

---

1

2        $\mu/\pi$  separation using

3       Convolutional Neural Networks

4       for the MicroBooNE

5       Charged Current Inclusive Cross Section

6       Measurement

---

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 February 1, 2018 - \* -

17                   December, 2017

18                   Syracuse University

19

Syracuse, New York

20

Copyright 2017

21

Jessica Nicole Esquivel

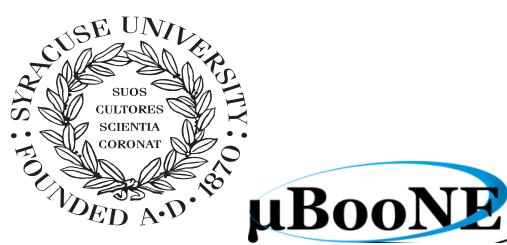
22

All Rights Reserved

23

Syracuse University

24



**μBooNE**





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...

# 41 Contents

42	<b>List of figures</b>	x
43	<b>List of tables</b>	xix
44	<b>1 Introduction</b>	1
45	<b>2 Neutrinos</b>	2
46	2.1 What are Neutrinos . . . . .	2
47	2.2 History of Neutrinos . . . . .	3
48	2.2.1 Solar Oscillations and the Solar Neutrino Problem . . . . .	4
49	2.2.2 Atmospheric Oscillations and the Atmospheric Neutrino Anomaly	6
50	2.3 Neutrino Oscillations . . . . .	8
51	2.3.1 Two Flavor Neutrino Oscillation Formulation . . . . .	8
52	2.3.2 Three Flavor Neutrino Oscillation Formulation . . . . .	11
53	2.4 Neutrino Interactions . . . . .	11
54	2.4.1 Weak Interactions . . . . .	11
55	2.4.2 Neutrino-Nucleon Interactions . . . . .	11
56	<b>3 The MicroBooNE Experiment</b>	13
57	3.1 Liquid argon time projection chambers . . . . .	13
58	3.2 The MicroBooNE Time Projection Chamber . . . . .	14
59	3.3 MicroBooNE’s Physics Goals . . . . .	16
60	3.3.1 The low-energy excess . . . . .	16
61	3.3.2 Cross sections . . . . .	17
62	3.3.3 Liquid argon detector development . . . . .	18
63	3.4 The Booster Neutrino Beam . . . . .	18
64	3.4.1 Creating the Booster Neutrino Beam . . . . .	19
65	3.5 Event Reconstruction . . . . .	20

66	<b>4 Neutrino Identification: Finding MicroBooNE's first Neutrinos</b>	23
67	4.1 Flash Finding . . . . .	23
68	4.1.1 Flash Reconstruction . . . . .	24
69	4.1.2 Beam Timing . . . . .	25
70	4.1.3 Event Rates . . . . .	26
71	4.2 TPC Topology Selection . . . . .	26
72	4.2.1 Cosmic Tagging . . . . .	27
73	4.2.2 2D Cluster Selection . . . . .	27
74	4.2.3 3D Tracks and vertices Selection . . . . .	29
75	4.2.4 TPC Updates . . . . .	31
76	4.3 Conclusion . . . . .	32
77	<b>5 CC-Inclusive Cross Section Selection Filter</b>	35
78	5.1 Data and MC Processing Chain . . . . .	36
79	5.2 Normalization of data and MC . . . . .	37
80	5.3 Optical Software Trigger and Reconstruction . . . . .	38
81	5.3.1 Software Trigger . . . . .	38
82	5.3.2 Flash Reconstruction . . . . .	39
83	5.3.3 Beam Window . . . . .	40
84	5.4 TPC Reconstruction . . . . .	42
85	5.4.1 Hit Reconstruction . . . . .	43
86	5.4.2 Clustering . . . . .	43
87	5.4.3 Pandora . . . . .	44
88	5.4.4 Trackkalmanhit . . . . .	44
89	5.4.5 Cosmic Hit Removal . . . . .	44
90	5.4.6 Projection Matching Algorithm . . . . .	45
91	5.5 Event Selection . . . . .	45
92	5.5.1 Expected Bakgrounds . . . . .	46
93	5.5.2 Truth Distributions . . . . .	49
94	<b>6 Background on Convolutional Neural Networks</b>	56
95	6.1 Image Classification . . . . .	56
96	6.2 CNN Structure . . . . .	57
97	6.2.1 Backpropagation . . . . .	60
98	6.3 Choosing Hyperparameters . . . . .	61

99	<b>7 Training process of Convolutional Neural Networks</b>	63
100	7.1 Hardware Configurations for Convolutional Neural Network Training	63
101	7.2 Creating images using LArTPC data for training/validation of CNNS	64
102	7.3 Convolutional Neural Network Training . . . . .	66
103	7.3.1 Training CNN1075 . . . . .	66
104	7.3.2 Training CNN10000 . . . . .	69
105	7.3.3 Training CNN100000 . . . . .	72
106	<b>8 Using Convolutional Neural Networks for <math>\nu_\mu</math> CC event classification</b>	82
107	8.1 Classification using CNN10000 . . . . .	82
108	8.1.1 Classification of MC data using Selection I CC-Inclusive Filter .	82
109	8.1.2 Conclusions of CNN10000 classification of MC data . . . . .	90
110	8.2 Classification using CNN100000 . . . . .	91
111	8.2.1 Classification of MC data using Selection I CC-Inclusive Filter .	93
112	8.2.2 Classification of MicroBooNE data using Selection I CC-Inclusive	
113	Filter . . . . .	93
114	8.2.3 Comparing two CC-Inclusive Cross Section Selection Filters . .	93
115	<b>9 Conclusion</b>	102
116	<b>Bibliography</b>	103

# <sup>117</sup> List of figures

<sup>118</sup>	2.1	The Standard Solar Model [?]	4
<sup>119</sup>	2.2	Solar Neutrino Experiments	6
<sup>120</sup>	2.2a	Ray Davis's Homestake Experiment [?]	6
<sup>121</sup>	2.2b	Kamiokande-II Experiment [?]	6
<sup>122</sup>	2.2c	SNO Experiment [?]	6
<sup>123</sup>	2.3	The flavor eigenstates are rotated by an angle $\theta$ with respect to the mass eigenstates	8
<sup>124</sup>			
<sup>125</sup>	3.1	Low Energy excess seen in MiniBooNE	17
<sup>126</sup>	3.2	Arial view of the Main Injector and the Booster Neutrino Beam at Fermilab	19
<sup>127</sup>	3.3	Arial view of the Main Injector and the Booster Neutrino Beam at Fermilab	21
<sup>128</sup>	3.4	Energy spectrum of the Booster Neutrino Beam at Fermi National Laboratories	22
<sup>129</sup>			
<sup>130</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.	24
<sup>131</sup>			
<sup>132</sup>	4.2a	Predicted distribution of flash times with respect to trigger time for 1 day of data taking at nominal rate and intensity	29
<sup>133</sup>			
<sup>134</sup>	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.	29
<sup>135</sup>			
<sup>136</sup>			
<sup>137</sup>			

138	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. . . . .	30
141	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. . . . .	30
144	4.5	First Neutrino Interaction Candidate Events from MicroBooNE . . . . .	33
145	4.6	First Neutrino Interaction Candidate Events from MicroBooNE . . . . .	34
146	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. . . . .	38
155	5.1a	Track range distribution of selection I . . . . .	38
156	5.1b	Selection efficiency as a function of the true muon momentum . . . . .	38
157	5.2	Time distribution of reconstructed optical flashes with a PE value of 50 or more for a sample of BNB unbiased triggered events. . . . .	40
159	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 . . . . .	41
164	5.4	Reconstruction chain run on both data and MC. The red stars mean that the algorithm returns reconstructed 3D vertices. . . . .	42
166	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. . . . .	47

169	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. . . . .	50
175	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. . . . .	51
180	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. . . . .	52
185	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. . . . .	53
192	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. . . . .	54

---

199	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. . . . .	55
206	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 . . . . .	58
209	6.2 Visualization of filters found in first layer of a CNN. . . . .	58
210	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 detec- tion mask that leaves only shape information which helps to distinguish between different types of clothes. . . . .	59
215	6.4 Pictorial Representation of Convolutional Neural Networks as well as a visual representation on CNN's complexity of layer feature extraction	59
217	6.5 Pictoral representation of the AlexNet model. The AlexNet model consists of 5 convolution layers and 3 fully connected layers. . . . .	62
219	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. . . . .	62
226	7.1 Accuracy vs. Loss of AlexNet 2-output $\mu/\pi$ sample consisting of 2,150 images each. . . . .	67

228	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 = $true\mu/(true\mu + false\mu)$ . . . . .	68
232	7.2a Confusion Matrix showing Accuracy of CNN1075 using training MC data . . . . .	68
234	7.2b Confusion Matrix showing Accuracy of CNN1075 using testing MC data . . . . .	68
236	7.3 Accuracy vs. Loss of AlexNet 2-output $\mu/\pi$ sample consisting of 10,000 images each. . . . .	70
238	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 = $true\mu/(true\mu + false\mu)$ . . . . .	71
242	7.4a Confusion Matrix showing Accuracy of CNN10000 using train- ing MC data . . . . .	71
244	7.4b Confusion Matrix showing Accuracy of CNN10000 using testing MC data . . . . .	71
246	7.5 Probability plot of muons and pions from testing set. Images surround- ing histogram are a random event from lowest bin and highest bin for each particle. . . . .	72
249	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 . . . . .	73
253	7.7 Training and testing loss of CNN trained on 100,000 images of $\mu/\pi/p/\gamma/e$	73
254	7.8 Confusion Matrix of all five particles . . . . .	74
255	7.9 t-SNE of CNN . . . . .	75
256	7.10 Probabilities of different particle classes as well as their contamination from other classes . . . . .	76

258	7.10a Muon Prob . . . . .	76
259	7.10b Pion Prob . . . . .	76
260	7.10c Proton Prob . . . . .	76
261	7.10d Electron Prob . . . . .	76
262	7.10e Gamma Prob . . . . .	76
263	7.11 Muon probability of true muons (blue) versus pions (red), protons (cyan), gammas (green) and electrons (magenta). . . . .	77
265	7.12 Kinematic distributions versus muon probability for true muons and true pions. . . . .	78
267	7.12a Track range versus muon probability for true muons (blue) and true pions (red). . . . .	78
269	7.12b Track range $\leq 75$ cm versus muon probability for true muons (blue) and true pions (red). . . . .	78
271	7.12c Momentum versus muon probability for true muons (blue) and true pions (red). . . . .	78
273	7.13 Kinematic distributions versus muon probability for true muons and true protons. . . . .	79
275	7.13a Track range versus muon probability for true muons (blue) and true protons (cyan). . . . .	79
277	7.13b Track range $\leq 75$ cm versus muon probability for true muons (blue) and true protons (cyan). . . . .	79
279	7.13c Momentum versus muon probability for true muons (blue) and true protons (cyan). . . . .	79
281	7.14 Kinematic distributions versus muon probability for true muons and true electrons. . . . .	80
283	7.14a Track range versus muon probability for true muons (blue) and true electrons (magenta). . . . .	80

285	7.14b Track range $\leq 75$ cm versus muon probability for true muons (blue) and true electrons (magenta). . . . .	80
287	7.14c Momentum versus muon probability for true muons (blue) and true electrons (magenta). . . . .	80
289	7.15 Kinematic distributions versus muon probability for true muons and true gammas. . . . .	81
291	7.15a Track range versus muon probability for true muons (blue) and true gammas (green). . . . .	81
293	7.15b Track range $\leq 75$ cm versus muon probability for true muons (blue) and true gammas (green). . . . .	81
295	7.15c Momentum versus muon probability for true muons (blue) and true gammas (green). . . . .	81
297	8.1 Confusion matrix and probability plot of events passing selection I cc-inclusive cuts right before 75cm track length cut . . . . .	83
299	8.1a Confusion Matrix for CNN10000 classified events from selection I	83
300	8.1b Probability plot for CNN10000 classified events from selection I	83
301	8.2 CNN10000 distributions of track candidate images output from Selection I cc-inclusive filter . . . . .	84
303	8.2a Track range distribution of events from Selection I passing CNN with 70% accuracy . . . . .	84
305	8.2b Stacked signal and background track range distributions from Selection I passing CNN with 70% accuracy . . . . .	84
307	8.2c Stacked signal and background track range distributions from Selection I passing 75 cm track length cut . . . . .	84
309	8.2d Stacked signal muons and background muons/pions of track range distributions from Selection I passing CNN with 70% accuracy . . . . .	84
312	8.2e Stacked signal muons and background muons/pions of track range distributions from Selection I passing 75 cm track length cut	84

314	8.3 CNN10000 stacked signal/background track range distributions of track candidate images output from Selection I cc-inclusive filter . . . . .	85
315		
316	8.3a Stacked signal muons and background muons/pions of track range distributions from Selection I passing CNN with 75% accuracy . . . . .	85
317		
318		
319	8.3b Stacked signal muons and background muons/pions of track range distributions from Selection I passing CNN with 80% accuracy . . . . .	85
320		
321		
322	8.3c Stacked signal muons and background muons/pions of track range distributions from Selection I passing CNN with 85% accuracy . . . . .	85
323		
324		
325	8.3d Stacked signal muons and background muons/pions of track range distributions from Selection I passing CNN with 90% accuracy . . . . .	85
326		
327		
328	8.4 CNN10000 momentum distributions of track candidate images output from Selection I cc-inclusive filter . . . . .	86
329		
330	8.4a Momentum distribution of events from Selection I passing CNN with 70% accuracy . . . . .	86
331		
332	8.4b Stacked signal and background momentum distributions from Selection I passing CNN with 70% accuracy . . . . .	86
333		
334	8.4c Stacked signal and background momentum distributions from Selection I passing 75 cm track length cut . . . . .	86
335		
336	8.5 Track distribution comparisons of true CC muons plotted vs true CC muons and pions plotted . . . . .	88
337		
338	8.5a Stacked signal $\mu$ /background $\mu$ and $\pi$ track range distribution of CNN @ 70% . . . . .	88
339		
340	8.5b Stacked signal $\mu\&\pi$ /background $\mu\&\pi$ track range distribution of CNN @ 70% . . . . .	88
341		
342	8.6 Images of true CC events where the pion was the tagged track candidate	88

---

343	8.6a	Pion reconstructed track range is less than 75 cm and longer than muon track due to dead wires . . . . .	88
344			
345	8.6b	Pion reconstructed track range is less than 75 cm and larger than muon reconstructed track . . . . .	88
346			
347	8.6c	Pion reconstructed track range is greater than 75 cm and larger than muon reconstructed track . . . . .	88
348			
349	8.7	CNN performance of classified muons and pions compared to the already implemented 75 cm track length cut . . . . .	91
350			
351	8.8	Image making steps used for classifying BNB+Cosmic events using CNN100000 . . . . .	92
352			
353	8.9	$\text{Cos}(\theta)$ distribution at CNN10000 $\geq 85\%$ . . . . .	93
354	8.10	$\text{Cos}(\theta)$ distribution at CNN10000 $\geq 85\%$ . . . . .	94
355	8.11	$\text{Cos}(\theta)$ distribution at CNN10000 $\geq 85\%$ . . . . .	95
356	8.12	$\text{Cos}(\theta)$ distribution at CNN10000 $\geq 85\%$ . . . . .	96
357	8.13	$\text{Cos}(\theta)$ distribution at CNN10000 $\geq 85\%$ . . . . .	97
358	8.14	$\text{Cos}(\theta)$ distribution at CNN10000 $\geq 85\%$ . . . . .	98
359	8.15	$\text{Cos}(\theta)$ distribution at CNN10000 $\geq 85\%$ . . . . .	99
360	8.16	$\text{Cos}(\theta)$ distribution at CNN10000 $\geq 85\%$ . . . . .	100
361	8.17	$\text{Cos}(\theta)$ distribution at CNN10000 $\geq 85\%$ . . . . .	101

# <sup>362</sup> List of tables

<sup>363</sup> 2.1	Measurements of the double ratio for various atmospheric neutrino experiments . . . . .	<sup>364</sup> 7
<sup>365</sup> 2.2	Current world knowledge of neutrino oscillation parameters [?] . . . . .	<sup>366</sup> 12
<sup>366</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 seccond line scaled with the flash finding factor of 0.008. All events are normalized to per day assuming we are running at 5 Hz. . . . .	<sup>367</sup> 29
<sup>372</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. . . . .	<sup>373</sup> 48

384	5.2	Passing rates for Selection I selection applied to on-beam and off-beam 385 data. THe numbers in brackets give the passing rate wrt the step before 386 (first percentage) and wrt the generated events (second percentage). 387 Off-beam data has been scaled with a factor 1.23 to normalize to the 388 on-beam data stream. . . . .	48
389	8.1	Comparing passing rates of CNN at different probabilities versus 75 390 cm track length cut: Numbers are absolute event counts and Cosmic 391 background is not scaled appropriately. The BNB+Cosmic sample con- 392 tains all events. The numbers in brackets give the passing rate wrt 393 the step before (first percentage) and wrt the generated events (second 394 percentage). In the BNB+Cosmic MC Truth column shows how many 395 true $\nu_\mu$ CC-inclusive events (in FV) are left in the sample. This number 396 includes possible mis-identifications where a cosmic track is picked by 397 the selection instead of the neutrino interaction in the same event.The 398 CNN MC True generated events were scaled wrt the MC True generated 399 events for the 75 cm cut passing rates due to only running over 188,880 400 generated events versus the 191362 generated events. The last column 401 Signal:Cosmic only gives an estimate of the $\nu_\mu$ CC events wrt the cosmic 402 only background at each step. For this number, the cosmic background 403 has been scaled as described in [?]. Note that these numbers are not a 404 purity, since other backgrounds can't be determined at this step. . .	89
405	8.2	Signal and background event numbers of selection I and selection I 406 with CNN cut estimated from a BNB+Cosmic sample and Cosmic only 407 sample normalized to $5 * 10^{19}$ PoT. The last column gives the fraction 408 of this signal or background type to the total selected events per CNN 409 probability. . . . .	89

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

— Shirley Chisholm



<sup>412</sup> **Chapter 1**

<sup>413</sup> **Introduction**

<sup>414</sup> This thesis will be a description of work done to further increase efficiency and purity  
<sup>415</sup> of the charged current inclusive cross section measurement using the MicroBooNE  
<sup>416</sup> detector. It will also describe the MicroBooNE detector, what neutrinos are, the  
<sup>417</sup> charged current inclusive cross section measurement and its importance as well as  
<sup>418</sup> convolutional neural networks and how they can be used in  $\mu/\pi$  separation. Chapter  
<sup>419</sup> **2** will talk about the background of neutrinos and the people and detectors that  
<sup>420</sup> discovered neutrinos as well as an in depth history of neutrino oscillation and the  
<sup>421</sup> discovery that neutrinos have mass.

<sup>422</sup> Chapter **3** will discuss the MicroBooNE experiment, specifically, how Liquid  
<sup>423</sup> Argon Time Projection Chambers work, the Light Collection System and the Electronic  
<sup>424</sup> and Readout Trigger systems. This chapter will also describe the Booster Neutrino  
<sup>425</sup> Beam sationed at Fermilab.

<sup>426</sup> Chapter **4** will discuss the work that was done to detect the first neutrinos seen in  
<sup>427</sup> the MicroBooNE detector and the software reconstruction efforts required to create an  
<sup>428</sup> automated neutrino ID filter that was used to find the first neutrinos and then was  
<sup>429</sup> later expanded on to create the charged current inclusive filter that will be discussed  
<sup>430</sup> in chapter **5**

<sup>431</sup> Chapter **6** will give a brief description of what Convolutional Neural Networks are  
<sup>432</sup> and how it will be used for  $\mu/\pi$  separation in this selection. Chapter **7** will discuss  
<sup>433</sup> the hardware frameworks and training methods used to train multiple Convolutional  
<sup>434</sup> Neural Networks for use in the charged current inclusive cross section measurement.  
<sup>435</sup> Chapters **8** and **??** will discuss the results of using Convolutional Neural Networks on  
<sup>436</sup> monte-carlo and data to sift out charged current inclusive neutrino events.

<sup>437</sup> **Chapter 2**

<sup>438</sup> **Neutrinos**

<sup>439</sup> **2.1 What are Neutrinos**

<sup>440</sup> Neutrinos are fundamental particles which help make up the universe. They are also  
<sup>441</sup> one of the least understood. Neutrinos are not affected by the electromagnetic force  
<sup>442</sup> because they do not have electric charge. Neutrinos are affected by a weak sub-atomic  
<sup>443</sup> force of much shorter range than electromagnetism, and are therefore able to pass  
<sup>444</sup> through great distances in matter without much possibility of being affected by it.  
<sup>445</sup> Until the late 1990's, neutrinos were thought to have no mass. Neutrinos are created  
<sup>446</sup> by radioactive decay such as the ones that happen in the sun, in nuclear reactors or  
<sup>447</sup> when cosmic rays hit atoms. There are three types of neutrinos,  $\nu_e$ ,  $\nu_\mu$  and  $\nu_\tau$  which  
<sup>448</sup> correspond to their charged lepton pairs.

<sup>449</sup> As previously stated, neutrinos are very weakly interacting; in fact, neutrinos can  
<sup>450</sup> pass unscathed through a wall of lead several hundred light-years thick. Because  
<sup>451</sup> neutrinos interact so rarely, studying neutrinos requires a massive detector and a  
<sup>452</sup> powerful neutrino source. With that being said, we can only "see" a neutrino when  
<sup>453</sup> they interact in a detector. In a collision, distinct charged particles are produced with  
<sup>454</sup> each type of neutrino because of the weak force. An electron neutrino will create an  
<sup>455</sup> electron, a muon neutrino will create a muon, and a tau neutrino will create a tau. The  
<sup>456</sup> charged lepton track the particle leaves in the detector is how one figures out what  
<sup>457</sup> type of neutrino interaction was "seen". Liquid Argon Time Projection Chambers are  
<sup>458</sup> being used to study neutrinos due to their excellent imaging and particle identification  
<sup>459</sup> capabilities.

## 460 2.2 History of Neutrinos

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

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

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

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

### 493 2.2.1 Solar Oscillations and the Solar Neutrino Problem

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

499 The standard solar model (SSM) is a mathematical model developed by John  
500 Bahcall of the sun as a spherical ball of gas with varying states of ionization. The solar  
501 neutrino flux derived from the SSM is shown in figure 2.1 [?]. Nuclear fusion and  
502 decay processes produce an abundant amount of neutrinos. The standard solar model  
503 predicts that these reactions produce several groups of neutrinos, each with differing  
504 fluxes and energy spectra. The figure also shows the ranges of detection of existing  
505 solar neutrino experiments in different shades of blue to illustrate that they sample  
506 different portions of the solar neutrino energy spectrum. Three of these experiments  
507 are discussed below.

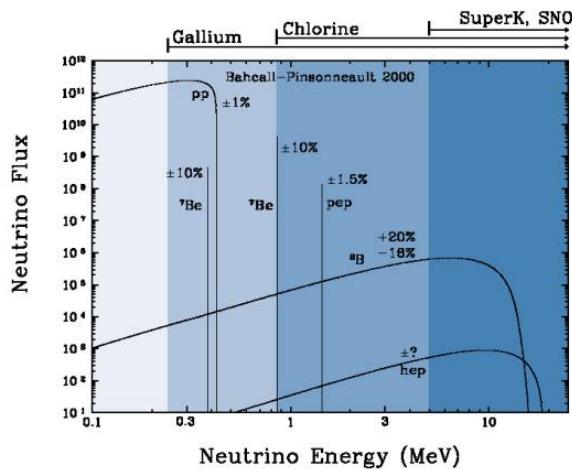


Figure 2.1: The Standard Solar Model [?]

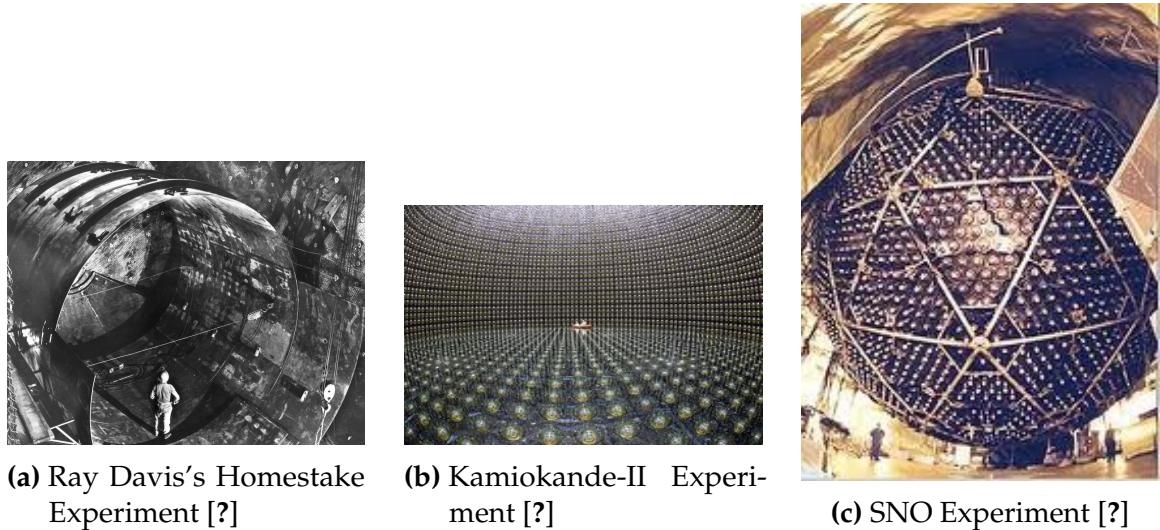
508 The first experiment to detect the effects of neutrino oscillation was the Ray Davis's  
509 Homestake Experiment. The detector was stationed in the Homestake Gold Mine in  
510 Lead, South Dakota. It was 1,478 meters underground and was  $380 \text{ m}^3$ . The detector  
511 was filled with perchloroethylene. Perchloroethylene was chosen because of its high  
512 concentrations of chlorine. When an  $\nu_e$  interacted with a chlorine-37 atom, the atom  
513 would transform to argon-37 which was then extracted and counted. The neutrino

514 capture reaction is shown in equation 2.1. Davis observed a deficit of about 1/3  
515 the flux of solar neutrinos that was predicted by Bahcall's Standard Solar Model.  
516 The unexplained difference between the measured solar neutrino flux and model  
517 predictions led to the Solar Neutrino Problem. [4]



518 While it is now known that the Homestake Experiment results were indicating neu-  
519 trino flavor oscillation, some physicists were weary of the results. Conclusive evidence  
520 of the Solar Neutrino Problem was provided by the Kamiokande-II experiment, a water  
521 cherenkov detector with a low enough energy threshold to detect neutrinos through  
522 neutrino-electron elastic scattering. In the elastic scattering interaction the electrons  
523 coming out of the point of reaction strongly point in the direction that the neutrino  
524 was traveling, away from the sun. While the neutrinos observed in Kamiokande-II  
525 were clearly from the sun, there was still a discrepancy between Kamiokande-II and  
526 Homestake; The Kamiokande-II experiment measured about 1/2 the predicted flux,  
527 rather than the 1/3 that the Homestake Experiment saw.

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



**Figure 2.2:** Solar Neutrino Experiments

## 542 2.2.2 Atmospheric Oscillations and the Atmospheric Neutrino 543 Anomaly

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

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

547

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

548 In general, these neutrinos have energies from 1 GeV to 100s of GeV and the ratio  
549 of  $\nu_\mu$ 's to  $\nu_e$ 's equals to 2 (see equation 2.4)

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

550 There have been two types of detectors used to study atmospheric neutrinos: Water  
551 Cherenkov detectors and tracking calorimeters. Super-Kamiokande is the detector we  
552 will focus on. These atmospheric detector experiments measure the ratio of  $\nu_\mu$  to  $\nu_e$ .

553 They also measure the zenith angle distribution of the neutrinos. These experiments  
 554 report a double ratio (shown in equation 2.5). This double ratio is the ratio measured  
 555 in the detector to the ratio that's expected which is 2. If the double ratio equals to 1, the  
 556 data agrees with the prediction. Various measurements from multiple experiments  
 557 are shown in figure 2.1. Except for Frejus, all R measurements are less than 1. This  
 558 discrepancy between the predicted R and the measured R became known as the  
 559 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 Cherenkov	$0.675 \pm 0.085$
Soudan2	Iron Tracking Calorimeter	$0.69 \pm 0.13$
IMB	Water Cherenkov	$0.54 \pm 0.12$
Kamiokande	Water Cherenkov	$0.60 \pm 0.07$
Frejus	Iron Tracking Calorimeter	$1.0 \pm 0.15$

**Table 2.1:** Measurements of the double ratio for various atmospheric neutrino experiments

560 Kamiokande-II has the capability of measuring the direction of the incoming  
 561 neutrinos. The expectation of atmospheric neutrino detection is that the flux will  
 562 be isotropic due to the fact that atmospheric neutrinos can reach the detector from  
 563 all directions. Kamiokande-II noticed that muon-like data did not agree well with  
 564 this expectation. At low energies approximately half of the  $\nu_\mu$  are missing over the  
 565 full range of zenith angles. At high energies the number of  $\nu_\mu$  coming down from  
 566 above the detector seems to agree with expectation, but half of the same  $\nu_\mu$  coming  
 567 up from below the detector are missing. This anomaly can be easily explained by  
 568 neutrino flavor oscillations. Due to the fact that the neutrino travels less distance  
 569 coming straight down into the detector (about 15 km) than coming up from the bottom  
 570 of the detector (13000 km) changes the probability of oscillation. The probability of  
 571 oscillation for the muon neutrinos coming down into the detector is roughly zero,  
 572 whereas for neutrinos coming up, the oscillation probability is  $\sin^2(2\theta)$ . Both the solar  
 573 and atmospheric neutrino problems can be explained by neutrino oscillation so it's  
 574 fitting to derive this phenomenon mathematically. In the next two sections, two flavor  
 575 and three flavor neutrino oscillation derivations will be explained.

