

1

2  $\mu/\pi$  separation using

3 Convolutional Neural Networks

4 for the MicroBooNE

5 Charged Current Inclusive Cross Section

6 Measurement

Jessica Nicole Esquivel

Bachelor of Science in Electrical Engineering and Applied Physics  
St. Mary's University  
San Antonio, TX, USA 2011

# DISSERTATION

Submitted in partial fulfillment  
of the requirements for the degree  
*Doctor of Philosophy in Physics*

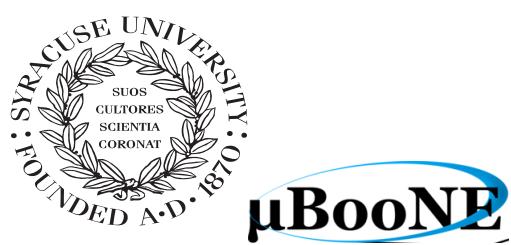
- \* - DRAFT December 30, 2017 - \* -

December, 2017  
Syracuse University  
Syracuse, New York

20  
21  
22  
23

Copyright 2017  
Jessica Nicole Esquivel  
All Rights Reserved  
Syracuse University

24



$\mu$ BooNE







26

## Abstract

27

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



33

## Dedication

34

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

37

Jessica Nicole Esquivel



38

## Acknowledgements

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



# **Contents**

41	<b>List of figures</b>	xv
43	<b>List of tables</b>	xxiii
44	<b>1 Introduction</b>	1
45	<b>2 Neutrinos</b>	3
46	2.1 What are Neutrinos . . . . .	3
47	2.2 History of Neutrinos . . . . .	4
48	2.3 Neutrino Oscillations . . . . .	4
49	2.3.1 Solar Oscillations and the Solar Neutrino Problem . . . . .	5
50	2.3.2 Atmospheric Oscillations and the Atmospheric Neutrino Anomaly	7
51	2.3.3 Two Flavor Neutrino Oscillation Formulation . . . . .	9
52	2.3.4 Three Flavor Neutrino Oscillation Formulation . . . . .	12
53	2.3.5 Reactor Oscillation . . . . .	13
54	<b>3 The MicroBooNE Experiment</b>	15
55	3.1 Liquid argon time projection chambers . . . . .	15
56	3.2 The MicroBooNE Time Projection Chamber . . . . .	16
57	3.3 MicroBooNE’s Physics Goals . . . . .	18
58	3.3.1 The low-energy excess . . . . .	18
59	3.3.2 Cross sections . . . . .	19
60	3.3.3 Liquid argon detector development . . . . .	20
61	3.4 The Booster Neutrino Beam . . . . .	20
62	3.4.1 Creating the Booster Neutrino Beam . . . . .	21
63	3.5 Event Reconstruction . . . . .	22

64	<b>4 Neutrino Identification: Finding MicroBooNE's first Neutrinos</b>	25
65	4.1 Flash Finding . . . . .	25
66	4.1.1 Flash Reconstruction . . . . .	26
67	4.1.2 Beam Timing . . . . .	27
68	4.1.3 Event Rates . . . . .	28
69	4.2 TPC Topology Selection . . . . .	28
70	4.2.1 Cosmic Tagging . . . . .	29
71	4.2.2 2D Cluster Selection . . . . .	29
72	4.2.3 3D Tracks and vertices Selection . . . . .	31
73	4.2.4 TPC Updates . . . . .	33
74	4.3 Conclusion . . . . .	34
75	<b>5 CC-Inclusive Cross Section Selection Filter</b>	35
76	5.1 Data and MC Processing Chain . . . . .	36
77	5.2 Normalization of data and MC . . . . .	37
78	5.3 Optical Software Trigger and Reconstruction . . . . .	39
79	5.3.1 Software Trigger . . . . .	39
80	5.3.2 Flash Reconstruction . . . . .	40
81	5.3.3 Beam Window . . . . .	40
82	5.4 TPC Reconstruction . . . . .	42
83	5.4.1 Hit Reconstruction . . . . .	43
84	5.4.2 Clustering . . . . .	43
85	5.4.3 Pandora . . . . .	44
86	5.4.4 Trackkalmanhit . . . . .	44
87	5.4.5 Cosmic Hit Removal . . . . .	44
88	5.4.6 Projection Matching Algorithm . . . . .	44
89	5.4.7 Calorimetry . . . . .	44
90	5.5 Event Selection . . . . .	44
91	<b>6 Background on Convolutional Neural Networks</b>	45
92	6.1 Image Classification . . . . .	45
93	6.2 CNN Structure . . . . .	46
94	6.2.1 Backpropagation . . . . .	49
95	6.3 Choosing Hyperparameters . . . . .	50

96	<b>7 Training Convolutional Neural Networks on particles WORKING TITLE</b>	53
97	7.1 Hardware Frameworks used for Training . . . . .	54
98	7.1.1 Syracuse CPU Machine setup . . . . .	54
99	7.1.2 Syracuse University GPU Cluster Setup . . . . .	54
100	7.2 Convolutional Neural Network Training . . . . .	54
101	7.2.1 Image Making Scheme . . . . .	54
102	7.2.2 Training CNN1075 . . . . .	55
103	7.2.3 Training CNN10000 . . . . .	55
104	7.2.4 Training CNN100000 . . . . .	57
105	<b>8 Results of Convolutional Neural Networks on particles WORKING TITLE</b>	71
106	8.1 Classification using CNN10000 . . . . .	71
107	8.1.1 Classification of MC data using Selection I Original CC-Inclusive Filter . . . . .	71
108	8.1.2 Classification of MC data using Selection I Modified CC-Inclusive Filter . . . . .	75
109	8.1.3 Conclusions and Future Work . . . . .	83
110	8.2 Classification using CNN100000 . . . . .	85
111	8.2.1 Classification of MC data using Selection I Modified CC-Inclusive Filter . . . . .	85
112	8.2.2 Classification of MicroBooNE data using Selection I Modified CC-Inclusive Filter . . . . .	85
113	8.2.3 Comparing two CC-Inclusive Cross Section Selection Filters . . . . .	85
114	<b>9 Conclusion</b>	87
115	<b>Bibliography</b>	89



