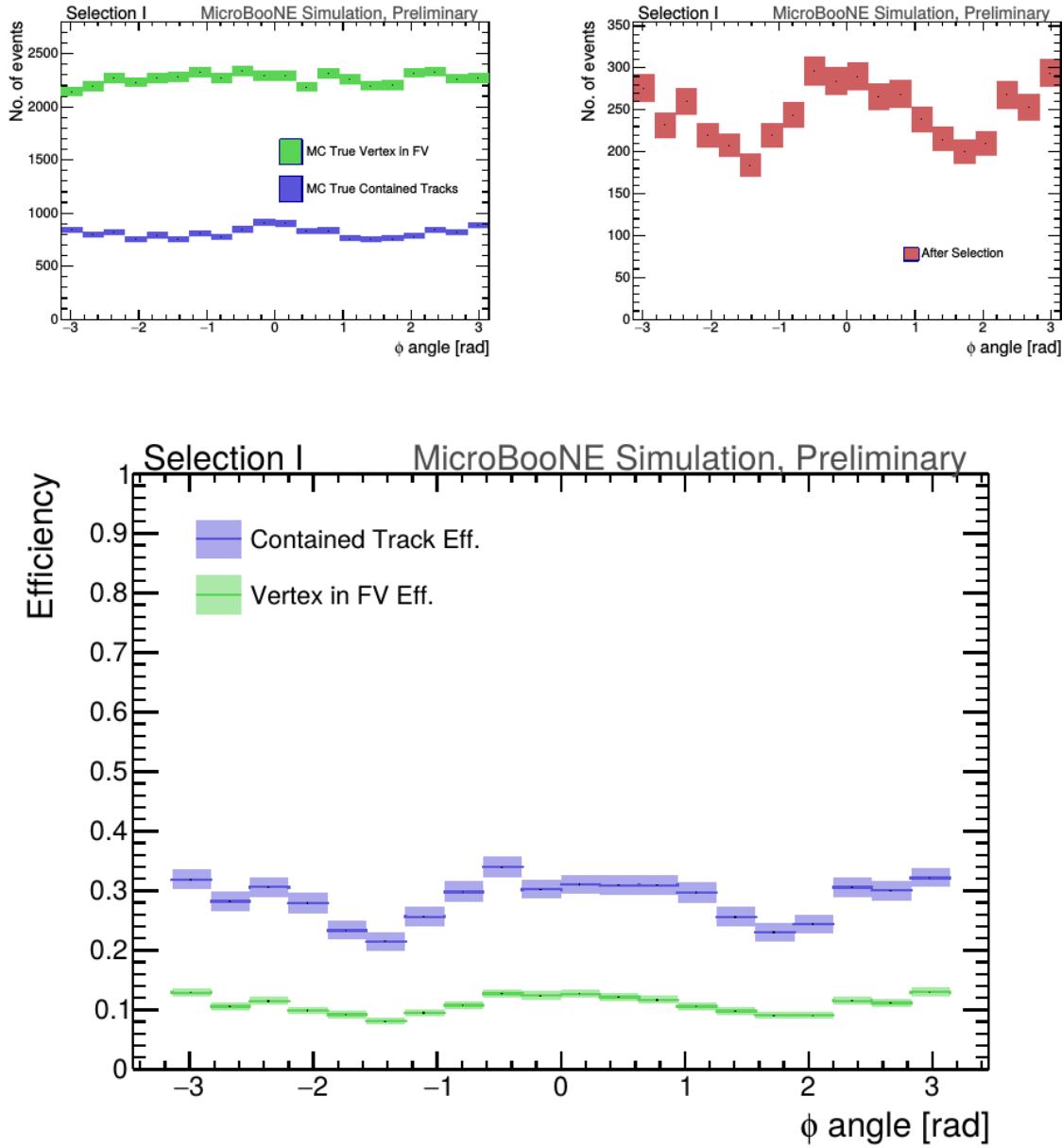
1318 The truth distributions of MC truth variables before and after the selection are detailed  
1319 in this section. The overall efficiencies are calculated for all  $\nu_\mu$  CC signal events  
1320 with a true interaction within the fiducial volume and a fully contained muon track  
1321 originating from said vertex. Figures 5.6 through 5.8 detail the truth distributions for  
1322 muon momentum,  $\cos(\theta)$  and  $\phi$  and figures 5.9 through 5.11 detail the total efficiency  
1323 of the selection for charged current quasi elastic (CCQE) events, charged current  
1324 resonant (CCRES) events, and charged current deep inelastic (CCDIS) events.



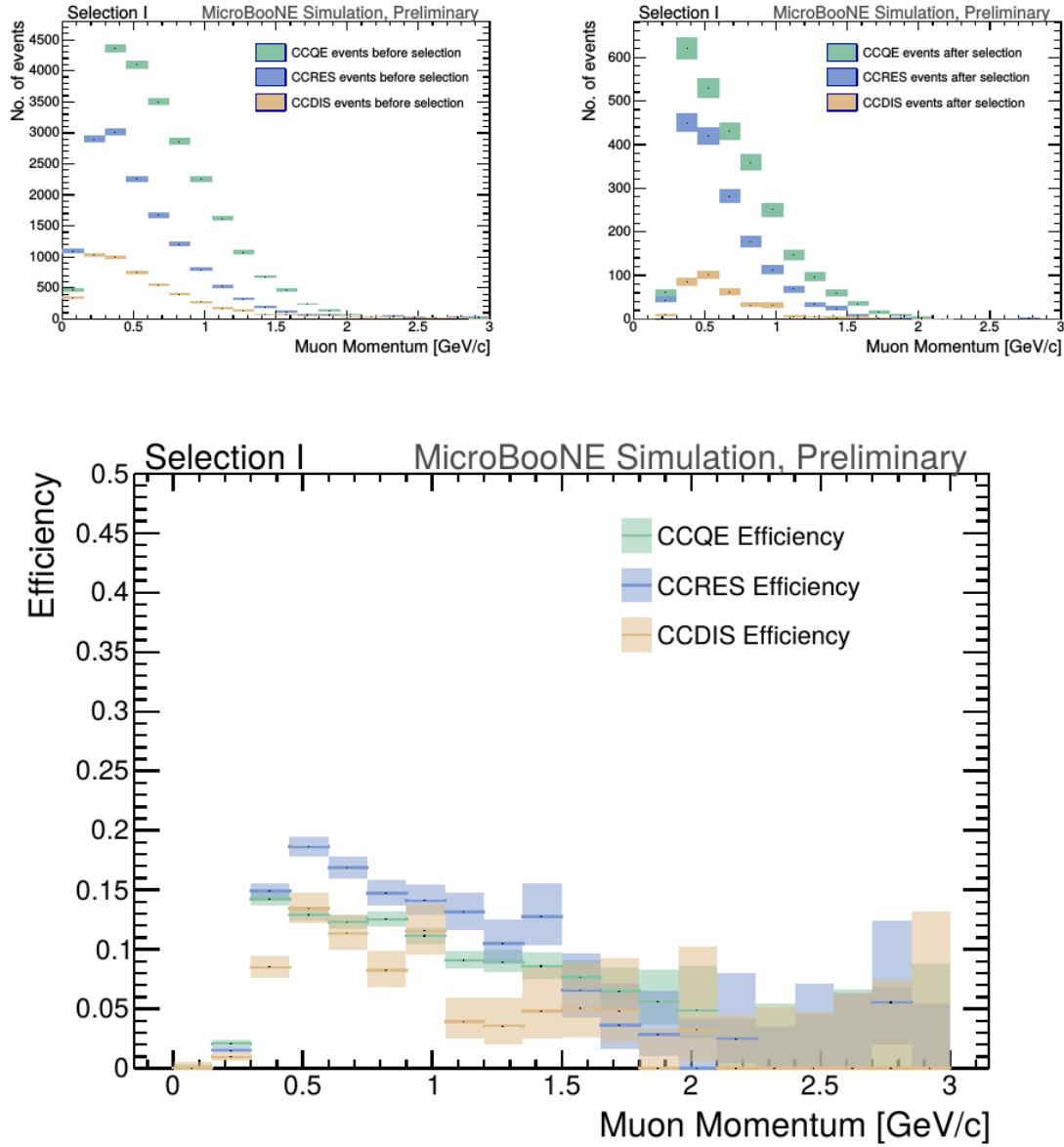
**Figure 5.6:** MC momentum distributions of the muon originating from a  $\nu_\mu$  CC interaction. Upper left is the momentum distributions of events with a vertex within the FV (green) and the events with a fully contained track (blue) before the selection. The upper right side is the momentum distribution after the selection (red). The lower plot is the selection efficiencies.



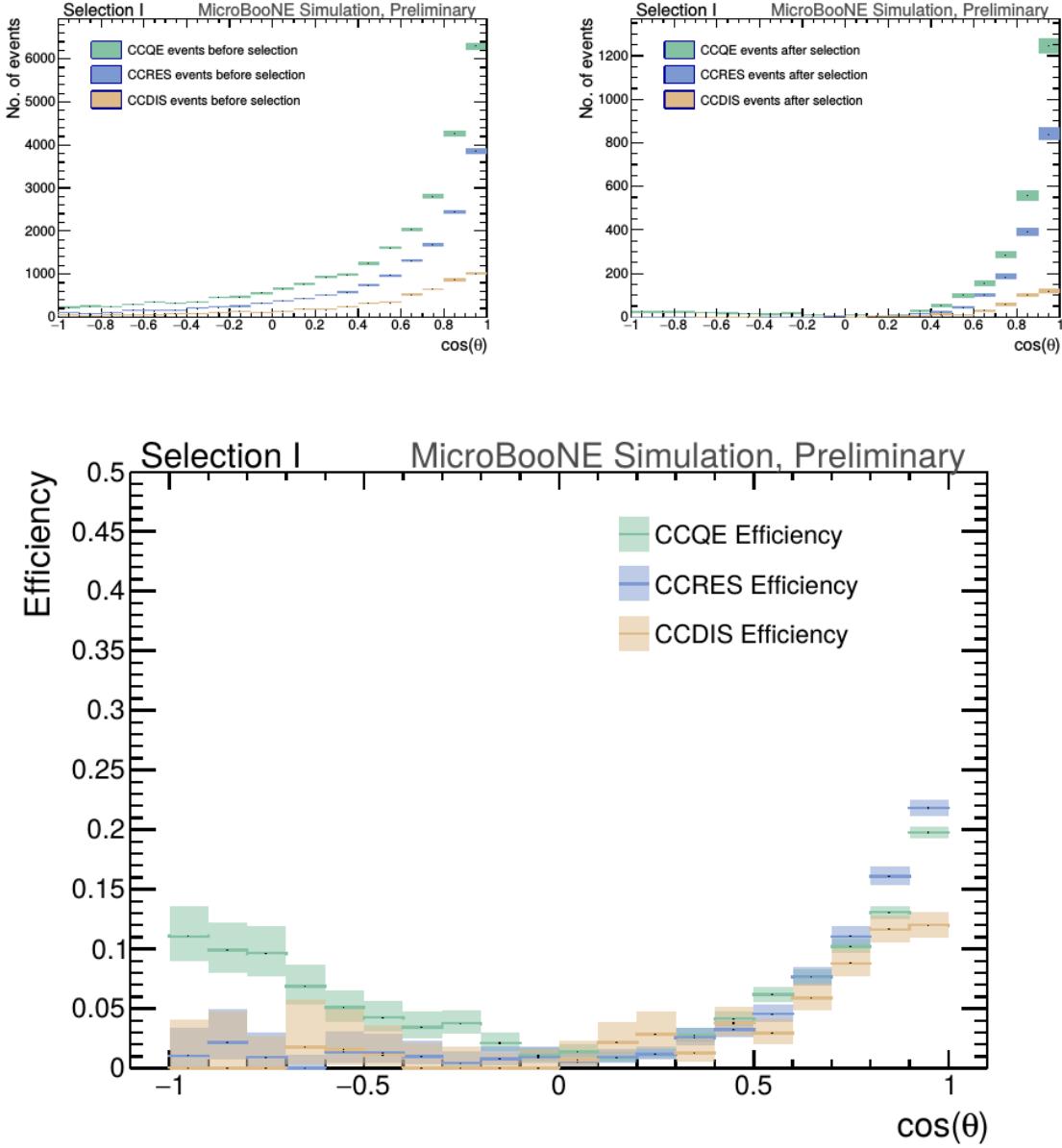
**Figure 5.7:** MC  $\cos(\theta)$  distributions of the muon originating from a  $\nu_\mu$  CC interaction. Upper left is the  $\cos(\theta)$  distributions of events with a vertex within the FV (green) and the events with a fully contained track (blue) before the selection. The upper right side is the  $\cos(\theta)$  distribution after the selection (red). The lower plot is the selection efficiencies.



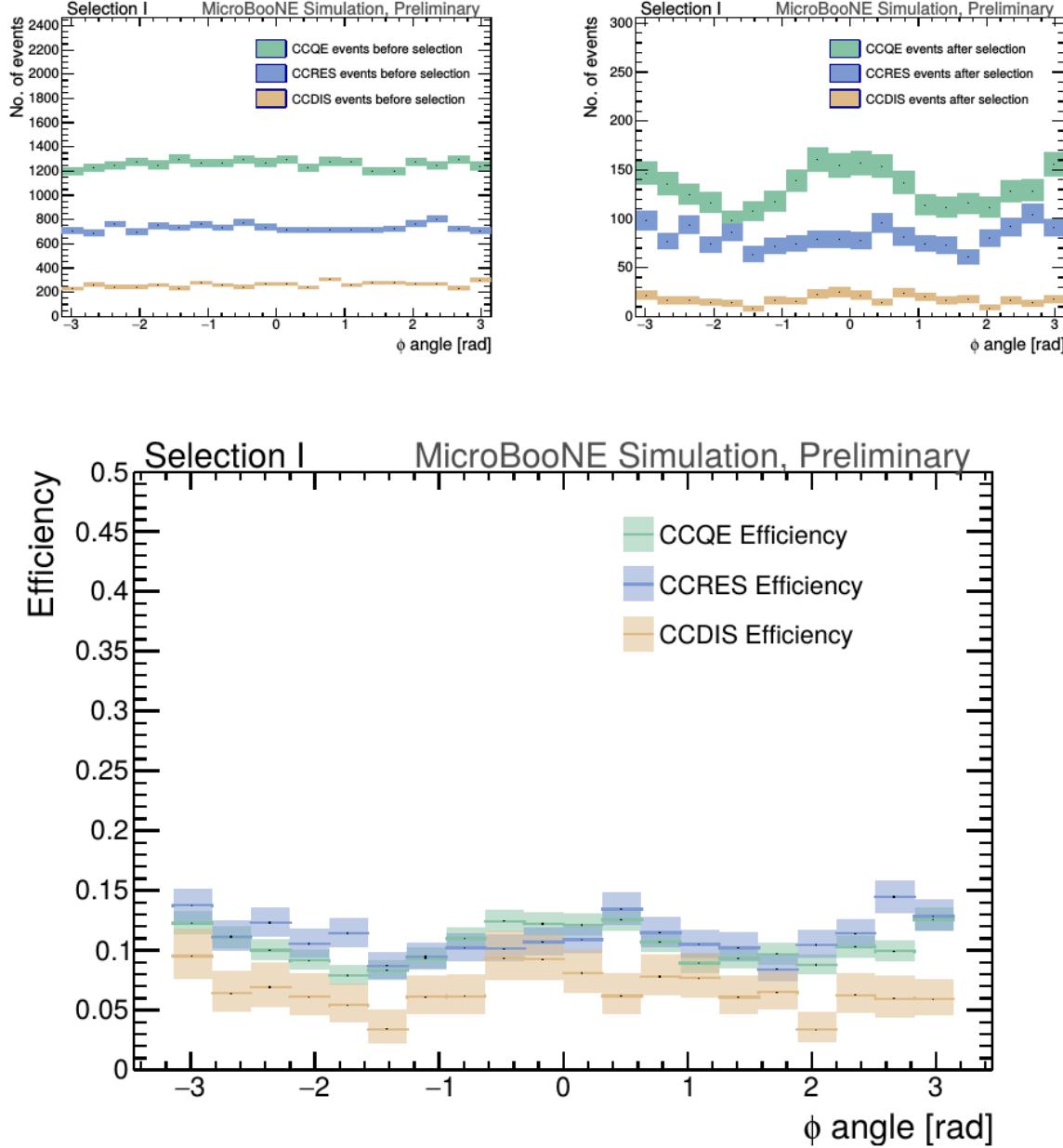
**Figure 5.8:** MC  $\phi$  distributions of the muon originating from a  $\nu_\mu$  CC interaction. Upper left is the  $\phi$  distributions of events with a vertex within the FV (green) and the events with a fully contained track (blue) before the selection. The upper right side is the  $\phi$  distribution after the selection (red). The lower plot is the selection efficiencies.



**Figure 5.9:** MC momentum distributions of the muon originating from a  $\nu_\mu$  CC interaction. Upper left is the momentum distributions of events with a vertex within the FV split up into CCQE (red), CCRES (yellow), and CCDIS (green) before selection. The upper right side is the momentum distribution after the selection with the same color schemes. The lower plot is the selection efficiencies for all three interaction types. The definition of QE, RES, and DIS is based on GENIE.



**Figure 5.10:** MC  $\cos(\theta)$  distributions of the muon originating from a  $\nu_\mu$  CC interaction. Upper left is the  $\cos(\theta)$  distributions of events with a vertex within the FV split up into CCQE (red), CCRES (yellow), and CCDIS (green) before selection. The upper right side is the  $\cos(\theta)$  distribution after the selection with the same color schemes. The lower plot is the selection efficiencies for all three interaction types. The definition of QE, RES, and DIS is based on GENIE.



**Figure 5.11:** MC  $\phi$  distributions of the muon originating from a  $\nu_\mu$  CC interaction. Upper left is the  $\phi$  distributions of events with a vertex within the FV split up into CCQE (red), CCRES (yellow), and CCDIS (green) before selection. The upper right side is the  $\phi$  distribution after the selection with the same color schemes. The lower plot is the selection efficiencies for all three interaction types. The definition of QE, RES, and DIS is based on GENIE.



# 1325 Chapter 6

## 1326 **Background on Convolutional Neural Networks**

1328 Convolutional neural networks (CNNs) have been one of the most influential innovations in the field of computer vision. Neural networks became popular in 2012 1329 when Alex Krizhevsky used them to win that year's ImageNet competition [?] by 1330 dropping the error from 26% to 15%. Since then, many companies are using deep 1331 learning including Facebook's tagging algorithms, Google for their photo search and 1332 Amazon for product recommendations. For the purpose of this thesis CNNs were 1333 used for image classification, specifically, images of varying particles created using 1334 LArTPC data.