## <sup>576</sup> 2.3 Neutrino Oscillations

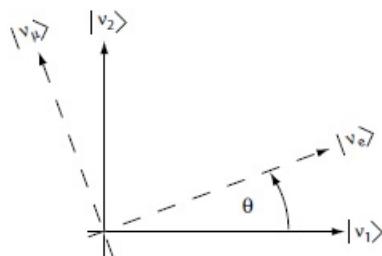
<sup>577</sup> Neutrino oscillation was first predicted by Bruno Pontecorvo. It describes the phe-  
<sup>578</sup> nomenon of a neutrino created with a specific lepton flavor (electron, muon or tau)  
<sup>579</sup> that is later measured to have a different flavor. Neutrino oscillation is important  
<sup>580</sup> theoretically and experimentally due to the fact that this observation implies that the  
<sup>581</sup> neutrino has a non-zero mass, which is not part of the original Standard Model of  
<sup>582</sup> particle physics. [2]

### <sup>583</sup> 2.3.1 Two Flavor Neutrino Oscillation Formulation

<sup>584</sup> The flavor eigenstates can oscillate between each other because they are composed  
<sup>585</sup> of an add mixture of mass eigenstates( $\nu_1, \nu_2$ ). Figure 2.3 shows the mass and flavor  
<sup>586</sup> eigenstates rotated by an angle  $\theta$  which is the mixing angle.

<sup>587</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 2.3:** The flavor eigenstates are rotated by an angle  $\theta$  with respect to the mass eigenstates

<sup>588</sup> Applying the time evolution operator to  $\nu_\mu$ :

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

<sup>589</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>590</sup> being, let us assume  $\hbar = c = 1$ . With this assumption:  $E_1 = \sqrt{p^2 + m_1^2}$  and  $E_2 =$   
<sup>591</sup>  $\sqrt{p^2 + m_2^2}$ . The next modifications is to assume neutrinos are relativistic:

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

<sup>592</sup> because of this,

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

<sup>593</sup>

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

<sup>594</sup> where the binomial expansion is used. Now  $E_1$  and  $E_2$  can be written as:

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

<sup>595</sup> Now applying all these assumptions back into equation 2.7 gives us:

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

<sup>596</sup>

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

<sup>597</sup> 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:

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

<sup>598</sup> Finding the Probability for a  $\nu_\mu \rightarrow \nu_e$ :

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

599 Remembering that  $\langle \nu_i | \nu_j \rangle = \delta_{ij}$

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

600 Taking the absolute value squared gives us:

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

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

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

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

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

605 This equations has three important variables.

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

### 617 2.3.2 Three Flavor Neutrino Oscillation Formulation

618 Seeing the quantum mechanics involved in deriving the probability of a two flavor  
 619 neutrino oscillation, it is now possible to formulate the three flavor neutrino oscillation.  
 620 The three flavor neutrino oscillation formulation begins similarly to the two flavor,  
 621 but there is the Pontecorvo-Maki-Nakagawa-Sakata matrix (PMNS) instead of the 2X2  
 622 matrix in the previous section. The PMNS matrix is show below:

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

623 where  $c_{ij} = \cos\theta_{ij}$  and  $s_{ij} = \sin\theta_{ij}$

624 Following the same steps as before we get:

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

625 The main things to notice here are  $\delta_{ij}$  which is the CP violating term and has not  
 626 been measured yet, and  $\theta_{13}$  which has just been measured. CP violation is a violation  
 627 of the postulated CP-symmetry. CP-symmetry states that the laws of physics should  
 628 be the same if a particle were to be exchanged with its antiparticle and then if the left  
 629 hand side of a decay were switched with the right hand side. Table 2.2 shows the  
 630 current world knowledge of the values of all the fundamental parameters for neutrino  
 631 oscillations [?].

## 632 2.4 Neutrino Interactions

### 633 2.4.1 Weak Interactions

### 634 2.4.2 Neutrino-Nucleon Interactions

---

Parameter	Value
$\theta_{12}$	$33.9 \pm 1.0^\circ$
$\theta_{23}$	$39^\circ < \theta_{23} < 51^\circ$
$\theta_{13}$	$9.1 \pm 0.6^\circ$
$\Delta m_{21}^2$	$(7.50 \pm 0.20) * 10^{-5} \text{ eV}^2$
$ \Delta m_{32}^2 $	$(2.32^{+0.12}_{-0.08}) * 10^{-3} \text{ eV}^2$
$\delta_{cp}$	unknown

**Table 2.2:** Current world knowledge of neutrino oscillation parameters [?]

# <sup>635</sup> Chapter 3

## <sup>636</sup> The MicroBooNE Experiment

<sup>637</sup> The purpose of this chapter is to discuss and understand the details of the MicroBooNE  
<sup>638</sup> detector. A thorough understanding of MicroBooNE and the technology behind liquid  
<sup>639</sup> argon time projection chambers is important for understanding results as well as  
<sup>640</sup> understanding how images were made for use in deep learning efforts that will be  
<sup>641</sup> outlined in later chapters.

### <sup>642</sup> 3.1 Liquid argon time projection chambers

<sup>643</sup> Liquid Argon Time Projection Chambers (LArTPCs) are an exciting detector technol-  
<sup>644</sup> ogy that provide excellent imaging and particle identification, and are now being  
<sup>645</sup> used to study neutrinos. The Time Projection Chamber (TPC) was first invented by  
<sup>646</sup> Nygren in 1974 [?] and the proposal for a LArTPC for neutrino physics was made  
<sup>647</sup> by Rubbia [?] in 1977 with the ICARUS collaboration implementing this concept [?].  
<sup>648</sup> A LArTPC is a three-dimensional imaging detector that uses planes of wires at the  
<sup>649</sup> edge of an active volume to read out an interaction. When a neutrino interacts with an  
<sup>650</sup> argon atom, the charged particles that are produced ionize the LAr as they travel away  
<sup>651</sup> from the interaction. By placing a uniform electric field throughout the LAr volume,  
<sup>652</sup> the ionization is made to drift towards a set of anode planes, which consist of wires  
<sup>653</sup> spaced very closely together collecting the ionized charge, which is subsequently read  
<sup>654</sup> out by electronics connected to the anode wires. The collected ionization creates a  
<sup>655</sup> spatial image of what happened in the detector on each anode plane. The position  
<sup>656</sup> resolution of the interaction along the beam direction (perpendicular to drift direction)  
<sup>657</sup> relies on the wire pitch, while the resolution in drift direction is dependent on the

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

## 683 3.2 The MicroBooNE Time Projection Chamber

684 MicroBooNE (Micro Booster Neutrino Experiment) is a 89 ton active volume (180 ton  
685 total mass) LArTPC which is then inserted into a cylindrical cryostat on axis of the  
686 Booster Neutrino Beam (BNB) stationed at Fermilab in Batavia, Illinois. Understanding  
687 LArTPC technology and detector physics is necessary to build a LArTPC the size of  
688 DUNE, and MicroBooNE has made many advances in developing this technology [7]  
689 [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.

725 The acrylic plates take in the scintillation light, at 128 nm, and re-emits it visible  
726 wavelengths visible to the PMTs, with a peak at 425 nm.

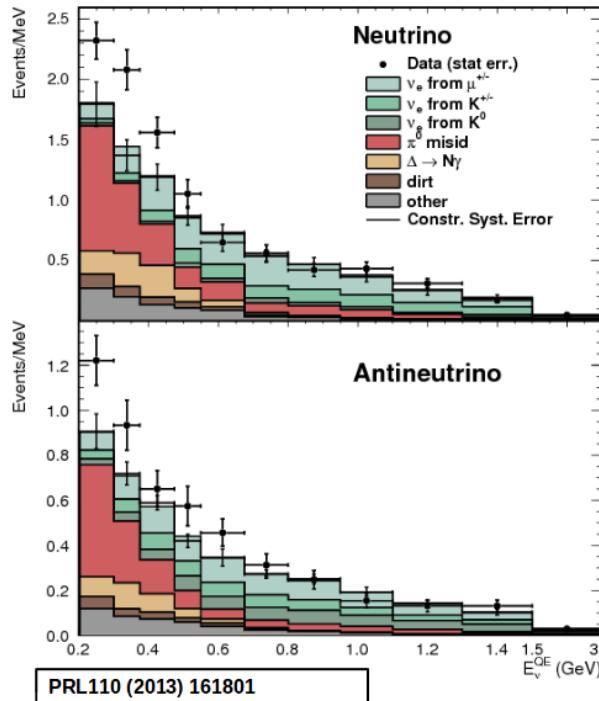
727 Both the light collection system and the TPC create analog signal that is read out and  
728 digitized by the electronics system. The process requires amplification and shaping of  
729 the signal which then goes to the data acquisition (DAQ) software for writing of the  
730 digitized data to disk.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

### 752 3.3 MicroBooNE's Physics Goals

#### 753 3.3.1 The low-energy excess

754 The primary goal of the MicroBooNE experiment is to study and investigate the low-  
755 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 1 GeV energy range provides information about short range nuclear correlations that affect the interpretations of neutrino oscillation experiment data.

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

789 **3.3.3 Liquid argon detector development**

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

794 **3.4 The Booster Neutrino Beam**

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

802 introduces to a measurement. An aerial view of fermilab as well as the BNB is shown  
 803 in figure 3.2

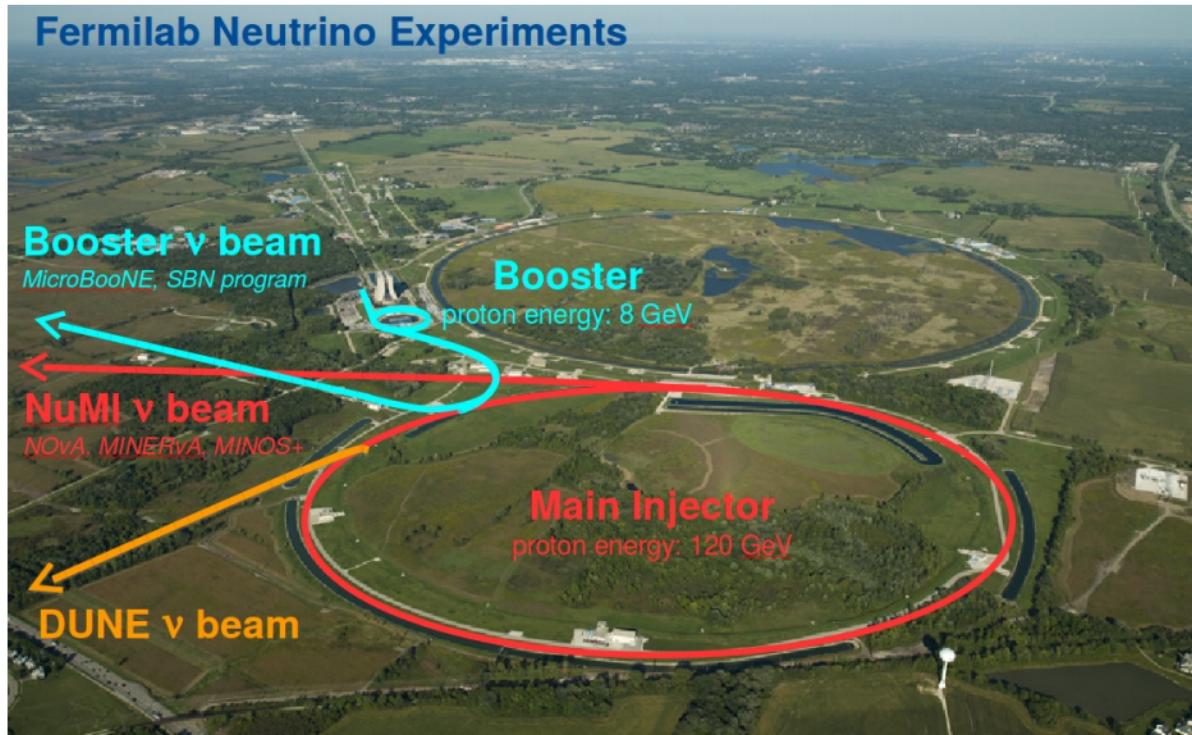


Figure 3.2: Arial view of the Main Injector and the Booster Neutrino Beam at Fermilab

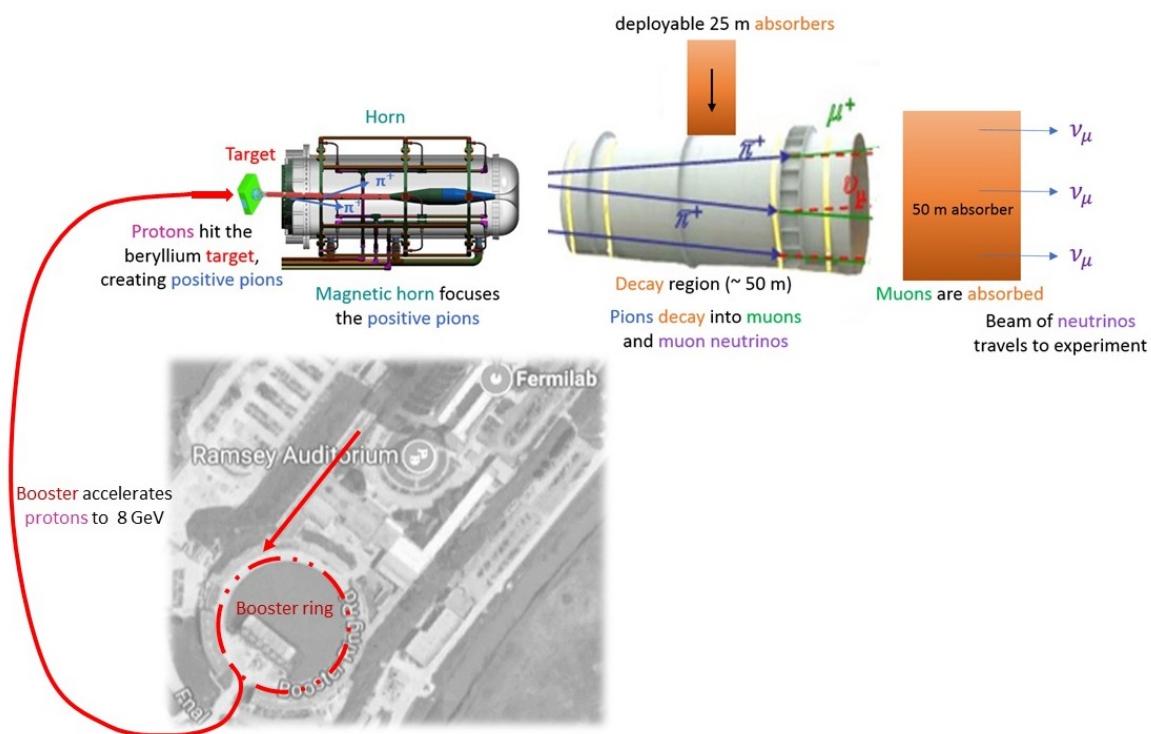
### 804 3.4.1 Creating the Booster Neutrino Beam

805 The BNB is a very pure  $\nu_\mu$  beam, with only 0.6% contamination from  $\nu_e$ s. The energy  
 806 also peaks around 700 MeV which is desired based on the probability of oscillation  
 807 equation which depends on the the value of  $L/E$ , where  $L$  is the distance of the  
 808 detector from the neutrino beam and  $E$  is the energy of the neutrino beam.  $L/E$  was  
 809 chosen to increase the probability of seeing neutrino oscillations in the MiniBooNE  
 810 Low Energy Excess (LEE) range based on the probability of oscillation equation, which  
 811 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  
 812 protons from the FNAL booster synchrotron into a beryllium target which produces a  
 813 high flux of neutrinos. The protons originate from  $H^2$  gas molecules that are turned  
 814 into  $H^-$  ions by a Cockcroft-Walton generator shown in figure ???. The  $H^-$  initially are  
 815 accelerated to 1MeV kinetic energy and are then passed to a linear accelerator using  
 816 alternating electromagnetic fields to increase their energy to 400MeV. The ions are  
 817 stripped of electrons by passing them through a carbon foil. The protons are bunched

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

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

839 **3.5 Event Reconstruction**



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

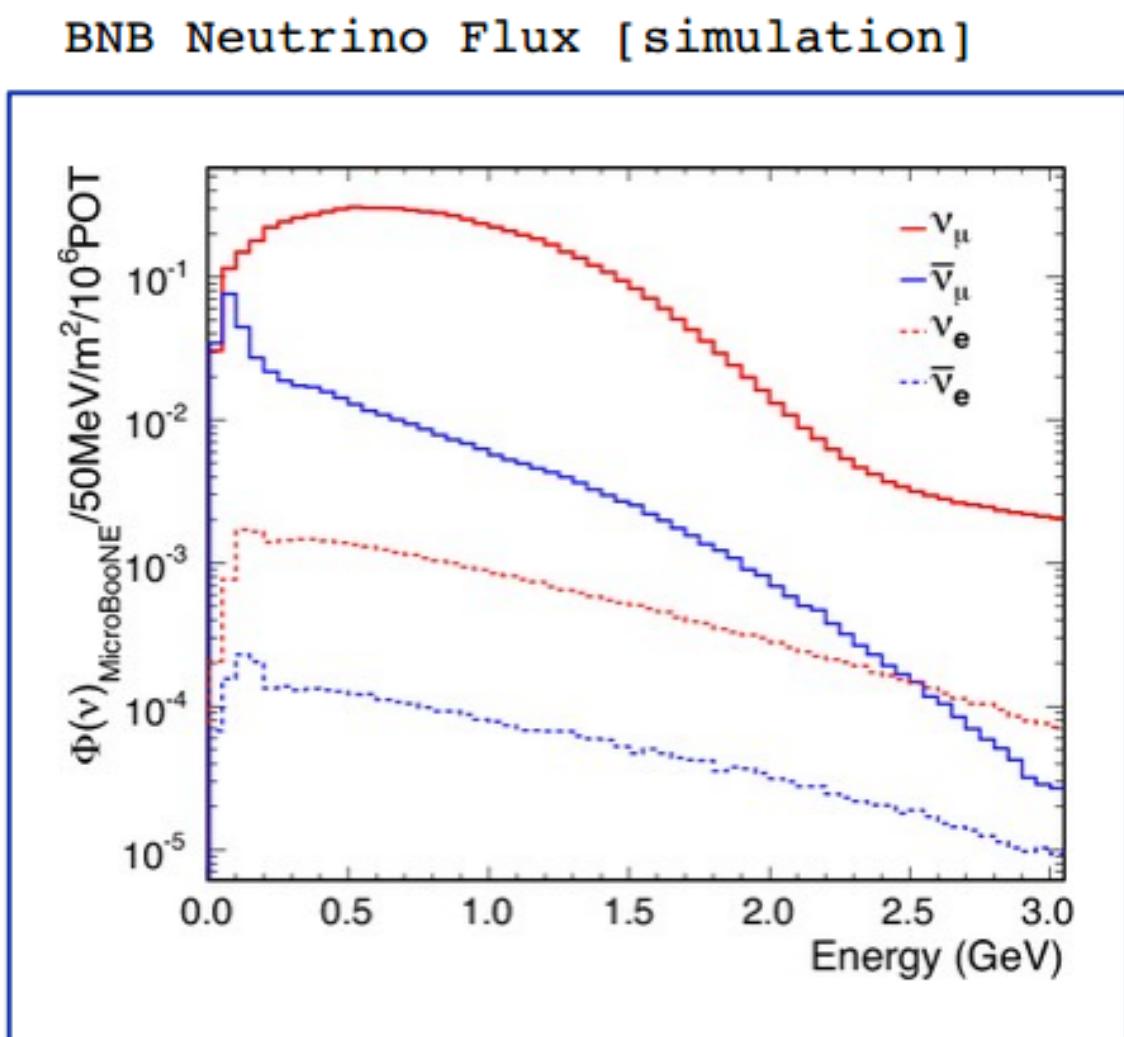


Figure 3.4: Energy spectrum of the Booster Neutrino Beam at Fermi National Laboratories

840 **Chapter 4**

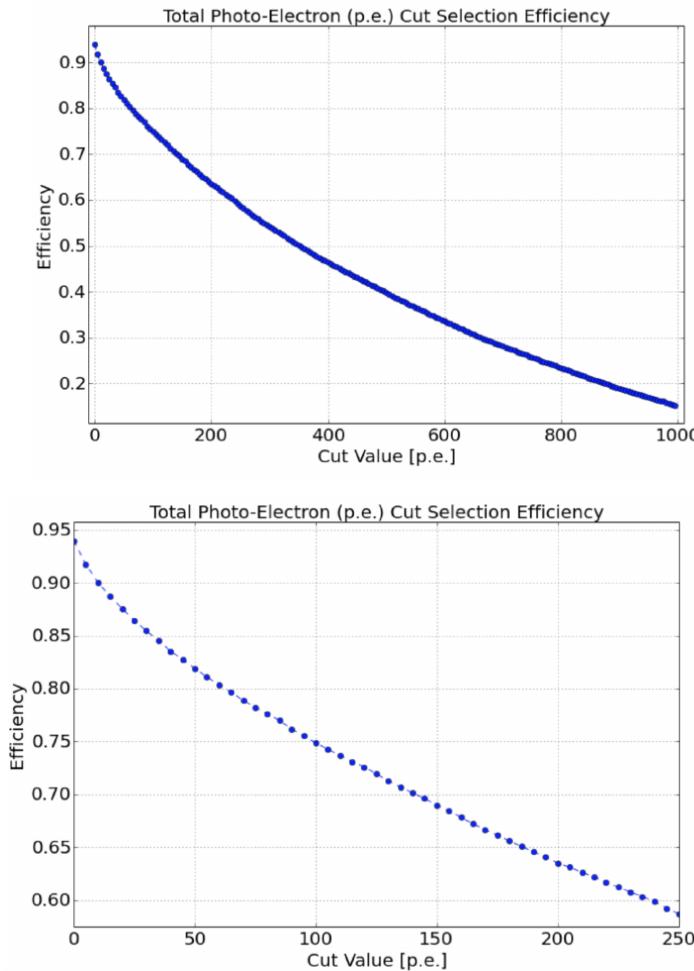
841 **Neutrino Identification: Finding**  
842 **MicroBooNE's first Neutrinos**

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

857 **4.1 Flash Finding**

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

861 **4.1.1 Flash Reconstruction**

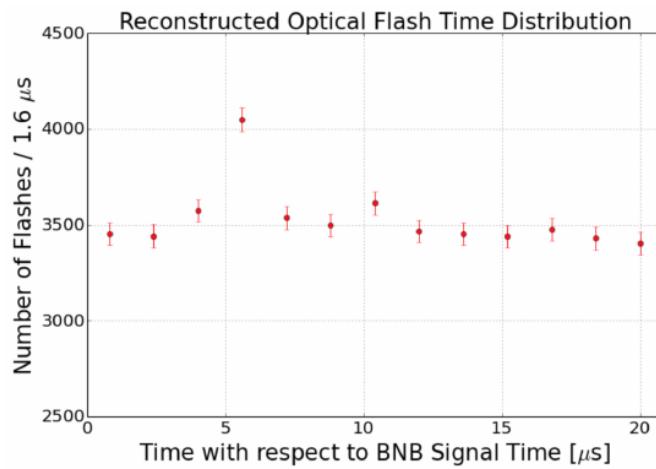


**Figure 4.1:** Efficiency for selecting beam events as a function of minimum total PE cut for all PE cuts as well as zoomed into interesting region.

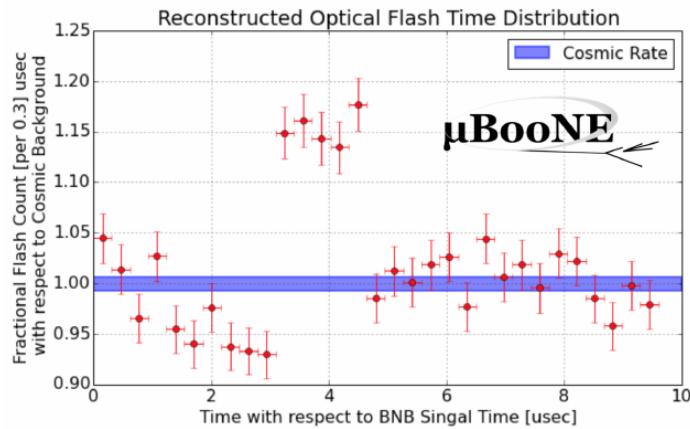
862 A flash is described as a collection of light seen at the same time within the detector.  
863 They are then reconstructed by identifying signal from the PMTs above a specific  
864 photoelectron (PE) threshold. These signals are called optical hits. Optical hits from  
865 all the PMTs are then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

<sup>872</sup> cut of 50 PE was deemed sufficient for this analysis. Figure 4.1 show the total PE  
<sup>873</sup> versus the selection efficency of selecting neutrino beam events.

#### <sup>874</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>875</sup> It is necessary to get the specific time from flashes if one uses flashes to filter out  
<sup>876</sup> neutrino interactions coincident with the neutrino beam spill period and background.  
<sup>877</sup> Before a filter can be applied, an understanding of the timing of the trigger and PMT

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

#### 4.1.3 Event Rates

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

## 4.2 TPC Topology Selection

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

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

### 904 4.2.1 Cosmic Tagging

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

- 909 • 1: The track is tagged as entering or exiting the TPC
- 910 • 0.95: The track is a delta ray associated with a tagged track
- 911 • 0.5: The track is either entering or exiting, but not both
- 912 • 0.4: The track is entering or exiting through the Z boundary
- 913 • 0: The track isn't tagged

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

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

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

### 929 4.2.2 2D Cluster Selection

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

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

939 **Primary Cuts**

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

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

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

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

959 **Secondary Cuts**

960 The secondary cuts look to match long, low-angle clusters with short, high-charge  
961 clusters. Only clusters that have passed previous cuts are used. First clusters with  
962 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.

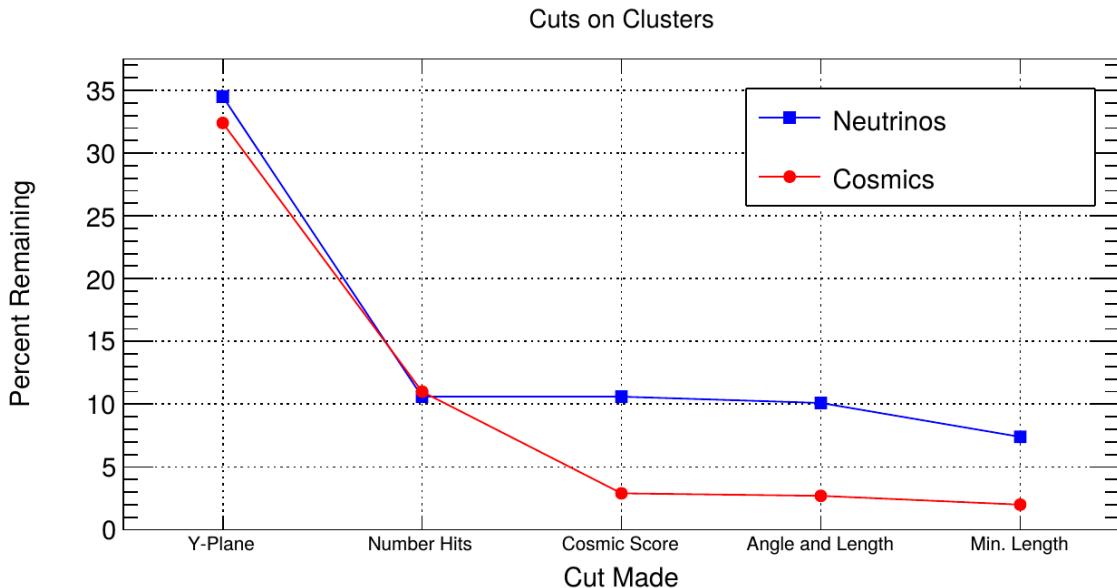
963 direction. Then we search for any cluster that is within approximately 3 cm ( 10 wires  
 964 and 30 time ticks) away from the low-z end of the long cluster. This cluster must also  
 965 be shorter than the first. In our reconstruction, the start and end point of a cluster can  
 966 be swapped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

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

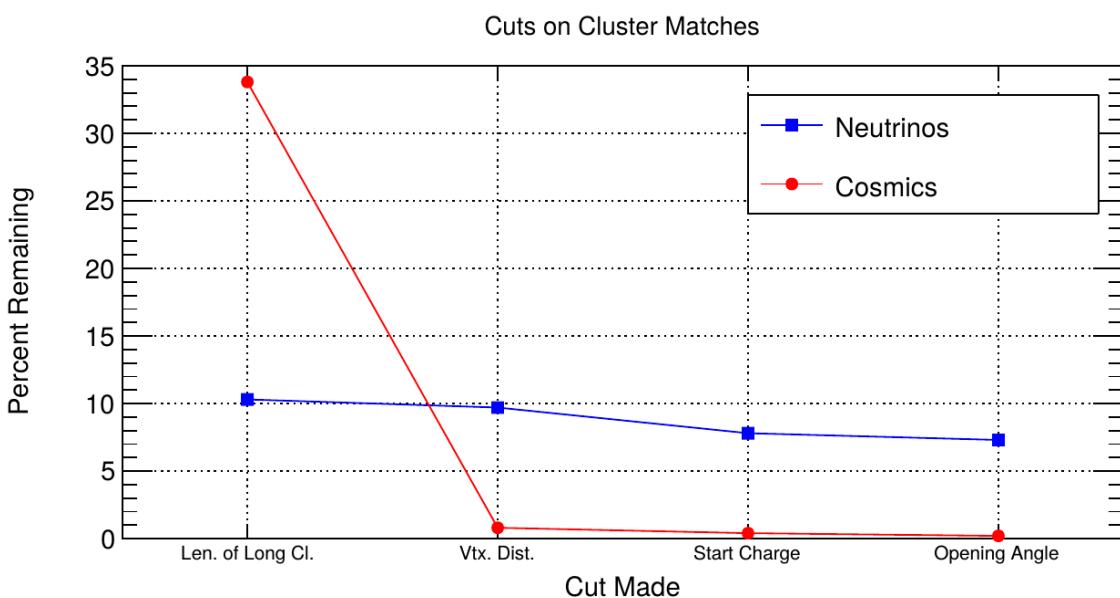
### 975 4.2.3 3D Tracks and vertices Selection

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

- 979 •  $d$ : distance between the start points of the two tracks.
- 980 •  $d_1$ : distance between vertex and start of track 1.
- 981 •  $d_2$ : 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.

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

- 989 • trackkalmanhit with cccluster  $\min(\max_{d,i}) < 3$  cm.
- 990 • trackkalmanhit with pandoraNu  $\min(\max_{d,i}) < 4.5$  cm.
- 991 • pandoraNu with cccluster  $\min(\max_{d,i}) < 5$  cm.