# <sup>120</sup> List of figures

<sup>121</sup>	2.1	The Standard Solar Model	5
<sup>122</sup>	2.2	Solar Neutrino Experiments	7
<sup>123</sup>	2.2a	Ray Davis's Homestake Experiment	7
<sup>124</sup>	2.2b	Kamiokande Experiment	7
<sup>125</sup>	2.2c	SNO Experiment	7
<sup>126</sup>	2.3	Cosmic Ray Shower	8
<sup>127</sup>	2.4	Measurements of the double ratio for various atmospheric neutrino experiments	9
<sup>129</sup>	2.5	The flavor eigenstates are rotated by an angle $\theta$ with respect to the mass eigenstates	10
<sup>131</sup>	3.1	Low Energy excess seen in MiniBooNE	19
<sup>132</sup>	3.2	Arial view of the Main Injector and the Booster Neutrino Beam at Fermilab	21
<sup>133</sup>	3.3	Arial view of the Main Injector and the Booster Neutrino Beam at Fermilab	23
<sup>134</sup>	3.4	Energy spectrum of the Booster Neutrino Beam at Fermi National Laboratories	24
<sup>136</sup>	4.1	Efficiency for selecting beam events as a function of minimum total PE cut for all PE cuts as well as zoomed into interesting region.	26
<sup>138</sup>	4.2a	Predicted distribution of flash times with respect to trigger time for 1 day of data taking at nominal rate and intensity	31

140	4.2b Measured distribution of flash times with a 50 PE threshold cut, with respect to trigger time. Shown as a ratio to the expected cosmic rate from off-beam data. A clear excess from neutrinos is visible between 3- 5 $\mu$ s after the trigger time. . . . .	31
144	4.3 Percent of good clusters remaining for neutrinos and cosmics after the primary cuts were applied. This is relative to total number of initial clusters. . . . .	32
147	4.4 Percent matched cluster pairs remaining for neutrinos and cosmics after secondary cuts applied. This is relative to the number of events that contain clusters which pass the primary cuts. . . . .	32
150	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
159	5.1a Track range distribution of selection I . . . . .	38
160	5.1b Selection efficiency as a function of the true muon momentum .	38
161	5.2 Time distribution of reconstructed optical flashes with a PE value of 50 or more for a sample of BNB unbiased triggered events. . . . .	41
163	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
168	5.4 Reconstruction chain run on both data and MC. The red stars mean that the algorithm returns reconstructed 3D vertices. . . . .	42

170	6.1 Pixel representation and visualization of a curve detector filter. As you 171 can see, in the pixel representation, the weights of this filter are greater 172 along a curve we are trying to find in the input image . . . . .	47
173	6.2 Visualization of filters found in first layer of a CNN. . . . .	47
174	6.3 Applying a feature mask over a set of fashion items to extract necessary 175 information for auto-encoding. Unnecessary information for example 176 color or brand emblems are not saved. This feature map is an edge detec- 177 tion mask that leaves only shape information which helps to distinguish 178 between different types of clothes. . . . .	48
179	6.4 Pictorial Representation of Convolutional Neural Networks as well as a 180 visual representation on CNN's complexity of layer feature extraction	48
181	6.5 Pictoral representation of the AlexNet model. The AlexNet model 182 consists of 5 convolution layers and 3 fully connected layers. . . . .	51
183	6.6 Pictoral representation of the GoogleNet model. The GoogleNet model 184 consists of 22 layers. The model implements 9 Inception modules which 185 performs covolution and pooling in parallel and strays away from 186 the basis that CNN layers need to be stacked up sequentially. The 187 GoogleNet model also doesn't use fully connected layers, instead it 188 uses average pooling which greatly reduces the amount of parameters. 189 GoogleNet has 12x fewer parameters than AlexNet. . . . .	51
190	7.1 Accuracy vs. Loss of ImageNet 2-output $\mu/\pi$ sample consisting of 191 10000 images each. . . . .	57
192	7.2 Description of confusion matrix variables: False pion rate = $false\pi/total\pi$ 193 True pion rate = $true\pi/total\pi$ Accuracy = $(true\pi rate + true\mu rate)/2$ 194 Pion prediction value = $true\pi/(true\pi + false\pi)$ Muon prediction value 195 = $true\mu/(true\mu + false\mu)$ 7.2c The probability plot includes muons and 196 pions that are classified as primary particles. . . . .	58
197	7.2a Confusion Matrix showing Accuracy of CNN using training data	58
198	7.2b Confusion Matrix showing Accuracy of CNN using testing data	58

199	7.2c 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. . . . .	58
200		
201		
202	7.3 Confusion Matrix of all five particles . . . . .	59
203		
204		
205		
206	7.4 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 . . . . .	60
207		
208	7.5 Training and testing loss of CNN trained on 100,000 images of $\mu/\pi/p/\gamma/e$ . . . . .	60
209		
210	7.6 t-SNE of CNN . . . . .	61
211		
212	7.7 Muon Prob . . . . .	62
213		
214	7.8 Pion Prob . . . . .	62
215		
216	7.9 Proton Prob . . . . .	63
217		
218	7.10 Electron Prob . . . . .	63
219		
220	7.11 Gamma Prob . . . . .	64
221		
222	7.12 Prob . . . . .	64
223		
224	7.13 mupi . . . . .	65
225		
226	7.14 mupi . . . . .	65
227		
228	7.15 mupi . . . . .	66
229		
230	7.16 mup . . . . .	66
231		
232	7.17 mup . . . . .	67
233		
234	7.18 mup . . . . .	67
235		
236	7.19 mue . . . . .	68
237		
238	7.20 mue . . . . .	68
239		
240	7.21 mue . . . . .	69
241		
242	7.22 mug . . . . .	69
243		