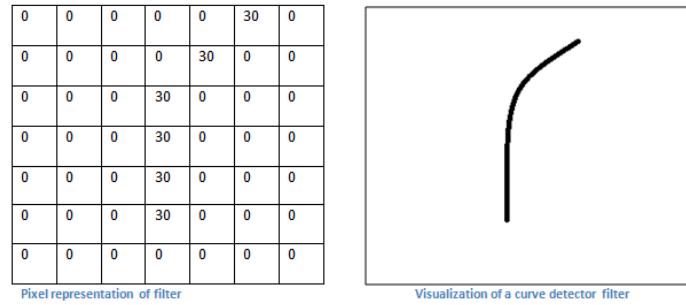
### 1336 6.1 Image Classification

1337 Image classification is the process of inputting an image into the CNN and receiving a 1338 probability of classes that best describes what is happening in the image. As humans, 1339 image classification is something that is learned at a very young age and is easy to 1340 do without much effort. This is also apparent when hand-scanning LArTPC images. 1341 After learning what a neutrino event looks like in MicroBooNE, it is relatively easy 1342 to recognize simple neutrino events from cosmic ray background as well as highly 1343 ionizing particles like protons from minimum ionizing particles like muons. The very 1344 detailed images LArTPC detectors output are prime candidates for input images into 1345 a CNN. CNNs mimic a human's ability to classify objects by creating an architecture 1346 that can learn differences between all the images it's given as well as figure out the

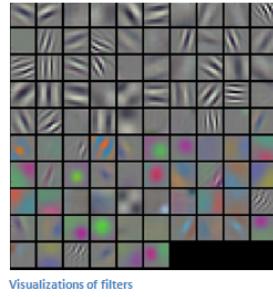
1347 unique features that make up each object. CNNs are modeled after the visual cortex.  
1348 Hubel and Wiesel [?] found that there are small regions of neuronal cells in the brain  
1349 that respond to specific regions of the visual field. They saw that some neurons fired  
1350 when exposed to vertical edges while others fired when shown horizontal or diagonal  
1351 edges. They also saw that these neurons were organized in columns. The idea of  
1352 specific neurons inside of the brain firing to specific characteristics is the basis behind  
1353 CNNs.

1354 **6.2 CNN Structure**

1355 When used for image recognition, convolutional neural networks consist of multiple  
1356 layers that extract different information on small portions of the input image. How  
1357 many layers is tunable to increase the accuracy. The output of these collections are  
1358 then tiled so that they overlap to gain a better representation of the original image  
1359 and allow for translation. The first of these layers is always a convolution layer. To  
1360 the CNN, an image is an array of pixel values. For a RGB color image with width  
1361 and height equal to 32x32 the corresponding array is 32x32x3. Filters, also known as  
1362 neurons, of any size set by the user is then convolved with the receptive field of the  
1363 image. If the filter is 5x5, the receptive field will be a patch of 5x5 on the input image.  
1364 The filter is also an array of numbers called weights. The convolution of the filter and  
1365 image are matrix multiplications of the weights and the pixel values. By stepping the  
1366 receptive field by 1 unit, for an input image of 32x32x3 and a filter of 5x5x3 you'd get  
1367 an output array of 28x28x1. This output array is called an activation map or feature  
1368 map. The use of more filters preserves the spatial dimensions better. The filters can  
1369 be described as feature identifiers. Examples of features in an image consist of edges,  
1370 curves, and changes in colors. The first filters in a CNN will primarily be straight  
1371 line and curve feature identifiers. An example of a curve filter is shown in figure 6.2.  
1372 When a curve in the same concavity is found in the input image, the corresponding  
1373 pixel in the output feature map will be activated. Going back to our example of a  
1374 32x32 input image and a 5x5 filter, if there were to be a curve in the top left corner  
1375 of the input image, our output feature map would have a high pixel value in the top  
1376 left. Therefore, feature maps tell us where a specific feature is located in the original  
1377 image. Figure ?? shows a visualization of filters found in the first layers of many CNN  
1378 architectures. These filters in the first layer convolve around the image and activate  
1379 when the specific feature it is looking for is in the receptive field.



**Figure 6.1:** Pixel representation and visualization of a curve detector filter. As you can see, in the pixel representation, the weights of this filter are greater along a curve we are trying to find in the input image

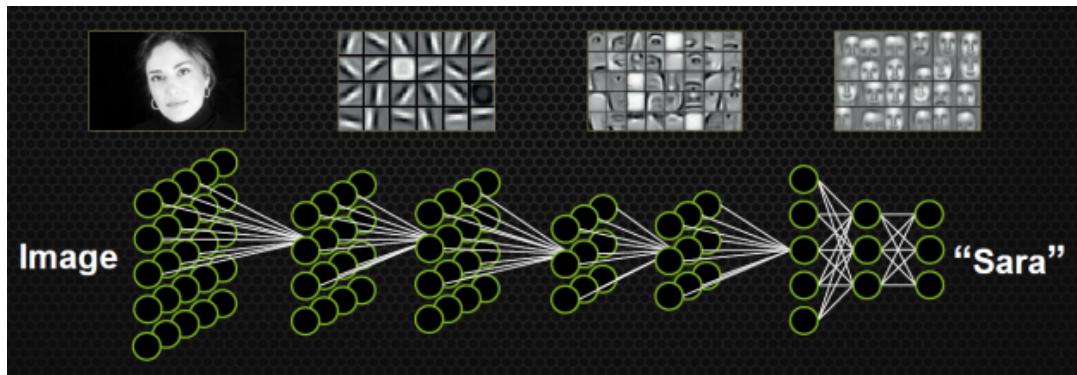


**Figure 6.2:** Visualization of filters found in first layer of a CNN.

1380 In figure 6.3 you can see how an edge detection filter is used to save only necessary  
 1381 information for recognizing different types of clothes. You can also see by having  
 1382 multiple filters you can get more detail or less detail from an image which can then  
 1383 simplify or complicate the object recognition task. Being able to distinguish between a  
 1384 shirt or a leg garment is as much information you want, having a filter that extracts  
 1385 outline edge or shape information would be all that you need. But if instead you  
 1386 wanted to distinguish between a formal cocktail dress or a summer dress, more  
 1387 information would need to be saved equating to many more filters for one image.  
 1388 Rather than trying to come up with how many filters and what features are important  
 1389 for detection, CNNs do this automatically. CNNs take input parameters, called  
 1390 hyperparameters, for example number of layers, number of filters per layers, number  
 1391 of weights per filter, and uses these to create the output feature maps. The layers build  
 1392 upon each-other, for example if we were creating a CNN for facial recognition the  
 1393 convolutional layers will start learning feature combinations off of the previous layers.  
 1394 The low level features like edges, gradients, and corners of the first layers become high  
 1395 level features like eyes, noses, and hairs. This process is visualized in figure 6.4



**Figure 6.3:** Applying a feature mask over a set of fashion items to extract necessary information for auto-encoding. Unnecessary information for example color or brand emblems are not saved. This feature map is an edge detection mask that leaves only shape information which helps to distinguish between different types of clothes.



**Figure 6.4:** Pictorial Representation of Convolutional Neural Networks as well as a visual representation on CNN's complexity of layer feature extraction

1396 There are other layers in a CNN architecture that will not be covered in the scope  
 1397 of this thesis but in a general sense, these layers are interspersed between convolution  
 1398 layers to preserve dimensionality and control overfitting of the network. The last layer  
 1399 is called a fully connected layer and it's job is to output an  $N$  dimensional vector where  
 1400  $N$  is the number of classes the network has been trained on. Each number in this vector  
 1401 represents the probability that the input image is a certain class. Fully connected layers  
 1402 use the feature maps of the high level features to compute the products between the  
 1403 weights of the previous layer to get the probabilities of each class. These weights are  
 1404 then adjusted through the training process using backpropagation.

---

### 1405 6.2.1 Backpropagation

1406 A CNN at it's onset has weights that are randomized. The filters themselves don't  
 1407 know how to pull out identifying information per class. For a neural network to learn,  
 1408 it must be trained on a training set that is labeled. Backpropagation has four seperate  
 1409 steps: foward pass, loss function, backward pass and updating weights. In the forward  
 1410 pass, a training image is passed through the whole network. All of our weights at this  
 1411 time are randomly initialized so the output for the first image will have no preference  
 1412 to a specific class. A common loss function is mean squared error (MSE):

$$E_{total} = \sum \frac{1}{2} (actualclass - predictedclass)^2 \quad (6.1)$$

1413 If we assume that the MSE is the loss of our CNN, the goal would be that our  
 1414 predicted label (output of CNN) is the same as our training label. To do this, we need  
 1415 to minimize the loss function. To do this, it is necessary to find out which weights most  
 1416 directly affect the loss of the network i.e  $\frac{dL}{dW}$  where L is our loss function and W are  
 1417 the weights of a specific layer. The next step is the backward pass which determines  
 1418 which weights contribute the most to the loss and finds ways to adjust these weights  
 1419 so that the loss decreases. After the derivative is computed, the last step updates the  
 1420 weights in the opposite direction of the gradient.

$$w = w_i - \eta \frac{dL}{dW} \quad (6.2)$$

$$w = \text{Weight} \quad (6.3)$$

$$w_i = \text{Initial Weight} \quad (6.4)$$

$$\eta = \text{Learning Rate} \quad (6.5)$$

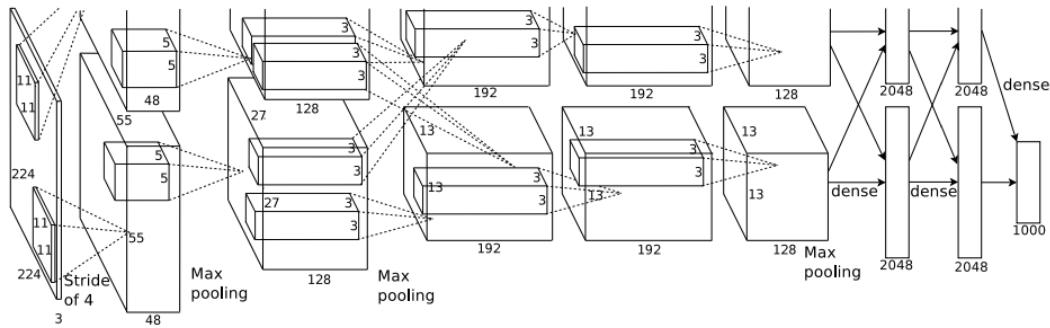
1421 The learning rate is a parameter given to the CNN and it describes the steps the  
 1422 network takes to update the weights. Higher learning rate equals large steps and a  
 1423 lower training time, but a learning rate that is too large can mean the CNN never  
 1424 converges.

1425 Going through backpropagation consists of one training iteration. Once the net-  
 1426 work completes a specific number of iterations, another parameter given, and runs  
 1427 over all training images that are split up into batches, the process is considered com-  
 1428 plete. User input parameters, called hyperparameters, help the network converge to

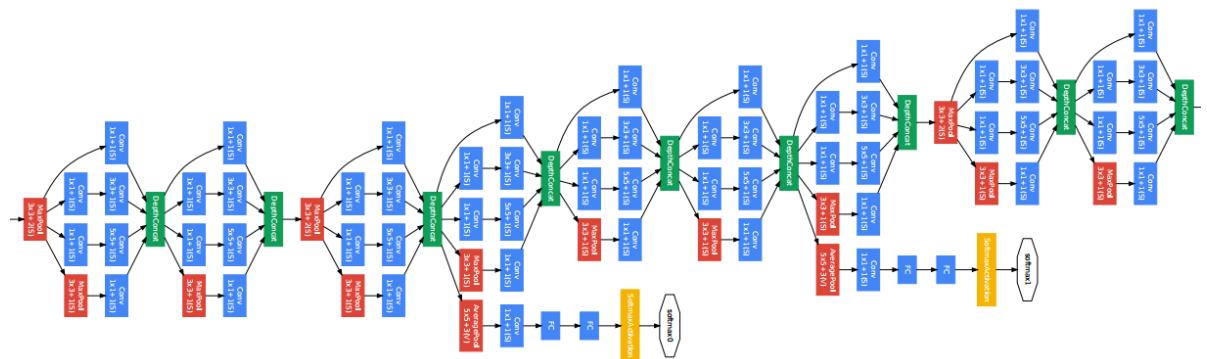
1429 optimal weights for each layer. Batch size, learning rate, and training iteration are just  
1430 some of the user input hyperparameters that help. Lastly, to check if the network has  
1431 learned, a different set of labeled images are fed to the CNN iteratively through the  
1432 training process to see how well it's learning. This process is especially important to  
1433 make sure the network architecture isn't being affected by overfitting (memorizing  
1434 training input rather than learning).

## 1435 6.3 Choosing Hyperparameters

1436 Convolutional neural networks are a relatively new tools in computer vision. Choosing  
1437 hyperparameters for your specific dataset is a non-trivial task. Hyperparameters can  
1438 range from the amount of layers and filters per layer in an CNN architecture to the  
1439 stride the receptive field of a filter takes, not to mention training hyperparameters  
1440 such as learning rate and batch size described above. They're ways to optimize these  
1441 hyperparameters via hyperparameter optimization using Bayesian Optimization [?]  
1442 but as you can imagine, optimizing an CNN architecture from scratch can be very  
1443 computationally intensive. For the purpose of this thesis, two well known CNN  
1444 architectures were used, AlexNet [?], which won the ImageNet Large-Scale Visual  
1445 Recognition Challenge (ILSVRC) in 2012 and therefore bringing awareness of CNNs,  
1446 and GoogleNet [?], which won the ILSVRC in 2014, giving rise to deep networks. Both  
1447 AlexNet and GoogleNet architectures were used to train on LArTPC images and their  
1448 low level filter weights. Higher level filter weights were randomly initialized before  
1449 training so the network can learn high level features of LArTPC image classes. The  
1450 AlexNet architecture is shown in figure 6.5 and the GoogleNet architecture is shown  
1451 in figure 6.6



**Figure 6.5:** Pictoral representation of the AlexNet model. The AlexNet model consists of 5 convolution layers and 3 fully connected layers.



**Figure 6.6:** Pictoral representation of the GoogleNet model. The GoogleNet model consists of 22 layers. The model implements 9 Inception modules which performs covolution and pooling in parallel and strays away from the basis that CNN layers need to be stacked up sequentially. The GoogleNet model also doesn't use fully connected layers, instead it uses average pooling which greatly reduces the amount of parameters. GoogleNet has 12x fewer parameters than AlexNet.



# <sup>1452</sup> Chapter 7

## <sup>1453</sup> **Training process of Convolutional 1454 Neural Networks**

<sup>1455</sup> Three Convolutional Neural Networks (CNNs) were trained throughout this analysis.  
<sup>1456</sup> There are differences to each CNN and will be described fully in the next sections but  
<sup>1457</sup> the main difference are the amount of particle images used for training and validation.  
<sup>1458</sup> CNN1075 used 1,075 muons and 1,075 pions for training and the same amount of  
<sup>1459</sup> each particle for validation. CNN10000 used 10,000 muons and 10,000 pions split in  
<sup>1460</sup> half for testing and training. Lastly CNN100000 had muons, pions, protons, electrons,  
<sup>1461</sup> and gammas in it's training and validation set. Each particle had 20,000 images and  
<sup>1462</sup> training and validation was split 90% training, 10% validation. This chapter will also  
<sup>1463</sup> describe the different hardware frameworks used for training beginning on a CPU  
<sup>1464</sup> and ending on a GPU cluster.

### <sup>1465</sup> **7.1 Hardware Configurations for Convolutional Neural 1466 Network Training**

<sup>1467</sup> The first training iteration, CNN1075, was a proof of concept. This CNN was trained  
<sup>1468</sup> on my local machine for  $\sim$  4-5 weeks. The batch size had to be very small as well as the  
<sup>1469</sup> image size due to the lack of computation resources. The second iteration of training,  
<sup>1470</sup> CNN10000, was trained on a Fermilab stationed Syracuse University machine. This  
<sup>1471</sup> machine had 6 TB of disk space, 6 cores at 2.1 GHz and 32 GB of RAM. The use of  
<sup>1472</sup> this machine allowed me to increase the training sample as well as the batch size and  
<sup>1473</sup> hence further increase the accuracy of the neural network. Lastly, the CNN100000 was

<sup>1474</sup> trained using two GTX 1080 Ti GPUs with 11GB of memory on a node on the Syracuse  
<sup>1475</sup> University GPU cluster, SUrge, that has 8 cores and 16GB of memory. This increase in  
<sup>1476</sup> memory as well as the capability to use 2 GPUs drastically cut down on training time  
<sup>1477</sup> from  $\sim$  4-5 weeks to  $\sim$  8 hours. SUrge also allowed for hyperparameter optimization  
<sup>1478</sup> by being able to run multiple training iterations over the two GPUs. Lastly, SUrge  
<sup>1479</sup> allowed for training over higher resolution images and a larger particle class of 5  
<sup>1480</sup> particles vs 2 particles.

<sup>1481</sup> **7.2 Creating images using LArTPC data for**  
<sup>1482</sup> **training/validation of CNNs**

<sup>1483</sup> The  $\mu/\pi$  image dataset used to train and validate CNN1075 was created using single  
<sup>1484</sup> generated isotropic muons and pions from 0-2 GeV energy range. 2,150 muons and  
<sup>1485</sup> 2,150 pions were used for training and testing split 50%. The images were created  
<sup>1486</sup> using LArSoft, a liquid argon software, and were based on wire number and time  
<sup>1487</sup> tick in the collection plane. Uboonecode reconstruction version v05\_08\_00 was used.  
<sup>1488</sup> The raw ADC value after noise filtering was the wire signal. Each collection plane  
<sup>1489</sup> greyscale image was 3456x1600x1 where 6 time ticks were pooled into 1 bin.

<sup>1490</sup> After the image was created, the region of interest (ROI) in the image was found by  
<sup>1491</sup> using Open CV, a image processing open source software package, to scan the image  
<sup>1492</sup> starting from the edges and stopping once a bright pixel is encountered. At this step,  
<sup>1493</sup> the ROI can be larger or smaller than the necessary size of a training image and the XY  
<sup>1494</sup> ratio of the image is not kept. This ROI is then resized to an image of 224x224x1.