#### 992 4.2.4 TPC Updates

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

#### 999 2D Filter Updates

1000 The main background observed in the 2D filter were Michel events, where the muon  
1001 and electron formed two connected clusters. These events were rejected by comparing  
1002 the start and end charge deposition of the long cluster (i.e muon particle). The start  
1003 charge deposition must be less than the end charge deposition. This cut is implemented  
1004 because muons have a higher ionizaiton loss at the end.

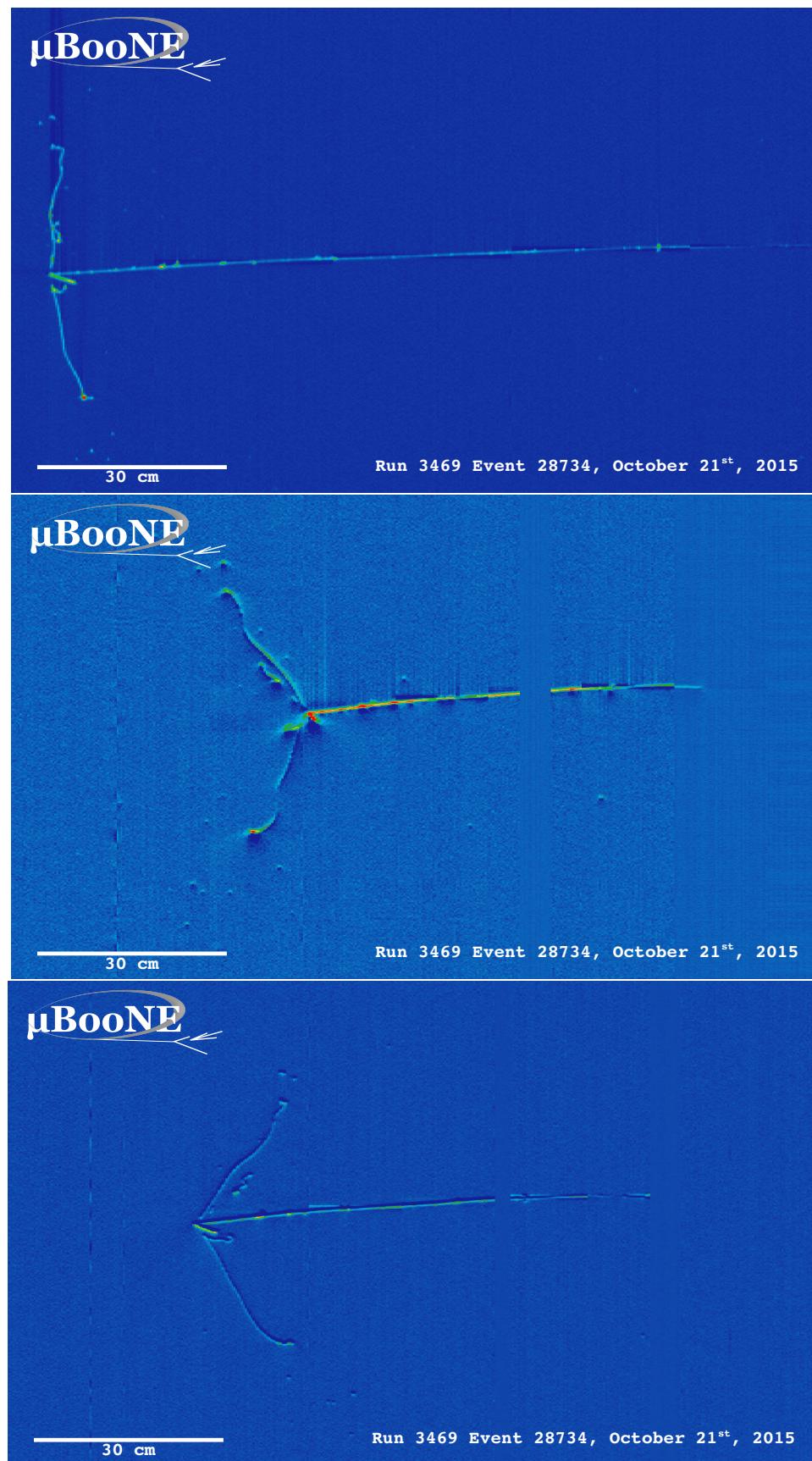
#### 1005 3D Filter Updates

1006 It was seen that cosmic tracks can often originate or end at the same point, therefore  
1007 faking a signal. Cosmic tracks, however, are mostly vertical. By requiring the angle  
1008 of the longer track have a cosine greater than 0.85 with respect to the z-axis as well

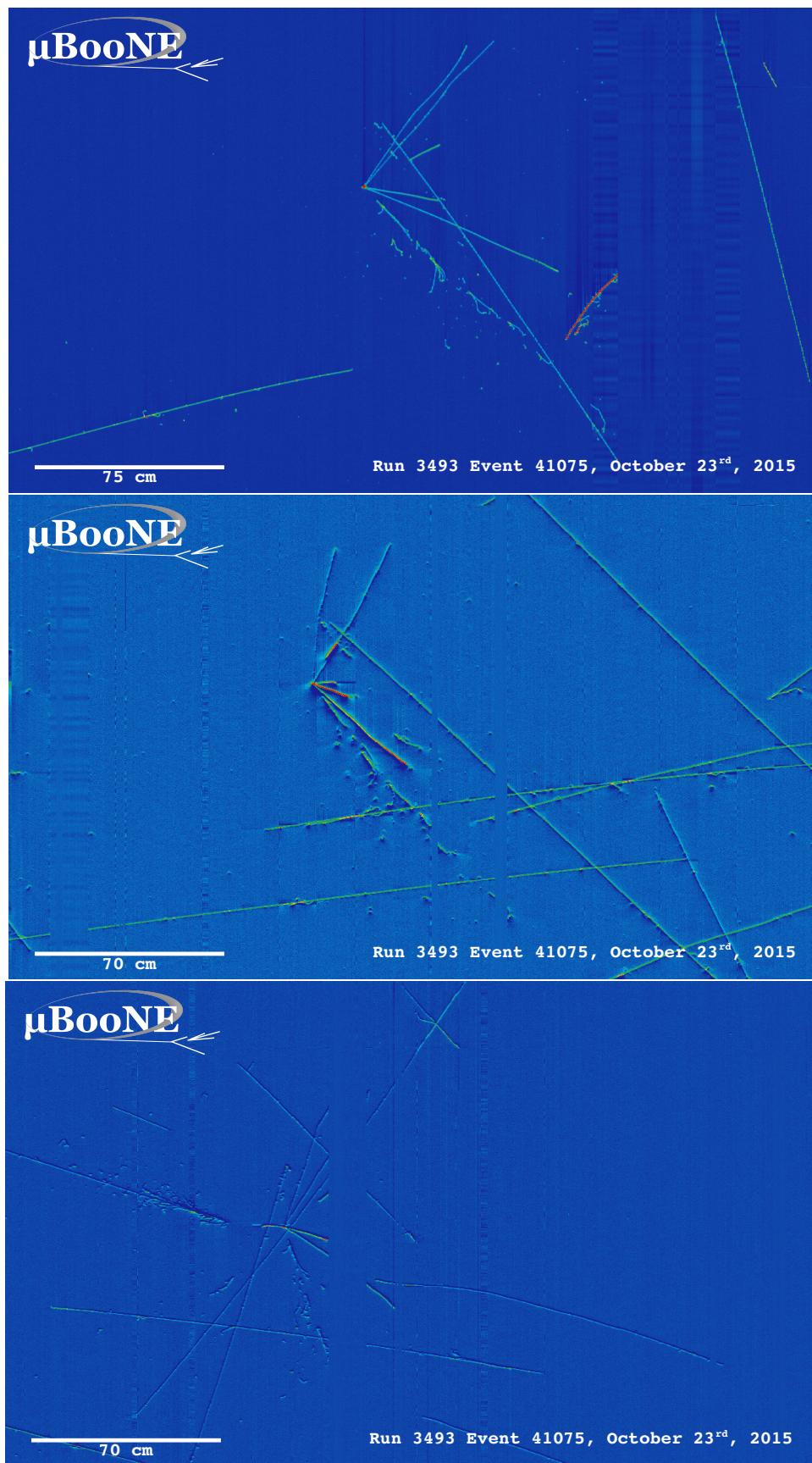
1009 as requiring the longer track to have a length greater than 10 cm, we can reduce this  
1010 background.

1011 **4.3 Conclusion**

1012 After proccesing these filters in parallel, it was shown that the 3D filter had a higher  
1013 purity than the 2D filter because of the higher cosmic rejection being used due to 3D  
1014 reconstruction. The 2D filter is blind to track entering/exiting from the top or bottom  
1015 of the TPC. Although the 3D filter had a higher purity, the 2D filter was still able to  
1016 find identifiable events in data that were used as public event displays. A sample of  
1017 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

1018 **Chapter 5**

1019 **CC-Inclusive Cross Section Selection**

1020 **Filter**

1021 The CC-Inclusive cross-section selection I and selection I modified filters used in this  
1022 analysis will be described in the following sections below. These filters are an expan-  
1023 sion of the Neutrino ID filter. The work done in this thesis was to further improve these  
1024 selections by increasing both efficiency and purity as well as increasing acceptance  
1025 without further affecting the kinematic distributions of the selected neutrino events.

1026 MicroBooNE requires fully automated event reconstruction and selection algo-  
1027 rithms for use in the many physics measurements being worked on to date due to  
1028 the large data rate MicroBooNE receives. Being able to automatically pluck out the  
1029 neutrino interaction among a sea of cosmics proved to be challenging but was accom-  
1030 plished. MicroBooNE has developed two complementary and preliminary selection  
1031 algorithms to select charged-current  $\nu_\mu - Ar$  interactions. Both are fully automated  
1032 and cut based. The results of this thesis will focus on selection I and selection I modi-  
1033 fied and will focus on further improving these algorithms using Convolutional Neural  
1034 Network (CNN) implementations. These selections identify the muon from a neutrino  
1035 interaction without biasing towards track multiplicity. To combat cosmic and neutral  
1036 current background, the analysis is strongly biased towards forward-going long tracks  
1037 which are contained. This limits phase space and reduces acceptance.

## 1038 5.1 Data and MC Processing Chain

1039 The data used for this analysis were based on hardware and software triggers. Events  
1040 used came from the *BNB\_INCLUSIVE* and *EXT\_BNB\_INCLUSIVE* streams and were  
1041 used for signal and background. The *BNB\_INCLUSIVE* stream is chosen by requiring  
1042 that the hardware trigger bit is fired and that the event passed an optical software  
1043 trigger within a BNB spill timing window. The *EXT\_BNB\_INCLUSIVE* stream requires  
1044 the EXT hardware trigger to fire as well as pass the same optical software trigger  
1045 within a BNB spill size timing window similar to the *BNB\_INCLUSIVE*.

1046 The two MC samples used in this analysis and for determining selection efficiencies  
1047 and purities were GENIE BNB neutrino interactions with CORSIKA cosmic ray overlay  
1048 within the readout window and inTime CORSIKA cosmic rays. The MC samples  
1049 generated used *uboonecode v04\_36\_00* and are based on the following packages:

- 1050 • larsoft v04\_36\_00
- 1051 • GEANT v04\_09\_06\_p04d
- 1052 • GENIE v02\_08\_06d
- 1053 • GENIE xsec v02\_08\_06a
- 1054 • pandora v02\_03\_0a
- 1055 • CORSIKA v07\_4003

1056 Both data and MC samples were processed using the same reconstruction release,  
1057 *uboonecode v05\_08\_00* and the fcl files used for reconstruction are listed below:

- 1058 • MC fcl files
  - 1059 – reco\_uboone\_mcc7\_driver\_stage1.fcl
  - 1060 – reco\_uboone\_mcc7\_driver\_stage2.fcl
- 1061 • Data fcl files
  - 1062 • reco\_uboone\_data\_Feb2016\_driver\_stage1.fcl
  - 1063 • reco\_uboone\_data\_Feb2016\_driver\_stage2.fcl

1064 On top of the hardware and software triggers, the data also had to pass more  
1065 criteria to be identified as part of the good run list. The criteria is detailed below.

- 1066     • **Detector conditions:** the detector has to be in a good operating condition. The  
 1067        detector conditions are read from the slow monitoring database and are required  
 1068        to be within the alarm thresholds. The variables of interest for events passing  
 1069        the good run list criteria include DAQ, PMT, HV, Drift HV, wire bias, electron  
 1070        lifetime and detector power. These conditions need to be met on a run-by-run  
 1071        basis in order to pass the selection.
- 1072     • **Data quality:** normal and stable behavior for basic reconstruction quantities.  
 1073        These reconstruction variables include average number of tracks, hits, and flashes  
 1074        in each event, the average length of tracks, the average amplitude and area of  
 1075        hits, the average PE and the average spread of each one of these quantities.
- 1076     • **Beam Conditions:** the BNB must be on and stable and the POT per spill needs  
 1077        to above the intensity threshold. Beam quality conditions include checking the  
 1078        fraction of proton beam interacting within the target, the horn current, and the  
 1079        intensity of protons per spill. The final sample is  $5 * 10^{19}$  and a per-spill intensity  
 1080        of  $4 * 10^{12}$
- 1081     • **Run processed:** the full run must be processed completely without missing  
 1082        subruns or crashes in the data processing.

## 1083 5.2 Normalization of data and MC

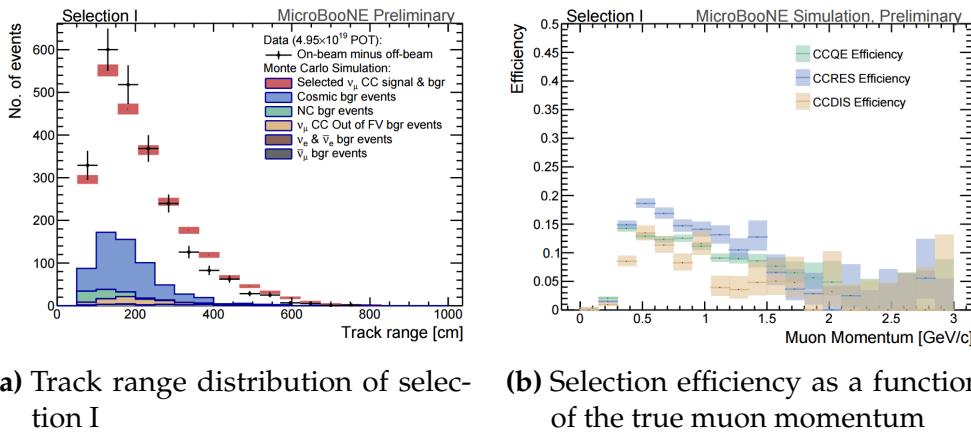
1084 The off-beam sample is used to measure beam unrelated backgrounds. For normalization  
 1085 one needs the total number of BNB spills ( $N_{BNB}$ ) and the total number of external  
 1086 triggers. The BNB spills used need to pass the beam quality cuts. The normalization  
 1087 factor is then  $N_{BNB}/N_{EXT}$  which is 1.23.

1088 To normalize generated BNB MC events to POT, we used the following:

- 1089     •  $5 * 10^{19} POT = 41524.3$  generated events

1090 where this scaling factor only applies to mcc7 generated events. The inTime cosmic  
 1091 sample is normalized with respect to the open cosmic sample so an understanding  
 1092 of both is necessary. The POT per beam spill for mcc7 BNB samples is  $5 * 10^{12}$ . To  
 1093 calculate how many spills are necessary to produce a specific POT one would multiply  
 1094 the total POT by the average 1/POT per spill. For a total POT of  $5 * 10^{19}$  the amount  
 1095 of spills necessary is  $\frac{5 * 10^{19}}{5 * 10^{12}} = 1 * 10^7$ . This is only one in  $\sim 241$  events therefore each

1096 cosmic event needs to be scaled up by a factor of 240.8 when comparing to BNB  
 1097 MC. For inTime cosmics however, two filters are applied to reduce computing and  
 1098 processing time and only leave cosmics that will interact within the detector. The  
 1099 passing rate after these two filters is 0.02125, therefore the total inTime cosmic scaling  
 1100 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.

## 1101 5.3 Optical Software Trigger and Reconstruction

### 1102 5.3.1 Software Trigger

1103 Most of the BNB spills from the accelerator do not have a neutrino interaction in  
 1104 MicroBooNE. To save computation resources and reduce data-rates, we require a  
 1105 burst of light in the light collection system in coincidence with the 1.6  $\mu\text{s}$  beam spill.  
 1106 Requiring light activity in coincidence with the beam spill eliminates the vast majority  
 1107 of triggers with no neutrino interaction in the detector, however, it doesn't guarantee  
 1108 the activity in the detector is a neutrino interaction since a cosmic ray can interact in  
 1109 coincidence with the beam spill as well.

1110 To implement this, a software trigger was used on the PMT waveforms to decide  
1111 whether or not to keep that event. The software trigger is implemented after the event  
1112 builder combines data from the PMTs and triggers into a single event. The software  
1113 trigger uses the digitized output of the 32 PMT channels in the light collection system.  
1114 Only the waveform region in coincidence with the beam spill is used to search for  
1115 possible triggers. For each PMT, a waveform is found by taking the difference of ADC  
1116 values is calculated between  $t$  and  $t + s$ . This waveform is then scanned for ADC  
1117 values above a threshold  $X_0$ . Once an ADC is above this threshold, a discriminator  
1118 window is opened for a fixed number of time ticks ( $W_0$ ). If the ADC count within this  
1119 window  $W_0$  is greater than a second larger threshold  $X_3$ , a final window of width  $W_3$   
1120 is opened. The max ADC value within this final window is set as the peak amplitude  
1121 for the PMT and then summed across all 32 PMTs and set to the variable PHMAX. The  
1122 software trigger places a final cut on the PHMAX variable to decide whether or not  
1123 to keep the event. The thresholds were found by the Trigger task force using Monte  
1124 Carlo Studies and are as follows:

- 1125 •  $X_0 = 5$  ADC
- 1126 •  $X_3 = 10$  ADC
- 1127 •  $W_0 = 6$  Ticks
- 1128 •  $W_3 = 6$  Ticks
- 1129 • PHMAX cut = 130 ADC

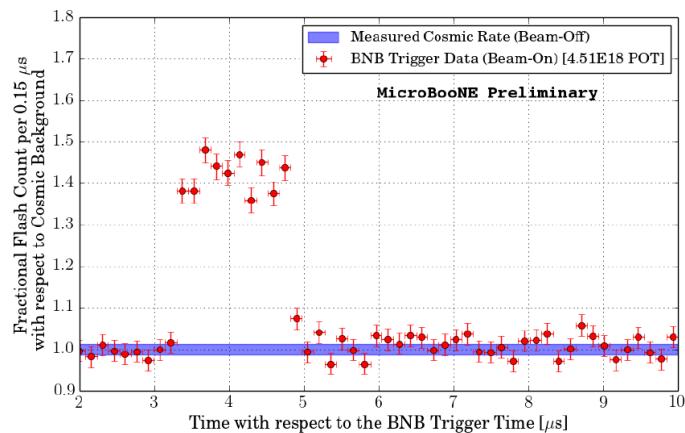
1130 **5.3.2 Flash Reconstruction**

1131 MicroBooNE collects light from each of the 32 PMTs either in a continuous readout  
1132 window of  $23.4 \mu\text{s}$  activated by a beam gate signal on the trigger board, or in discrimi-  
1133 nated pulses of  $\sim 1 \mu\text{s}$  duration activated if the ADC count for any PMT goes above 80  
1134 ADC count. These two formats are saved as output waveforms and put onto an event.  
1135 Additionally, each PMT can provide two output streams, high-gain ( $\sim 20$  ADC/PE)  
1136 and low-gain ( $\sim 2$  ADC/PE) channels. The first step in the reconstruction is to merge  
1137 both these channels into a “saturation corrected waveform” which uses information  
1138 from the low-gain waveform to correct for saturating high-gain pulses.

1139 The saturation corrected waveform in the continuous readout window is used to  
1140 reconstruct optical hits. Each PMT’s waveform is scanned for hits then a threshold

1141 based hit reconstruction algorithm is applied which requires pulses of a minimum  
1142 area in order to be reconstructed. Each reconstructed hit is associated to a PMT, a time  
1143 in  $\mu\text{s}$ , and a PE count.

1144 Once hits are reconstructed for all 32 PMTs, all PMT information is then combined  
1145 into optical flashes which represent optical information seen by the PMTs from interac-  
1146 tions in the detector. Each flash has information on total light seen per interaction, the  
1147 distribution of the light across all 32 PMTs, the flash time with respect to the trigger  
1148 time of the flash, and lastly, the spacial information of the flash in Y-Z plane of the  
1149 detector. These flashes are reconstructed by requiring that there is a  $\sim 1 \mu\text{s}$  coincidence  
1150 between the reconstructed hits in all 32 PMTs. The total PE is summed up among  
1151 all coincident hits across the PMTs and if the total PE is greater than 2 PE, a flash is  
1152 reconstructed. There are also safe guards in place to take care of late scintillation light.

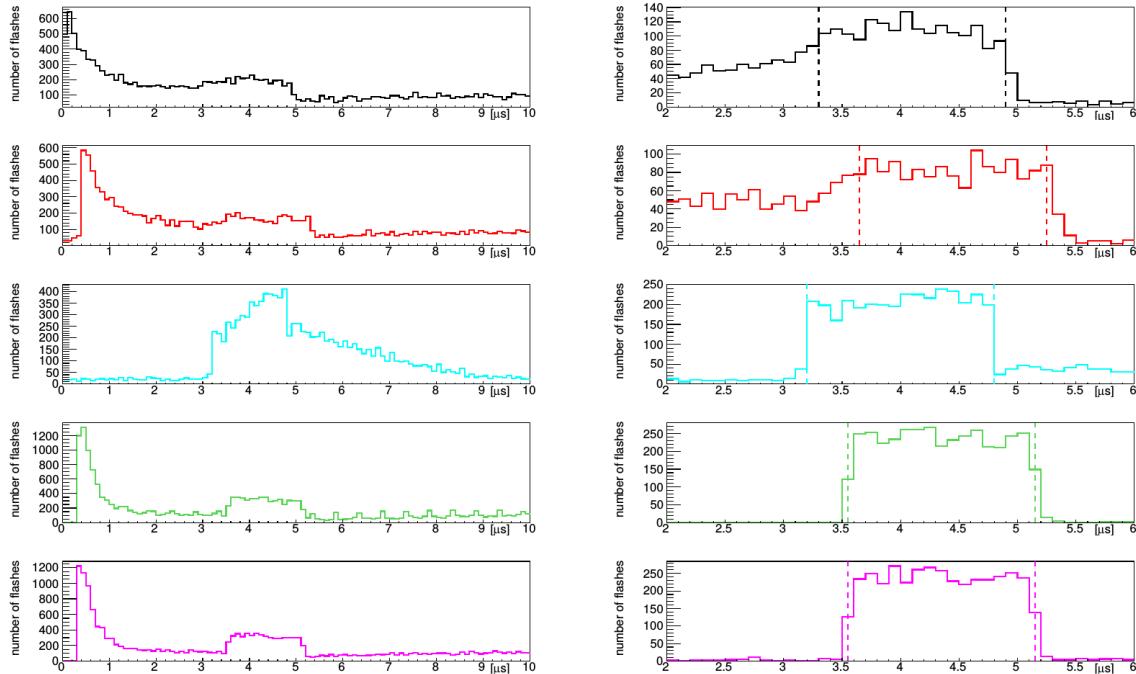


**Figure 5.2:** Time distribution of reconstructed optical flashes with a PE value of 50 or more for a sample of BNB unbiased triggered events.

1153 Figure 5.2 shows the time distribution of reconstructed optical flashes using the  
1154 BNB continuous stream. You can see a clear excess in coincidence with the expected  
1155 arrival time of neutrinos. The same flash reconstruction that was used in the cc-  
1156 inclusive filter detailed here was used to create this plot in data.

### 1157 5.3.3 Beam Window

1158 Figure 5.3 shows the distribution of flashes for on-beam, off-beam and various MC  
1159 samples. The software trigger has been applied to these samples. The pile-up seen just  
1160 after 0  $\mu\text{s}$  is a feature of the flash finding algorithm and consists of low PE flashes and



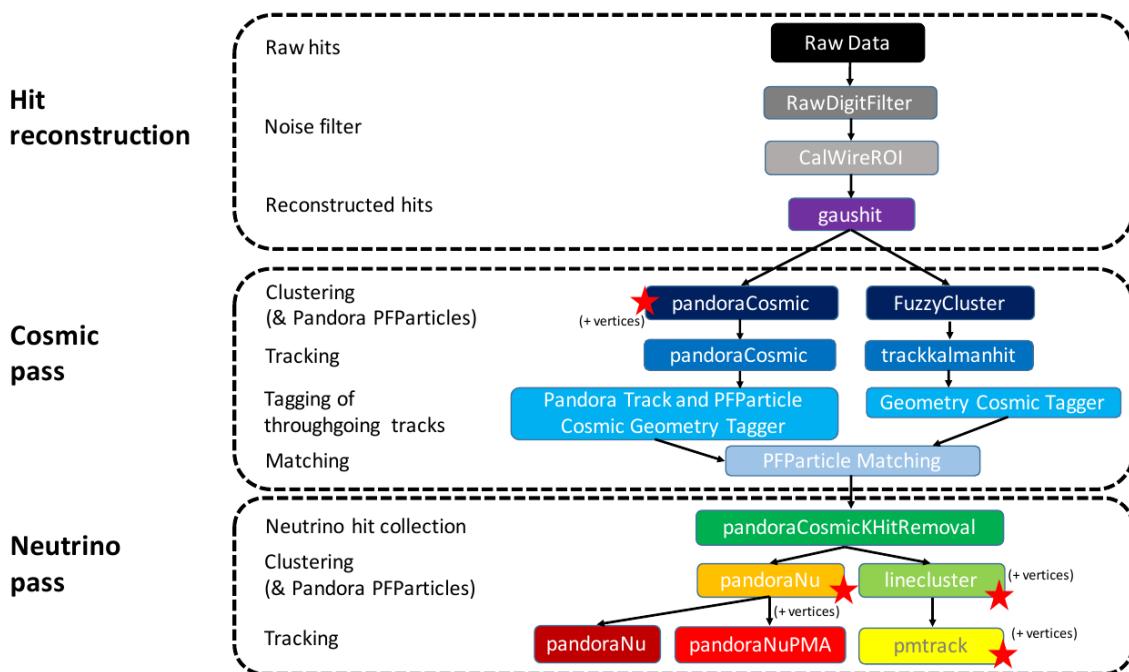
**Figure 5.3:** Flash time distribution for all flashes (left plot) and flashes  $> 20\text{PE}$  (right plot). The different curves are as follows: on-beam data (black), off-beam data (red), CORSIKA inTime MC (light blue), BNB only MC (green), and BNB+Cosmic MC (purple). The dashed vertical lines mark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

1161 is removed in the second column of distributions with a low 20 PE threshold cut. The  
1162 plots show that the time window for the distributions are shifted a small amount from  
1163 each-other. This is caused by different hardware configurations per sample. Using  
1164 these distributions, the windows chosen per sample are as follows:

- 1165 • On-Beam: 3.3 to 4.9  $\mu$ s
- 1166 • Off-Beam: 3.65 to 5.25  $\mu$ s
- 1167 • CORSIKA inTime: 3.2 to 4.8  $\mu$ s
- 1168 • BNB only: 3.55 to 5.15  $\mu$ s
- 1169 • BNB+Cosmic: 3.55 to 5.15  $\mu$ s

1170 Each window has a width of 1.6  $\mu$ s.

## 1171 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.

1172     Figure 5.4 summarizes the reconstruction chain applied to both MC and data for  
1173     this analysis. After the hit reconstruction, a cosmic pass is applied which removes all  
1174     hits associated to through-going tracks. A description of these TPC reconstruction  
1175     algorithms will be detailed below.

1176     **5.4.1 Hit Reconstruction**

1177     The waveforms used for hit reconstruction consist of charge deposited on the sense  
1178     wire in drift time. The first step in hit reconstruction is to pass the waveforms through  
1179     a filtering algorithm to filter out the noise introduced from the electronics. The input  
1180     waveforms are also truncated from 9600 time ticks to 6400 time ticks in this first step  
1181     to reduce the data footprint of these waveforms.

1182     Once noise filtering is complete, a deconvolution algorithm is applied to the wave-  
1183     forms to remove the drift field and electronics response, therefore leaving only the  
1184     ionized electrons kicked off the argon atoms by an incident track. During this process,  
1185     Region of Interests (ROI) are identified and cut out of the waveforms to further reduce  
1186     the data volume.

1187     The hit finding algorithm then finds candidate peaks in these ROI's and fits the  
1188     peaks to Gaussian curves. These Gaussian shaped peaks are now called hits and  
1189     represent the charge deposition on a wire by the incoming track. These hit objects  
1190     have a peak time and width and are the basic object input to further algorithms down  
1191     the reconstruction chain.

1192     **5.4.2 Clustering**

1193     There are multiple clustering algorithms used in this analysis. The main purpose of all  
1194     the clustering algorithms is to associate hits together in 2D space to create objects like  
1195     tracks, vertices and showers. For the fuzzy cluster algorithm, three steps are used to  
1196     achieve this. The first step is to associate hits to each-other using a fuzzy clustering  
1197     algorithm which gives each hit a degree of belonging to the cluster. Second, a Hough  
1198     transform is used to find hits associated to candidate tracks and showers within each  
1199     of the clusters found in the first step. The last step merges smaller candidate tracks  
1200     and showers into large clusters. The last step also associates unclustered hits into

1201 nearby objects which helps shower reconstruction. The result is a set of clusters made  
1202 up of associate hits that represent tracks or showers per plane.

1203 The pandora algorithm utilizes it's own clustering algorithm and will be detailed  
1204 in the next section. The last clustering algorithm is called linecluster. The linecluster  
1205 algorithm reconstructs 2D linear clusters per plane by fitting a line onto nearby hits  
1206 which is then extrapolated to neighboring wires. 2D vertices are found per plane by  
1207 using the intersection points of the ends of nearby clusters. These 2D vertices are then  
1208 matched in time across all three planes to get a 3D vertex in space.