---

225	7.23 mug . . . . .	70
226	7.24 mug . . . . .	70
227	8.1 Snapshot of passing rates of Selection I from CC-Inclusive Filter . . . . .	72
228	8.2 Results of CNN10000 classification of track candidate images output from cc-inclusive filter. . . . .	73
230	8.2a Confusion Matrix showing Accuracy of CNN using data with wrong normilazion . . . . .	73
232	8.2b Probability plot showing $\mu/\pi$ separation of CNN using wrong normalization . . . . .	73
234	8.2c Confusion Matrix showing Accuracy of CNN using data with correct normilazion . . . . .	73
236	8.2d Probability plot showing $\mu/\pi$ separation of CNN using correct normalization . . . . .	73
238	8.3 CNN10000 distributions of track candidate images output from Selec- tion I Original cc-inclusive filter with different image data normalizations	74
240	8.3a Track range distribution of events from Selection I Original passing CNN with 70% accuracy using image data with wrong normilazion . . . . .	74
243	8.3b Track range distribution of events from Selection I Original pass- ing CNN with 70% accuracy using image data with correct normilazion . . . . .	74
246	8.3c Momentum distribution of events from Selection I Original pass- ing CNN with 70% accuracy using image data with wrong normilazion . . . . .	74
249	8.3d Momentum distribution of events from Selection I Original pass- ing CNN with 70% accuracy using image data with correct normilazion . . . . .	74
252	8.4 Snapshot of passing rates of all cuts from Selection I Modified cc- inclusive filter . . . . .	75

---

254	8.5 Confusion matrix and probability plot of events passing selection I modified cc-inclusive cuts right before 75cm track length cut . . . . .	76
255		
256	8.5a Confusion Matrix for CNN10000 classified events from selection I modified . . . . .	76
257		
258	8.5b Probability plot for CNN10000 classified events from selection I modified . . . . .	76
259		
260	8.6 CNN10000 distributions of track candidate images output from Selection I Modified cc-inclusive filter . . . . .	77
261		
262	8.6a Track range distribution of events from Selection I Modified passing CNN with 70% accuracy . . . . .	77
263		
264	8.6b Stacked signal and background track range distributions from Selection I Modified passing CNN with 70% accuracy . . . . .	77
265		
266	8.6c Stacked signal and background track range distributions from Selection I Modified passing 75 cm track length cut . . . . .	77
267		
268	8.6d Stacked signal muons and background muons/pions of track range distributions from Selection I Modified passing CNN with 70% accuracy . . . . .	77
269		
270	8.6e Stacked signal muons and background muons/pions of track range distributions from Selection I Modified passing 75 cm track length cut . . . . .	77
271		
272		
273		
274	8.7 CNN10000 stacked signal/background track range distributions of track candidate images output from Selection I Modified cc-inclusive filter	78
275		
276	8.7a Stacked signal muons and background muons/pions of track range distributions from Selection I Modified passing CNN with 75% accuracy . . . . .	78
277		
278	8.7b Stacked signal muons and background muons/pions of track range distributions from Selection I Modified passing CNN with 80% accuracy . . . . .	78
279		
280		
281		

282	8.7c	Stacked signal muons and background muons/pions of track range distributions from Selection I Modified passing CNN with 85% accuracy . . . . .	78
285	8.7d	Stacked signal muons and background muons/pions of track range distributions from Selection I Modified passing CNN with 90% accuracy . . . . .	78
288	8.8	CNN10000 momentum distributions of track candidate images output from Selection I Modified cc-inclusive filter . . . . .	79
290	8.8a	Momentum distribution of events from Selection I Modified passing CNN with 70% accuracy . . . . .	79
292	8.8b	Stacked signal and background momentum distributions from Selection I Modified passing CNN with 70% accuracy . . . . .	79
294	8.8c	Stacked signal and background momentum distributions from Selection I Modified passing 75 cm track length cut . . . . .	79
296	8.9	Track distribution comparisons of true CC muons plotted vs true CC muons and pions plotted . . . . .	81
298	8.9a	Stacked signal $\mu$ /background $\mu$ and $\pi$ track range distribution of CNN @ 70% . . . . .	81
300	8.9b	Stacked signal $\mu\&\pi$ /background $\mu\&\pi$ track range distribution of CNN @ 70% . . . . .	81
302	8.10	Images of true CC events where the pion was the tagged track candidate . . . . .	81
303	8.10a	Pion reconstructed track range is less than 75 cm and longer than muon track due to dead wires . . . . .	81
305	8.10b	Pion reconstructed track range is less than 75 cm and larger than muon reconstructed track . . . . .	81
307	8.10c	Pion reconstructed track range is greater than 75 cm and larger than muon reconstructed track . . . . .	81
309	8.11	Snapshot of signal and background event numbers of Selection I modified from cc-inclusive note [?] . . . . .	84

311	8.12 CNN performance of classified muons and pions compared to the al-	
312	ready implemented 75 cm track length cut . . . . .	84

# <sup>313</sup> List of tables

<sup>314</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. . . . .	31
<sup>320</sup> 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 canâŽt be determined at this step. . .	82

---

336	8.2 Signal and background event numbers at modified selection level with	
337	CNN cut estimated from a BNB+Cosmic sample and Cosmic only sam-	
338	ple normalized to $5 * 10^{19}$ PoT. The last column gives the fraction of	
339	this signal or background type to the total selected events per CNN	
340	probability. . . . .	82

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

— Shirley Chisholm



<sup>343</sup> **Chapter 1**

<sup>344</sup> **Introduction**

<sup>345</sup> This thesis will be a description of work done to further increase efficiency and purity  
<sup>346</sup> of the charged current inclusive cross section measurement using the MicroBooNE  
<sup>347</sup> detector. It will also describe the MicroBooNE detector, what neutrinos are, the  
<sup>348</sup> charged current inclusive cross section measurement and its importance as well as  
<sup>349</sup> convolutional neural networks and how they can be used in  $\mu/\pi$  separation. Chapter  
<sup>350</sup> **2** will talk about the background of neutrinos and the people and detectors that  
<sup>351</sup> discovered neutrinos as well as an in depth history of neutrino oscillation and the  
<sup>352</sup> discovery that neutrinos have mass.