<sup>1495</sup> The greyscale color standard is 8bit therefore the ADC value of wire and time tick  
<sup>1496</sup> was also downsampled due to the 12bit ADC value MicroBooNE has. To do this,  
<sup>1497</sup> the highest ADC pixel in the image was found and then this was divided by the rest  
<sup>1498</sup> placing all pixel values between 0-1. From there, all pixel values are then multiplied  
<sup>1499</sup> by 255.

<sup>1500</sup> The  $\mu/\pi$  image dataset used to train and validate the CNN10000 was also created  
<sup>1501</sup> using single generated isotropic muons and pions from 0-2 GeV energy range. 10,000  
<sup>1502</sup> muons and 10,000 pions were used for training and testing split 50%. Uboonecode  
<sup>1503</sup> v06\_23\_00 was used instead of v05\_08\_00. Each collection plane greyscale image was  
<sup>1504</sup> 3456x1280x1 where 5 time ticks were pooled into 1 bin which is different than the

1505 previous dataset and was implemented due to the fact that the time ticks of an event  
1506 went from 9400 to 6400 with the change of uboonecode version. Issues that arose in  
1507 CNN1075 that were fixed in CNN10000 include zero-padding images in X and Y that  
1508 are smaller than 224X224 to eliminate over-zooming effect and fixing a bug that shifted  
1509 pixels separated by a dead-wire region.

1510 The  $\mu/\pi/p/e/\gamma$  image dataset used to train and validate the CNN100000 were  
1511 created using single generated isotropic particles with energy range from 0-2 GeV.  
1512 20,000 of each particle were used for training and were split 90/10 between training  
1513 and testing sets. Uboonecode v06\_23\_00 was used for these images. The collection  
1514 plane greyscale iamge had the same dimensions as CNN10000, 3456x1280x1 and the  
1515 ROI algorithm was the same except for resizing these images to 576x576.

1516 A major change other than the higher resolution images was the treatment of the  
1517 ADC values. In the first two image making schemes, the highest pixel value was found  
1518 per image and the image was then normalized by that. The issue arising from this  
1519 ADC normalization wasn't inherent in  $\mu/\pi$  training due to the fact that both particles  
1520 are minimum ionizing particles in liquid argon, however, when dealing with a larger  
1521 particle class, it was necessary to try and make sure energy deposition by each particle  
1522 was preserved. The energy deposition in a particle image corresponds to the ADC  
1523 value or pixel brightness. To preserve energy deposition, the ADC float value was  
1524 passed straight to the image rather than doing any image normalization. This then  
1525 makes sure that minimum ionizing particles like muons and pions appear dimmer  
1526 than highly ionizing particles like protons.

1527 Images were also made from BNB+Cosmic events that passed the cc-inclusive  
1528 selection 1 filter right before the 75 cm track length cut and were classified using  
1529 the CNN10000. The dataset used to create these images is the same one used in  
1530 [?], *prodgenie\_bnb\_nu\_cosmic\_uboone\_mcc7\_reco2*. These images were created using  
1531 information from the track candidate that passed the filter. Only wire number and  
1532 time ticks associated to the track candidate were drawn on the image to mimic a single  
1533 particle generated image.

1534 These images were then classified using CNN10000. Two approaches were taken  
1535 in making these images. The first was using the image normalization above where  
1536 the maximum pixel in each image is used as a normalization constant to get all pixels  
1537 between 0-1 then multiply all pixels by 255. As described above, this is the incorrect

1538 way to normalize. The second way the images were created was by passing the ADC  
1539 float to the image. The results of CNN10000 performance are shown in section [7.1](#).

1540 Lastly, multiple BNB+Cosmic images per event were made for CNN100000 by  
1541 reducing many of selection I cuts to try and let the CNN do particle as well as event  
1542 identification. This image making scheme used for CNN100000 will be described in  
1543 more detail in later sections.

## 1544 7.3 Convolutional Neural Network Training

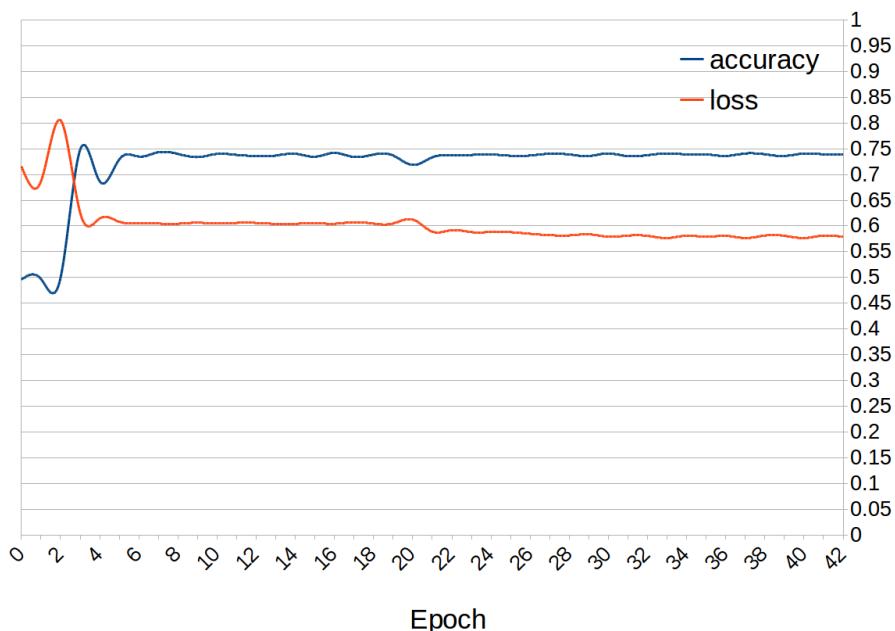
### 1545 7.3.1 Training CNN1075

1546 The results of CNN1075 are described in this section. The accuracy is how well  
1547 CNN1075 is doing by epoch and was 74.5%. The loss is gradient descent or mini-  
1548 mization of the error of the weights and biases used in each neuron of each layer of  
1549 CNN1075 and was 58% with a trend sloping downwards on the loss curve as well as a  
1550 trend sloping upward in the accuracy curve. The accuracy and loss of CNN1075 are  
1551 shown in figure [7.1](#). Due to the depth of the neural network framework, it was neces-  
1552 sary to train with a larger dataset and for more epochs, however, the downward slope  
1553 of the loss curve is an indication that once trained for longer with a higher training  
1554 sample, neural networks can be used for  $\mu/\pi$  separation. The hyperparameters used  
1555 to train CNN1075 are detailed below:

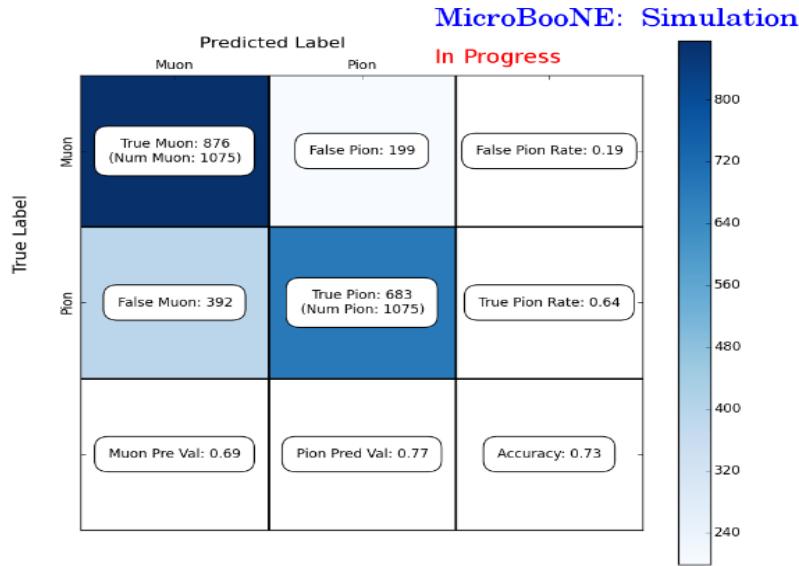
- 1556     • *train\_batch\_size*: 50    1561     • *lr\_policy*: "step"           1566     • *momentum*: 0.9
- 1557     • *test\_batch\_size*: 50    1562     • *gamma*: 0.1                   1567     • *weight\_decay*: 0.0005
- 1558     • *test\_iter*: 50            1563     • *stepsize*: 200                1568     • *snapshot*: 100
- 1559     • *test\_interval*: 50      1564     • *display*: 50
- 1560     • *base\_lr*: 0.01          1565     • *max\_iter*: 5000

1569 The confusion matrices shown in figure [7.2](#) show the accuracy for both the training  
1570 and testing datasets. The fact that these two have similar accuracies is important  
1571 because if the training dataset had a much higher accuracy, that indicates an over-  
1572 training of the training sample which means the neural network didn't learn features  
1573 to separate muons from pions, it just memorized what was in the training dataset.

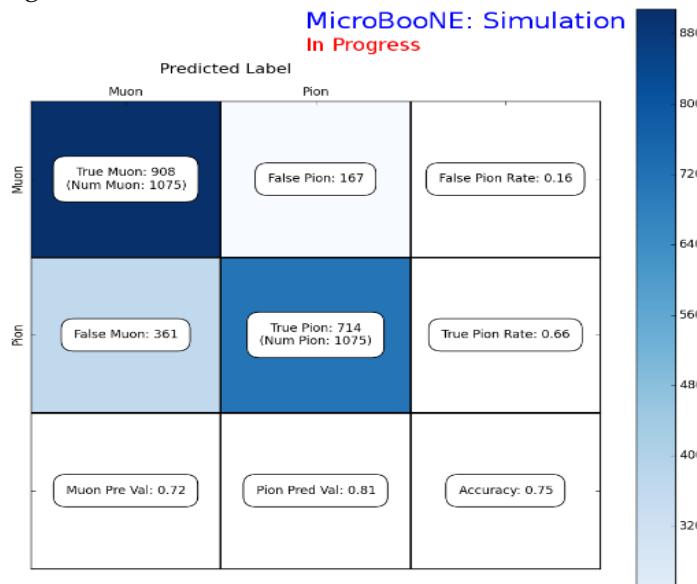
Loss/Accuracy of CNN trained on 2150 images



**Figure 7.1:** Accuracy vs. Loss of AlexNet 2-output  $\mu/\pi$  sample consisting of 2,150 images each.



**(a)** Confusion Matrix showing Accuracy of CNN1075 using training MC data



**(b)** Confusion Matrix showing Accuracy of CNN1075 using testing MC data

**Figure 7.2:** Description of confusion matrix variables: False pion rate =  $false\pi / total\pi$  True pion rate =  $true\pi / total\pi$  Accuracy =  $(true\pi rate + true\mu rate) / 2$  Pion prediction value =  $true\pi / (true\pi + false\pi)$  Muon prediction value =  $true\mu / (true\mu + false\mu)$

1574 Also note that the neural network does a better job of identifying muons than pions.  
1575 This can be attributed to the more complex event scenes pions tend to leave in the  
1576 detector due to pion interacting more in LAr than muons do. The CNN may do better  
1577 at identifying pions with a larger training sample.

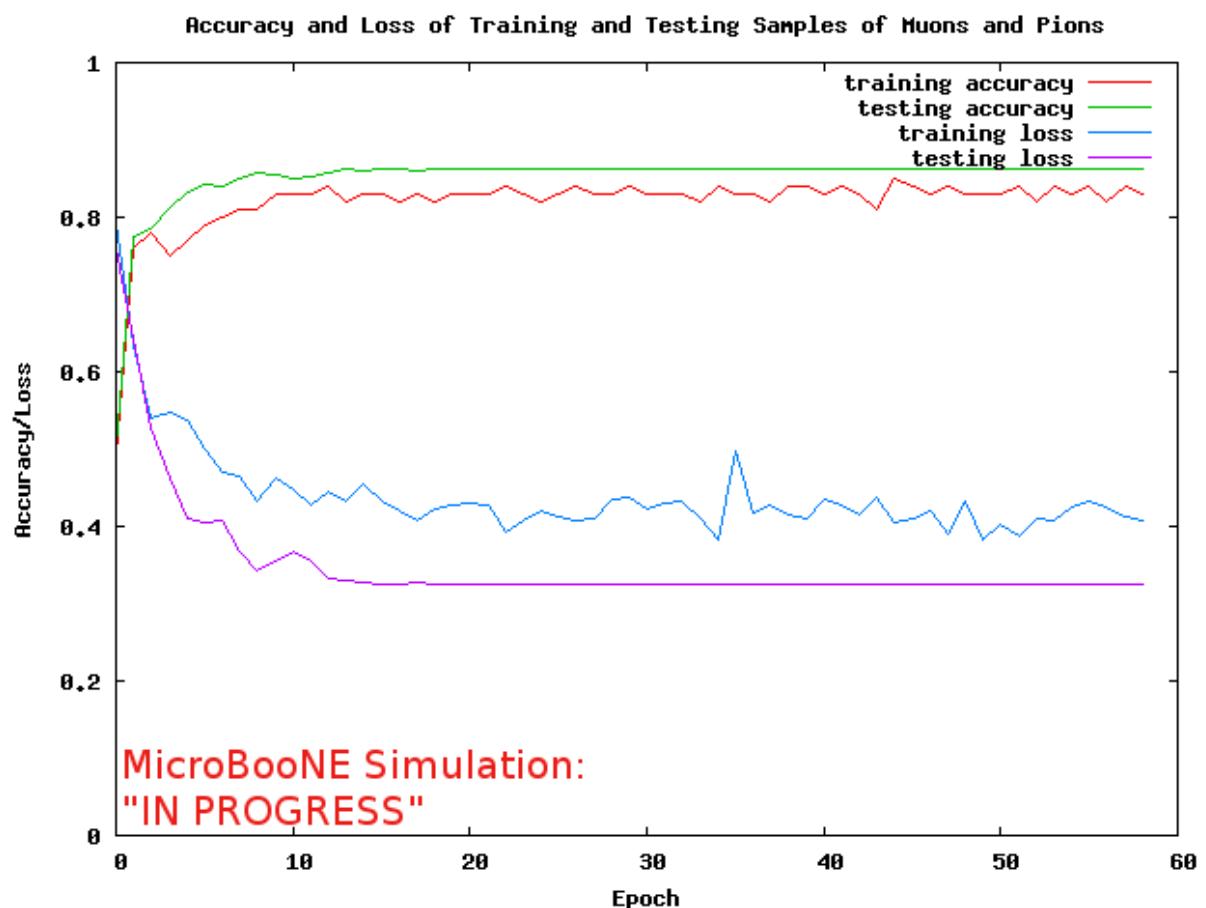
1578 **7.3.2 Training CNN10000**

1579 The hyperparameters used for CNN10000 are shown below. The batch size for the  
1580 training and testing as well as the test\_iter were chosen to encompass the whole  
1581 training/testing image set when doing accuracy/loss calculations. To do this, multi-  
1582 plying the test\_iter by the test batch size gives you the amount of images used when  
1583 calculating accuracy/loss curves.

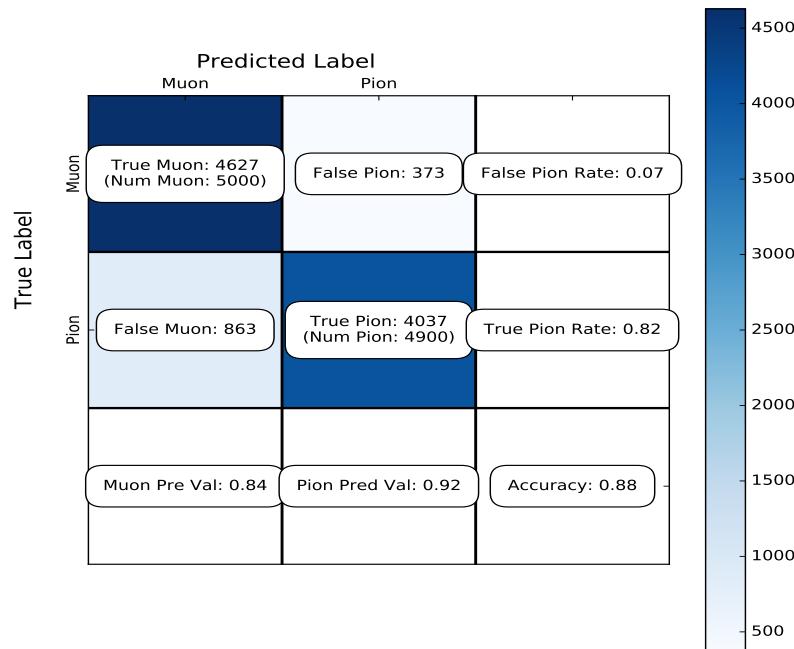
- 1584     • *train\_batch\_size*: 100<sub>1589</sub>     • *lr\_policy*: "step"     1594     • *momentum*: 0.99
- 1585     • *test\_batch\_size*: 100<sub>1590</sub>     • *gamma*: 0.1     1595     • *weight\_decay*: 0.0005
- 1586     • *test\_iter*: 100<sub>1591</sub>     • *stepsize*: 1000     1596     • *snapshot*: 100
- 1587     • *test\_interval*: 100<sub>1592</sub>     • *display*: 100
- 1588     • *base\_lr*: 0.001<sub>1593</sub>     • *max\_iter*: 10000

1597     The same architecture that was used to train CNN1075 was employed on CNN10000,  
1598 AlexNet. Caffe [?] was the software package used for both CNNs. The differences  
1599 include batch size and test\_iter and momentum to account for the larger dataset. Fig-  
1600 ure 7.3 shows the loss and accuracy of CNN10000. There is around a 10% increase in  
1601 accuracy from CNN1075 to CNN10000, 85%, and around a 20% decrease in loss, 36%.