### 1209 5.4.3 Pandora

### 1210 5.4.4 Trackkalmanhit

1211 The trackkalmanhit algorithm takes 2D clusters returned from the fuzzy cluster algo-  
1212 rithm and outputs track objects. There are no hierarchy structure as it is in pandora,  
1213 each track is independent. There also is no vertex reconstruction with this algorithm  
1214 as well.

### 1215 5.4.5 Cosmic Hit Removal

1216 The Pandora algorithm is applied to the events twice, the first to remove downward  
1217 going tracks primarily from cosmic ray muon like particles. The second pass only runs  
1218 on a subset of hits that aren't associated with cosmic ray muon tracks.

1219 After the first pass, the output of PFParticle hierarchy is then passed to a cosmic  
1220 ray tagger to look through all hits to determine start and end points. If the start or  
1221 end point trajectories are consistent with entering or exiting the TPC, then these hits  
1222 are removed from the second pass. Hits are considered entering or exiting the TPC  
1223 if the drift time are outside of the neutrino drift window or outside of the fiducial  
1224 volume of the TPC. The fiducial volume was based on a montecarlo study and is 20  
1225 cm from the top or bottom of the TPC and 10 cm from the TPC ends. Hits associated  
1226 with candidate cosmic ray tracks are removed from the input hit collection and the  
1227 remaining hits are passed to the neutrino optimized pass of Pandora.

---

**1228 5.4.6 Projection Matching Algorithm**

1229 The projection matching algorithm (PMA) was inherited from ICARUS and has been  
1230 implemented in LArSoft. PMA differs from traditional LArSoft 3D reconstruction  
1231 algorithms. Most 3D reconstruction attempts to match 2D objects from all three planes  
1232 by drift time, while the PMA algorithm projects a track hypothesis on each plane  
1233 then the distance between this projection and the hits on each plane is minimized  
1234 simultaneously. More information can be found in [?].

---

**1235 5.5 Event Selection**

1236 The first requirement for selecting  $\nu_\mu$  CC events is that the event has at least one  
1237 scintillation light flash in the beam trigger window with more than 50 PE on all PMTs  
1238 combined. From the flashes that pass, the most intense is chosen and considered to be  
1239 originating from a neutrino interaction and will be the only flash used in further cuts.

1240 Vertices are then required to have at least one reconstructed track start or endpoint  
1241 within a 5 cm radius. Showers associated with a vertex do not pass this cut. All  
1242 tracks associated with a vertex are then used to calculate a track length weighted  
1243 average of the  $\theta$ -angle. Of all the vertices that do pass, only the vertex with the most  
1244 forward going  $\theta$ -angle average of all associated tracks is considered the neutrino vertex  
1245 candidate. The most forward going  $\theta$ -angle average is chosen by picking the largest  
1246 track range weighted average of  $|\cos(\theta)|$ , seeing as  $\cos(\theta) = 1$  is the beam direction.  
1247 Next, it is required that the reconstructed neutrino vertex candidate be within the  
1248 fiducial volume as well as within the drift time starting at  $t_0$ . The fiducial volume  
1249 boundaries chosen are 10 cm from the edges of the TPC in x and z which is the drift  
1250 direction and beam direction respectively, and 20 cm from the edges of the TPC in y  
1251 which is the vertical direction. For all further cuts, only the longest track associated  
1252 with the neutrino vertex candidate and this track is assumed to be the muon candidate  
1253 of the neutrino event.

1254 The next cut requires the position of the flash in the z-direction and the track z-  
1255 projection to be compared. This basic flash matching algorithm is rudimentary and a  
1256 placeholder for a more sophisticated algorithm. The z-position of the flash needs to be  
1257 within 80 cm to the z-positions of track start or endpoints. If the flash is between the  
1258 track start and endpoint, the distance of the flash to the track is considered to be 0 cm.

1259     Lastly, the track needs to be fully contained within the fiducial volume and have a  
 1260     track range greater than 75 cm. The range is the 3D distance between the track's start  
 1261     and endpoint. The length cut was optimized to remove NC background that contain  
 1262     a pion due to the pion interaction rate to be  $\sim 70$  cm. A track that makes all the cuts  
 1263     is considered to be the muon of a  $\nu_\mu$  CC event. The list of cuts for this selection is  
 1264     described below:

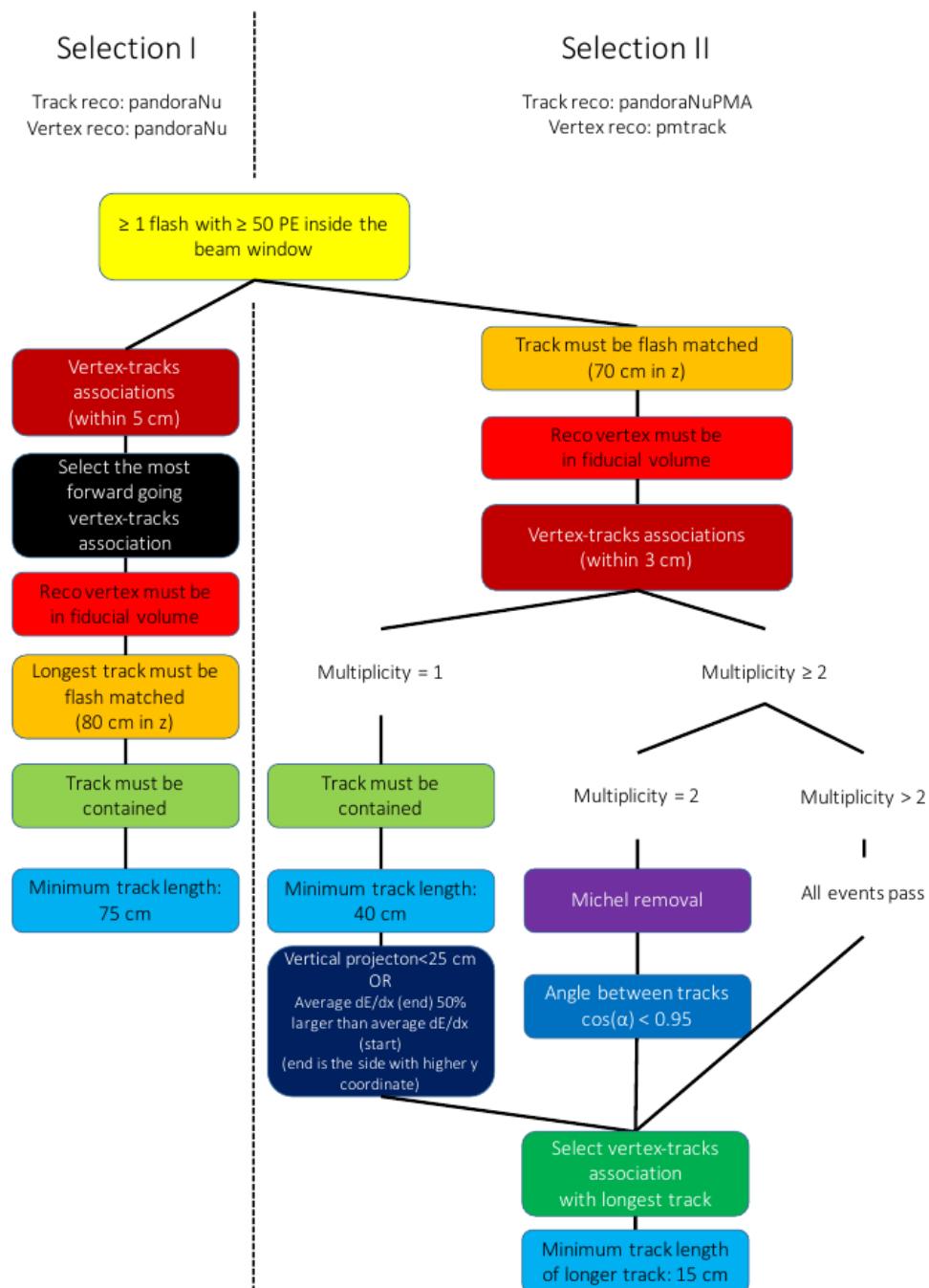
- 1265     1. At least one flash > 50 PE within the beam gate.
- 1266     2. At least one track within 5 cm around a vertex.
- 1267     3. Vertex with flattest tracks is chosen to be vertex candidate.
- 1268     4. Vertex candidate in fiducial volume.
- 1269     5. Longest track associated with vertex candidate is chosen to be track candidate.
- 1270     6. Longest track is within 80 cm (z-axis only) of the flash.
- 1271     7. Longest track is fully contained.
- 1272     8. Longest track is greater than 75 cm.

1273     The event selection scheme can also be seen in figure 5.5. Table 5.1 lists the passing  
 1274     rates for MC events for the selection scheme described above. Table 5.2 lists the passing  
 1275     rates for on-beam and off-beam data for the selection scheme. The normalization  
 1276     factors applied between on-beam and off-beam data are described in section 5.2.

### 1277    5.5.1 Expected Backgrounds

1278    Most of the selected background events will be of cosmic origin. There are two types  
 1279    of cosmic background, one triggered by a cosmic-ray event occurring in the beam  
 1280    gate time window, the other triggered by a beam induced interaction in the cryostat  
 1281    followed by a misidentification of a cosmic event as a neutrino event. The first  
 1282    cosmic background can be subtracted from the selected events using the off-beam  
 1283    BNBEXT sample normalized to the on-beam. The second cosmic background events  
 1284    are modeled by MC by using BNB+Cosmic MC sample.

1285    Other backgrounds originate from neutrino beam contaminants. A major contribu-  
 1286    tion in this sector is by neutral current neutrino events for example a charged pion track  
 1287    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.

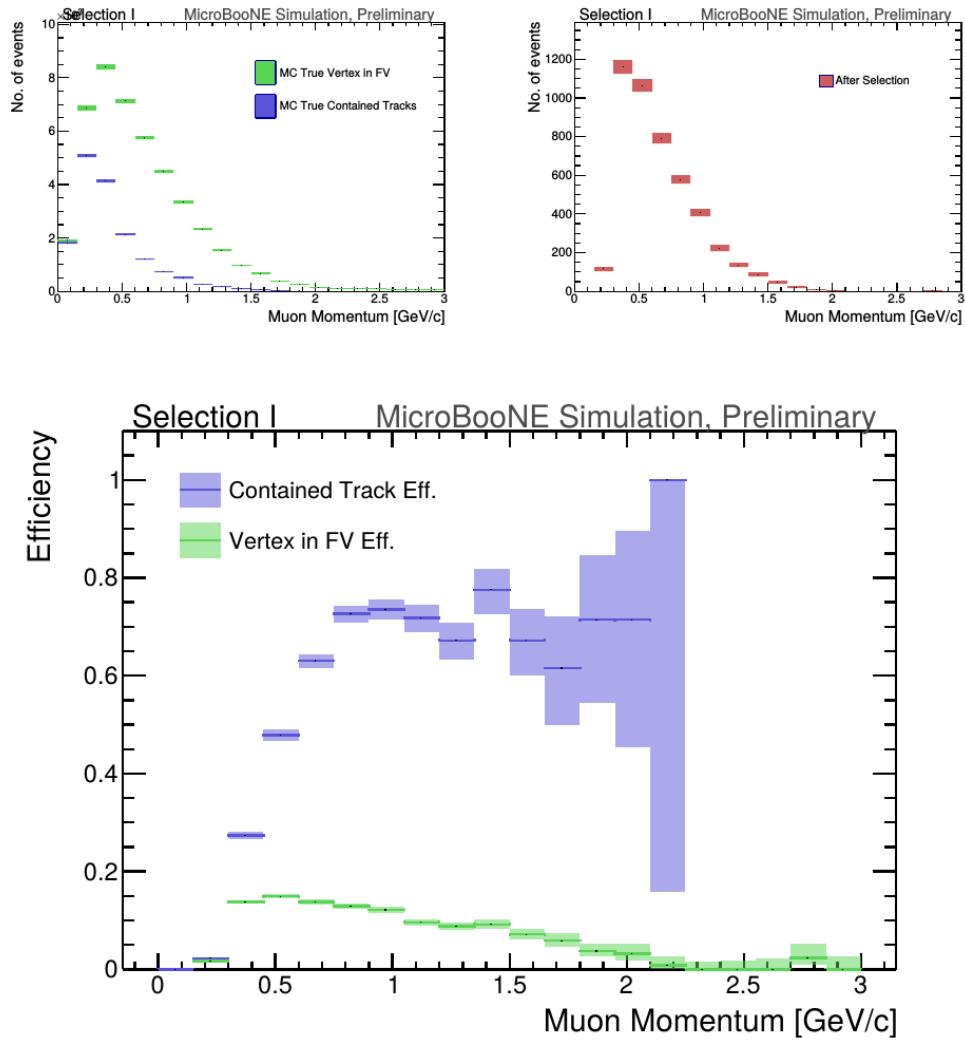
1288 events. These beam related backgrounds are an order of magnitude smaller than the  
1289 cosmic misidentification backgrounds. These backgrounds can not be subtracted and  
1290 are estimated using MC truth.

1291 The efficiency and purity of Selection I are calculated below:

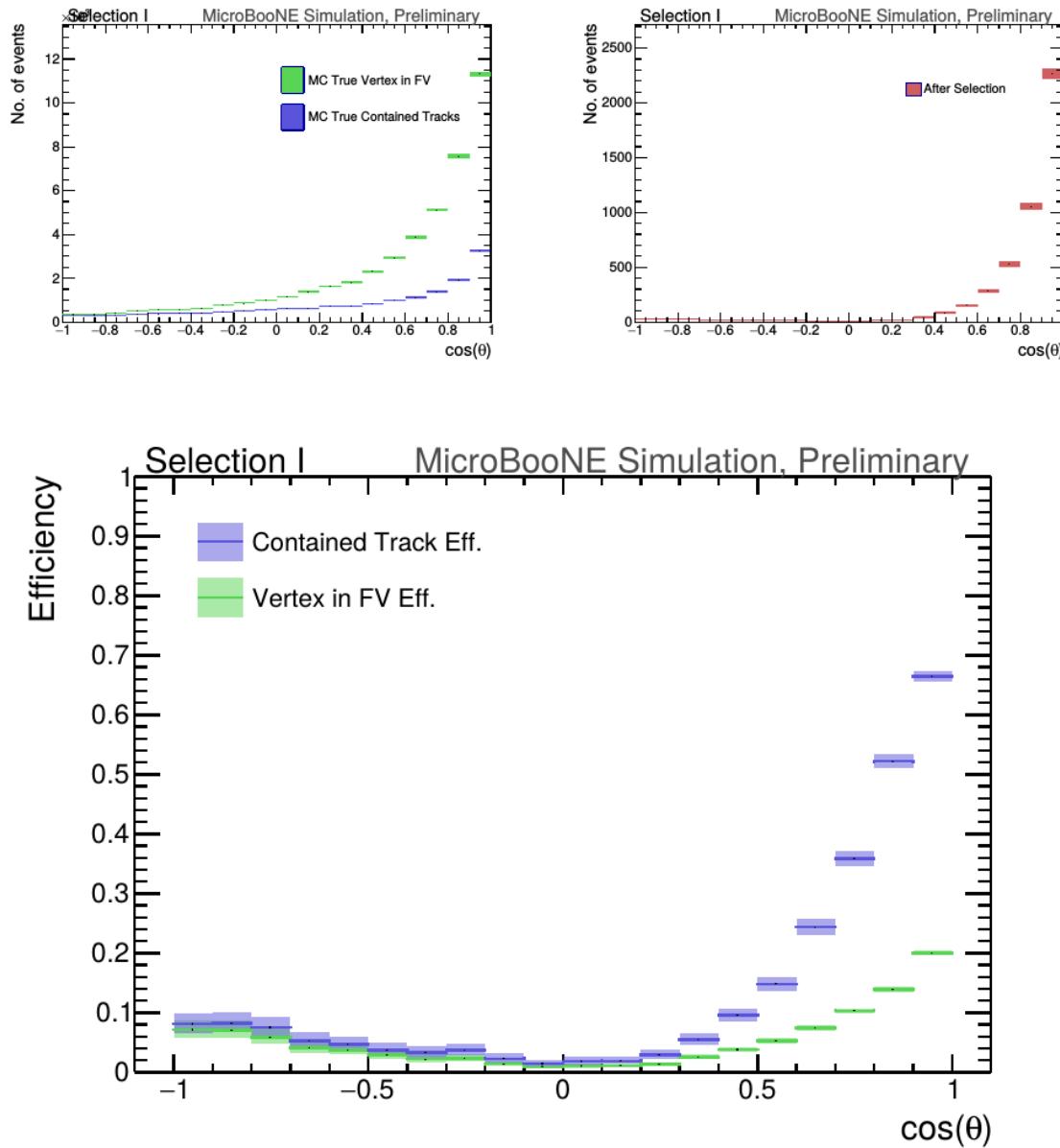
- 1292 • Efficiency: Number of selected true  $\nu_\mu$  CC events divided by the number of  
1293 expected true  $\nu_\mu$  CC events with interaction in the FV.  
1294 –  $(12.3 \pm 3.4) \%$
- 1295 • Purity: Number of selected true  $\nu_\mu$  CC events divided by the sum of itself and  
1296 the number of all backgrounds.  
1297 –  $(53.8 \pm 4.4) \%$

## 1298 5.5.2 Truth Distributions

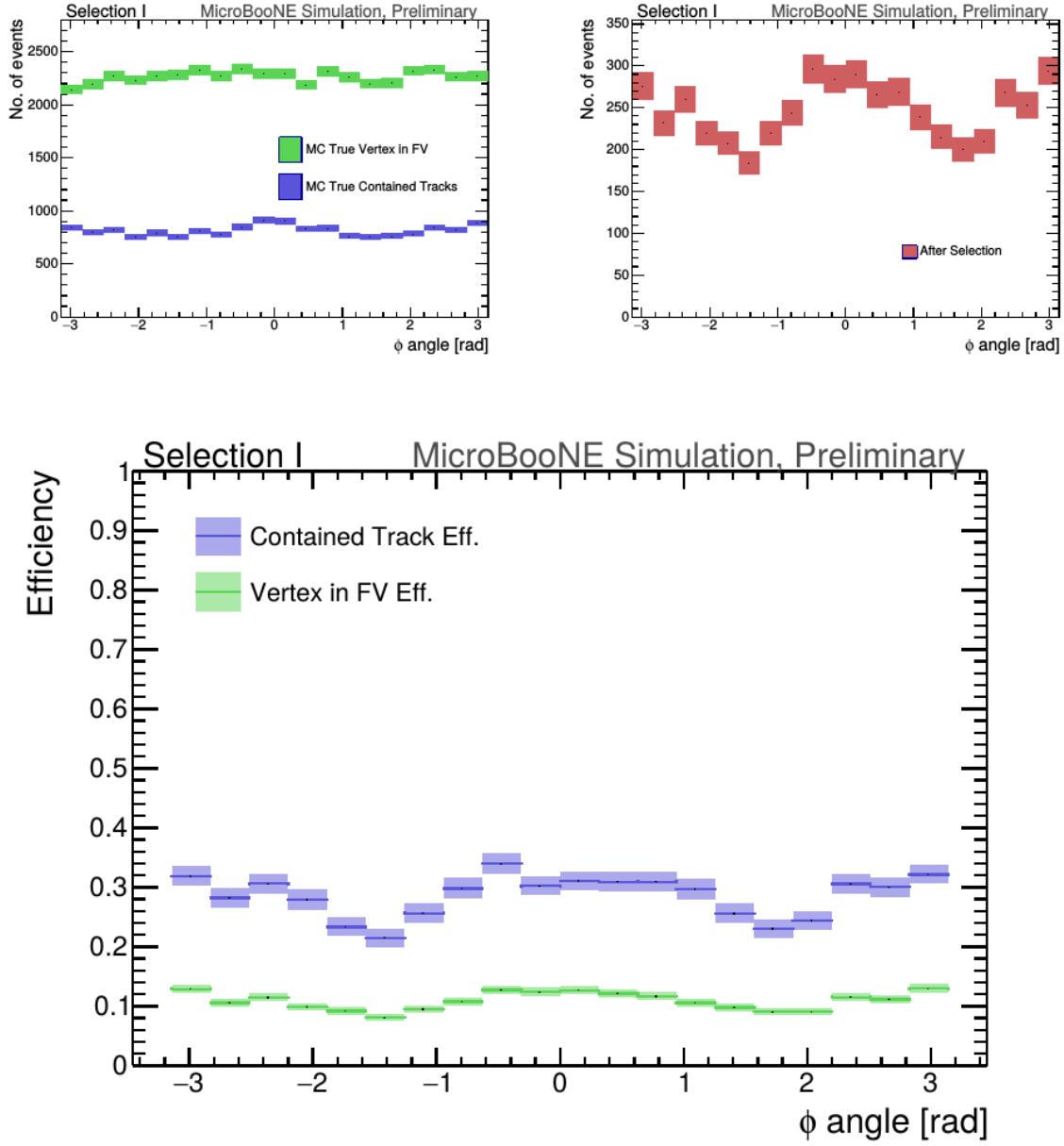
1299 The truth distributions of MC truth variables before and after the selection are detailed  
1300 in this section. The overall efficiencies are calculated for all  $\nu_\mu$  CC signal events  
1301 with a true interaction within the fiducial volume and a fully contained muon track  
1302 originating from said vertex. Figures 5.6 through 5.8 detail the truth distributions for  
1303 muon momentum,  $\cos(\theta)$  and  $\phi$  and figures 5.9 through 5.11 detail the total efficiency  
1304 of the selection for charged current quasi elastic (CCQE) events, charged current  
1305 resonant (CCRES) events, and charged current deep inelastic (CCDIS) events.



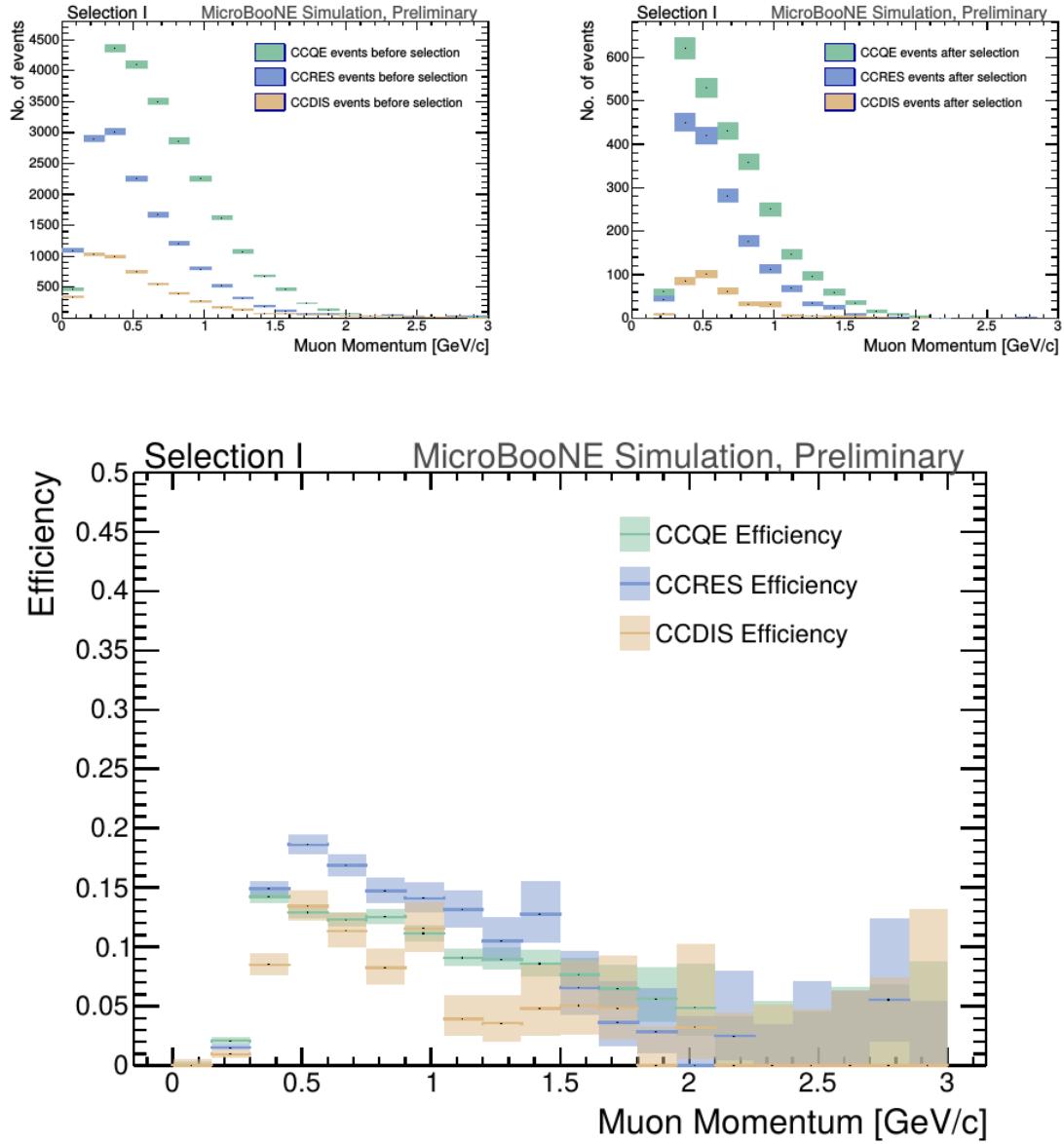
**Figure 5.6:** MC momentum distributions of the muon originating from a  $\nu_\mu$  CC interaction. Upper left is the momentum distributions of events with a vertex within the FV (green) and the events with a fully contained track (blue) before the selection. The upper right side is the momentum distribution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



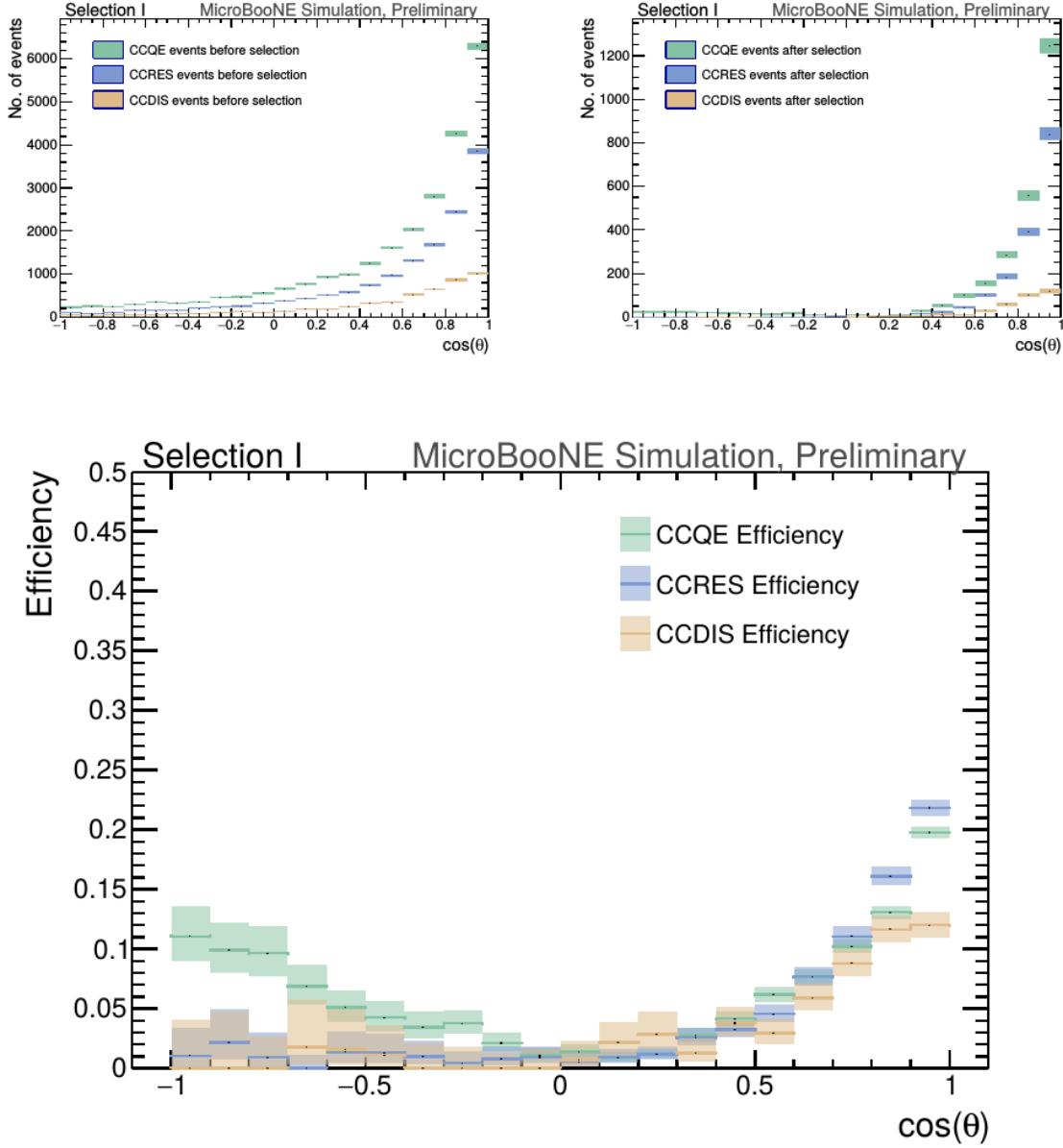
**Figure 5.7:** MC  $\cos(\theta)$  distributions of the muon originating from a  $\nu_\mu$  CC interaction. Upper left is the  $\cos(\theta)$  distributions of events with a vertex within the FV (green) and the events with a fully contained track (blue) before the selection. The upper right side is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



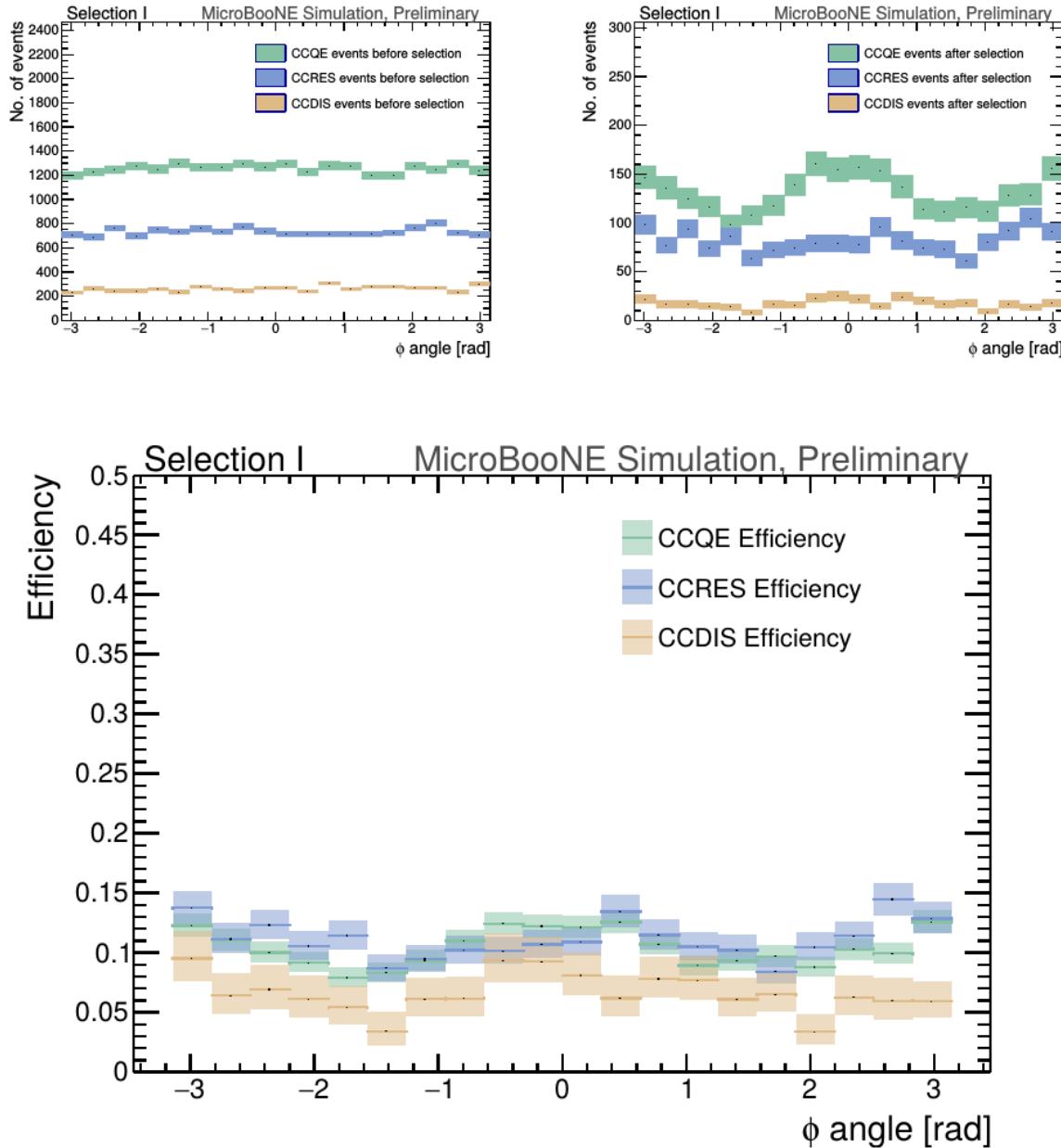
**Figure 5.8:** MC  $\phi$  distributions of the muon originating from a  $\nu_\mu$  CC interaction. Upper left is the  $\phi$  distributions of events with a vertex within the FV (green) and the events with a fully contained track (blue) before the selection. The upper right side is the  $\phi$  distribution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

1306 **Chapter 6**

1307 **Background on Convolutional Neural  
1308 Networks**

1309 Convolutional neural networks (CNNs) have been one of the most influential inno-  
1310 vations in the field of computer vision. Neural networks became popular in 2012  
1311 when Alex Krizhevsky used them to win that year's ImageNet competition [?] by  
1312 dropping the error from 26% to 15%. Since then, many companies are using deep  
1313 learning including Facebook's tagging algorithms, Google for their photo search and  
1314 Amazon for product recommendations. For the purpose of this thesis CNNs were  
1315 used for image classification, specifically, images of varying particles created using  
1316 LArTPC data.