<sup>353</sup> Chapter **3** will discuss the MicroBooNE experiment, specifically, how Liquid  
<sup>354</sup> Argon Time Projection Chambers work, the Light Collection System and the Electronic  
<sup>355</sup> and Readout Trigger systems. This chapter will also describe the Booster Neutrino  
<sup>356</sup> Beam sationed at Fermilab.

<sup>357</sup> Chapter **4** will discuss the work that was done to detect the first neutrinos seen in  
<sup>358</sup> the MicroBooNE detector and the software reconstruction efforts required to create an  
<sup>359</sup> automated neutrino ID filter that was used to find the first neutrinos and then was  
<sup>360</sup> later expanded on to create the charged current inclusive filter that will be discussed  
<sup>361</sup> in chapter **5**

<sup>362</sup> Chapter **6** will give a brief description of what Convolutional Neural Networks are  
<sup>363</sup> and how it will be used for  $\mu/\pi$  separation in this selection. Chapter **7** will discuss  
<sup>364</sup> the hardware frameworks and training methods used to train multiple Convolutional  
<sup>365</sup> Neural Networks for use in the charged current inclusive cross section measurement.  
<sup>366</sup> Chapters **8** and **??** will discuss the results of using Convolutional Neural Networks on  
<sup>367</sup> monte-carlo and data to sift out charged current inclusive neutrino events.



<sup>368</sup> **Chapter 2**

<sup>369</sup> **Neutrinos**

<sup>370</sup> **2.1 What are Neutrinos**

<sup>371</sup> Neutrinos are one of the fundamental particles which make up the universe. They are  
<sup>372</sup> also one of the least understood. Neutrinos are not affected by the electromagnetic  
<sup>373</sup> forces because they do not have electric charge. Neutrinos are affected by a "weak"  
<sup>374</sup> sub-atomic force of much shorter range than electromagnetism, and are therefore able  
<sup>375</sup> to pass through great distances in matter without being affected by it. Until the late  
<sup>376</sup> 90's, neutrinos were thought to have no mass. Due to their mass, neutrinos are also  
<sup>377</sup> affected by gravity. Neutrinos are created by radioactive decay or nuclear reactions  
<sup>378</sup> such as the ones that happen in the sun, in nuclear reactors or when cosmic rays hit  
<sup>379</sup> atoms. There are three types of neutrinos,  $\nu_e$ ,  $\nu_\mu$  and  $\nu_\tau$  which correspond to their  
<sup>380</sup> charged lepton pairs.

<sup>381</sup> As previously stated, neutrinos are very weakly interacting; in fact, neutrinos can  
<sup>382</sup> pass unscathed through a wall of lead several hundred light-years thick. Because  
<sup>383</sup> neutrinos interact so rarely, studying neutrinos requires a massive detector and a  
<sup>384</sup> powerful neutrino source. With that being said, we can only infer their existence when  
<sup>385</sup> they interact in a detector. In a collision, distinct charged particles are produced with  
<sup>386</sup> each type of neutrino. An electron neutrino will create an electron, a muon neutrino  
<sup>387</sup> will create a muon, and a tau neutrino will create a tau. The track the particle leaves  
<sup>388</sup> in the detector is how one figures out what type of neutrino interaction was "seen".  
<sup>389</sup> Liquid Argon Time Projection Chambers are the newest type of detectors being used to  
<sup>390</sup> study neutrinos due to their excellent imaging and particle identification capabilities.

## <sup>391</sup> 2.2 History of Neutrinos

<sup>392</sup> The neutrino was first postulated by Wolfgang Pauli in 1931 to explain how beta  
<sup>393</sup> decay could resolve the conservation of energy, momentum and angular momentum  
<sup>394</sup> problem. Pauli suggested that this missing energy might be carried off, unseen, by a  
<sup>395</sup> neutral particle (he called neutron) which was escaping detection. James Chadwick  
<sup>396</sup> discovered a much heavier nuclear particle in 1932 that he also named neutron, leaving  
<sup>397</sup> two particles with the same name. Enrico Fermi was the first person to coin the  
<sup>398</sup> term neutrino (which means little neutral one in latin) in 1933 to fix this confusion.  
<sup>399</sup> Fermi's paper, which was published in 1934, unified Pauli's neutrino with Paul Dirac's  
<sup>400</sup> positron and Werner Heisenberg's neutron-proton model and his theory accurately  
<sup>401</sup> explained many experimentally observed results. Wang Ganchang first proposed the  
<sup>402</sup> use of beta capture to experimentally detect neutrinos and in 1959 Clyde Cowan and  
<sup>403</sup> Frederick Reines published their work stating that they had detected the neutrino.  
<sup>404</sup> The experiment called for antineutrinos created in a nuclear reactor by beta decay that  
<sup>405</sup> reacted with protons producing neutrons and positrons:  $\nu_e + p^+ \rightarrow n^0 + e^+$ . Once  
<sup>406</sup> this happens, the positron finds an electron and they annihilate each other and the  
<sup>407</sup> resulting gamma rays are detectable. The neutron is detected by neutron capture and  
<sup>408</sup> the releasing of another gamma ray. In 1962 Leon M. Lenderman, Melvin Schwartz  
<sup>409</sup> and Jack Steinberger were the first to detect interactions of the muon neutrino. The  
<sup>410</sup> first detection of the tau neutrino was announced in the summer of 2000 by the  
<sup>411</sup> DONUT collaboration at Fermilab. In the late 1960s, many experiments found that the  
<sup>412</sup> number of electron neutrinos arriving from the sun was around 1/3 to 1/2 the number  
<sup>413</sup> predicted by the Standard Solar Model. This became known as the solar neutrino  
<sup>414</sup> problem and remained unresolved for around thirty years. This problem was resolved  
<sup>415</sup> by the discovery of neutrino oscillation and mass. [1]