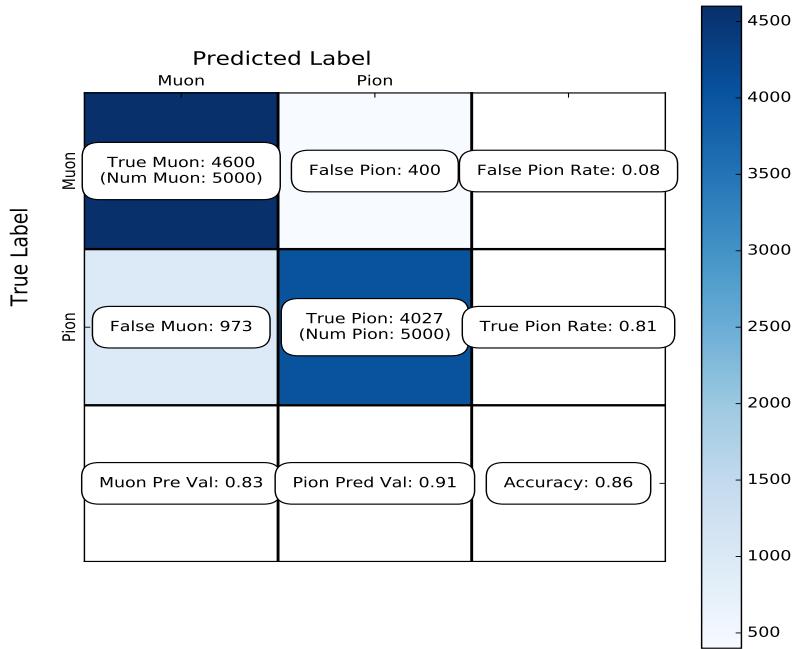
1602     Figure 7.4 show a breakdown of  $\mu/\pi$  separation for CNN10000. It also shows  
1603 the network is not being over-trained due to the Accuracy of both the training and  
1604 testing datasets being within .01% of each-other. Figure 7.5 shows how well the neural  
1605 network is doing at  $\mu/\pi$  separation with respect to muon probability. The red bins  
1606 corresponds to true pions and the blue bins correspond to true muons. There is  
1607 still pion contamination in the high muon probability bins but by choosing a muon  
1608 probability of  $\geq 80\%$  we can reduce this. The CNNs increase in total accuracy can be  
1609 attributed to an increase in accurately classifying pions as pions as seen in both the  
1610 confusion matrix in figure 7.4 and the large number of events in the zero bin of the  
1611 muon probability plot seen in figure 7.5 that corresponds to high probability pions.



**Figure 7.3:** Accuracy vs. Loss of AlexNet 2-output  $\mu/\pi$  sample consisting of 10,000 images each.

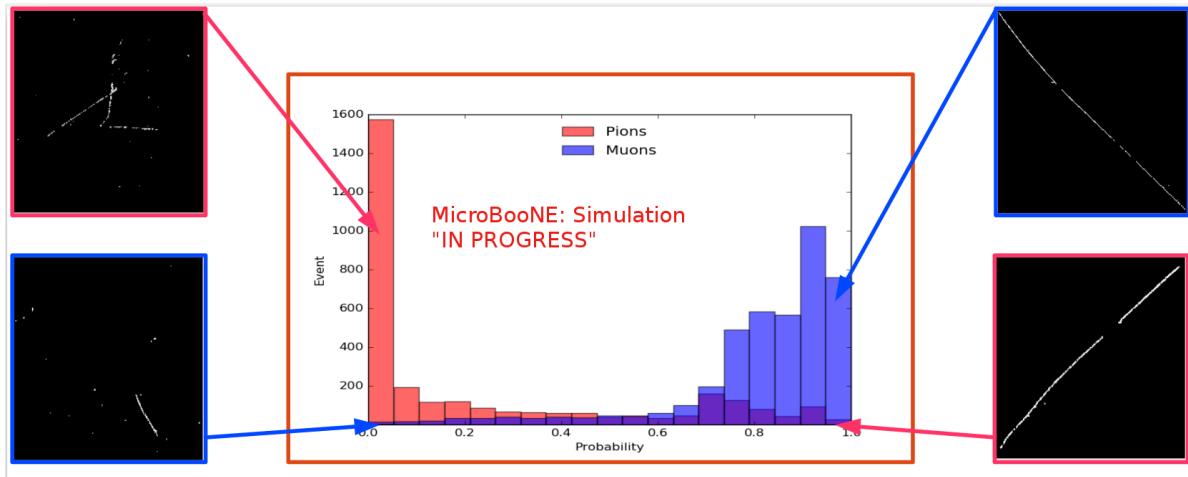


**(a)** Confusion Matrix showing Accuracy of CNN10000 using training MC data



**(b)** Confusion Matrix showing Accuracy of CNN10000 using testing MC data

**Figure 7.4:** Description of confusion matrix variables: False pion rate =  $false\pi / total\pi$ ; True pion rate =  $true\pi / total\pi$ ; Accuracy =  $(true\pi rate + true\mu rate) / 2$ ; Pion prediction value =  $true\pi / (true\pi + false\pi)$ ; Muon prediction value =  $true\mu / (true\mu + false\mu)$



**Figure 7.5:** Probability plot of muons and pions from testing set. Images surrounding histogram are a random event from lowest bin and highest bin for each particle.

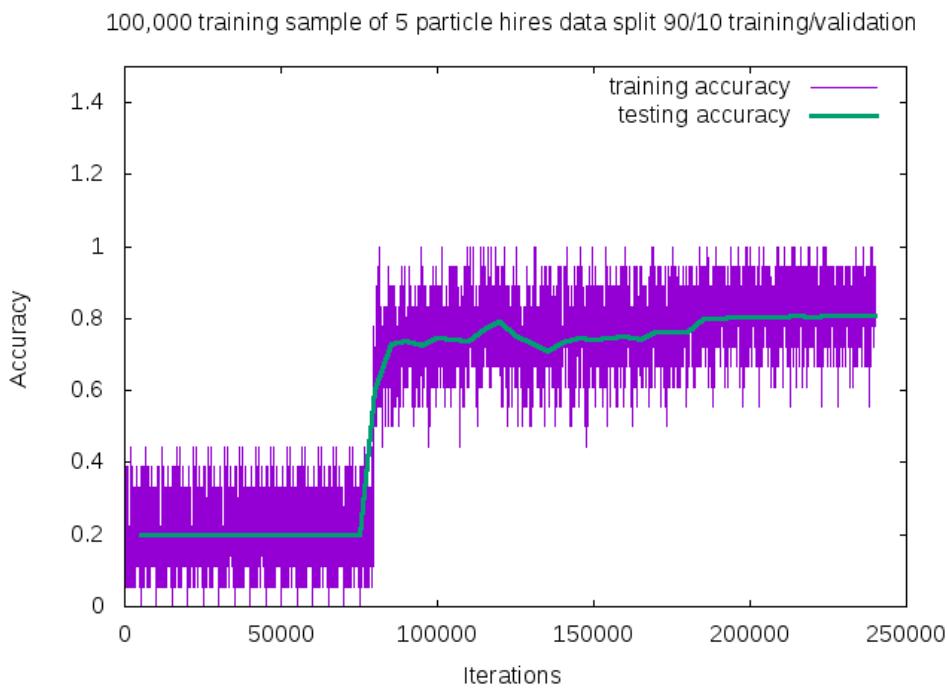
### 1612 7.3.3 Training CNN100000

1613 CNN100000 used the GoogleNet architecture rather than the AlexNet architecture  
 1614 used in the two previous trained CNNs. This is the first time the neural network was  
 1615 trained on a larger particle class,  $\mu/\pi/p/\gamma/e$ , and on higher resolution images. This  
 1616 CNN also employed GPUs during the training process. The hyperparameters are  
 1617 shown below:

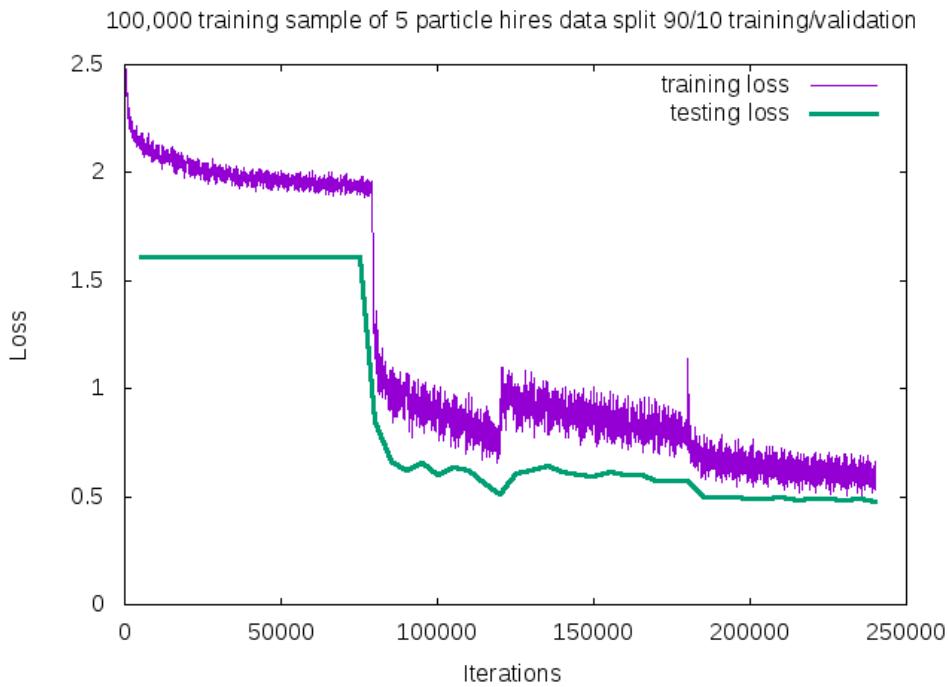
- 1618     • *train\_batch\_size*: 18     1623     • *lr\_policy*: "step"     1628     • *max\_iter*: 10000
- 1619     • *test\_batch\_size*: 2     1624     • *gamma*: 0.96     1629     • *momentum*: 0.99
- 1620     • *test\_iter*: 2000     1625     • *stepsize*: 10000     1630     • *weight\_decay*: 0.0002
- 1621     • *test\_interval*: 2000     1626     • *average\_loss*: 40     1631     • *snapshot*: 50000
- 1622     • *base\_lr*: 0.001     1627     • *display*: 40

1632     The accuracy and loss for CNN100000 are shown in figures 7.6 and 7.7. The jumps  
 1633 shown in both figures are when the training was stopped to fine-tune the weight decay  
 1634 and the learning rate. The accuracy leveled off at  $\sim 80\%$  and the loss was at  $\sim 0.48$ .

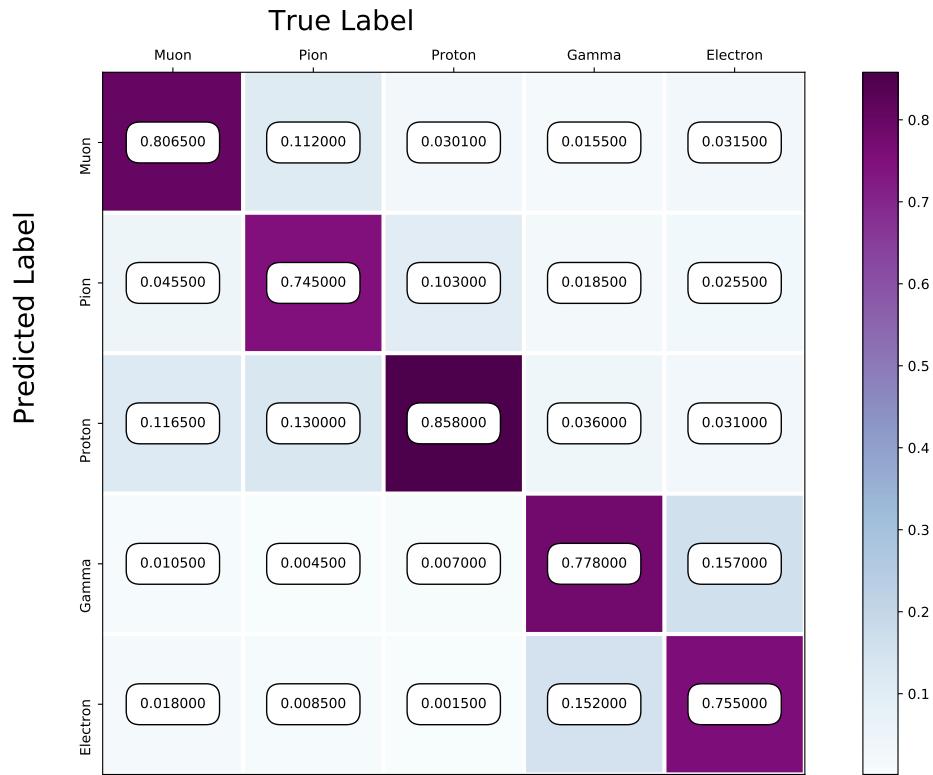
1635     Figure 7.8 shows the confusion matrix of CNN100000. The proton identification of  
 1636 the neural network is at 85% and the highest out of all five particles. One thing to note  
 1637 is clear separation between particles that leave track like objects in the MicroBooNE  
 1638 detector,  $\mu/\pi/p$ , versus particles that leave shower like objects in MicroBooNE,  $e/\gamma$ .



**Figure 7.6:** Training and testing accuracy of CNN trained on 100,000 images of  $\mu/\pi/p/\gamma/e$  with 20,000 images of each particle. Each image was a size of 576x576 and the images per particle were split 90% use for training and 10% used for testing the network



**Figure 7.7:** Training and testing loss of CNN trained on 100,000 images of  $\mu/\pi/p/\gamma/e$



**Figure 7.8:** Confusion Matrix of all five particles

Another visualization of how the neural network is learning is shown in 7.9. t-SNEs [?] is a technique used for dimensionality reduction developed for use in visualizing high-dimensional datasets. Each datapoint is given a location in a two or three-dimensional map by using stochastic neighbor embedding to convert high-dimensional euclidean distances between datapoints into conditional probabilities that represent the similarities between these datapoints. For datapoints close together on the map, their conditional probabilities are high, for datapoints with a wide separation between them, their conditional probabilities are very small. Figure 7.9 is a t-SNE of the final training iteration of a subset of the training sample used in CNN100000. You can see a clear separation between track like objects and shower like objects. You can also see that electrons and gammas are not as separated as muons, pions, and protons. For the purpose of this thesis, this isn't an issue but later iterations of training could include more images for the gamma and electron classes to help the CNN further separate these classes.

Figure 7.10 shows the probability of each particle class and the highest probability misidentification for each class. For muons, the largest misidentification is from

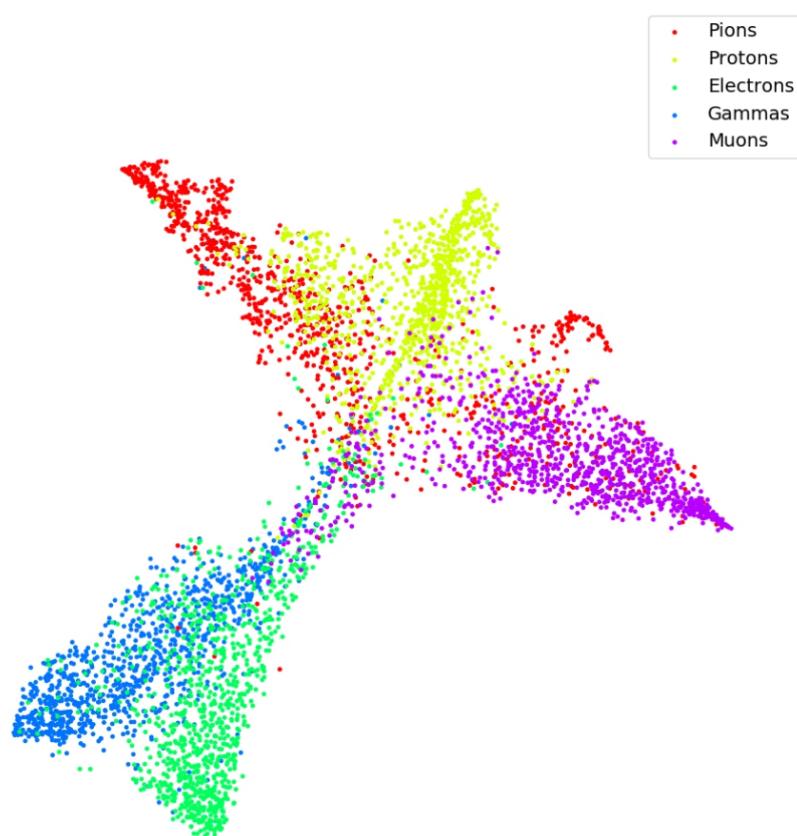
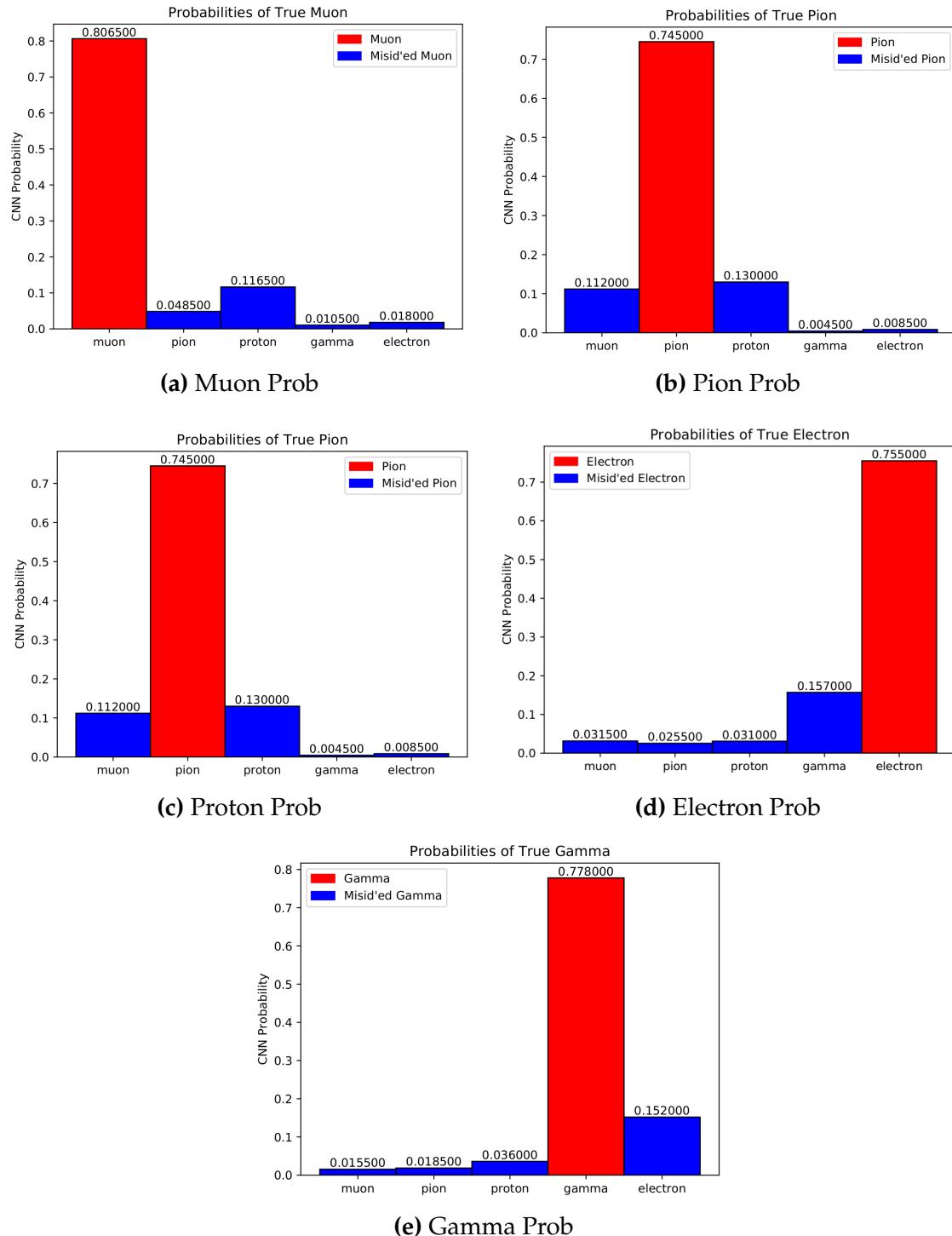
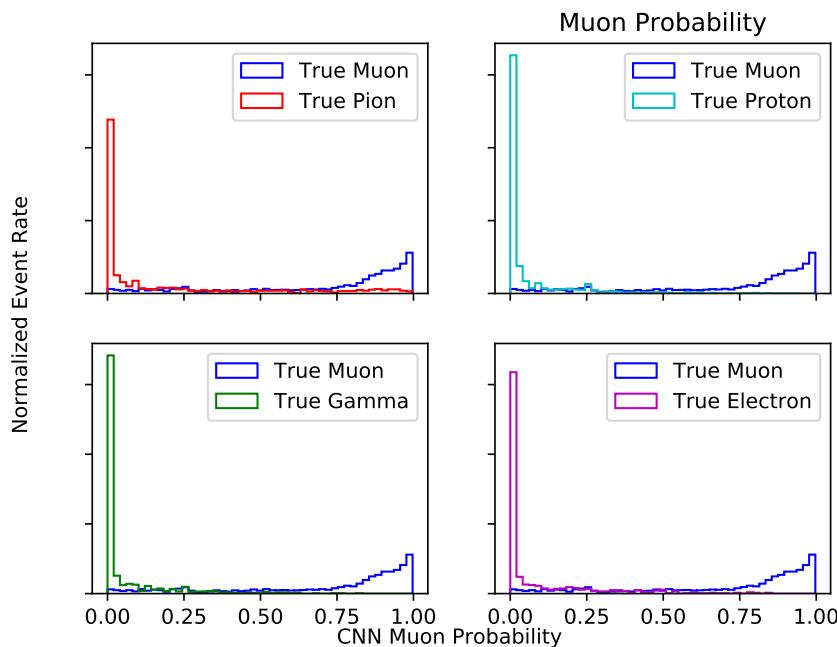