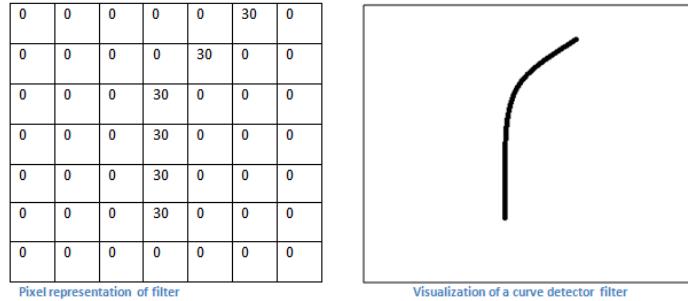
1317 **6.1 Image Classification**

1318 Image classification is the process of inputting an image into the CNN and receiving a  
1319 probability of classes that best describes what is happening in the image. As humans,  
1320 image classification is something that is learned at a very young age and is easy to  
1321 do without much effort. This is also apparent when hand-scanning LArTPC images.  
1322 After learning what a neutrino event looks like in MicroBooNE, it is relatively easy  
1323 to recognize simple neutrino events from cosmic ray background as well as highly  
1324 ionizing particles like protons from minimum ionizing particles like muons. The very  
1325 detailed images LArTPC detectors output are prime candidates for input images into  
1326 a CNN. CNNs mimic a human's ability to classify objects by creating an architecture  
1327 that can learn differences between all the images it's given as well as figure out the

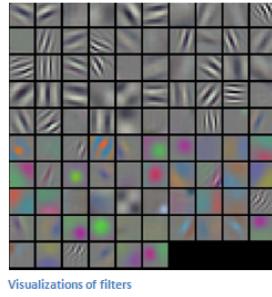
1328 unique features that make up each object. CNNs are modeled after the visual cortex.  
1329 Hubel and Wiesel [?] found that there are small regions of neuronal cells in the brain  
1330 that respond to specific regions of the visual field. They saw that some neurons fired  
1331 when exposed to vertical edges while others fired when shown horizontal or diagonal  
1332 edges. They also saw that these neurons were organized in columns. The idea of  
1333 specific neurons inside of the brain firing to specific characteristics is the basis behind  
1334 CNNs.

## 1335 6.2 CNN Structure

1336 When used for image recognition, convolutional neural networks consist of multiple  
1337 layers that extract different information on small portions of the input image. How  
1338 many layers is tunable to increase the accuracy. The output of these collections are  
1339 then tiled so that they overlap to gain a better representation of the original image  
1340 and allow for translation. The first of these layers is always a convolution layer. To  
1341 the CNN, an image is an array of pixel values. For a RGB color image with width  
1342 and height equal to 32x32 the corresponding array is 32x32x3. Filters, also known as  
1343 neurons, of any size set by the user is then convolved with the receptive field of the  
1344 image. If the filter is 5x5, the receptive field will be a patch of 5x5 on the input image.  
1345 The filter is also an array of numbers called weights. The convolution of the filter and  
1346 image are matrix multiplications of the weights and the pixel values. By stepping the  
1347 receptive field by 1 unit, for an input image of 32x32x3 and a filter of 5x5x3 you'd get  
1348 an output array of 28x28x1. This output array is called an activation map or feature  
1349 map. The use of more filters preserves the spatial dimensions better. The filters can  
1350 be described as feature identifiers. Examples of features in an image consist of edges,  
1351 curves, and changes in colors. The first filters in a CNN will primarily be straight  
1352 line and curve feature identifiers. An example of a curve filter is shown in figure 6.2.  
1353 When a curve in the same concavity is found in the input image, the corresponding  
1354 pixel in the output feature map will be activated. Going back to our example of a  
1355 32x32 input image and a 5x5 filter, if there were to be a curve in the top left corner  
1356 of the input image, our output feature map would have a high pixel value in the top  
1357 left. Therefore, feature maps tell us where a specific feature is located in the original  
1358 image. Figure ?? shows a visualization of filters found in the first layers of many CNN  
1359 architectures. These filters in the first layer convolve around the image and activate  
1360 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

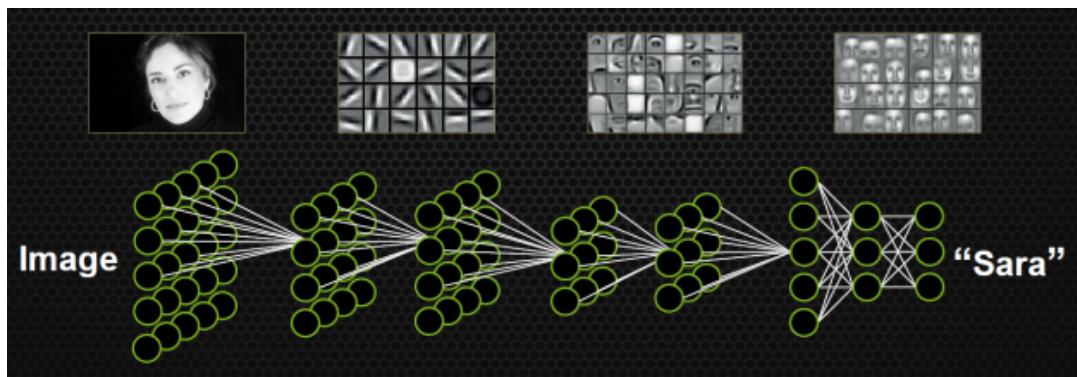


**Figure 6.2:** Visualization of filters found in first layer of a CNN.

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

1377 There are other layers in a CNN architecture that will not be covered in the scope  
 1378 of this thesis but in a general sense, these layers are interspersed between convolution  
 1379 layers to preserve dimensionality and control overfitting of the network. The last layer  
 1380 is called a fully connected layer and it's job is to output an  $N$  dimensional vector where  
 1381  $N$  is the number of classes the network has been trained on. Each number in this vector  
 1382 represents the probability that the input image is a certain class. Fully connected layers  
 1383 use the feature maps of the high level features to compute the products between the  
 1384 weights of the previous layer to get the probabilities of each class. These weights are  
 1385 then adjusted through the training process using backpropagation.

### 1386 6.2.1 Backpropagation

1387 A CNN at it's onset has weights that are randomized. The filters themselves don't  
1388 know how to pull out identifying information per class. For a neural network to learn,  
1389 it must be trained on a training set that is labeled. Backpropagation has four seperate  
1390 steps: foward pass, loss function, backward pass and updating weights. In the forward  
1391 pass, a training image is passed through the whole network. All of our weights at this  
1392 time are randomly initialized so the output for the first image will have no preference  
1393 to a specific class. A common loss function is mean squared error (MSE):

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

1394 If we assume that the MSE is the loss of our CNN, the goal would be that our  
1395 predicted label (output of CNN) is the same as our training label. To do this, we need  
1396 to minimize the loss function. To do this, it is necessary to find out which weights most  
1397 directly affect the loss of the network i.e  $\frac{dL}{dW}$  where L is our loss function and W are  
1398 the weights of a specific layer. The next step is the backward pass which determines  
1399 which weights contribute the most to the loss and finds ways to adjust these weights  
1400 so that the loss decreases. After the derivative is computed, the last step updates the  
1401 weights in the opposite direction of the gradient.

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

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

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

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

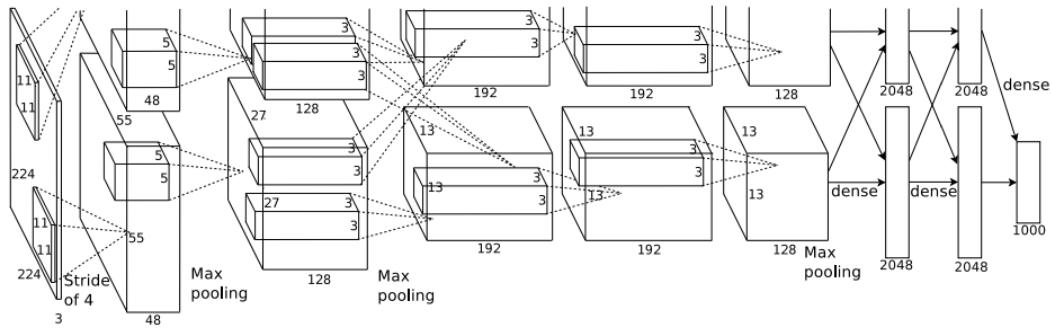
1402 The learning rate is a parameter given to the CNN and it describes the steps the  
1403 network takes to update the weights. Higher learning rate equals large steps and a  
1404 lower training time, but a learning rate that is too large can mean the CNN never  
1405 converges.

1406 Going through backpropagation consists of one training iteration. Once the net-  
1407 work completes a specific number of iterations, another parameter given, and runs  
1408 over all training images that are split up into batches, the process is considered com-  
1409 plete. User input parameters, called hyperparameters, help the network converge to

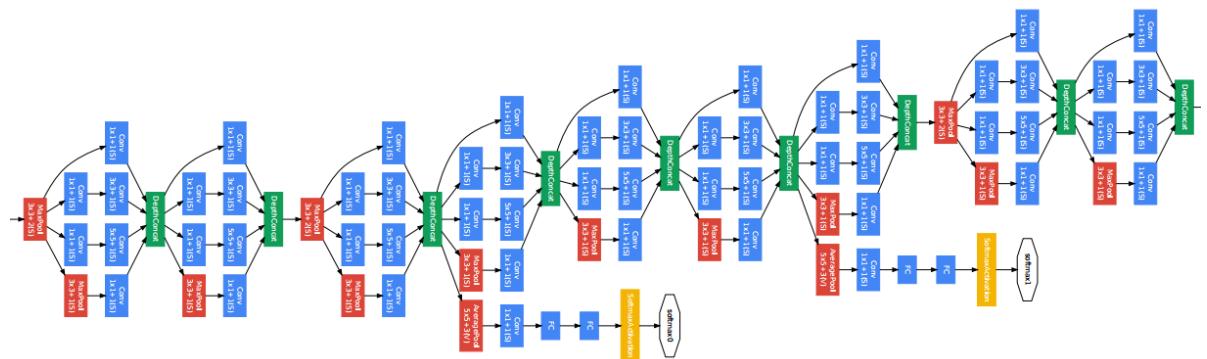
1410 optimal weights for each layer. Batch size, learning rate, and training iteration are just  
1411 some of the user input hyperparameters that help. Lastly, to check if the network has  
1412 learned, a different set of labeled images are fed to the CNN iteratively through the  
1413 training process to see how well it's learning. This process is especially important to  
1414 make sure the network architecture isn't being affected by overfitting (memorizing  
1415 training input rather than learning).

## 1416 6.3 Choosing Hyperparameters

1417 Convolutional neural networks are a relatively new tool in computer vision. Choosing  
1418 hyperparameters for your specific dataset is a non-trivial task. Hyperparameters can  
1419 range from the amount of layers and filters per layer in an CNN architecture to the  
1420 stride the receptive field of a filter takes, not to mention training hyperparameters  
1421 such as learning rate and batch size described above. They're ways to optimize these  
1422 hyperparameters via hyperparameter optimization using Bayesian Optimization [?]  
1423 but as you can imagine, optimizing an CNN architecture from scratch can be very  
1424 computationally intensive. For the purpose of this thesis, two well known CNN  
1425 architectures were used, AlexNet [?], which won the ImageNet Large-Scale Visual  
1426 Recognition Challenge (ILSVRC) in 2012 and therefore bringing awareness of CNNs,  
1427 and GoogleNet [?], which won the ILSVRC in 2014, giving rise to deep networks. Both  
1428 AlexNet and GoogleNet architectures were used to train on LArTPC images and their  
1429 low level filter weights. Higher level filter weights were randomly initialized before  
1430 training so the network can learn high level features of LArTPC image classes. The  
1431 AlexNet architecture is shown in figure 6.5 and the GoogleNet architecture is shown  
1432 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>1433</sup> Chapter 7

## <sup>1434</sup> Training process of Convolutional <sup>1435</sup> Neural Networks

<sup>1436</sup> Three Convolutional Neural Networks (CNNs) were trained throughout this analysis.  
<sup>1437</sup> There are differences to each CNN and will be described fully in the next sections but  
<sup>1438</sup> the main difference are the amount of particle images used for training and validation.  
<sup>1439</sup> CNN1075 used 1,075 muons and 1,075 pions for training and the same amount of  
<sup>1440</sup> each particle for validation. CNN10000 used 10,000 muons and 10,000 pions split in  
<sup>1441</sup> half for testing and training. Lastly CNN100000 had muons, pions, protons, electrons,  
<sup>1442</sup> and gammas in it's training and validation set. Each particle had 20,000 images and  
<sup>1443</sup> training and validation was split 90% training, 10% validation. This chapter will also  
<sup>1444</sup> describe the different hardware frameworks used for training beginning on a CPU  
<sup>1445</sup> and ending on a GPU cluster.

### <sup>1446</sup> 7.1 Hardware Configurations for Convolutional Neural <sup>1447</sup> Network Training

<sup>1448</sup> The first training iteration, CNN1075, was a proof of concept. This CNN was trained  
<sup>1449</sup> on my local machine for  $\sim$  4-5 weeks. The batch size had to be very small as well as the  
<sup>1450</sup> image size due to the lack of computation resources. The second iteration of training,  
<sup>1451</sup> CNN10000, was trained on a Fermilab stationed Syracuse University machine. This  
<sup>1452</sup> machine had 6 TB of disk space, 6 cores at 2.1 GHz and 32 GB of RAM. The use of  
<sup>1453</sup> this machine allowed me to increase the training sample as well as the batch size and  
<sup>1454</sup> hence further increase the accuracy of the neural network. Lastly, the CNN100000 was

<sup>1455</sup> trained using two GTX 1080 Ti GPUs with 11GB of memory on a node on the Syracuse  
<sup>1456</sup> University GPU cluster, SUrge, that has 8 cores and 16GB of memory. This increase in  
<sup>1457</sup> memory as well as the capability to use 2 GPUs drastically cut down on training time  
<sup>1458</sup> from  $\sim$  4-5 weeks to  $\sim$  8 hours. SUrge also allowed for hyperparameter optimization  
<sup>1459</sup> by being able to run multiple training iterations over the two GPUs. Lastly, SUrge  
<sup>1460</sup> allowed for training over higher resolution images and a larger particle class of 5  
<sup>1461</sup> particles vs 2 particles.

<sup>1462</sup> **7.2 Creating images using LArTPC data for**  
<sup>1463</sup> **training/validation of CNNs**

<sup>1464</sup> The  $\mu/\pi$  image dataset used to train and validate CNN1075 was created using single  
<sup>1465</sup> generated isotropic muons and pions from 0-2 GeV energy range. 2,150 muons and  
<sup>1466</sup> 2,150 pions were used for training and testing split 50%. The images were created  
<sup>1467</sup> using LArSoft, a liquid argon software, and were based on wire number and time  
<sup>1468</sup> tick in the collection plane. Uboonecode reconstruction version v05\_08\_00 was used.  
<sup>1469</sup> The raw ADC value after noise filtering was the wire signal. Each collection plane  
<sup>1470</sup> greyscale image was 3456x1600x1 where 6 time ticks were pooled into 1 bin.

<sup>1471</sup> After the image was created, the region of interest (ROI) in the image was found by  
<sup>1472</sup> using Open CV, a image processing open source software package, to scan the image  
<sup>1473</sup> starting from the edges and stopping once a bright pixel is encountered. At this step,  
<sup>1474</sup> the ROI can be larger or smaller than the necessary size of a training image and the XY  
<sup>1475</sup> ratio of the image is not kept. This ROI is then resized to an image of 224x224x1.

<sup>1476</sup> The greyscale color standard is 8bit therefore the ADC value of wire and time tick  
<sup>1477</sup> was also downsampled due to the 12bit ADC value MicroBooNE has. To do this,  
<sup>1478</sup> the highest ADC pixel in the image was found and then this was divided by the rest  
<sup>1479</sup> placing all pixel values between 0-1. From there, all pixel values are then multiplied  
<sup>1480</sup> by 255.

<sup>1481</sup> The  $\mu/\pi$  image dataset used to train and validate the CNN10000 was also created  
<sup>1482</sup> using single generated isotropic muons and pions from 0-2 GeV energy range. 10,000  
<sup>1483</sup> muons and 10,000 pions were used for training and testing split 50%. Uboonecode  
<sup>1484</sup> v06\_23\_00 was used instead of v05\_08\_00. Each collection plane greyscale image was  
<sup>1485</sup> 3456x1280x1 where 5 time ticks were pooled into 1 bin which is different than the

1486 previous dataset and was implemented due to the fact that the time ticks of an event  
1487 went from 9400 to 6400 with the change of uboonecode version. Issues that arose in  
1488 CNN1075 that were fixed in CNN10000 include zero-padding images in X and Y that  
1489 are smaller than 224X224 to eliminate over-zooming effect and fixing a bug that shifted  
1490 pixels separated by a dead-wire region.

1491 The  $\mu/\pi/p/e/\gamma$  image dataset used to train and validate the CNN100000 were  
1492 created using single generated isotropic particles with energy range from 0-2 GeV.  
1493 20,000 of each particle were used for training and were split 90/10 between training  
1494 and testing sets. Uboonecode v06\_23\_00 was used for these images. The collection  
1495 plane greyscale iamge had the same dimensions as CNN10000, 3456x1280x1 and the  
1496 ROI algorithm was the same except for resizing these images to 576x576.

1497 A major change other than the higher resolution images was the treatment of the  
1498 ADC values. In the first two image making schemes, the highest pixel value was found  
1499 per image and the image was then normalized by that. The issue arising from this  
1500 ADC normalization wasn't inherent in  $\mu/\pi$  training due to the fact that both particles  
1501 are minimum ionizing particles in liquid argon, however, when dealing with a larger  
1502 particle class, it was necessary to try and make sure energy deposition by each particle  
1503 was preserved. The energy deposition in a particle image corresponds to the ADC  
1504 value or pixel brightness. To preserve energy deposition, the ADC float value was  
1505 passed straight to the image rather than doing any image normalization. This then  
1506 makes sure that minimum ionizing particles like muons and pions appear dimmer  
1507 than highly ionizing particles like protons.

1508 Images were also made from BNB+Cosmic events that passed the cc-inclusive  
1509 selection 1 filter right before the 75 cm track length cut and were classified using  
1510 the CNN10000. The dataset used to create these images is the same one used in  
1511 [?], *prodgenie\_bnb\_nu\_cosmic\_uboone\_mcc7\_reco2*. These images were created using  
1512 information from the track candidate that passed the filter. Only wire number and  
1513 time ticks associated to the track candidate were drawn on the image to mimic a single  
1514 particle generated image.

1515 These images were then classified using CNN10000. Two approaches were taken  
1516 in making these images. The first was using the image normalization above where  
1517 the maximum pixel in each image is used as a normalization constant to get all pixels  
1518 between 0-1 then multiply all pixels by 255. As described above, this is the incorrect

1519 way to normalize. The second way the images were created was by passing the ADC  
1520 float to the image. The results of CNN10000 performance are shown in section [7.1](#).

1521 Lastly, multiple BNB+Cosmic images per event were made for CNN100000 by  
1522 reducing many of selection I cuts to try and let the CNN do particle as well as event  
1523 identification. This image making scheme used for CNN100000 will be described in  
1524 more detail in later sections.

## 1525 7.3 Convolutional Neural Network Training

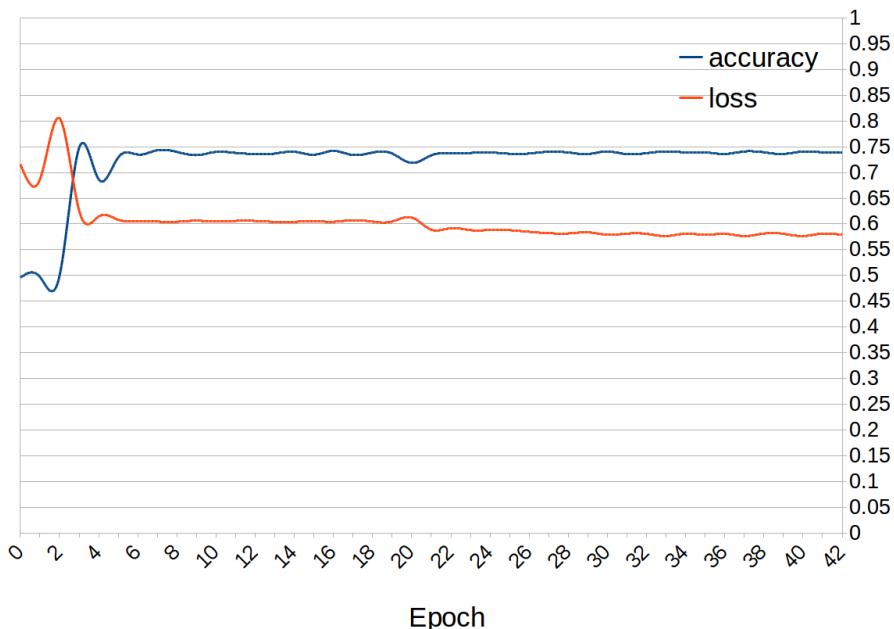
### 1526 7.3.1 Training CNN1075

1527 The results of CNN1075 are described in this section. The accuracy is how well  
1528 CNN1075 is doing by epoch and was 74.5%. The loss is gradient descent or mini-  
1529 mization of the error of the weights and biases used in each neuron of each layer of  
1530 CNN1075 and was 58% with a trend sloping downwards on the loss curve as well as a  
1531 trend sloping upward in the accuracy curve. The accuracy and loss of CNN1075 are  
1532 shown in figure [7.1](#). Due to the depth of the neural network framework, it was neces-  
1533 sary to train with a larger dataset and for more epochs, however, the downward slope  
1534 of the loss curve is an indication that once trained for longer with a higher training  
1535 sample, neural networks can be used for  $\mu/\pi$  separation. The hyperparameters used  
1536 to train CNN1075 are detailed below:

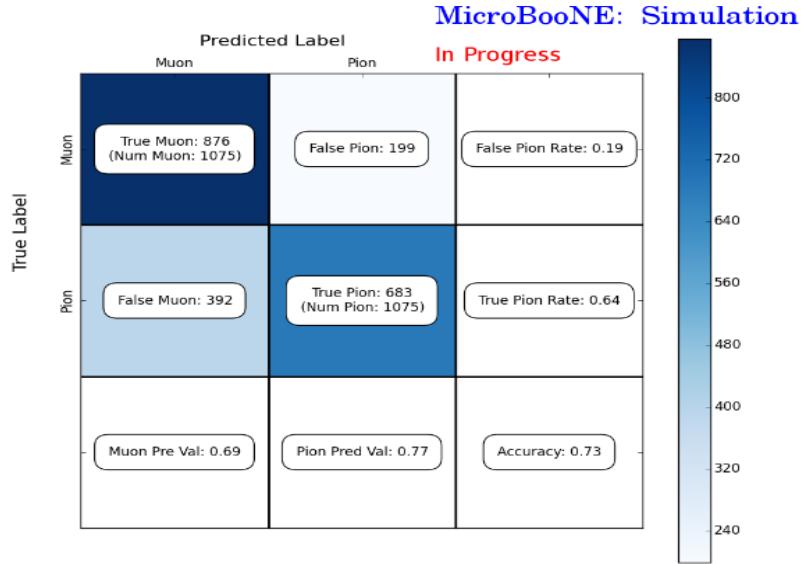
- 1537     • *train\_batch\_size*: 50   1542
- *lr\_policy*: "step"      1547
- *momentum*: 0.9
- 1538     • *test\_batch\_size*: 50   1543
- *gamma*: 0.1              1548
- *weight\_decay*: 0.0005
- 1539     • *test\_iter*: 50          1544
- *stepsize*: 200          1549
- *snapshot*: 100
- 1540     • *test\_interval*: 50      1545
- *display*: 50
- 1541     • *base\_lr*: 0.01        1546
- *max\_iter*: 5000

1550 The confusion matrices shown in figure [7.2](#) show the accuracy for both the training  
1551 and testing datasets. The fact that these two have similar accuracies is important  
1552 because if the training dataset had a much higher accuracy, that indicates an over-  
1553 training of the training sample which means the neural network didn't learn features  
1554 to separate muons from pions, it just memorized what was in the training dataset.

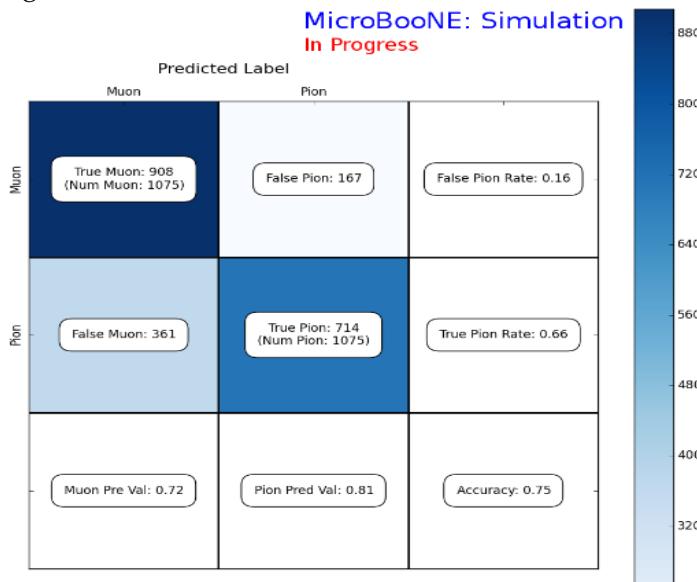
Loss/Accuracy of CNN trained on 2150 images



**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 =  $true\mu / (true\mu + false\mu)$

1555 Also note that the neural network does a better job of identifying muons than pions.  
1556 This can be attributed to the more complex event scenes pions tend to leave in the  
1557 detector due to pion interacting more in LAr than muons do. The CNN may do better  
1558 at identifying pions with a larger training sample.