## <sup>416</sup> 2.3 Neutrino Oscillations

<sup>417</sup> Neutrino oscillation was first predicted by Bruno Pontecorvo. It describes the phe-  
<sup>418</sup> nomenon of a neutrino created with a specific lepton flavor (electron, muon or tau)  
<sup>419</sup> that is later measured to have a different flavor. Neutrino oscillation is important  
<sup>420</sup> theoretically and experimentally due to the fact that this observation implies that the  
<sup>421</sup> neutrino has a non-zero mass, which is not part of the original Standard Model of  
<sup>422</sup> particle physics. [2]

### 423 2.3.1 Solar Oscillations and the Solar Neutrino Problem

424 The solar neutrino flux derived from Bahcall's Standard Solar Model is shown in figure  
 425 2.1. Nuclear fusion and decay processes produce an abundant amount of neutrinos.  
 426 The standard solar model predicts that these reactions produce several groups of  
 427 neutrinos, each with differing fluxes and energy spectra. The figure also shows the  
 428 ranges of detection of existing solar neutrino experiments in different shades of blue  
 429 to illustrate that they sample different portions of the solar neutrino energy spectrum.  
 430 Three of these experiments, plus a new one, are discussed below.

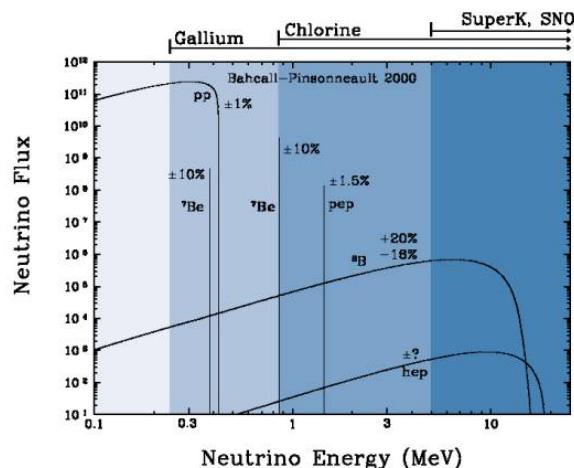


Figure 2.1: The Standard Solar Model

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

435 The first experiment to detect the effects of neutrino oscillation was the Ray Davis's  
 436 Homestake Experiment. The detector was stationed in the Homestake Gold Mine in  
 437 Lead, South Dakota. It was 1,478 meters underground and was  $380\ m^3$ . The detector  
 438 was filled with perchloroethylene. Perchloroethylene was chosen because of its high  
 439 concentrations of chlorine. When an  $\nu_e$  interacted with chlorine-37 atom, the atom  
 440 would transform to argon-37 which was then extracted and counted. The neutrino  
 441 capture reaction is shown in equation 2.1. Davis observed a deficit of about 1/3  
 442 the flux of solar neutrinos that was predicted by Bahcall's Standard Solar Model.

- <sup>443</sup> The unexplained difference between the measured solar neutrino flux and model  
<sup>444</sup> predictions lead to the Solar Neutrino Problem. [4]