Figure 7.9: t-SNE of CNN



**Figure 7.10:** Probabilities of different particle classes as well as their contamination from other classes

1655 protons. For pions, both protons and muons get misidentified as pions at around the  
 1656 same probability. Similar behavior is also seen for proton identification. Electrons and  
 1657 gammas are misidentified as each-other with similar probabilities.

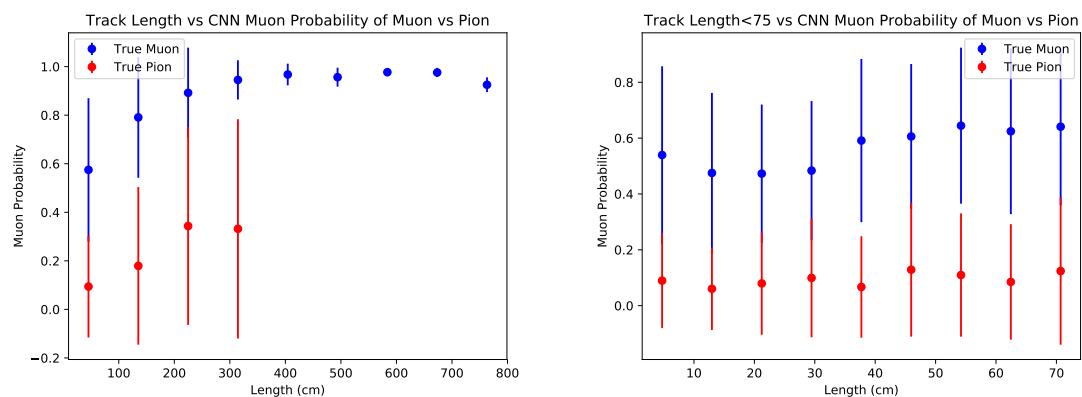


**Figure 7.11:** Muon probability of true muons (blue) versus pions (red), protons (cyan), gammas (green) and electrons (magenta).

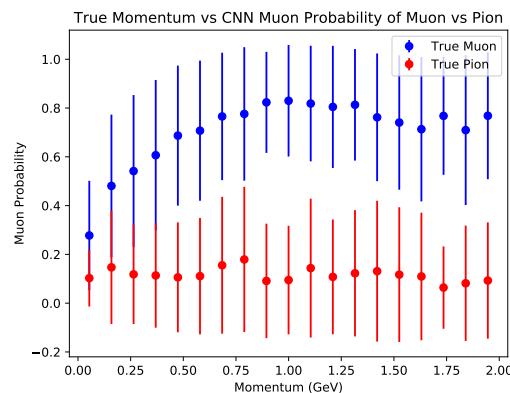
1658 To see what type of background contamination one would be dealing with when  
 1659 doing muon identification, muon probabilities for each particle class was plotted  
 1660 against the probability of true muons to see how well muon signal vs other particle  
 1661 background separation can be done with CNN100000. Figure 7.11 is showing the  
 1662 true muon probability for true muons, versus the rest of the particle classes. This plot  
 1663 describes which muon probability value should be chosen for the least amount of  
 1664 other particle contamination. For electrons and gammas, a muon probability of  $\sim 75\%$   
 1665 would eliminate  $e/\gamma$  contamination. For pions and protons, there is contamination at  
 1666 all values of muon probability, but the contamination is drastically reduced at a muon  
 1667 probability  $\geq 75\%$ .

1668 One of the main concerns with training a neural network was that the features the  
 1669 network would learn to separate muons from pions would be track range, which is  
 1670 what was used to begin with in selection I. To make sure that wasn't the case, the next  
 1671 thing that was looked at was the muon probability versus track range and momentum

of the track. Figures 7.12 through 7.15 show the muon probability in blue for all plots against all other particles. A zoomed in version of track range for all particles was also plotted to make sure there is separation between the particles at low track range. The  $\mu/\pi$  separation in track range and momentum is less than for  $p/e/\gamma$  but that was to be expected. Although the separation isn't as good as the other particles, there still is separation at low momentum and low track range which cannot be done by using a track range cut like selection I does.

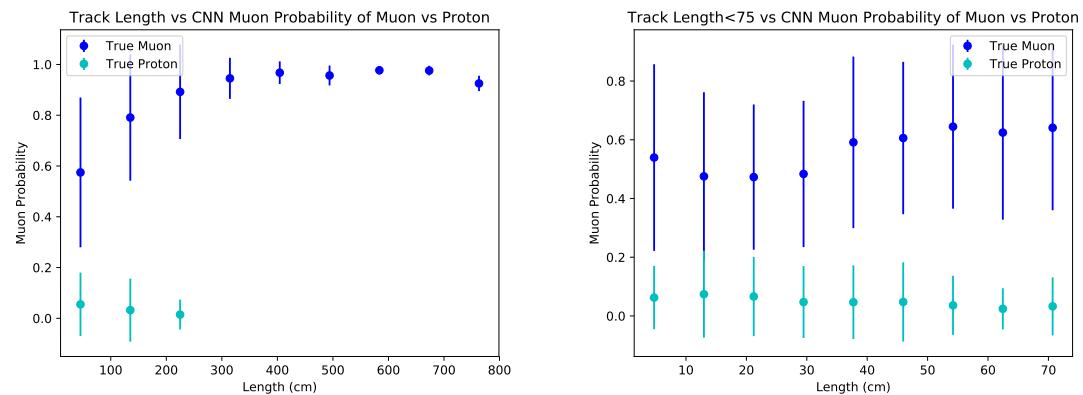


(a) Track range versus muon probability for true muons (blue) and true pions (red). (b) Track range  $\leq 75$  cm versus muon probability for true muons (blue) and true pions (red).

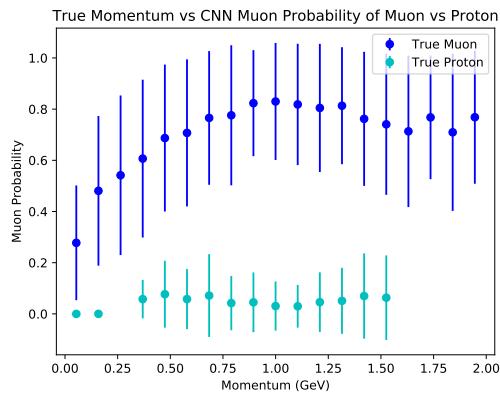


(c) Momentum versus muon probability for true muons (blue) and true pions (red).

**Figure 7.12:** Kinematic distributions versus muon probability for true muons and true pions.

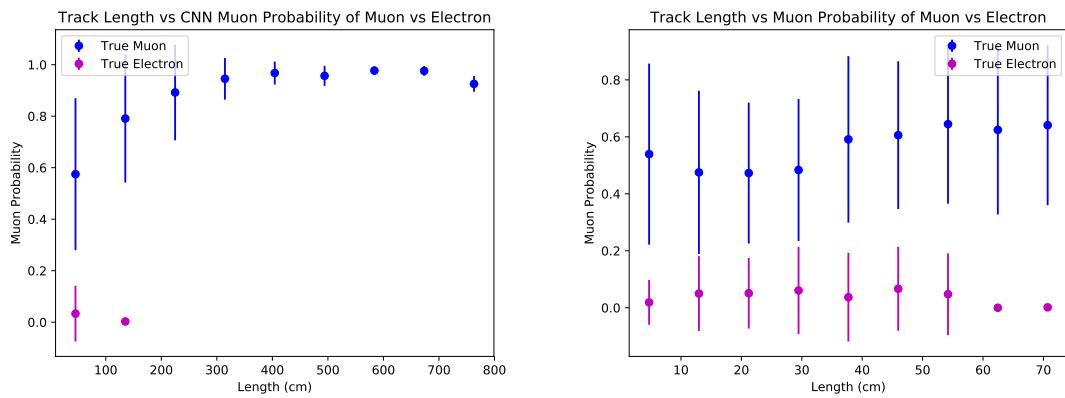


(a) Track range versus muon probability for true muons (blue) and true protons (cyan). (b) Track range  $\leq 75$  cm versus muon probability for true muons (blue) and true protons (cyan).

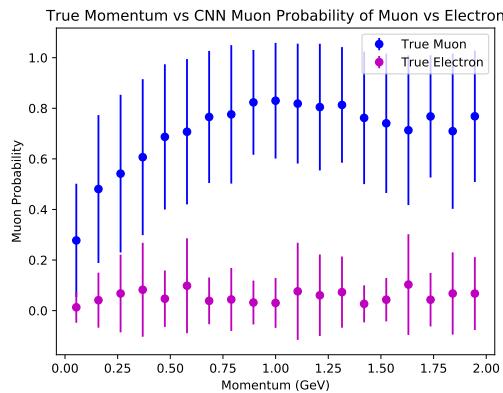


(c) Momentum versus muon probability for true muons (blue) and true protons (cyan).

**Figure 7.13:** Kinematic distributions versus muon probability for true muons and true protons.

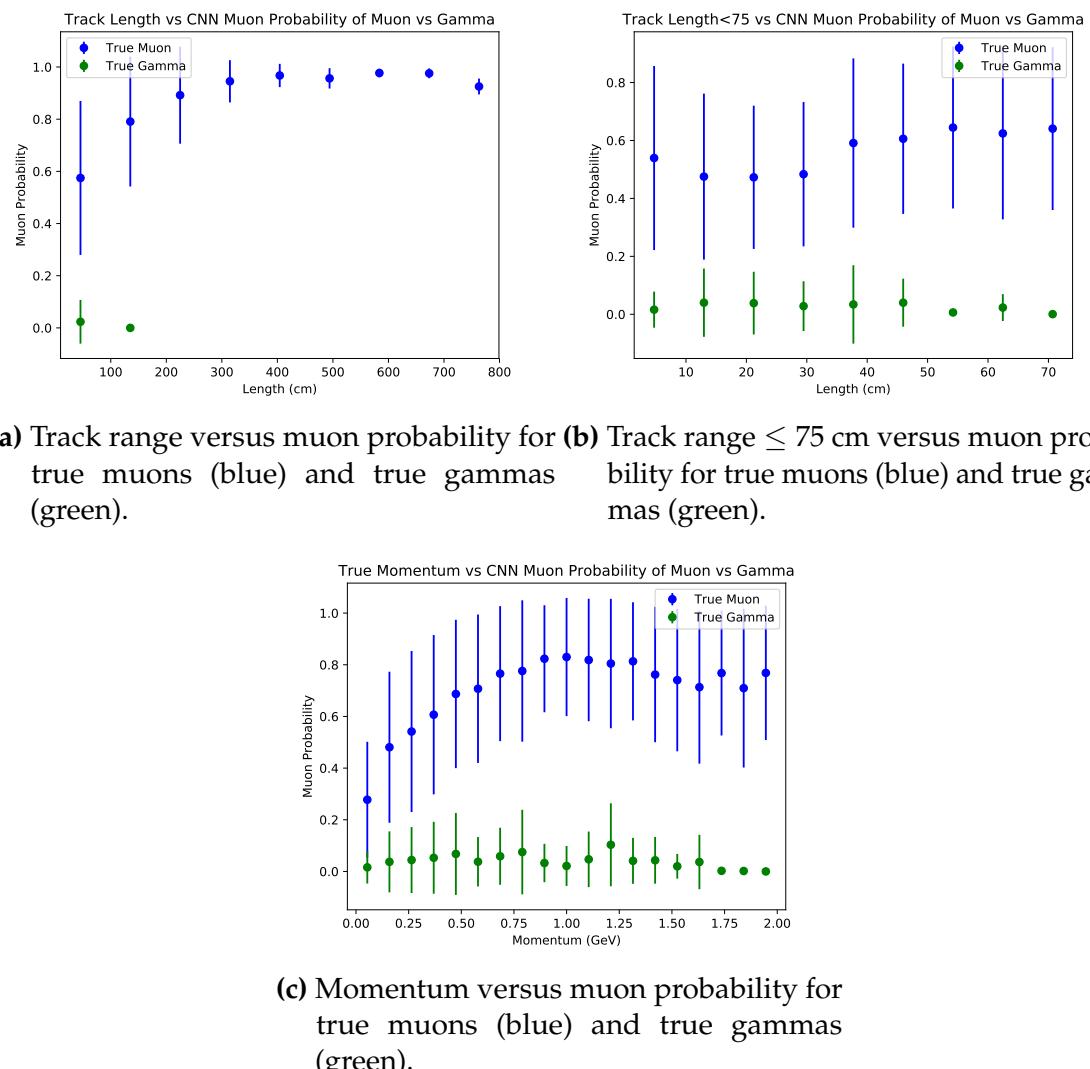


- (a) Track range versus muon probability for true muons (blue) and true electrons (magenta).
- (b) Track range  $\leq 75$  cm versus muon probability for true muons (blue) and true electrons (magenta).



- (c) Momentum versus muon probability for true muons (blue) and true electrons (magenta).

**Figure 7.14:** Kinematic distributions versus muon probability for true muons and true electrons.



**Figure 7.15:** Kinematic distributions versus muon probability for true muons and true gammas.



1679 **Chapter 8**

1680 **Using Convolutional Neural Networks  
1681 for  $\nu_\mu$  CC event classification**

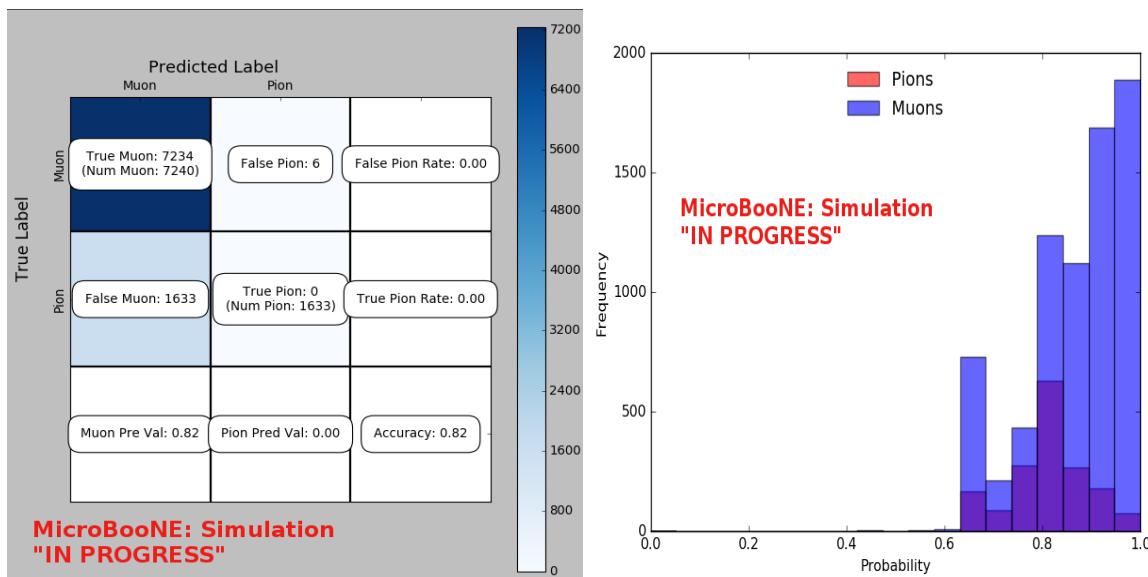
1682 **8.1 Classification using CNN10000**

1683 **8.1.1 Classification of MC data using Selection I CC-Inclusive Filter**

1684 CNN10000 was also used to classify track candidate images that were identified by  
1685 the selection I cc-inclusive filter described in [?]. Passing rates for each cut in this filter  
1686 are shown in figure ???. As seen in section ???, wrong image normalization had a higher  
1687 muon classification probability so all work done using selection I cc-inclusive filter  
1688 was done using this normalization. Out of 188,880 events, 19,112 passed the cut right  
1689 before the 75 cm track length cut which is a 10.1% passing rate and comparable to  
1690 the 10% passing rate shown in figure ???. In time cosmics were also run over, out of  
1691 14,606 in time cosmics events, 302 passed the cut right before the 75 cm track length  
1692 cut which is a 2.1% passing rate comparable to the 2.7% passing rate in the cc-inclusive  
1693 tech-note. Figures 8.1a and 8.1b show the accuracy and  $\mu/\pi$  separation. Both plots  
1694 are only composed of muons and pions and like selection I original data, all other  
1695 particles were id'ed as muons. Also like selection I original data, muons are being  
1696 identified at a very high rate. Figure 8.2a shows the track range distributions of all  
1697 events from selection I being classified by the CNN as a muon with a probability of  
1698 70% regardless of true particle type. We get entries for the CNN curve in the lowest  
1699 bin and none for the 75 cm curve. To see how many true CC events were identified  
1700 by CNN10000 breaking down figure 8.2a by event type was necessary. Figures 8.2b  
1701 and 8.2c show track range distributions separated by signal and various backgrounds.