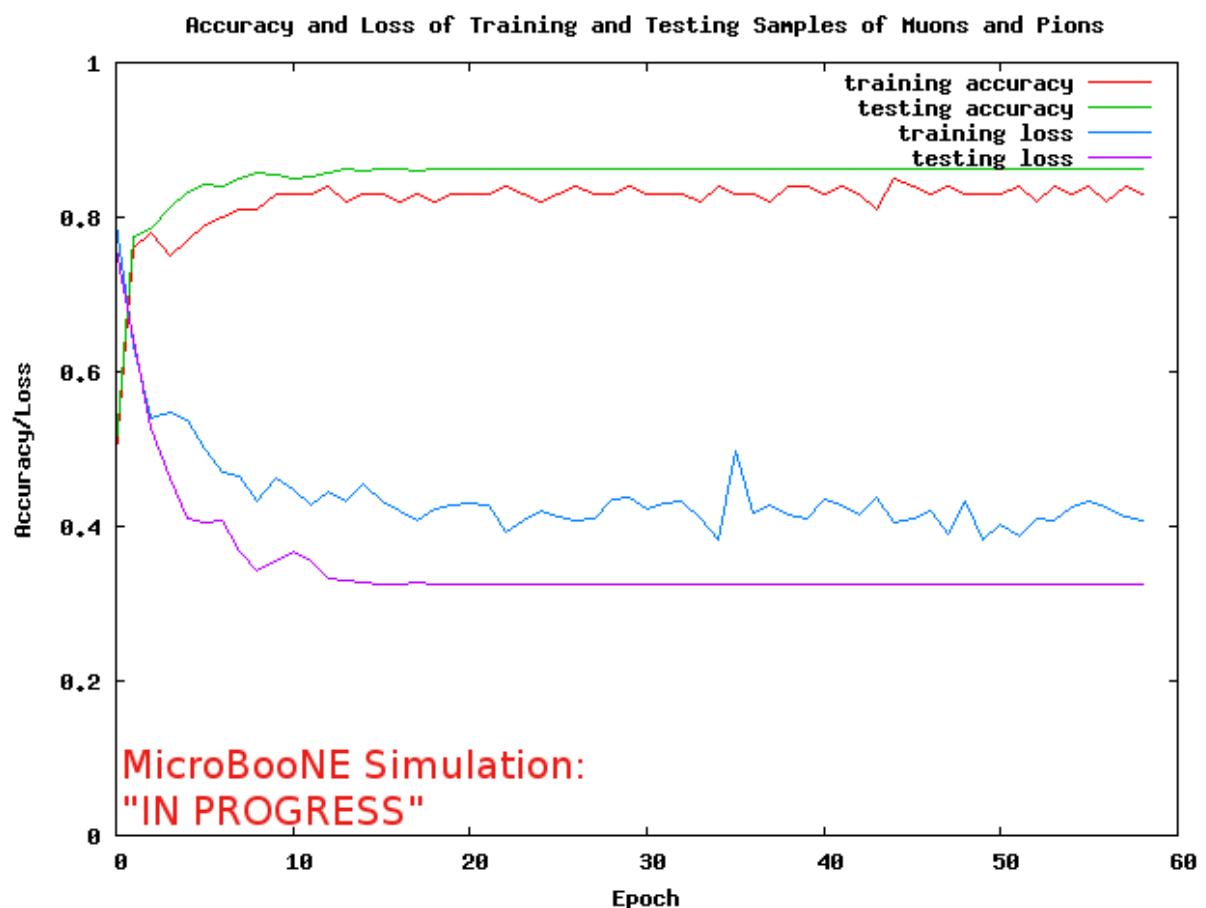
1559 **7.3.2 Training CNN10000**

1560 The hyperparameters used for CNN10000 are shown below. The batch size for the  
1561 training and testing as well as the test\_iter were chosen to encompass the whole  
1562 training/testing image set when doing accuracy/loss calculations. To do this, multi-  
1563 plying the test\_iter by the test batch size gives you the amount of images used when  
1564 calculating accuracy/loss curves.

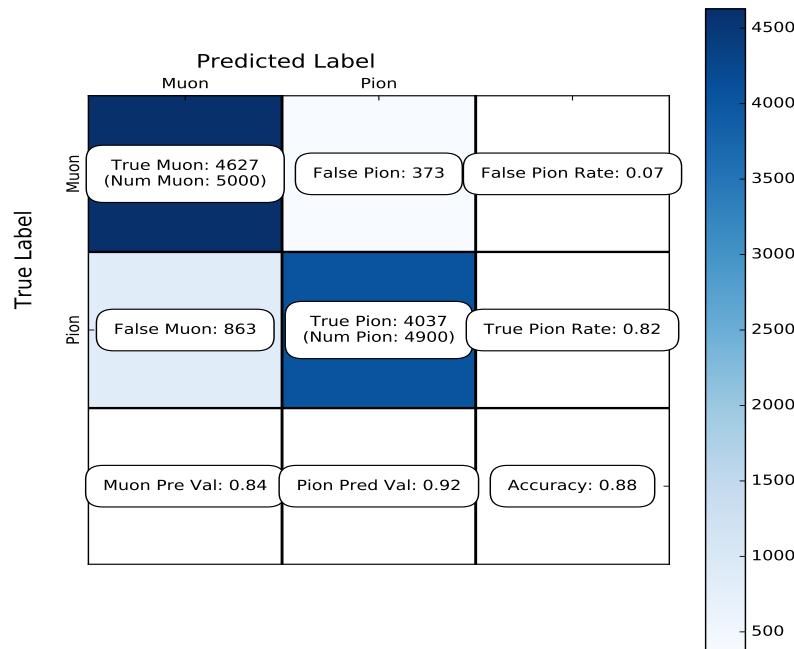
- 1565     • *train\_batch\_size*: 100<sub>1570</sub>     • *lr\_policy*: "step"<sub>1575</sub>     • *momentum*: 0.99
- 1566     • *test\_batch\_size*: 100<sub>1571</sub>     • *gamma*: 0.1<sub>1576</sub>     • *weight\_decay*: 0.0005
- 1567     • *test\_iter*: 100<sub>1572</sub>     • *stepsize*: 1000<sub>1577</sub>     • *snapshot*: 100
- 1568     • *test\_interval*: 100<sub>1573</sub>     • *display*: 100
- 1569     • *base\_lr*: 0.001<sub>1574</sub>     • *max\_iter*: 10000

1578     The same architecture that was used to train CNN1075 was employed on CNN10000,  
1579 AlexNet. Caffe [?] was the software package used for both CNNs. The differences  
1580 include batch size and test\_iter and momentum to account for the larger dataset. Fig-  
1581 ure 7.3 shows the loss and accuracy of CNN10000. There is around a 10% increase in  
1582 accuracy from CNN1075 to CNN10000, 85%, and around a 20% decrease in loss, 36%.

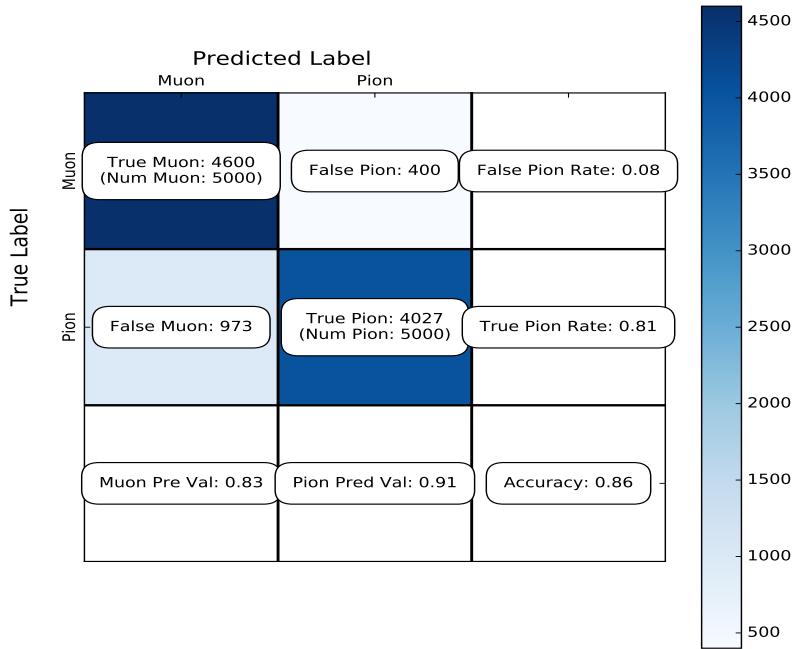
1583     Figure 7.4 show a breakdown of  $\mu/\pi$  separation for CNN10000. It also shows  
1584 the network is not being over-trained due to the Accuracy of both the training and  
1585 testing datasets being within .01% of each-other. Figure 7.5 shows how well the neural  
1586 network is doing at  $\mu/\pi$  separation with respect to muon probability. The red bins  
1587 corresponds to true pions and the blue bins correspond to true muons. There is  
1588 still pion contamination in the high muon probability bins but by choosing a muon  
1589 probability of  $\geq 80\%$  we can reduce this. The CNNs increase in total accuracy can be  
1590 attributed to an increase in accurately classifying pions as pions as seen in both the  
1591 confusion matrix in figure 7.4 and the large number of events in the zero bin of the  
1592 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.

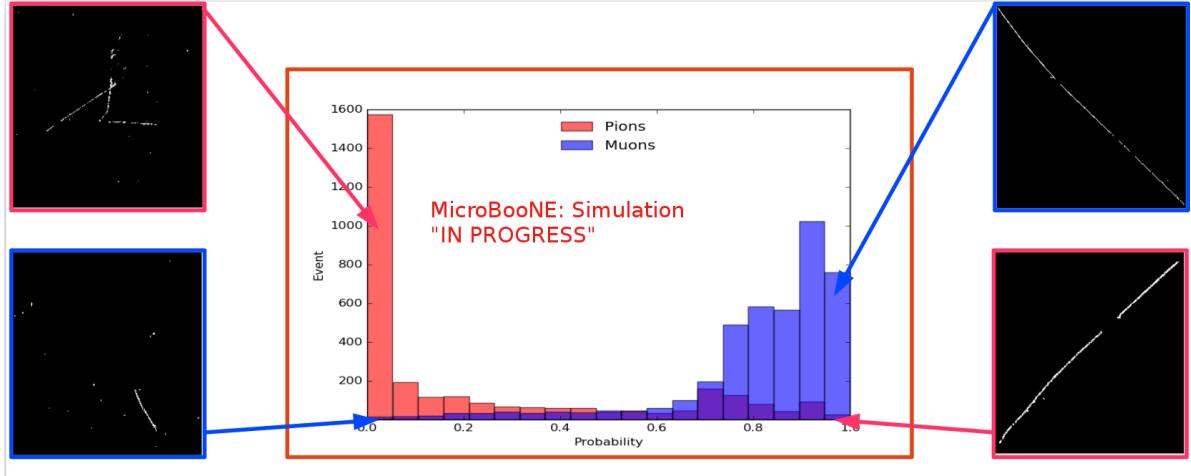


**(a)** Confusion Matrix showing Accuracy of CNN10000 using training MC data



**(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.

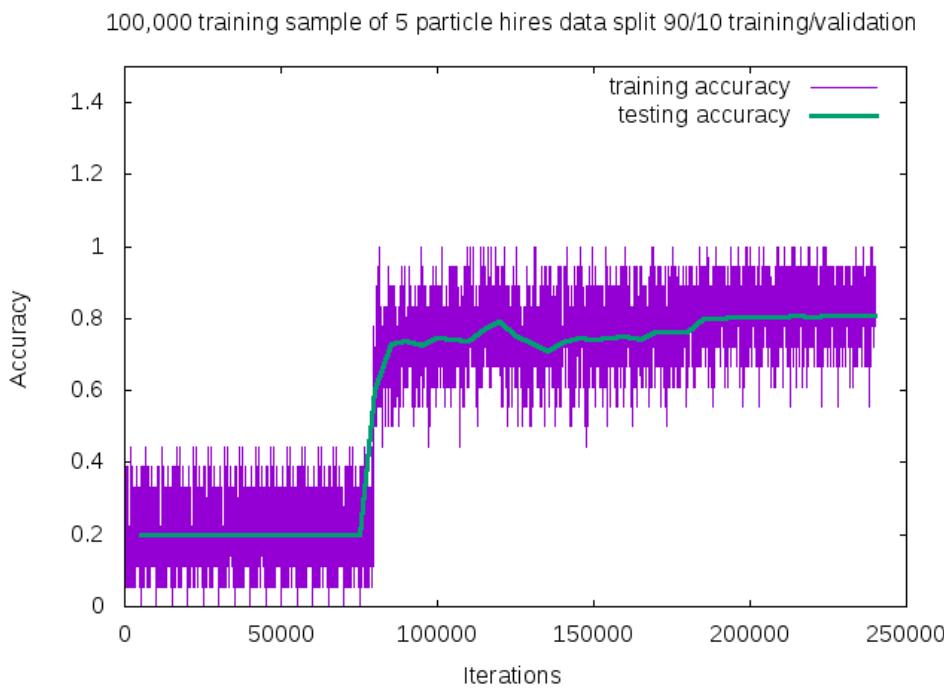
### 1593 7.3.3 Training CNN100000

1594 CNN100000 used the GoogleNet architecture rather than the AlexNet architecture  
 1595 used in the two previous trained CNNs. This is the first time the neural network was  
 1596 trained on a larger particle class,  $\mu/\pi/p/\gamma/e$ , and on higher resolution images. This  
 1597 CNN also employed GPUs during the training process. The hyperparameters are  
 1598 shown below:

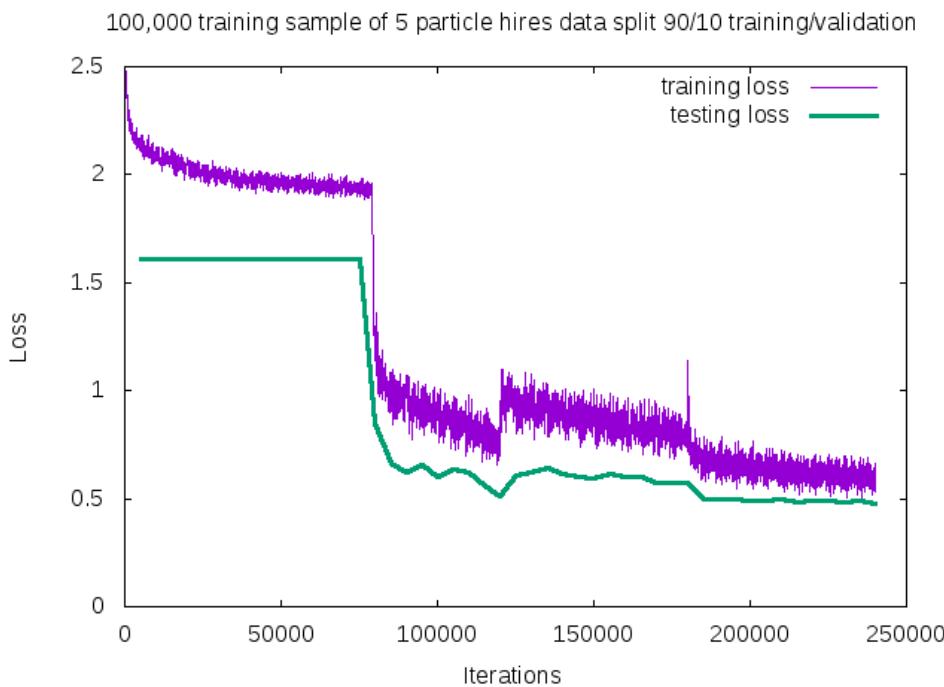
- 1599     • *train\_batch\_size*: 18   1604     • *lr\_policy*: "step"                 1609     • *max\_iter*: 10000
- 1600     • *test\_batch\_size*: 2   1605     • *gamma*: 0.96                         1610     • *momentum*: 0.99
- 1601     • *test\_iter*: 2000   1606     • *stepsize*: 10000                         1611     • *weight\_decay*: 0.0002
- 1602     • *test\_interval*: 2000   1607     • *average\_loss*: 40                         1612     • *snapshot*: 50000
- 1603     • *base\_lr*: 0.001   1608     • *display*: 40

1613     The accuracy and loss for CNN100000 are shown in figures 7.6 and 7.7. The jumps  
 1614     shown in both figures are when the training was stopped to fine-tune the weight decay  
 1615     and the learning rate. The accuracy leveled off at  $\sim 80\%$  and the loss was at  $\sim 0.48$ .

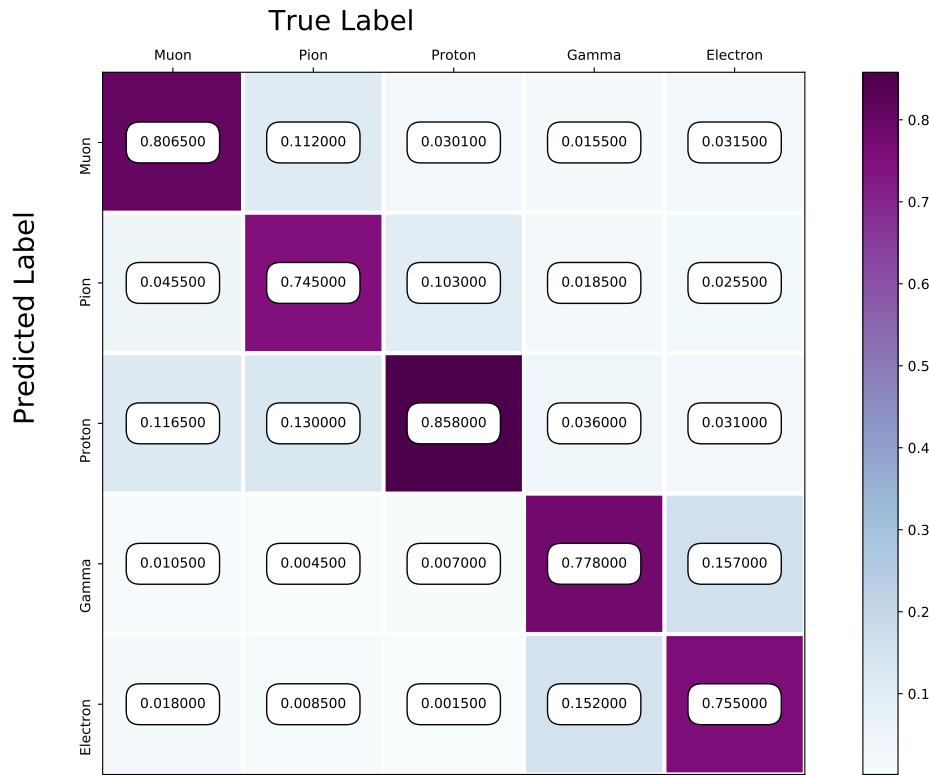
1616     Figure 7.8 shows the confusion matrix of CNN100000. The proton identification of  
 1617     the neural network is at 85% and the highest out of all five particles. One thing to note  
 1618     is clear separation between particles that leave track like objects in the MicroBooNE  
 1619     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

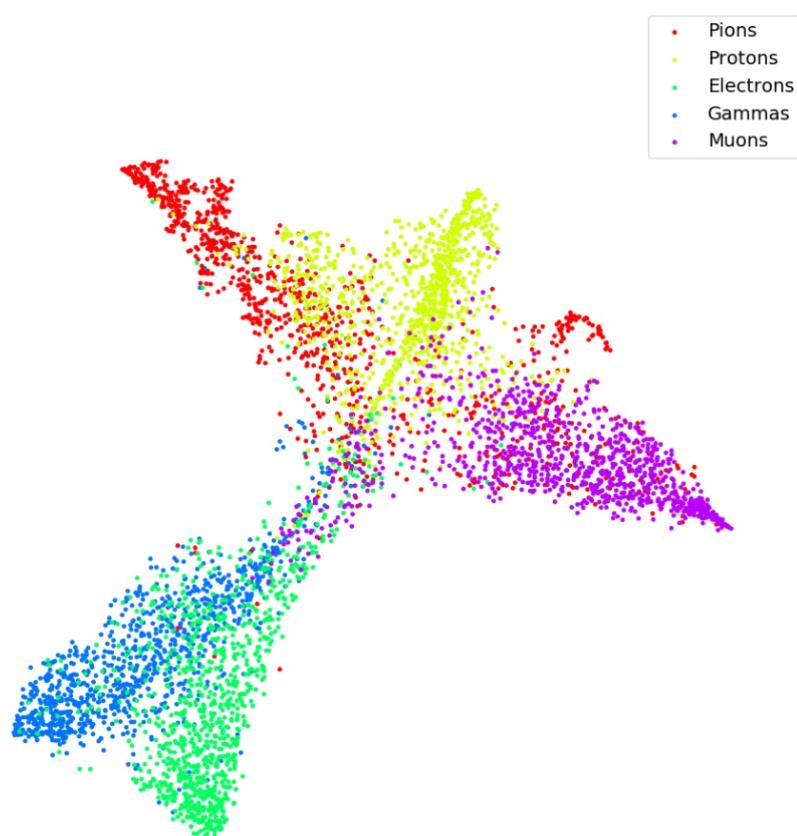
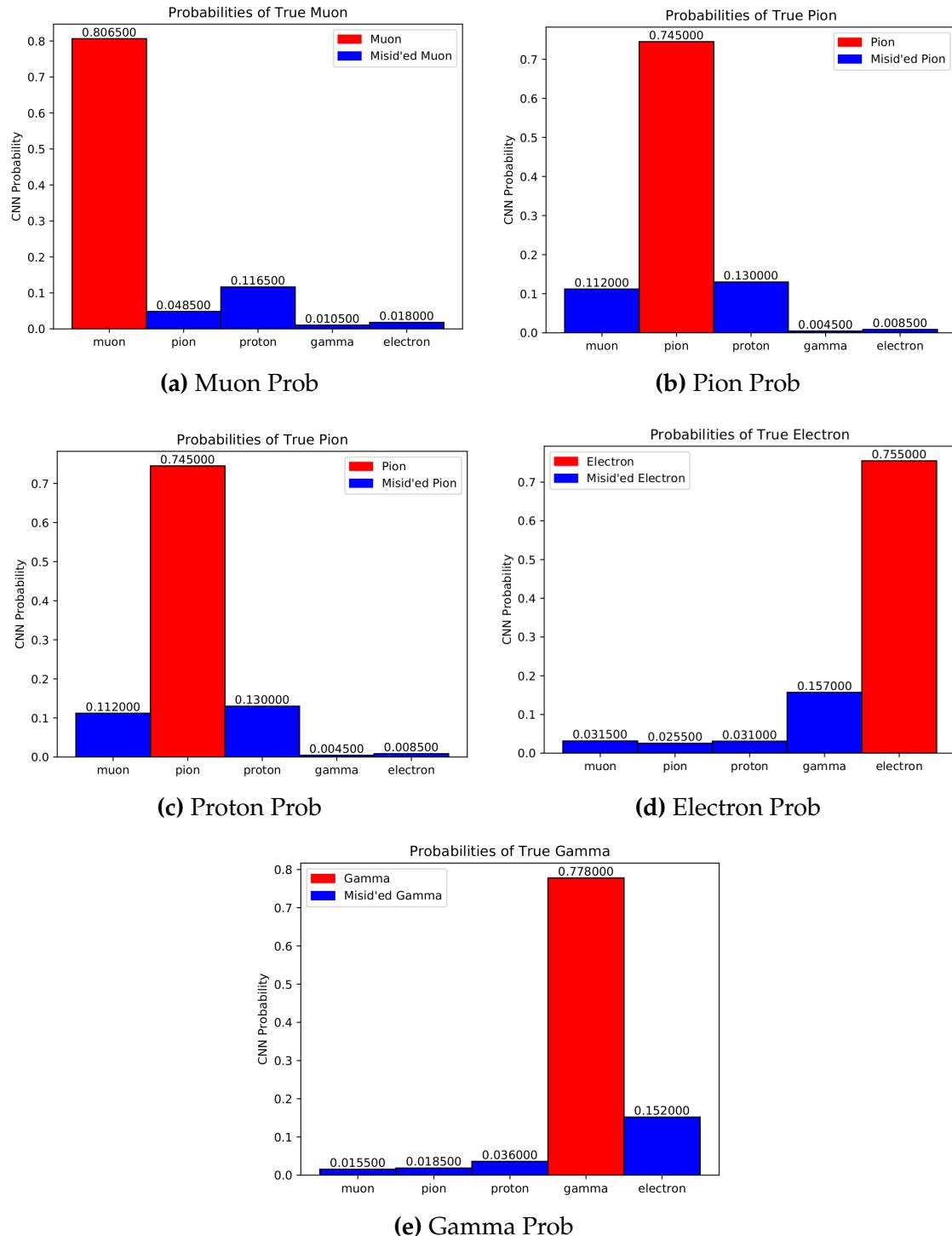
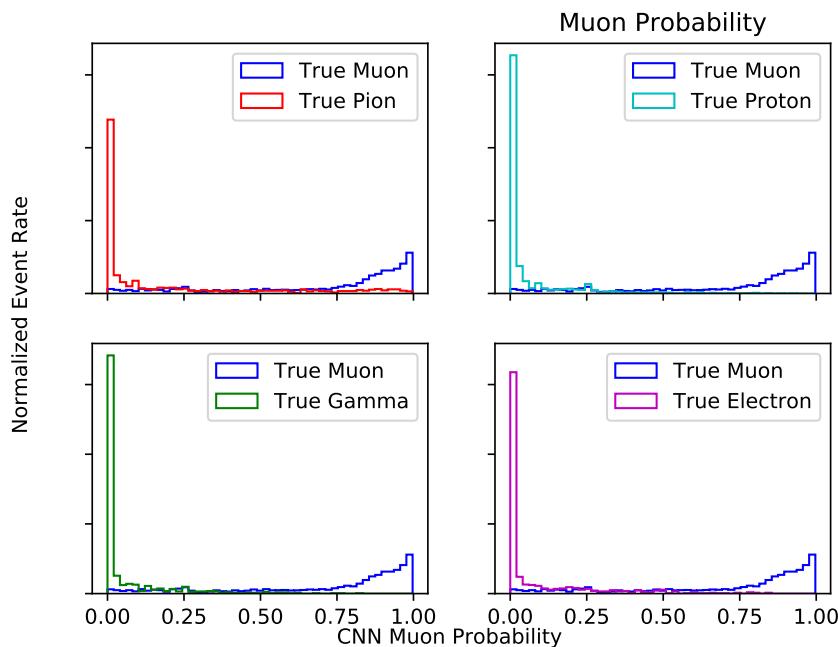


Figure 7.9: t-SNE of CNN



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

<sup>1636</sup> protons. For pions, both protons and muons get misidentified as pions at around the  
<sup>1637</sup> same probability. Similar behavior is also seen for proton identification. Electrons and  
<sup>1638</sup> gammas are misidentified as each-other with similar probabilities.

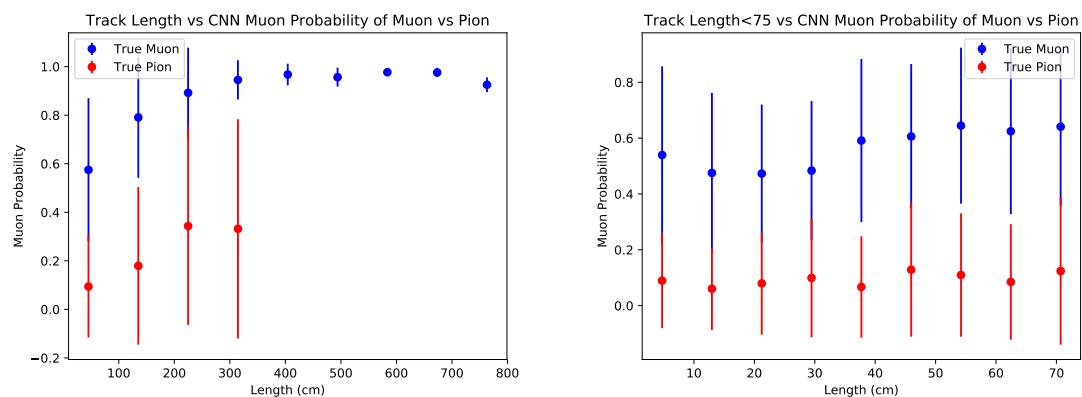


**Figure 7.11:** Muon probability of true muons (blue) versus pions (red), protons (cyan), gammas (green) and electrons (magenta).

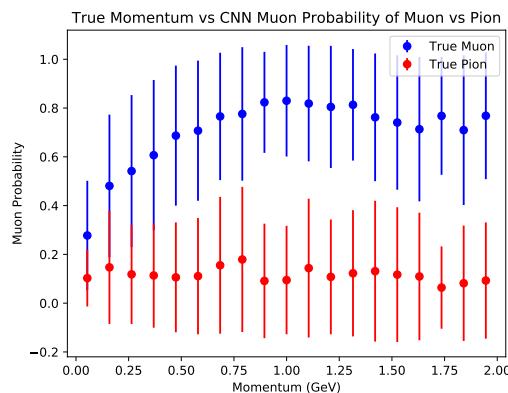
<sup>1639</sup> To see what type of background contamination one would be dealing with when  
<sup>1640</sup> doing muon identification, muon probabilities for each particle class was plotted  
<sup>1641</sup> against the probability of true muons to see how well muon signal vs other particle  
<sup>1642</sup> background separation can be done with CNN100000. Figure 7.11 is showing the  
<sup>1643</sup> true muon probability for true muons, versus the rest of the particle classes. This plot  
<sup>1644</sup> describes which muon probability value should be chosen for the least amount of  
<sup>1645</sup> other particle contamination. For electrons and gammas, a muon probability of  $\sim 75\%$   
<sup>1646</sup> would eliminate  $e/\gamma$  contamination. For pions and protons, there is contamination at  
<sup>1647</sup> all values of muon probability, but the contamination is drastically reduced at a muon  
<sup>1648</sup> probability  $\geq 75\%$ .

<sup>1649</sup> One of the main concerns with training a neural network was that the features the  
<sup>1650</sup> network would learn to separate muons from pions would be track range, which is  
<sup>1651</sup> what was used to begin with in selection I. To make sure that wasn't the case, the next  
<sup>1652</sup> 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.

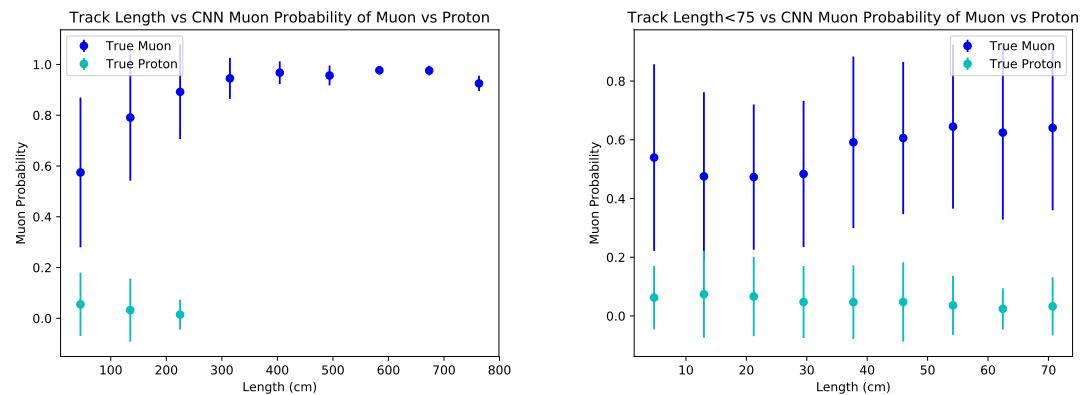


(a) Track range versus muon probability for true muons (blue) and true pions (red). (b) Track range  $\leq 75$  cm versus muon probability for true muons (blue) and true pions (red).