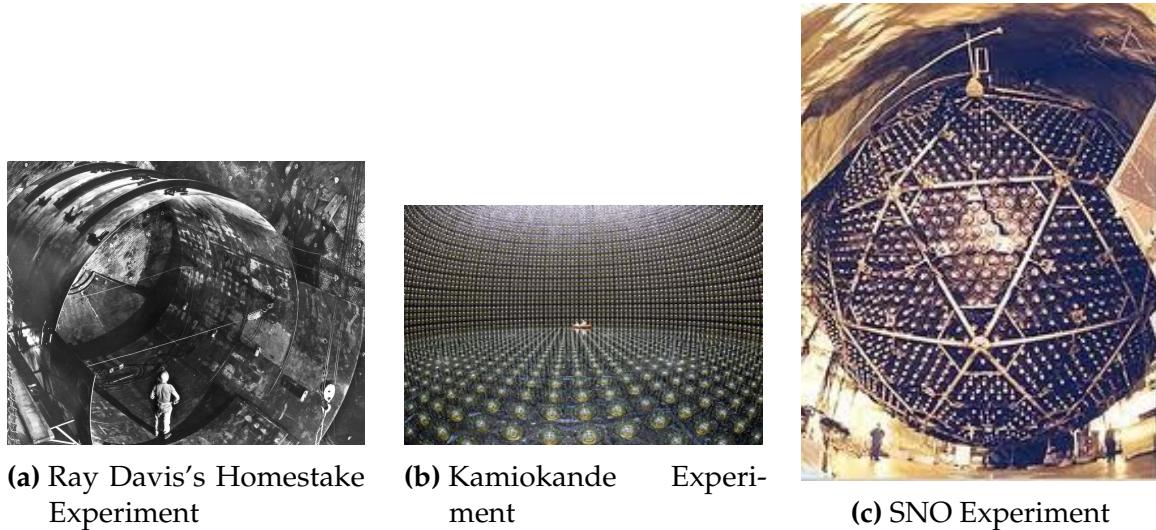


<sup>445</sup> While it is now known that the Homestake Experiment detected neutrinos, some  
<sup>446</sup> physicist were weary of the results. Conclusive evidence of the Solar Neutrino Problem  
<sup>447</sup> was provided by the Kamiokande-II experiment, a water Cherenkov detector with  
<sup>448</sup> a low enough energy threshold to detect neutrinos through neutrino-electron elastic  
<sup>449</sup> scattering. In the elastic scattering interaction the electrons coming out of the point of  
<sup>450</sup> reaction strongly point in the direction that the neutrino was traveling, away from the  
<sup>451</sup> sun. While the neutrinos observed in Kamiokande-II were clearly from the sun, there  
<sup>452</sup> was still a discrepancy between Kamiokande-II and Homestake; The Kamiokande-  
<sup>453</sup> II experiment measured about 1/2 the predicted flux, rather than the 1/3 that the  
<sup>454</sup> Homestake Experiment saw.

<sup>455</sup> The solution to the solar neutrino problem was finally experimentally determined  
<sup>456</sup> by the Sudbury Neutrino Observatory(SNO). The Ray Davis's Homestake Experiment  
<sup>457</sup> was only sensitive to electron neutrinos, and the Kamiokande-II Experiment was  
<sup>458</sup> dominated by the electron neutrino signal. The SNO experiment had the capability to  
<sup>459</sup> see all three neutrino flavors. Because of this, it was possible to measure the electron  
<sup>460</sup> neutrinos and total neutrino flux. The experiment demonstrated that the deficit was  
<sup>461</sup> due to the MSW effect, the conversion of electron neutrinos from their pure flavor  
<sup>462</sup> state into the second neutrino mass eigenstate as they passed through a resonance  
<sup>463</sup> due to the changing density of the sun. The resonance is energy dependent, and is  
<sup>464</sup> visible near 2MeV. The water cherenkov detectors only detect neutrinos above about  
<sup>465</sup> 5MeV, while the radiochemical experiments were sensitive to lower energy (0.8MeV  
<sup>466</sup> for chlorine, 0.2MeV for gallium), and this turned out to be the source of the difference  
<sup>467</sup> in the observed neutrino rates at the two types of experiments. Figure 2.2 shows  
<sup>468</sup> Homestake, Kamiokande-II and SNO experiments.

### <sup>469</sup> MSW Effect

<sup>470</sup> The Mikheyev-Smirnov-Wolfenstein effect is a process which acts to modify neu-  
<sup>471</sup> trino oscillations in matter. The presence of electrons in matter changes the energy



**Figure 2.2:** Solar Neutrino Experiments

472 levels of the mass eigenstates of neutrinos due to charged current coherent forward  
 473 scattering of the electron neutrinos. This coherent forward scattering is similar to  
 474 the electromagnetic process with respect to the refractive index of light in a medium.  
 475 Because of this MSW Effect, neutrinos in vacuum have a different effective mass than  
 476 neutrinos in matter and because neutrino oscillations depend on the squared mass  
 477 difference of the neutrinos, the neutrino oscillations are different in matter than in  
 478 vacuum. This effect is important at the sun where electron neutrinos are produced.  
 479 The neutrinos of high energy leaving the sun are in a vacuum propagation eigenstate  
 480  $\nu_2$  that has a very small overlap with the electron neutrino  $\nu_e = \nu_1\cos(\theta) + \nu_2\sin(\theta)$   
 481 seen by the charged current reactions in Kamiokande-II and SNO. The discrepancy of  
 482 the deficit between SNO, Kamiokande-II and Homestake is due to the energy of the  
 483 solar neutrinos. The MSW effect "turns on" at about 2MeV and at lower energies, this  
 484 MSW effect is negligible. [5]

485 **2.3.2 Atmospheric Oscillations and the Atmospheric Neutrino  
 486 Anomaly**

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

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

490

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



**Figure 2.3:** Cosmic Ray Shower

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

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

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

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

503     Kamiokande-II has the the capability of measuring the direction of the incoming  
 504     neutrinos. The expectation of atmospheric neutrino detection is that the flux be

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

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

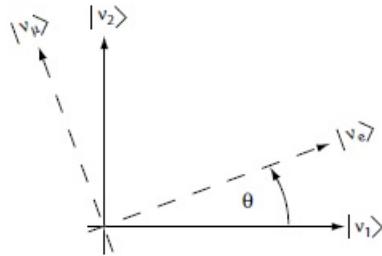
isotropic due to the fact that atmospheric neutrinos can reach the detector from all directions. Kamiokande-II noticed that muon-like data did not agree well with this expectation. At low energies approximately half of the  $\nu_\mu$  are missing over the full range of zenith angles. At high energies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

### 2.3.3 Two Flavor Neutrino Oscillation Formulation

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

In matrix form the wavefunctions are:

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



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

524 Applying the time evolution operator to  $\nu_\mu$ :

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

525 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  
526 being, let us assume  $\hbar = c = 1$ . With this assumption:  $E_1 = \sqrt{p^2 + m_1^2}$  and  $E_2 =$   
527  $\sqrt{p^2 + m_2^2}$ . The next modifications is to assume neutrinos are relativistic:

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

528 because of this,

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

529

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

530 where the binomial expansion is used. Now  $E_1$  and  $E_2$  can be written as:

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

531 Now applying all these assumptions back into equation 2.7 gives us:

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

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

<sup>532</sup> Substituting  $\Delta m^2 = m_1^2 - m_2^2$  and  $t = \frac{x}{c} = x$  and  $e^{-iz} = e^{-i\left(p+\frac{1}{2}\frac{m_1^2}{p}\right)t}$  gives us:

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

<sup>533</sup> Finding the Probability for a  $\nu_\mu \rightarrow \nu_e$ :

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

<sup>534</sup> Remembering that  $\langle \nu_i | \nu_j \rangle = \delta_{ij}$

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

<sup>535</sup> Taking the absolute value squared gives us:

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

<sup>536</sup> Since the neutrino is relativistic we can set  $p = E_\nu$  and change  $x = L$ . Also <sup>537</sup> recognizing the trigonometric relation  $(1 - \cos 2\theta)/2 = \sin^2\theta$  the above equation <sup>538</sup> becomes:

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

<sup>539</sup> All that's left to do now is re-introduce  $\hbar$  and  $c$  doing this we get:

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

<sup>540</sup> This equations has three important variables.