1702 Particle type was not taken into consideration in these plots so true CC event images  
 1703 can be any track candidate particle passing selection I cut right before track length cut  
 1704 including pions and protons.

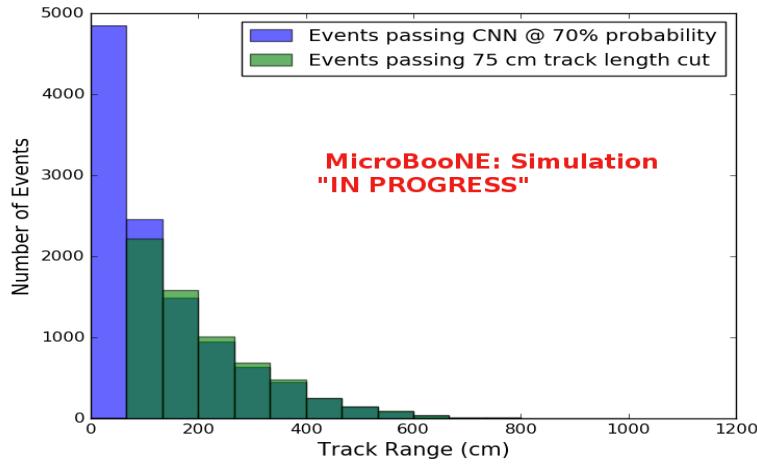
1705 To gain an even deeper understanding on how CNN10000 is performing, plotting  
 1706 these distributions with only muons and pions was done due to the fact that CNN10000  
 1707 was trained with only those particles for  $\mu/\pi$  separation. Figures 8.2d-8.3d show the  
 1708 stacked histograms of signal and background of the track range distributions with  
 1709 varying CNN probabilities starting from 70% and ending at 90% probability. With  
 1710 higher probabilities we get a purer sample in the lower bin but we end up losing  
 1711 events as well. Momentum distributions for all signal/background events are shown  
 1712 in figure 8.4.



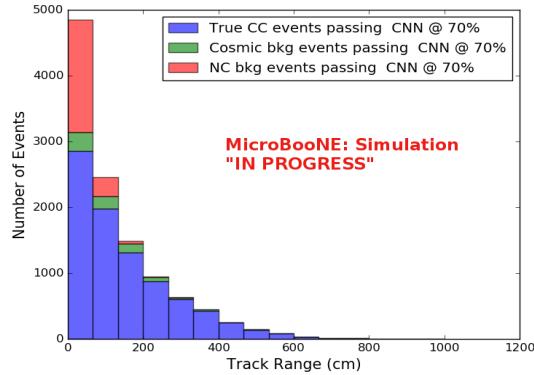
(a) Confusion Matrix for CNN10000 classified events from selection I      (b) Probability plot for CNN10000 classified events from selection I

**Figure 8.1:** Confusion matrix and probability plot of events passing selection I cc-inclusive cuts right before 75cm track length cut

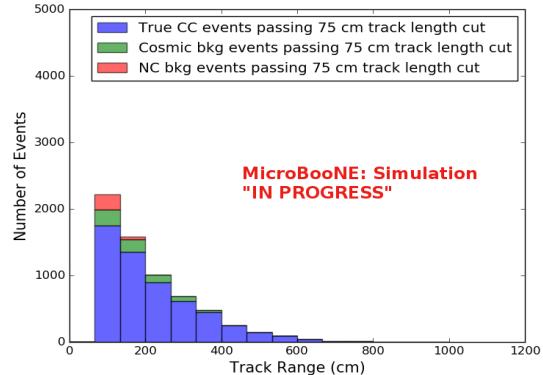
1713 Another check was to see if any true CC pions were passing through the cut right  
 1714 before the 75 cm track length cut. Figure 8.5 shows the comparison of the stacked track  
 1715 range distribution with only true CC muon signal versus the stacked distribution with  
 1716 true CC muons and pions signal. As you can see, we gain more events when plotting  
 1717 CC events with a particle type of either muons or pions due to the CNN classifying  
 1718 all pions in this dataset as muons. This is an interesting scenario and a sample of  
 1719 topologies of these images are represented in figure 8.6, at least 3 tracks are coming out



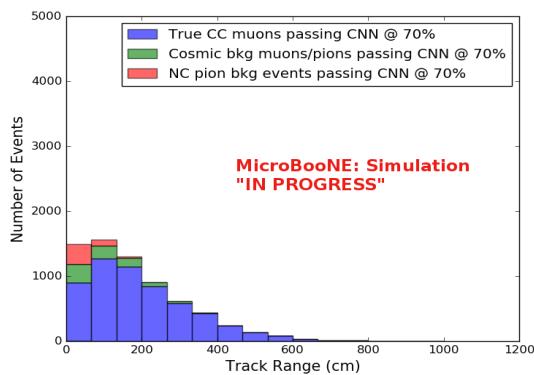
(a) Track range distribution of events from Selection I Modified passing CNN with 70% accuracy



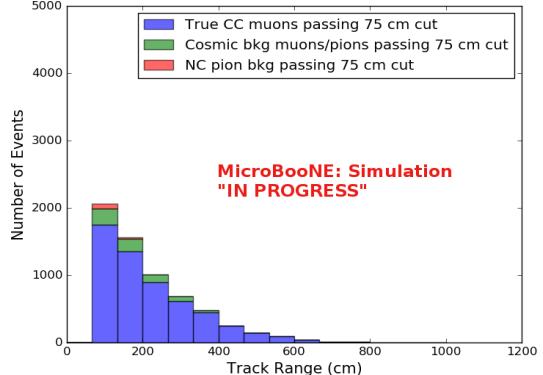
(b) Stacked signal and background track range distributions from Selection I Modified passing CNN with 70% accuracy



(c) Stacked signal and background track range distributions from Selection I Modified passing 75 cm track length cut

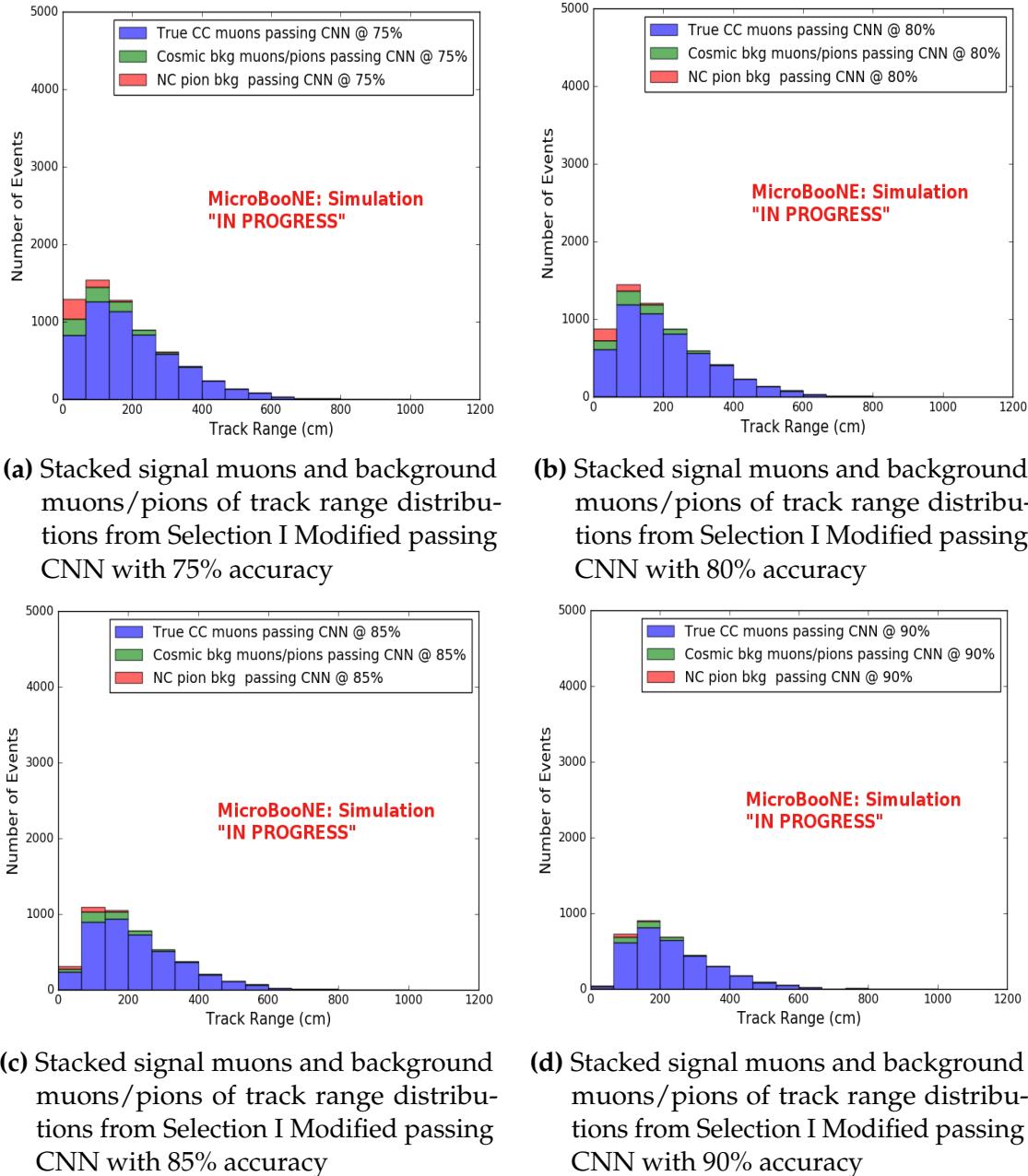


(d) Stacked signal muons and background muons/pions of track range distributions from Selection I Modified passing CNN with 70% accuracy

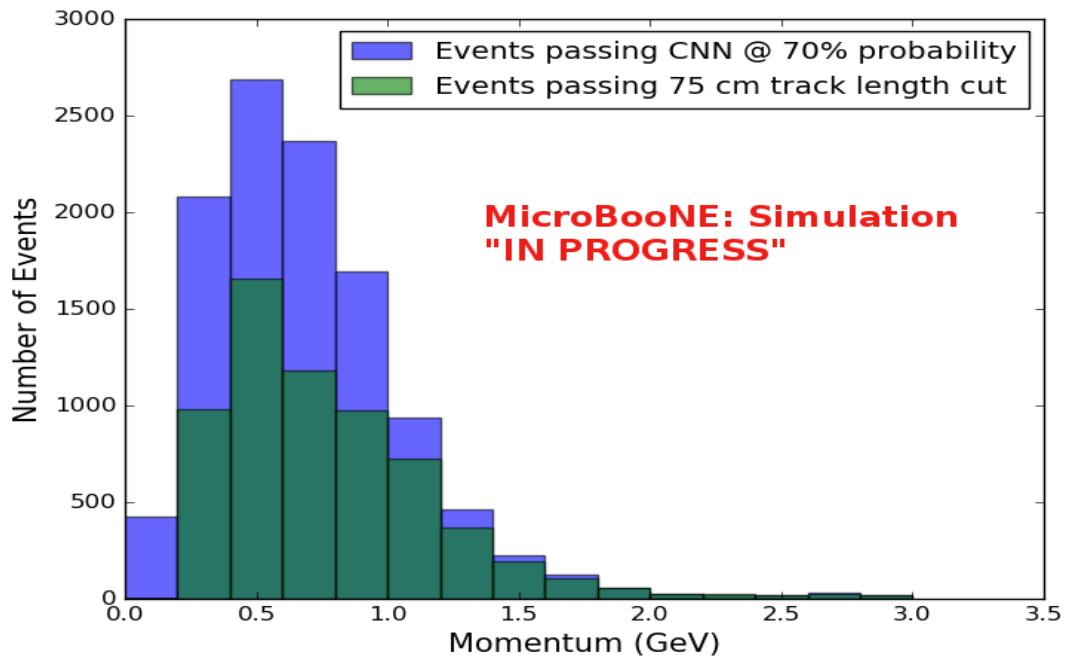


(e) Stacked signal muons and background muons/pions of track range distributions from Selection I Modified passing 75 cm track length cut

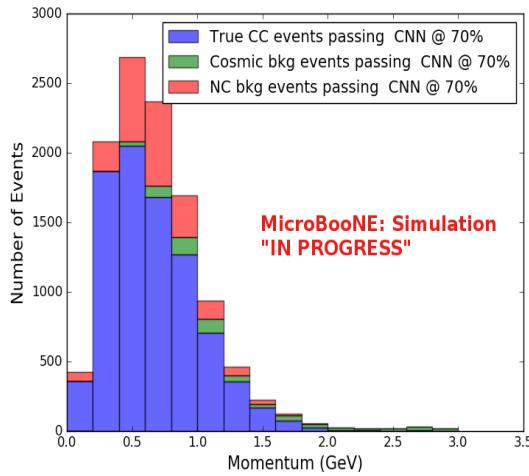
**Figure 8.2:** CNN10000 distributions of track candidate images output from Selection I Modified cc-inclusive filter



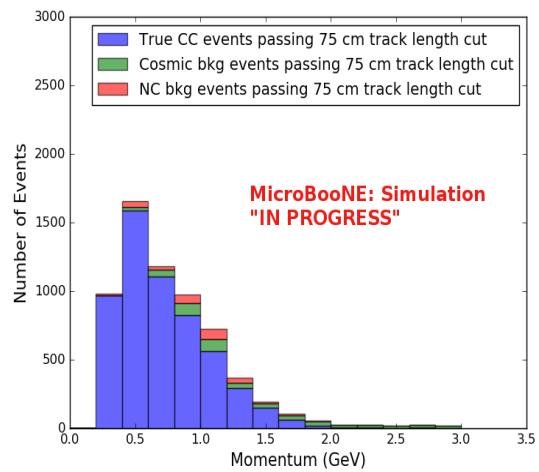
**Figure 8.3:** CNN10000 stacked signal/background track range distributions of track candidate images output from Selection I Modified cc-inclusive filter



(a) Momentum distribution of events from Selection I Modified passing CNN with 70% accuracy



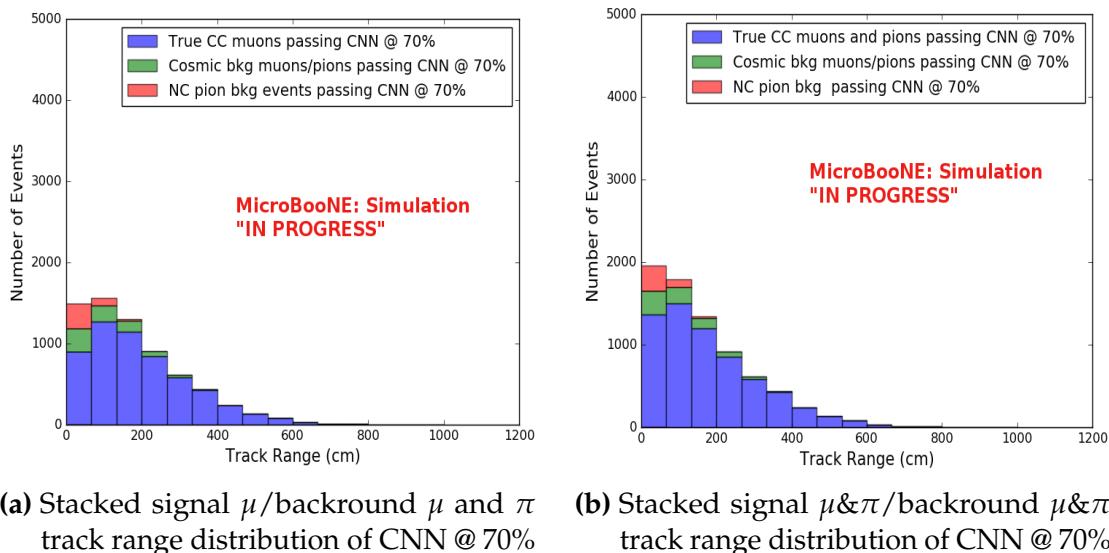
(b) Stacked signal and background momentum distributions from Selection I Modified passing CNN with 70% accuracy



(c) Stacked signal and background momentum distributions from Selection I Modified passing 75 cm track length cut

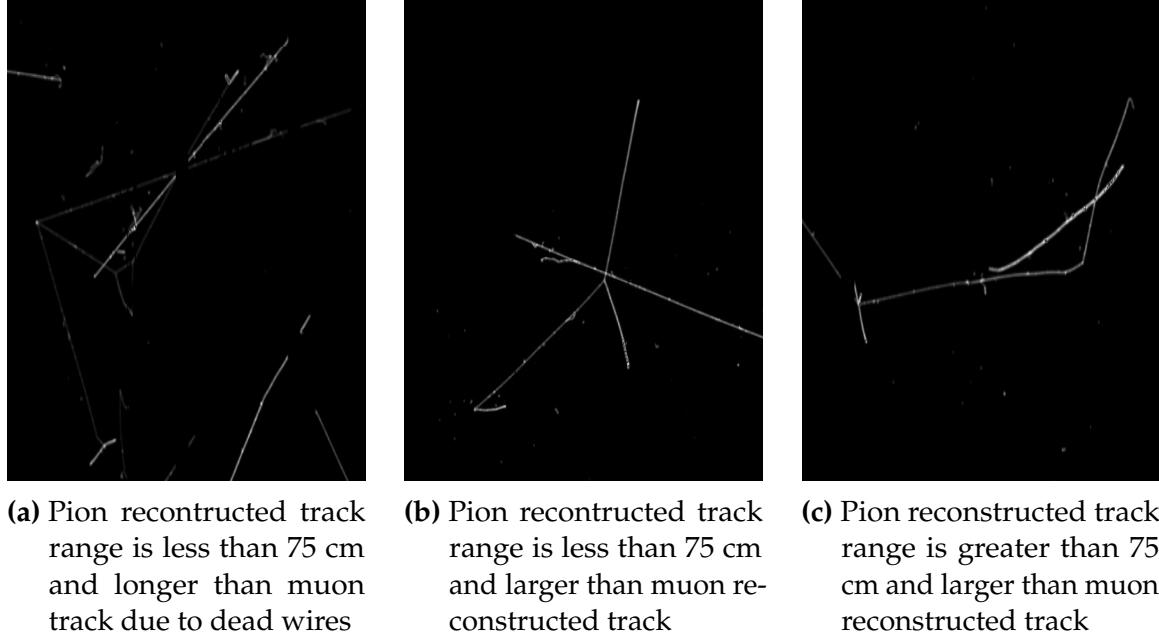
**Figure 8.4:** CNN10000 momentum distributions of track candidate images output from Selection I Modified cc-inclusive filter

of the vertex for these types of events. With the 75 cm track length cut, the selection is cutting event topologies like this where the pion is the tagged track candidate. Figure 8.6a has a defined longer muon track, but because of dead wires through the track, the reconstructed range is 1. less than 75 cm and 2. shorter than the reconstructed pion whose length is also less than 75 cm. This is a very interesting event, but because of issues with the tracking algorithm, the 75 cm cut would get rid of this event. The CNN was able to recover this event only because it has classified all pions as muons. Figure 8.6b shows the second case to think about, the pion, while still less than 75 cm has a reconstructed track length longer than the muon. Again, the CNN recovered this event due to pions being classified as muons. Lastly, figure 8.6c shows a pion with a reconstructed track length greater than 75 cm and the muon. These three cases show that a broader question must be asked when training the network other than is it a muon or pion. There are different routes to recover interesting events like these. One route is to ask the network “Is it a CC event or is it an NC event?” and obtain an image dataset consisting of whole CC/NC events that will train the network to answer this question. The other route is to ask the network “Is this a  $\mu/\pi/p/$  from a CC event or NC event and obtain an image dataset consisting of primary particles from a CC/NC event. Both these paths will be explored in future work.