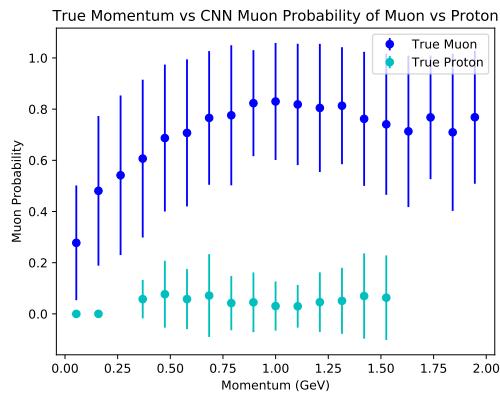


(c) Momentum versus muon probability for true muons (blue) and true pions (red).

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

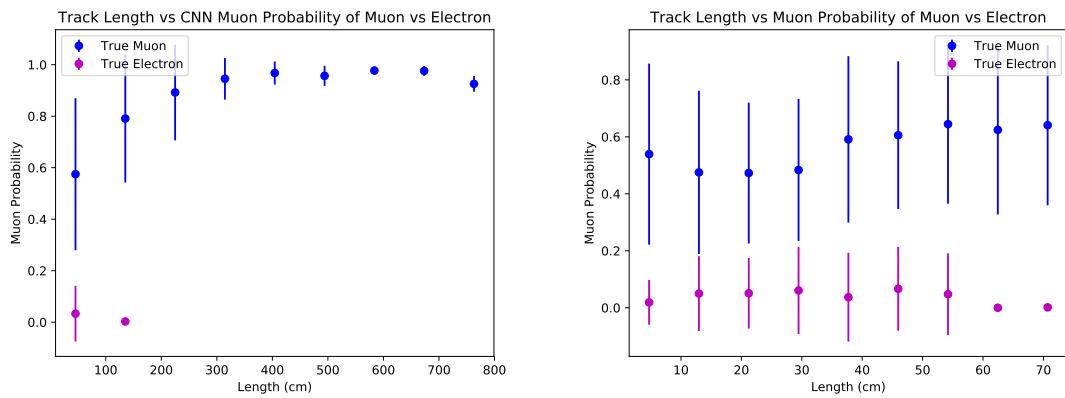


(a) Track range versus muon probability for true muons (blue) and true protons (cyan). (b) Track range  $\leq 75$  cm versus muon probability for true muons (blue) and true protons (cyan).

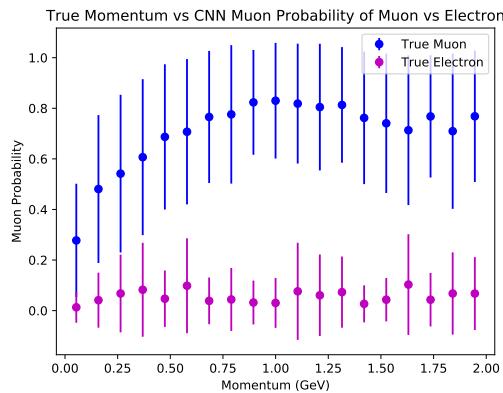


(c) Momentum versus muon probability for true muons (blue) and true protons (cyan).

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

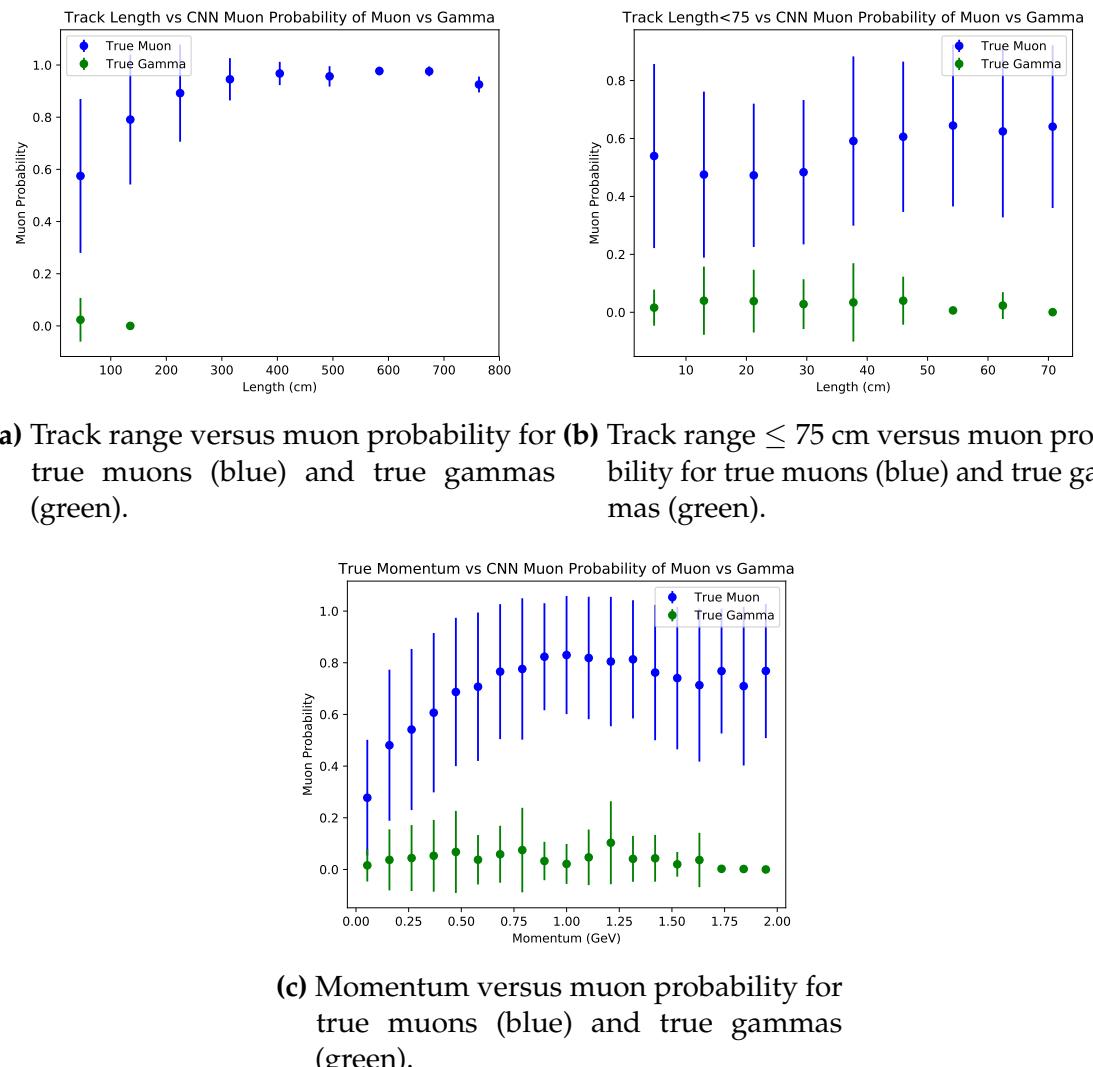


- (a) Track range versus muon probability for true muons (blue) and true electrons (magenta).
- (b) Track range  $\leq 75$  cm versus muon probability for true muons (blue) and true electrons (magenta).



- (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.

1660 **Chapter 8**

1661 **Using Convolutional Neural Networks  
1662 for  $\nu_\mu$  CC event classification**

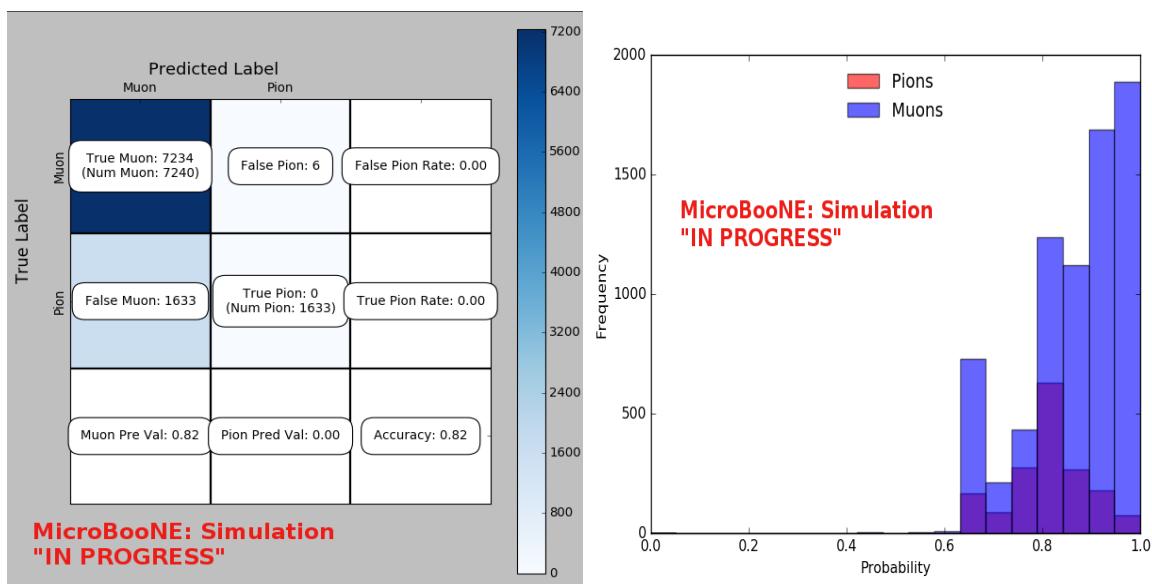
1663 **8.1 Classification using CNN10000**

1664 **8.1.1 Classification of MC data using Selection I CC-Inclusive Filter**

1665 After training CNN10000, it was then used to classify track candidate images that  
1666 were identified by the selection I cc-inclusive filter described in chapter 5. Passing  
1667 rates for each cut in this filter are shown in table 5.1. Out of 188,880 events, 19,112  
1668 passed the cut right before the 75 cm track length cut which is a 10.1% passing rate and  
1669 comparable to the 10% passing rate shown in table 5.1. In time cosmics were also run  
1670 over, out of 14,606 in time cosmics events, 302 passed the cut right before the 75 cm  
1671 track length cut which is a 2.1% passing rate comparable to the 2.7% passing rate in the  
1672 cc-inclusive tech-note. Figures 8.1a and 8.1b show the accuracy and  $\mu/\pi$  separation.  
1673 Both plots are only composed of muons and pions due to the focus on  $\mu/\pi$  separation  
1674 and the fact that CNN10000 was only trained on muons and pions, however, for  
1675 reference, all other particles that did pass selection I were mis-id'ed as muons. Muons  
1676 are being identified at a very high rate, while pions are all being mis-id'ed as muons.  
1677 This is due in part because the pion track candidate that does pass the cc-inclusive  
1678 filter right before the 75 cm track length cut has already been identified as a muon  
1679 candidate, hence, at a higher muon probability. Another reason for the pion mis'id  
1680 can be attributed to the training/classifying dataset difference. For training, the pion  
1681 images include the whole pion interaction in argon, including any decays or nucleon  
1682 scattering. The image created from a BNB+Cosmic event used for classification only

1683 includes the track candidate that passed the cc-inclusive filter. Figure 8.2a shows the  
 1684 track range distributions of all events from selection I being classified by the CNN as a  
 1685 muon with a probability of 70% regardless of true particle type. We get entries for the  
 1686 CNN curve in the lowest bin and none for the 75 cm curve. To see how many true CC  
 1687 events were identified by CNN10000 breaking down figure 8.2a by event type was  
 1688 necessary. Figures 8.2b and 8.2c show track range distributions separated by signal  
 1689 and various backgrounds. Particle type was not taken into consideration in these plots  
 1690 so true CC event images can be any track candidate particle passing selection I cut  
 1691 right before track length cut including pions and protons.

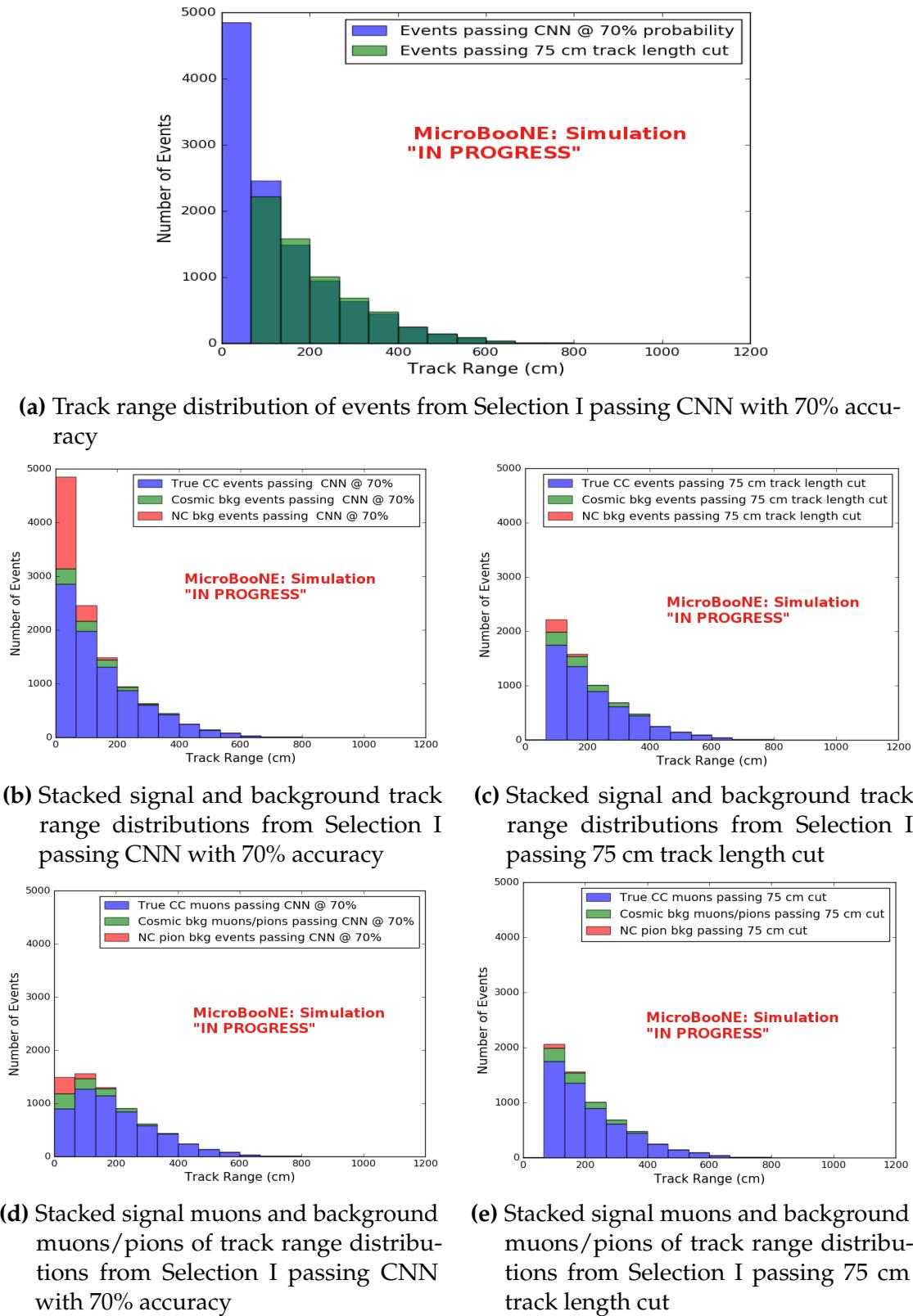
1692 To gain an even deeper understanding on how CNN10000 is performing, plotting  
 1693 these distributions with only muons and pions was done due to the fact that CNN10000  
 1694 was trained with only those particles for  $\mu/\pi$  separation. Figures 8.2d-8.3d show the  
 1695 stacked histograms of signal and background of the track range distributions with  
 1696 varying CNN probabilities starting from 70% and ending at 90% probability. With  
 1697 higher probabilities we get a purer sample in the lower bin but we end up losing  
 1698 events as well. Momentum distributions for all signal/background events are shown  
 1699 in figure 8.4.



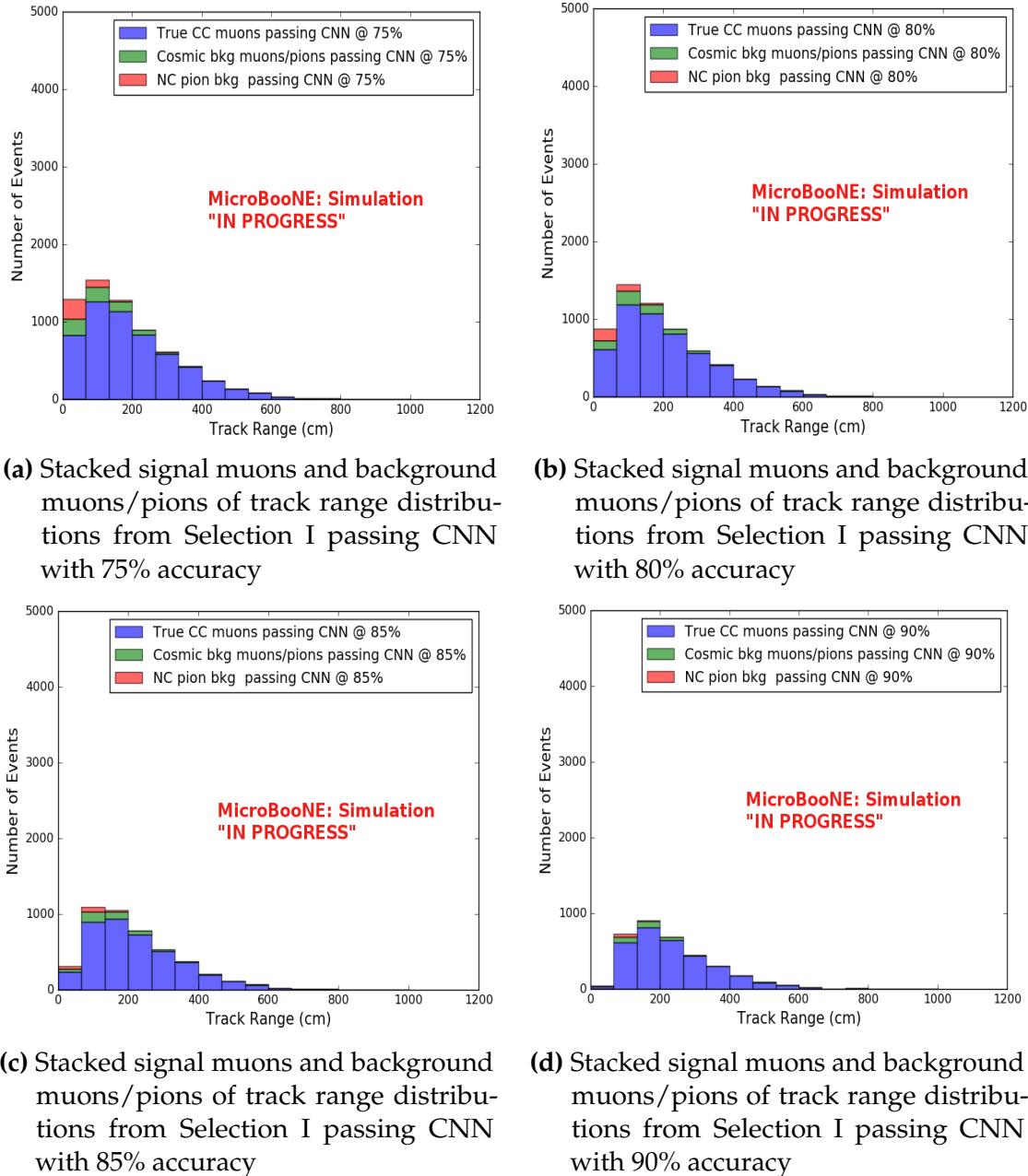
(a) Confusion Matrix for CNN10000 classified events from selection I

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

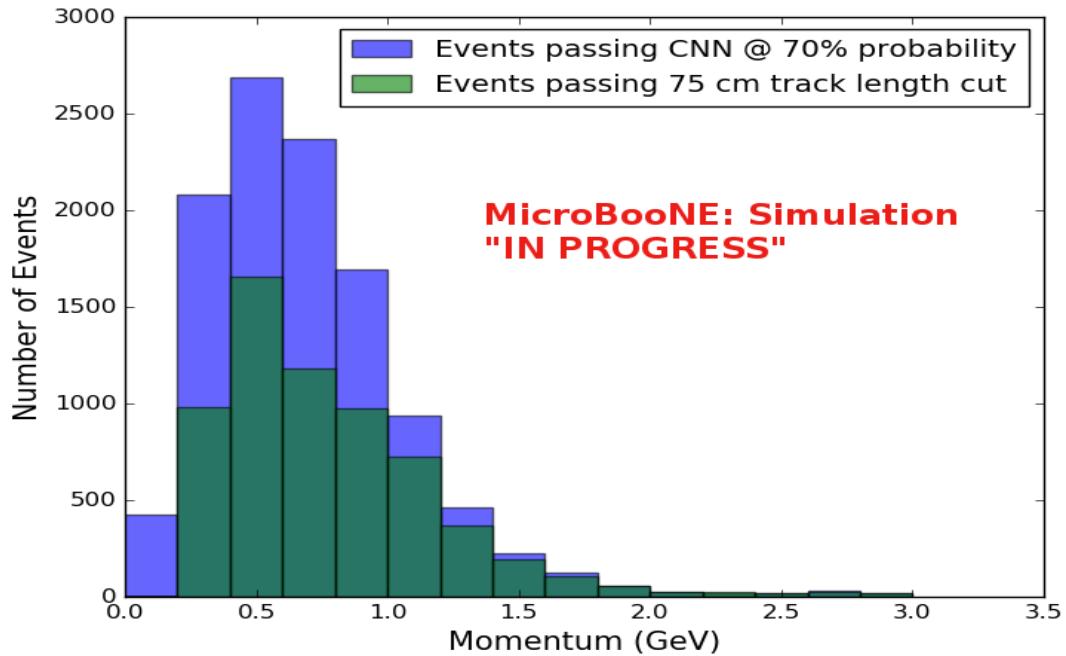
**Figure 8.1:** Confusion matrix and probability plot of events passing selection I cc-inclusive cuts right before 75cm track length cut



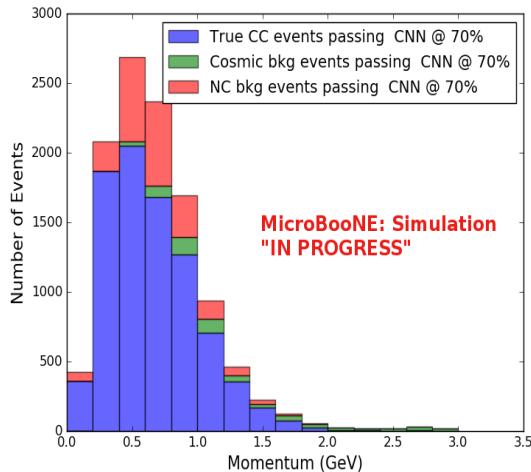
**Figure 8.2:** CNN10000 distributions of track candidate images output from Selection I cc-inclusive filter



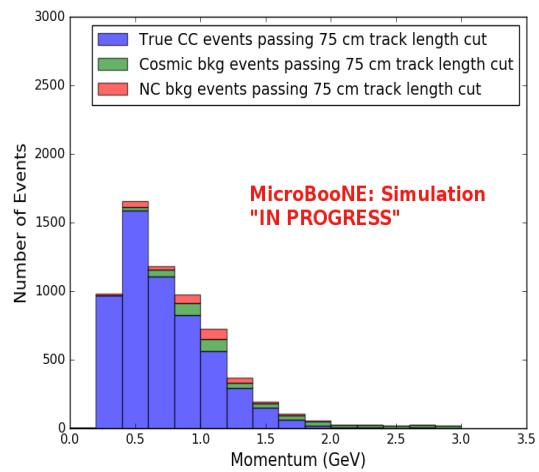
**Figure 8.3:** CNN10000 stacked signal/background track range distributions of track candidate images output from Selection I cc-inclusive filter



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



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

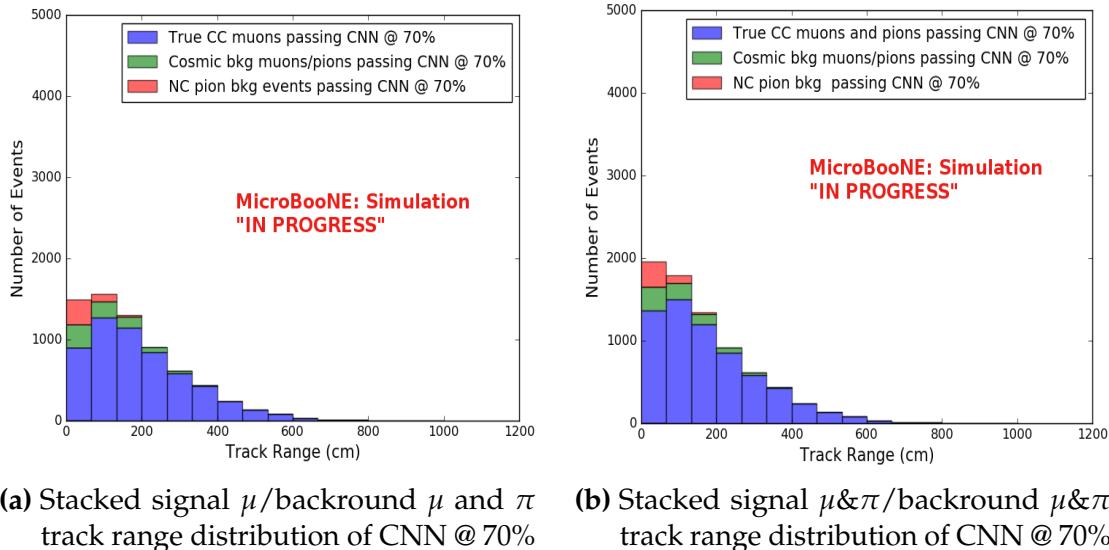


(c) Stacked signal and background momentum distributions from Selection I passing 75 cm track length cut

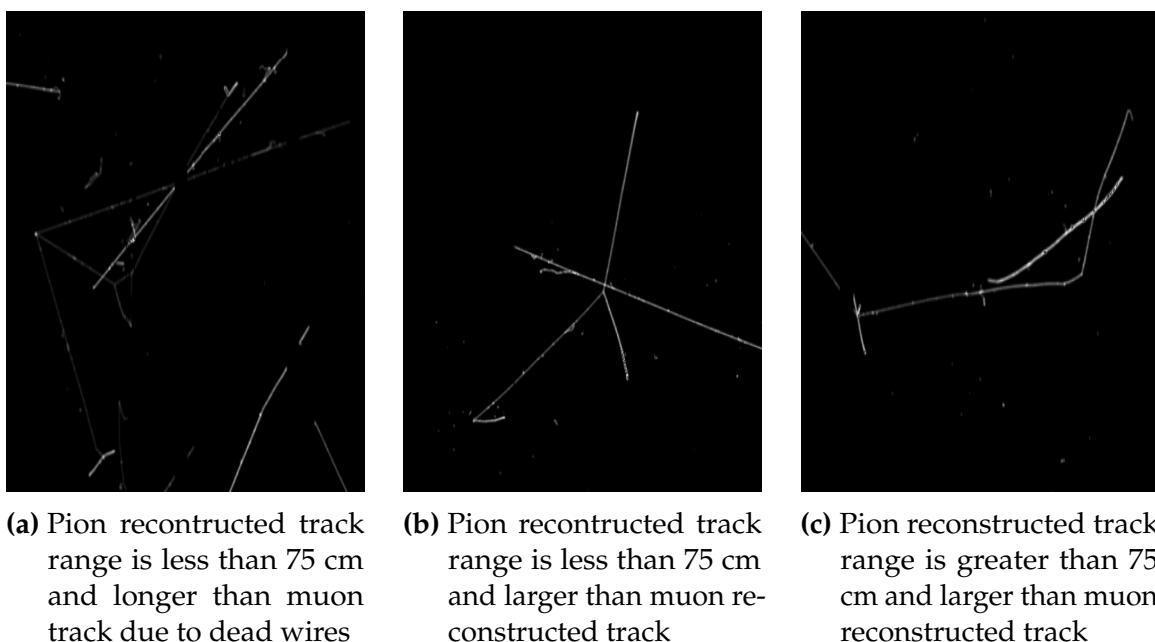
**Figure 8.4:** CNN10000 momentum distributions of track candidate images output from Selection I cc-inclusive filter

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

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 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



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



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

		BNB + Cosmics		Cosmic Only	Signal: Cosmic Only
		Selection	MC Truth		
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
		1168(53.8%)	6142(61%)
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.

<sup>1736</sup> of signal and backgrounds for the CNN at the different probabilities. We have a 61%  
<sup>1737</sup> signal passing rate with the CNN cut @ 83% versus the 53.8% signal passing rate of  
<sup>1738</sup> the 75 cm cut.