- The angle  $\theta$ : This angle, as mentioned before, is called the mixing angle. It defines the difference between the flavor and the mass eigenstates. When  $\theta = 0$  the mass and flavor eigenstates are identical and now oscillations occur.
- The mass squared difference,  $\Delta m^2$ : Again  $\Delta m^2 = m_1^2 - m_2^2$ . The reason this is an important variable is because it implies that for neutrinos to oscillate, neutrinos must have mass. Furthermore, the mass squared difference also tells us that the neutrino mass eigenstates must be different.
- L/E: This is the variable that is of most interest to experimental physicists due to the fact that it is the variable that we set. L is the distance between the source and detector and E is the energy of the neutrino. For a given  $\Delta m^2$ , the probability of oscillation changes with respect to L/E.

### 2.3.4 Three Flavor Neutrino Oscillation Formulation

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

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

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

Following the same steps as before we get:

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

The main things to notice here are  $\delta_{ij}$  which is the CP violating term and has not been measured yet, and  $\theta_{13}$  which has just been measured. CP violation is a violation

562 of the postulated CP-symmetry. CP-symmetry states that the laws of physics should  
563 be the same if a particle were to be exchanged with its antiparticle and then if the left  
564 hand side of a decay were switched with the right hand side.

565 **2.3.5 Reactor Oscillation**

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



<sup>573</sup> **Chapter 3**

<sup>574</sup> **The MicroBooNE Experiment**

<sup>575</sup> The purpose of this chapter is to discuss and understand the details of the MicroBooNE  
<sup>576</sup> detector. A thorough understanding of MicroBooNE and the technology behind liquid  
<sup>577</sup> argon time projection chambers is important for understanding results as well as  
<sup>578</sup> understanding how images were made for use in deep learning efforts that will be  
<sup>579</sup> outlined in later chapters.

<sup>580</sup> **3.1 Liquid argon time projection chambers**

<sup>581</sup> Liquid Argon Time Projection Chambers (LArTPCs) are an exciting detector technol-  
<sup>582</sup> ogy that provide excellent imaging and particle identification, and are now being  
<sup>583</sup> used to study neutrinos. The Time Projection Chamber (TPC) was first invented by  
<sup>584</sup> Nygren in 1974 [?] and the proposal for a LArTPC for neutrino physics was made  
<sup>585</sup> by Rubbia [?] in 1977 with the ICARUS collaboration implementing this concept [?].  
<sup>586</sup> A LArTPC is a three-dimensional imaging detector that uses planes of wires at the  
<sup>587</sup> edge of an active volume to read out an interaction. When a neutrino interacts with an  
<sup>588</sup> argon atom, the charged particles that are produced ionize the LAr as they travel away  
<sup>589</sup> from the interaction. By placing a uniform electric field throughout the LAr volume,  
<sup>590</sup> the ionization is made to drift towards a set of anode planes, which consist of wires  
<sup>591</sup> spaced very closely together collecting the ionized charge, which is subsequently read  
<sup>592</sup> out by electronics connected to the anode wires. The collected ionization creates a  
<sup>593</sup> spatial image of what happened in the detector on each anode plane. The position  
<sup>594</sup> resolution of the interaction along the beam direction (perpendicular to drift direction)  
<sup>595</sup> relies on the wire pitch, while the resolution in drift direction is dependent on the

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

## 621 3.2 The MicroBooNE Time Projection Chamber

622 MicroBooNE (Micro Booster Neutrino Experiment) is a 89 ton active volume (180 ton  
623 total mass) LArTPC which is then inserted into a cylindrical cryostat on axis of the  
624 Booster Neutrino Beam (BNB) stationed at Fermilab in Batavia, Illinois. Understanding  
625 LArTPC technology and detector physics is necessary to build a LArTPC the size of  
626 DUNE, and MicroBooNE has made many advances in developing this technology [7]  
627 [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.