**Figure 8.5:** Track distribution comparisons of true CC muons plotted vs true CC muons and pions plotted

Table 8.1 shows the passing rates for the 75 cm track length cut and the CNN cut at 70% and 83%. The passing rates at the track containment level for the 75 cm track length cut compared to the CNN are comparable with only a 0.6% difference in the



**Figure 8.6:** Images of true CC events where the pion was the tagged track candidate

		BNB + Cosmics Selection	BNB + Cosmics MC Truth	Cosmic Only	Signal: Cosmic Only
75 cm Cut passing rates	Generated Events	191362	45723	4804	1:22
	Track Containment	19391 (48%/10%)	11693 (45%/26%)	129 (38%/2.7%)	1:2.3
	track $\geq$ 75 cm	6920 (36%/3.6%)	5780 (49%/13%)	17 (13%/0.4%)	1:0.6
CNN passing rates	Generated Events	188880	44689	14606	1:21
	Track Containment	19112 ( /10%)	11554 ( /26%)	302 ( /2.1%)	1:1.73
	CNN cut @ 70% Probability	16502 (86%/8.7%)	10605 (92%/23%)	205 (68%/14%)	1:1.28
	CNN cut @ 83% Probability	7511 (46%/4.0%)	6142 (58%/14%)	32 (16%/0.2%)	1:0.4

**Table 8.1:** Comparing passing rates of CNN at different probabilities versus 75 cm track length cut: Numbers are absolute event counts and Cosmic background is not scaled appropriately. The BNB+Cosmic sample contains all events. The numbers in brackets give the passing rate wrt the step before (first percentage) and wrt the generated events (second percentage). In the BNB+Cosmic MC Truth column shows how many true  $\nu_\mu$  CC-inclusive events (in FV) are left in the sample. This number includes possible mis-identifications where a cosmic track is picked by the selection instead of the neutrino interaction in the same event. The CNN MC True generated events were scaled wrt the MC True generated events for the 75 cm cut passing rates due to only running over 188,880 generated events versus the 191362 generated events. The last column Signal:Cosmic only gives an estimate of the  $\nu_\mu$  CC events wrt the cosmic only background at each step. For this number, the cosmic background has been scaled as described in [?]. Note that these numbers are not a purity, since other backgrounds can't be determined at this step.

Signal	$\nu_\mu$ CC events with true vertex in FV	#Events(Fraction)	#Events(Fraction)
		passing Sel I	passing CNN @ 83% Probability
Backgrounds	Cosmics Only Events	725(33.4%)	2582(26%)
	Cosmics in BNB Events	144(6.6%)	492(4.9%)
	NC Events	75(3.5%)	778(7.7%)
	$\nu_e$ and $\bar{\nu}_e$ Events	4(0.2%)	32(0.3%)
	$\bar{\nu}_\mu$ Events	40(1.8%)	67(0.7%)

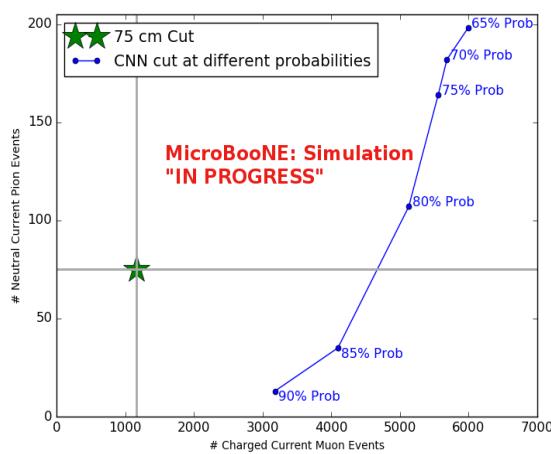
**Table 8.2:** Signal and background event numbers of selection I and selection I with CNN cut estimated from a BNB+Cosmic sample and Cosmic only sample normalized to  $5 * 10^{19}$  PoT. The last column gives the fraction of this signal or background type to the total selected events per CNN probability.

in time cosmic bin which may be due in part to the larger in time cosmic statistics used for the CNN dataset. These passing rates need to be comparable to then be able to compare the passing rates after the CNN cut to the 75 cm cut. Again, the same BNB+Cosmic sample was used for both selection I with 75 cm cut and selection I with CNN cut. As it stands, a CNN cut at 83% probability has a MC true CC event passing rate of 14% compared to the 13% passing rate of the 75 cm track length cut. The Signal:Cosmic Only background is also reduced from 1:0.6 to 1:0.4. The total passing rate is also higher than the 75 cm cut, 3.6% vs 4.0%. Table 8.2 shows the breakdown of signal and backgrounds for the CNN at the different probabilities. We have a 61% signal passing rate with the CNN cut @ 83% versus the 53.8% signal passing rate of the 75 cm cut.

Based on these numbers, the following performance values of the selection with 75 cm cut versus selection with CNN @ 83% probability cut were calculated:

- Efficiency: Number of selected true  $\nu_\mu$  CC events divided by the number of expected true  $\nu_\mu$  CC events with interaction in the FV.
  - Selection I: 12.3%
  - Selection I with CNN10000 cut @ 83% probability: 14%
- Purity: Number of selected true  $\nu_\mu$  CC events divided by sum of itself and the number of all backgrounds.
  - Selection I: 53.8%
  - Selection I with CNN10000 cut @ 83% probability: 61%

1762 Lastly, figure 8.7 shows a more representative performance of the CNN. Due to the  
 1763 fact that the CNN was trained on muons and pions, showing the performance of CC  
 1764 muon events versus NC pion events with respect to CNN probability gives a better  
 1765 picture of how the network is performing. Figure 8.7 shows that at 83% we are below  
 1766 the 75 cm cut NC pion threshold and still above the CC muon threshold. Using 83%  
 1767 probability not only reduced the NC pion background, it also dramatically reduced  
 1768 the in time cosmics and cosmics in the BNB.



**Figure 8.7:** CNN performance of classified muons and pions compared to the already implemented 75 cm track length cut

### 1769 8.1.2 Conclusions of CNN10000 classification of MC data

1770 It was shown that even though CNN10000 was trained with single particle generated  
 1771 muons and pions, it performs fairly well at classifying track candidate images from  
 1772 BNB+Cosmic events. Events have been regained below the 75 cm track length cut and  
 1773 the momentum and track range distributions have similar shapes to the distributions  
 1774 of Selection I. Efficiencies and purities were calculated for selection I events before 75  
 1775 cm track length cut with the CNN at 83% probability and are 14% and 62% respectively.  
 1776 Although the CNN doesn't have separation between muons and pions and although  
 1777 all particles passing CNN are classified as muon, increasing CNN probability allows  
 1778 us to increase the purity as well as maintain an efficiency comparable to the 75 cm track  
 1779 length cut all while recovering events below that 75 cm cut. Out of the 6142 events  
 1780 that passed the CNN @ 83% 1470 events were below the 75 cm cut, a recovery of 3.3%  
 1781 of data with a purity of 15%. Although these numbers are low, it is an improvement

<sub>1782</sub> from the selection I in both total efficiency and purity and an increase in phase space  
<sub>1783</sub> by recovering these events.

## <sub>1784</sub> 8.2 Classification using CNN100000

<sub>1785</sub> CNN100000 was trained using 20,000 images of each  $\mu/\pi/p/\gamma/e$ . For classification  
<sub>1786</sub> of BNB+Cosmics and data, images were made from track candidates that passed the  
<sub>1787</sub> Selection I filter, however, unlike for classifying BNB+Cosmics using CNN10000, the  
<sub>1788</sub> classification of CNN100000 went further up Selection I's cut chain. For CNN100000,  
<sub>1789</sub> steps 1,2,3 and 4 seen in 5.5. The results of using CNN100000 to classify BNB+Cosmics  
<sub>1790</sub> will be outlined below.

### <sub>1791</sub> 8.2.1 Classification of MC data using Selection I CC-Inclusive Filter

<sub>1792</sub> Kinematic truth distributions of BNB+Cosmic events passing Selection  
<sub>1793</sub> I+CNN10000

### <sub>1794</sub> 8.2.2 Classification of MicroBooNE data using Selection I <sub>1795</sub> CC-Inclusive Filter

### <sub>1796</sub> 8.2.3 Comparing two CC-Inclusive Cross Section Selection Filters

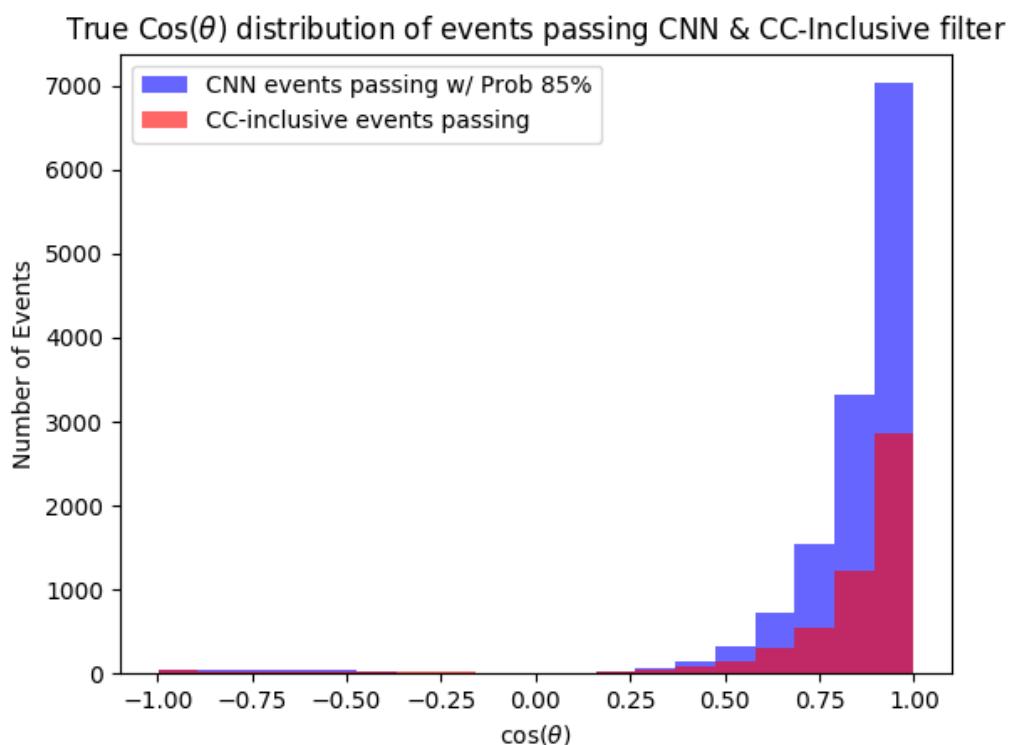
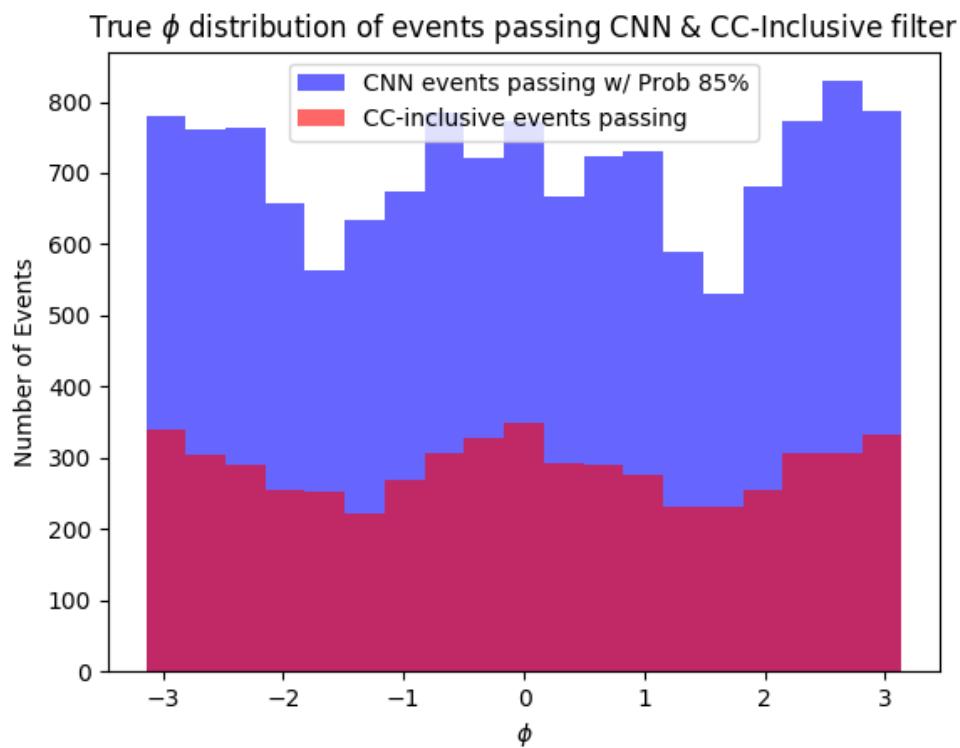
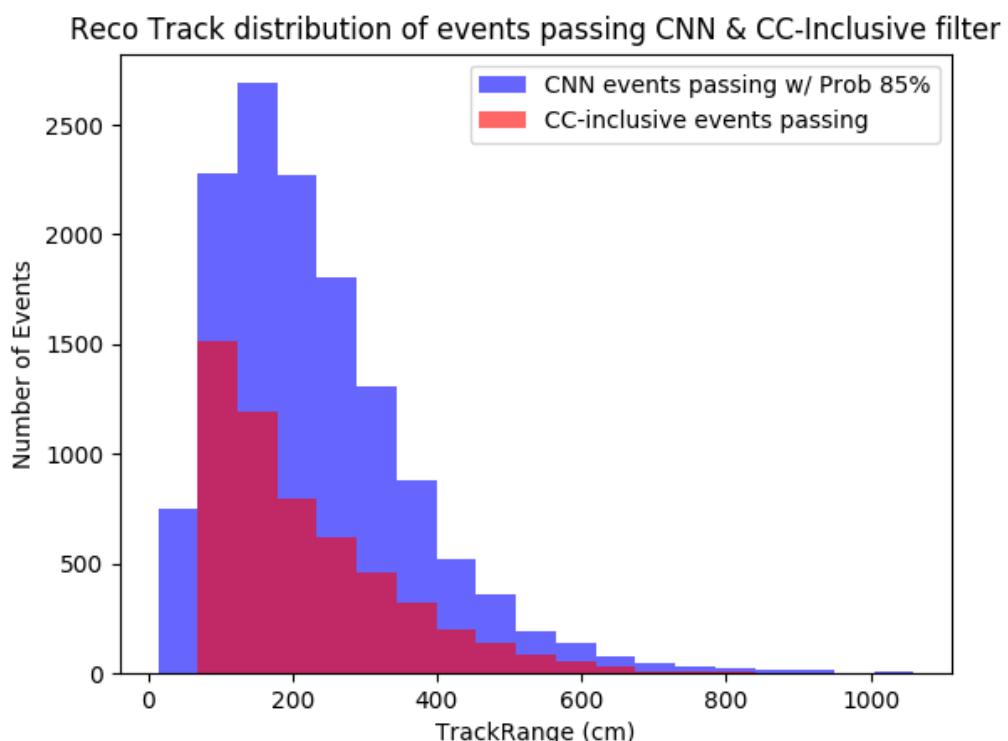