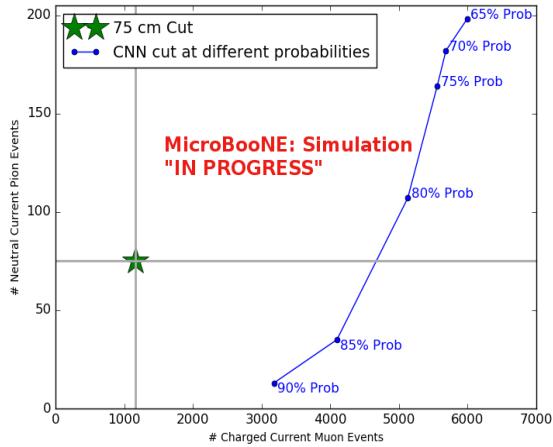
<sup>1739</sup> Based on these numbers, the following performance values of the selection with 75  
<sup>1740</sup> cm cut versus selection with CNN @ 83% probability cut were calculated:

- <sup>1741</sup> • Efficiency: Number of selected true  $\nu_\mu$  CC events divided by the number of  
<sup>1742</sup> expected true  $\nu_\mu$  CC events with interaction in the FV.
  - <sup>1743</sup> – Selection I: 12.3%
  - <sup>1744</sup> – Selection I with CNN10000 cut @ 83% probability: 14%
- <sup>1745</sup> • Purity: Number of selected true  $\nu_\mu$  CC events divided by sum of itself and the  
<sup>1746</sup> number of all backgrounds.
  - <sup>1747</sup> – Selection I: 53.8%
  - <sup>1748</sup> – Selection I with CNN10000 cut @ 83% probability: 61%

<sup>1749</sup> Lastly, figure 8.7 shows a more representative performance of the CNN. Due to the  
<sup>1750</sup> fact that the CNN was trained on muons and pions, showing the performance of CC  
<sup>1751</sup> muon events versus NC pion events with respect to CNN probability gives a better  
<sup>1752</sup> picture of how the network is performing. Figure 8.7 shows that at 83% we are below  
<sup>1753</sup> the 75 cm cut NC pion threshold and still above the CC muon threshold. Using 83%  
<sup>1754</sup> probability not only reduced the NC pion background, it also dramatically reduced  
<sup>1755</sup> the in time cosmics and cosmics in the BNB.

### <sup>1756</sup> 8.1.2 Conclusions of CNN10000 classification of MC data

<sup>1757</sup> It was shown that even though CNN10000 was trained with single particle generated  
<sup>1758</sup> muons and pions, it performs fairly well at classifying track candidate images from  
<sup>1759</sup> BNB+Cosmic events. Events have been regained below the 75 cm track length cut and  
<sup>1760</sup> the momentum and track range distributions have similar shapes to the distributions  
<sup>1761</sup> of Selection I. Efficiencies and purities were calculated for selection I events before 75  
<sup>1762</sup> cm track length cut with the CNN at 83% probability and are 14% and 62% respectively.  
<sup>1763</sup> Although the CNN doesn't have separation between muons and pions and although  
<sup>1764</sup> all particles passing CNN are classified as muon, increasing CNN probability allows  
<sup>1765</sup> us to increase the purity as well as maintain an efficiency comparable to the 75 cm track



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

length cut all while recovering events below that 75 cm cut. Out of the 6142 events that passed the CNN @ 83% 1470 events were below the 75 cm cut, a recovery of 3.3% of data with an purity of 15%. Although these numbers are low, it is an improvement from the selection I in both total efficiency and purity and an increase in phase space by recovering these events.

## 8.2 Classification using CNN100000

For classification of BNB+Cosmics and data using CNN100000, images were made from track candidates that passed the Selection I filter, however, unlike for classifying BNB+Cosmics using CNN10000, the classification of CNN100000 went further up Selection I's cut chain. For CNN100000, steps 5 through 8 seen in section 5.5 were removed. The image making algorithm would then create multiple images per event of pixels corresponding to each track associated to the flattest vertex candidate in the fiducial volume. One of the findings of CNN10000 was the possibility of recovering interesting events in which a pion from a cc-inclusive event is tagged as the track candidate of interest. This was the reason for trying to expand on what a convolutional neural network could accomplish. By allowing the CNN to particle ID all track associated with the vertex candidate, we allow the selection to contain the interesting events that were cut out in selection I due to the cc pion track being chosen as the track candidate. Figure 8.8 shows the image making algorithm for BNB+Cosmic images. The results of using CNN100000 to classify BNB+Cosmics will be outlined below.

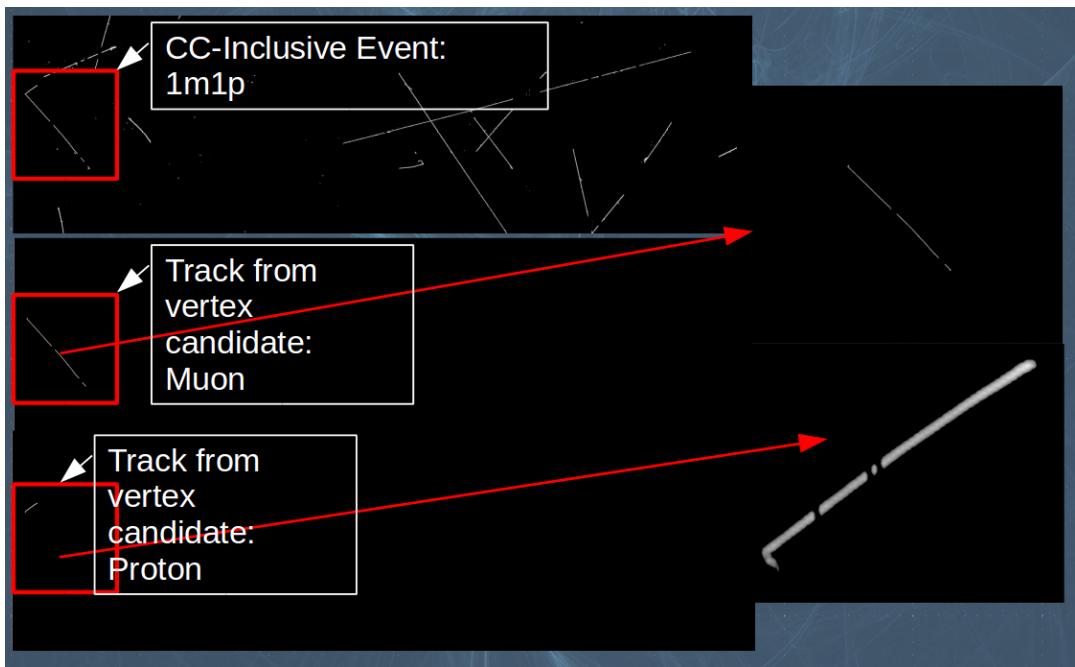


Figure 8.8: Image making steps used for classifying BNB+Cosmic events using CNN100000

<sub>1786</sub> **8.2.1 Classification of MC data using Selection I CC-Inclusive Filter**

<sub>1787</sub> Kinematic truth distributions of BNB+Cosmic events passing Selection  
<sub>1788</sub> I+CNN10000

<sub>1789</sub> **8.2.2 Classification of MicroBooNE data using Selection I  
1790 CC-Inclusive Filter**

<sub>1791</sub> **8.2.3 Comparing two CC-Inclusive Cross Section Selection Filters**

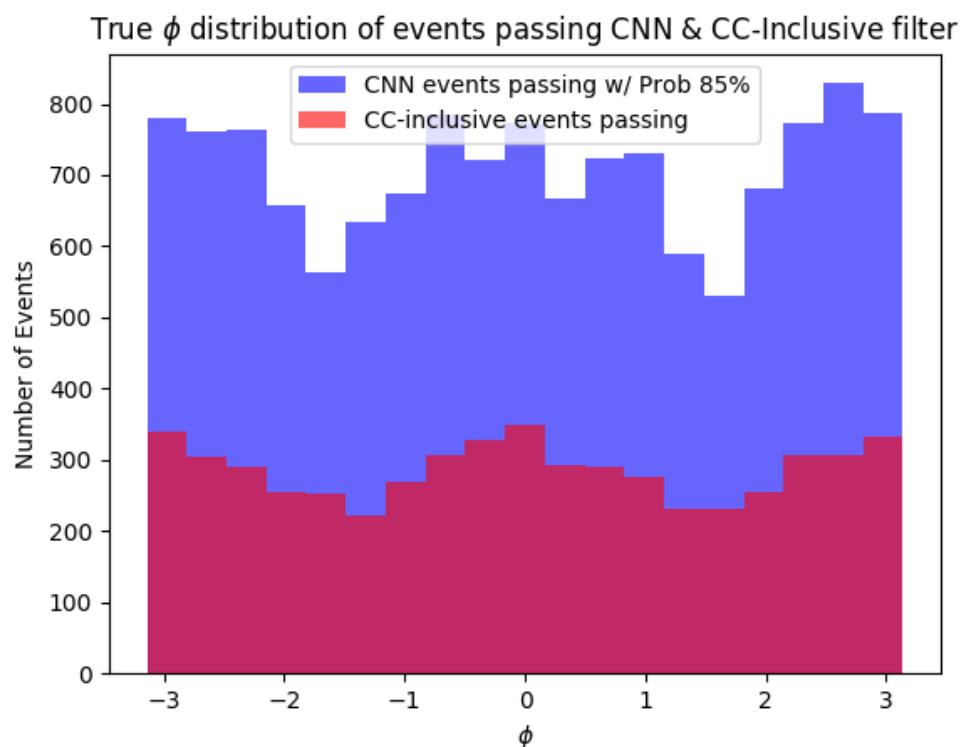
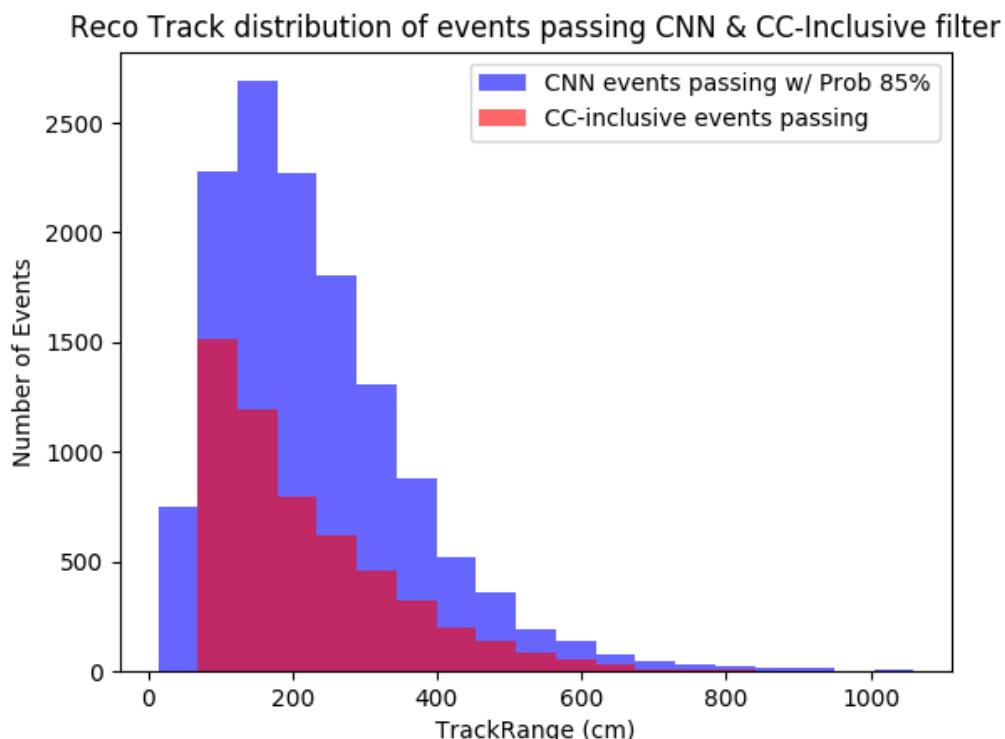
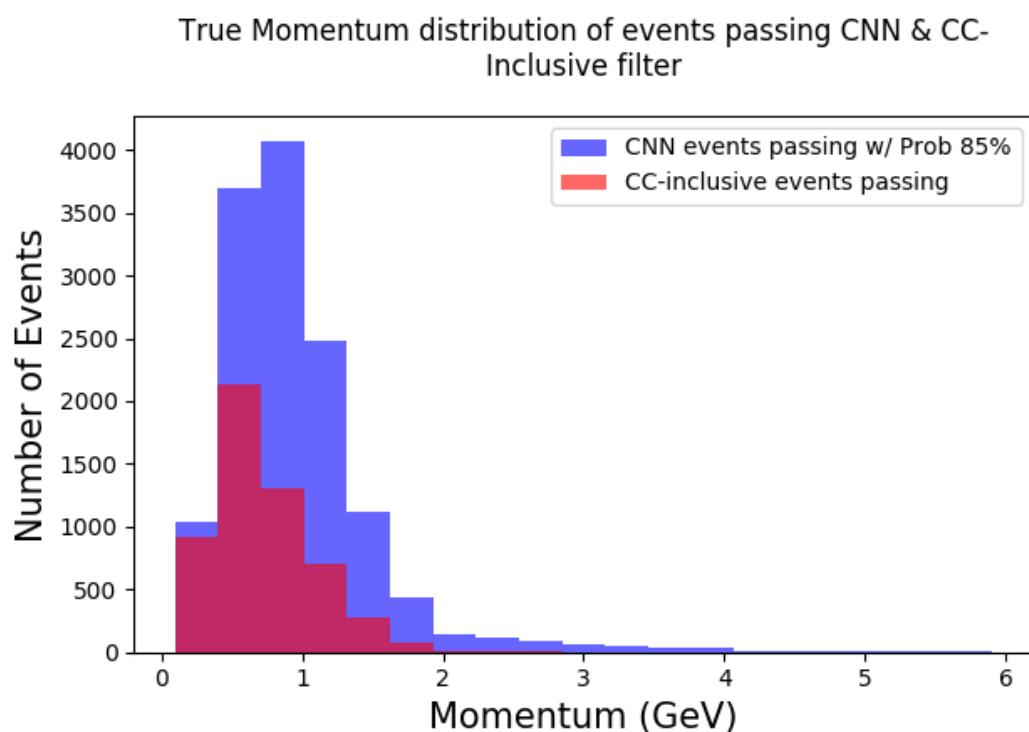


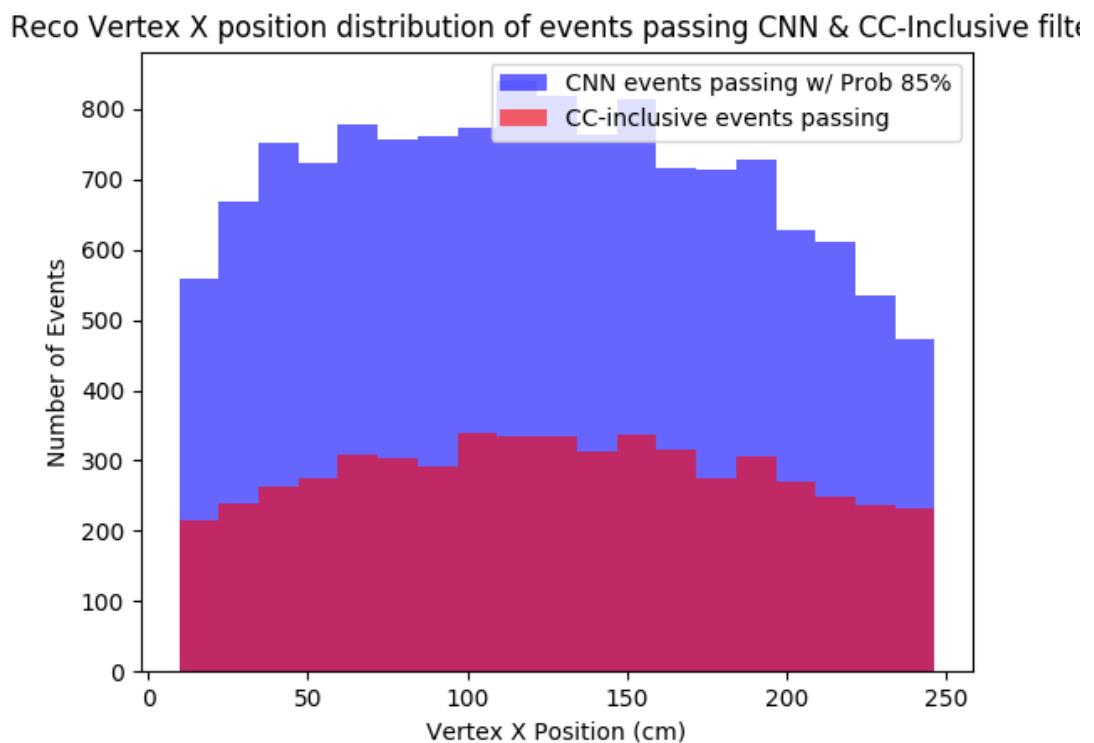
Figure 8.9:  $\cos(\theta)$  distribution at  $\text{CNN10000} \geq 85\%$



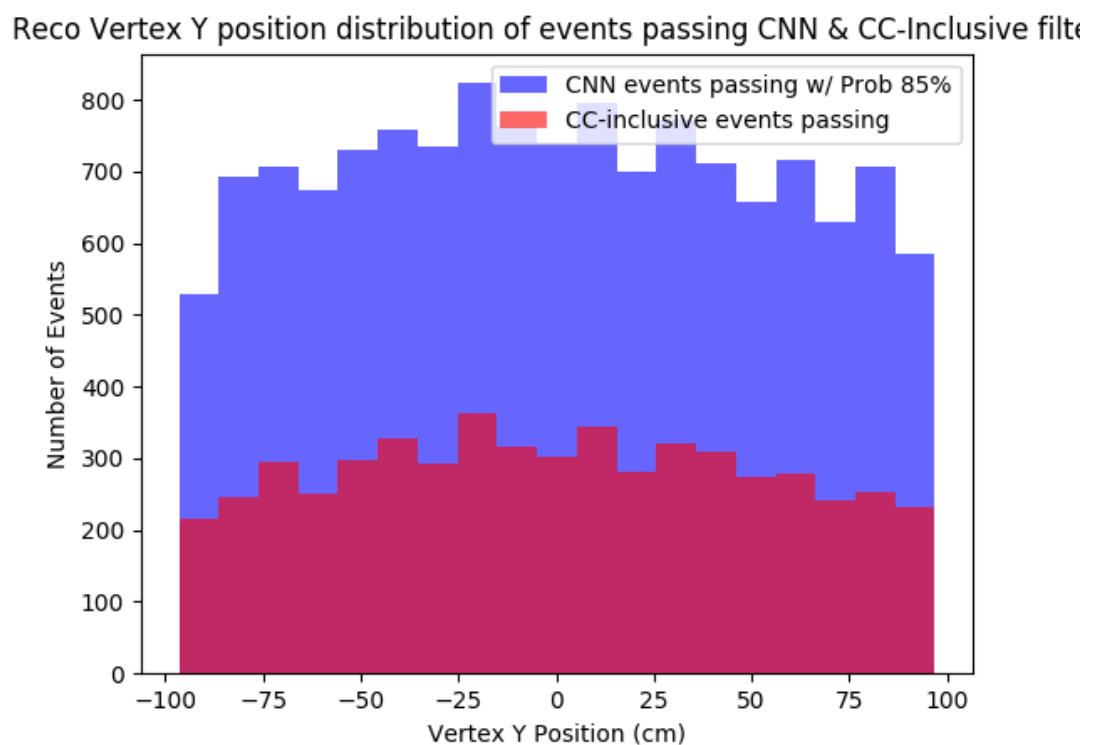
**Figure 8.10:**  $\cos(\theta)$  distribution at  $\text{CNN10000} \geq 85\%$



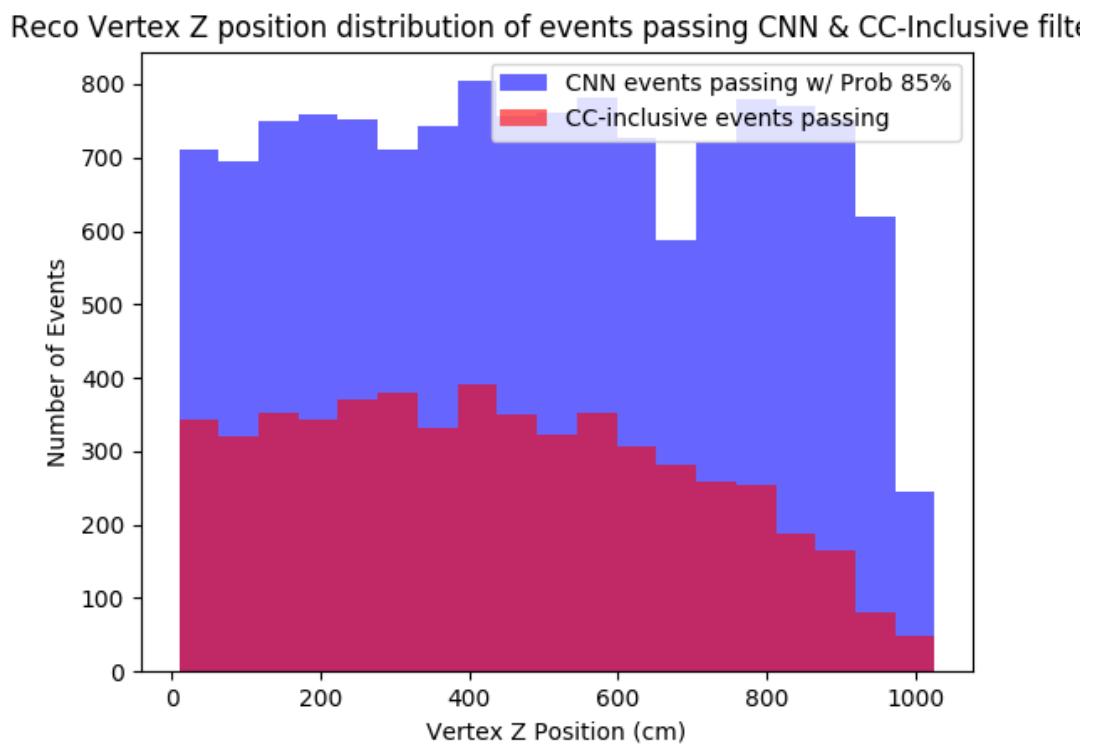
**Figure 8.11:**  $\cos(\theta)$  distribution at  $\text{CNN10000} \geq 85\%$



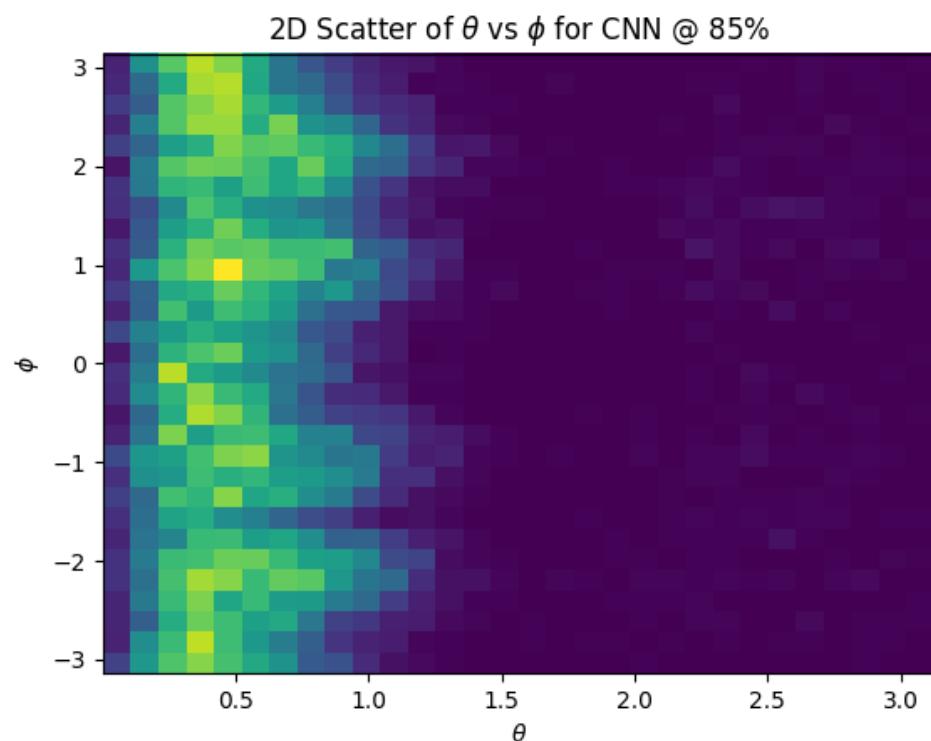
**Figure 8.12:**  $\cos(\theta)$  distribution at  $\text{CNN10000} \geq 85\%$



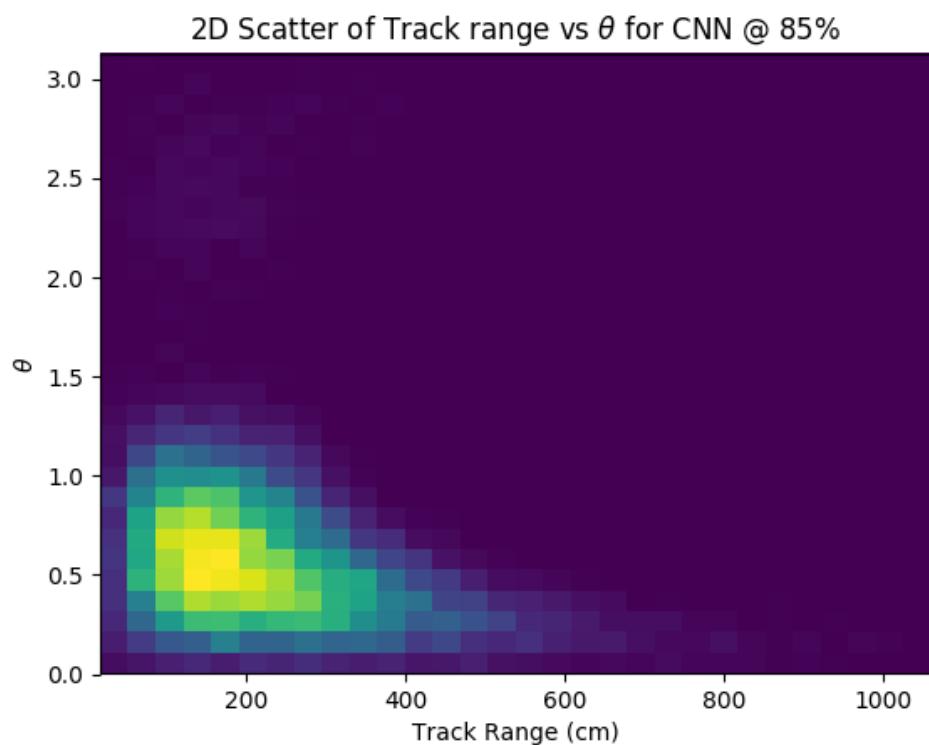
**Figure 8.13:**  $\cos(\theta)$  distribution at  $\text{CNN10000} \geq 85\%$



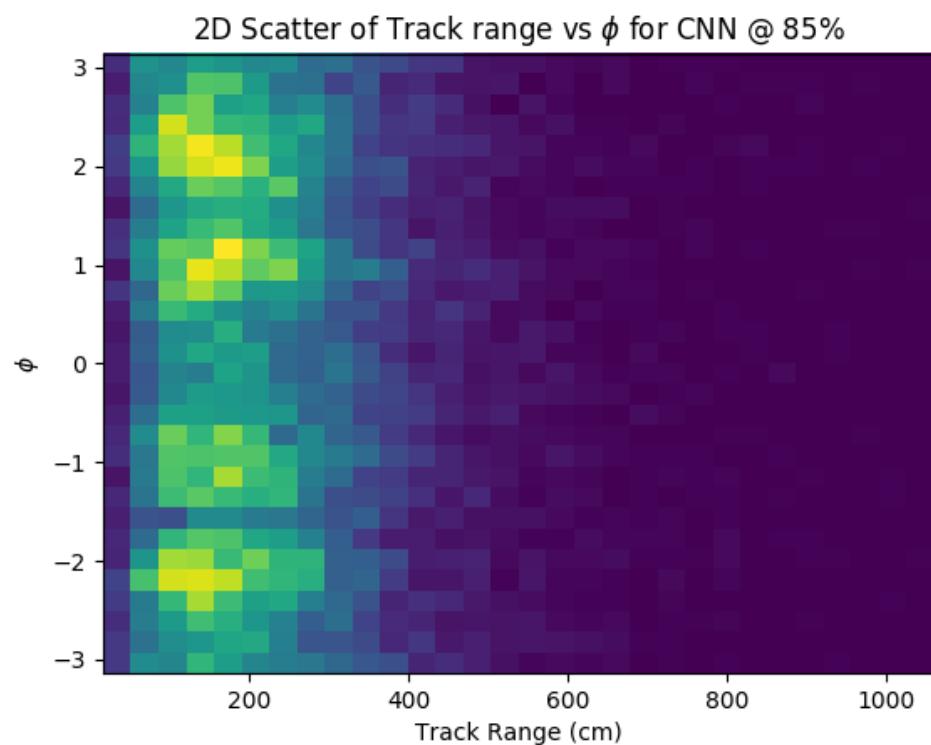
**Figure 8.14:**  $\cos(\theta)$  distribution at  $\text{CNN10000} \geq 85\%$



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



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



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

<sup>1792</sup> **Chapter 9**

<sup>1793</sup> **Conclusion**

<sup>1794</sup> Your Conclusions here.

<sup>1795</sup>

<sub>1796</sub> **Bibliography**

- <sub>1797</sub> [1] Wikipedia, Neutrino, <http://wikipedia.org/wiki/Neutrino>, 2013.
- <sub>1798</sub> [2] Wikipedia, Neutrino oscillation, [http://en.wikipedia.org/wiki/Neutrino\\_oscillation](http://en.wikipedia.org/wiki/Neutrino_oscillation), 2013.
- <sub>1799</sub>
- <sub>1800</sub> [3] B. N. Laboratory, Neutrinos and nuclear chemistry, <http://www.chemistry.bnl.gov/sciandtech/sn/default.htm>, 2010.
- <sub>1801</sub>
- <sub>1802</sub> [4] K. Heeger, Big world of small neutrinos, <http://conferences.fnal.gov/lp2003/forthepublic/neutrinos/>.
- <sub>1803</sub>
- <sub>1804</sub> [5] A. Y. Smirnov, p. 23 (2003), hep-ph/0305106.
- <sub>1805</sub> [6] The MicroBooNE Collaboration, R. e. a. Acciari, (2015), arXiv:1512.06148.
- <sub>1806</sub> [7] The MicroBooNE Collaboration, R. e. a. Acciari, (2016), arXiv:1705.07341.
- <sub>1807</sub> [8] The MicroBooNE Collaboration, R. e. a. Acciari, (2017), arXiv:1704.02927.
- <sub>1808</sub> [9] The MicroBooNE Collaboration, R. e. a. Acciari, JINST **12** (2017).