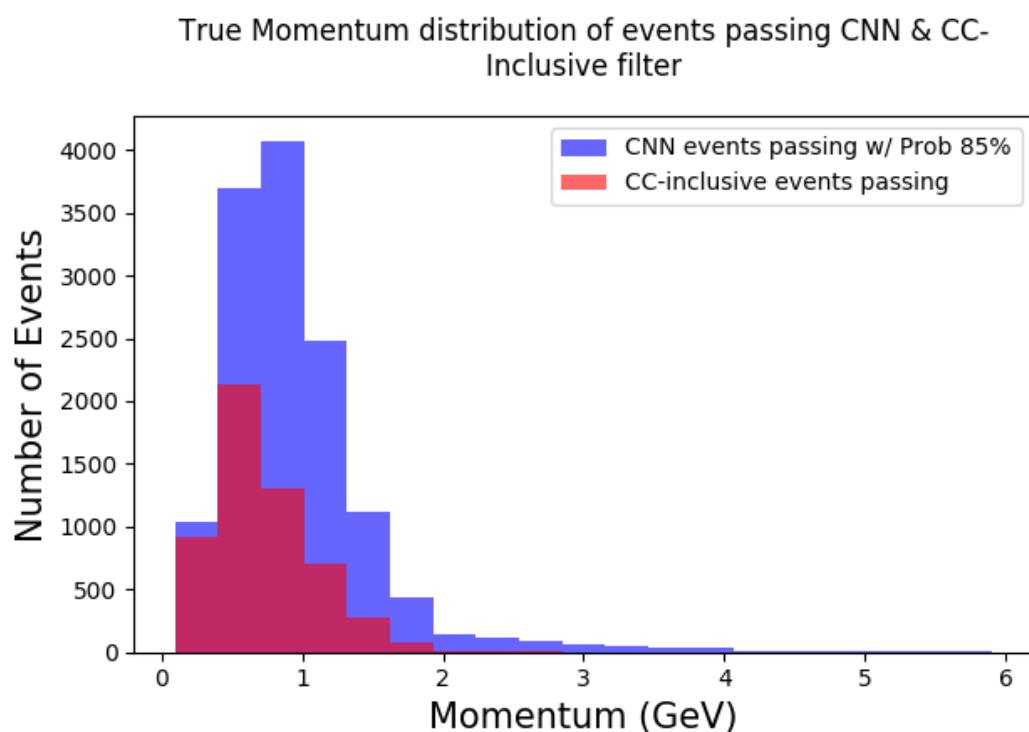
Figure 8.8:  $\text{Cos}(\theta)$  distribution at  $\text{CNN10000} \geq 85\%$



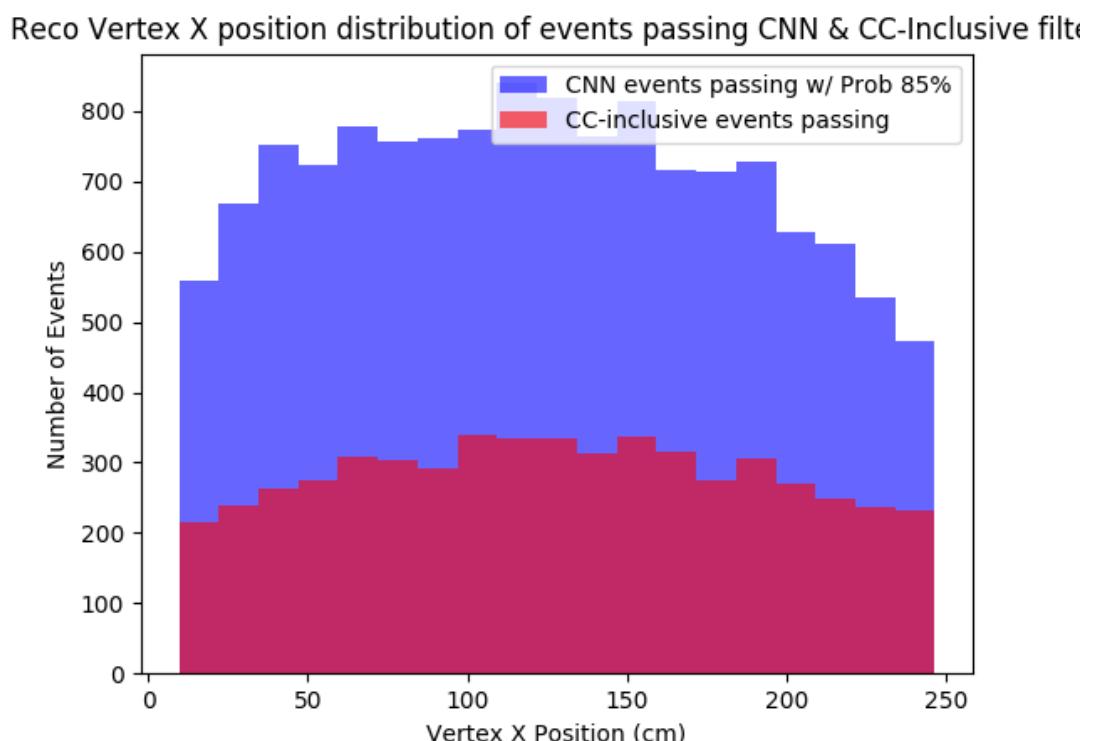
**Figure 8.9:**  $\cos(\theta)$  distribution at  $\text{CNN10000} \geq 85\%$



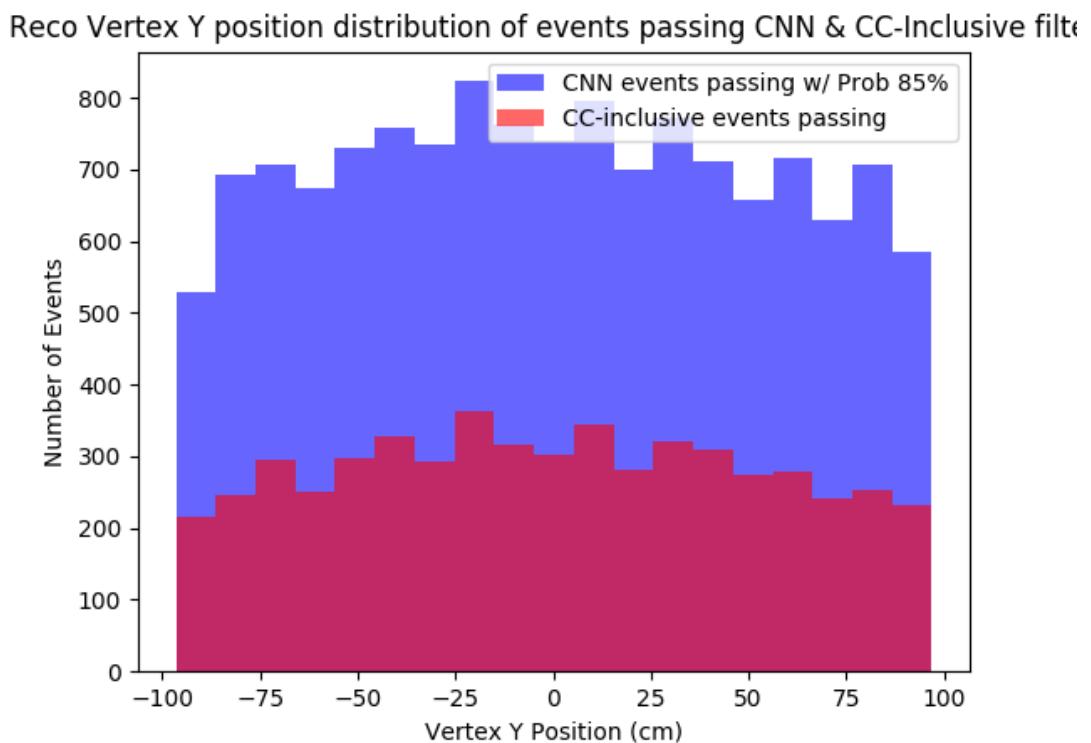
**Figure 8.10:**  $\cos(\theta)$  distribution at  $\text{CNN10000} \geq 85\%$



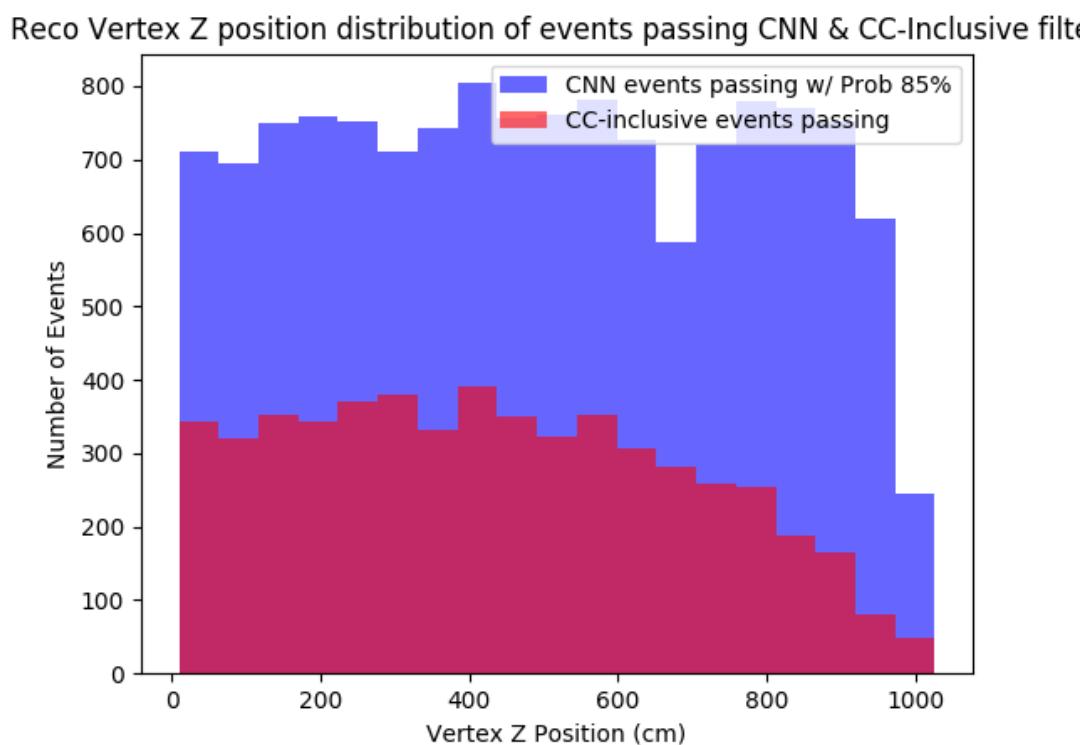
**Figure 8.11:**  $\cos(\theta)$  distribution at  $\text{CNN10000} \geq 85\%$



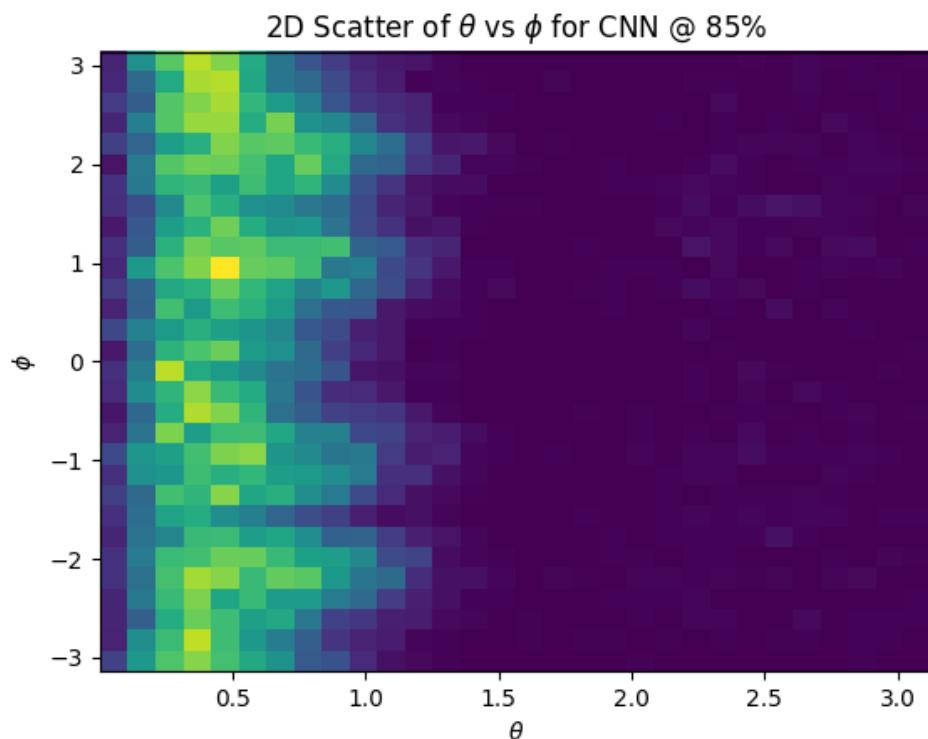
**Figure 8.12:**  $\cos(\theta)$  distribution at  $\text{CNN10000} \geq 85\%$



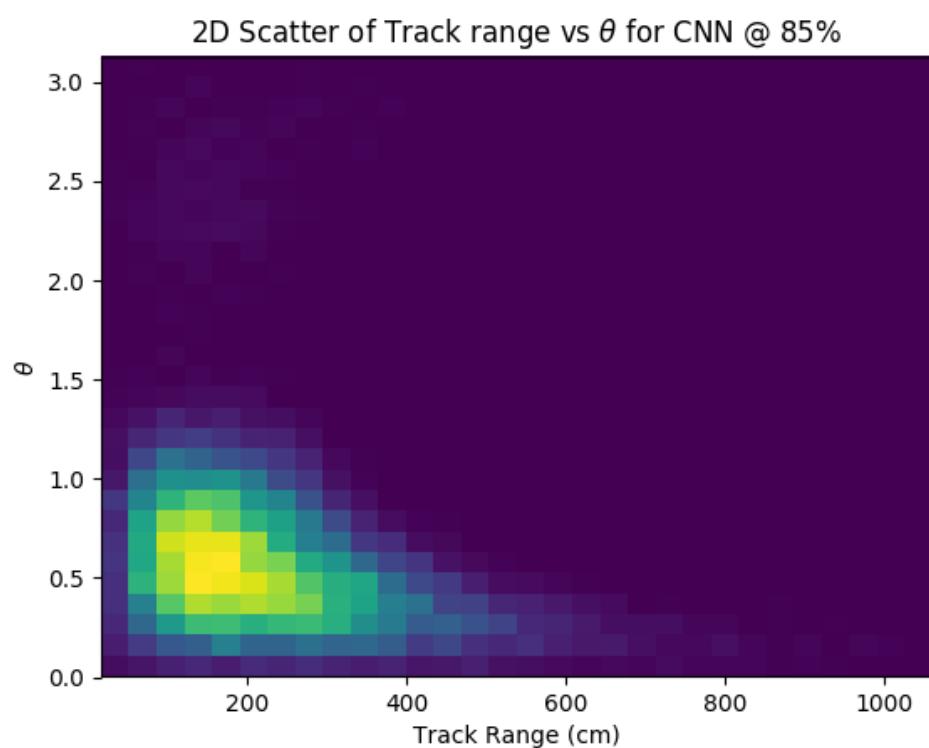
**Figure 8.13:**  $\cos(\theta)$  distribution at  $\text{CNN10000} \geq 85\%$



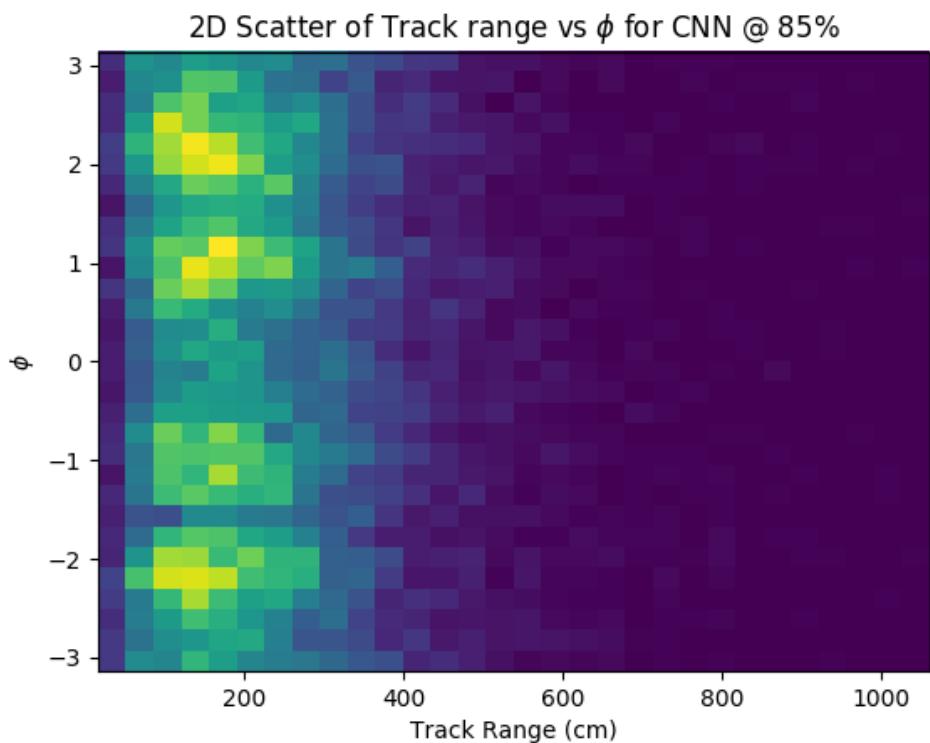
**Figure 8.14:**  $\cos(\theta)$  distribution at  $\text{CNN10000} \geq 85\%$



**Figure 8.15:**  $\cos(\theta)$  distribution at  $\text{CNN10000} \geq 85\%$



**Figure 8.16:**  $\cos(\theta)$  distribution at  $\text{CNN10000} \geq 85\%$



**Figure 8.17:**  $\cos(\theta)$  distribution at  $\text{CNN10000} \geq 85\%$

<sup>1797</sup> **Chapter 9**

<sup>1798</sup> **Conclusion**

<sup>1799</sup> Your Conclusions here.

<sup>1800</sup>



# <sup>1801</sup> Bibliography

- <sup>1802</sup> [1] Wikipedia, Neutrino, <http://wikipedia.org/wiki/Neutrino>, 2013.
- <sup>1803</sup> [2] Wikipedia, Neutrino oscillation, [http://en.wikipedia.org/wiki/Neutrino\\_oscillation](http://en.wikipedia.org/wiki/Neutrino_oscillation), 2013.
- <sup>1804</sup>
- <sup>1805</sup> [3] B. N. Laboratory, Neutrinos and nuclear chemistry, <http://www.chemistry.bnl.gov/sciandtech/sn/default.htm>, 2010.
- <sup>1806</sup>
- <sup>1807</sup> [4] K. Heeger, Big world of small neutrinos, <http://conferences.fnal.gov/lp2003/forthepublic/neutrinos/>.
- <sup>1808</sup>
- <sup>1809</sup> [5] A. Y. Smirnov, p. 23 (2003), hep-ph/0305106.
- <sup>1810</sup> [6] The MicroBooNE Collaboration, R. e. a. Acciari, (2015), arXiv:1512.06148.
- <sup>1811</sup> [7] The MicroBooNE Collaboration, R. e. a. Acciari, (2016), arXiv:1705.07341.
- <sup>1812</sup> [8] The MicroBooNE Collaboration, R. e. a. Acciari, (2017), arXiv:1704.02927.
- <sup>1813</sup> [9] The MicroBooNE Collaboration, R. e. a. Acciari, JINST **12** (2017).