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

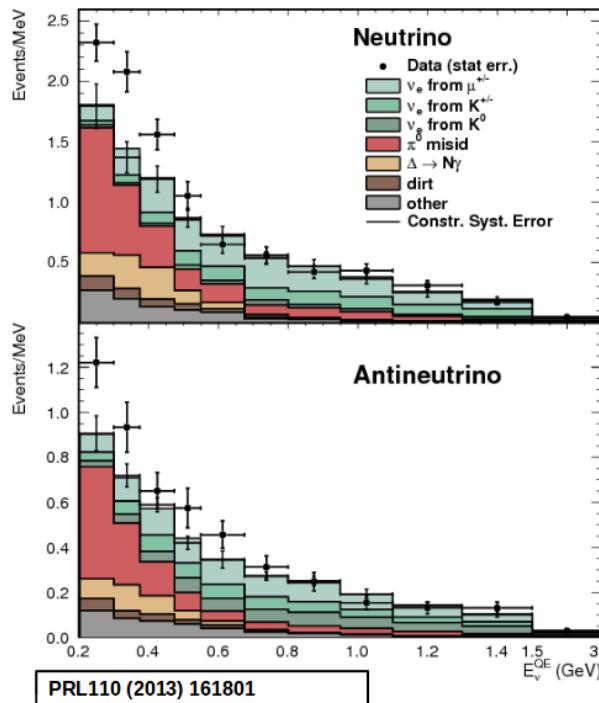
665 Both the light collection system and the TPC create analog signal that is read out and  
666 digitized by the electronics system. The process requires amplification and shaping of  
667 the signal which then goes to the data acquisition (DAQ) software for writing of the  
668 digitized data to disk. The anode plane wires are connected to detector specific circuit  
669 boards (ASICS) that are submerged and operate inside the liquid argon volume. These  
670 ASICS send amplified signal to 11 feed-throughs where further amplification of the  
671 signal happens outside the cryostat. The signal is received by custom LArTPC readout  
672 modules distributed over nine readout crates which do the digitization. The TPC wires  
673 are digitized at 16 MHz then downsampled to 2 MHz. The TPC system reads out 4  
674 frames of wire signal data per event, 1 frame before a trigger and 2 frames after the  
675 triggered frame. The four frames allows for identification of a neutrino interaction as  
676 well as cosmic background rejection. The process of digitization is similar for the light  
677 collection system. Each PMT signal undergoes a shaping with a 60 ns peaking time  
678 for digitization of multiple samples. The digitization occurs at 64 MHz but are not  
679 read out continuously during the TPC readout time. Only shaped PMT signal samples  
680 above a small threshold are read out and saved. Both the TPC and PMT readouts are  
681 initiated via triggers on a separate trigger board located in a warm electronics crate.  
682 The timing trigger is created by a timing signal from the BNB accelerator which is  
683 shaped and sent to the trigger board. The PMT trigger is generated when the PMT  
684 signal multiplicity is greater than 1 and the summed PMT pulse-height is more than 2  
685 photo-electrons summed up over all PMT channels. When the trigger board gets both  
686 a timing trigger and a PMT trigger in coincidence, at BNB trigger is then generated by  
687 the board. This signal is then passed to all readout crates initiating the readout of data.  
688 The data is then sent to the DAQ software which then saves the data to disk into one  
689 event memory.

690 **3.3 MicroBooNE's Physics Goals**

691 **3.3.1 The low-energy excess**

692 The primary goal of the MicroBooNE experiment is to study and investigate the low-  
693 energy excess seen in MiniBooNE 3.1. MicroBooNE has the capability of confirming or

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



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

### 3.3.2 Cross sections

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

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

### 727 3.3.3 Liquid argon detector development

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

## 732 3.4 The Booster Neutrino Beam

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

<sup>740</sup> introduces to a measurement. An aerial view of fermilab as well as the BNB is shown  
<sup>741</sup> in figure 3.2

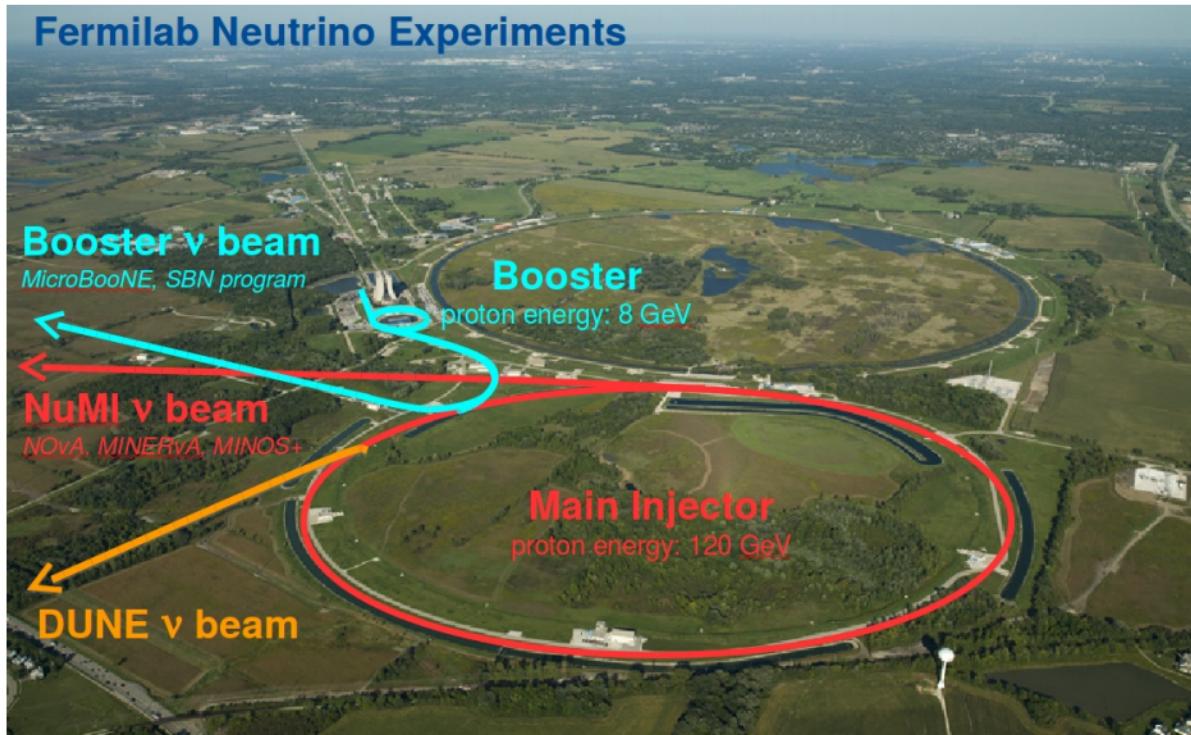


Figure 3.2: Arial view of the Main Injector and the Booster Neutrino Beam at Fermilab

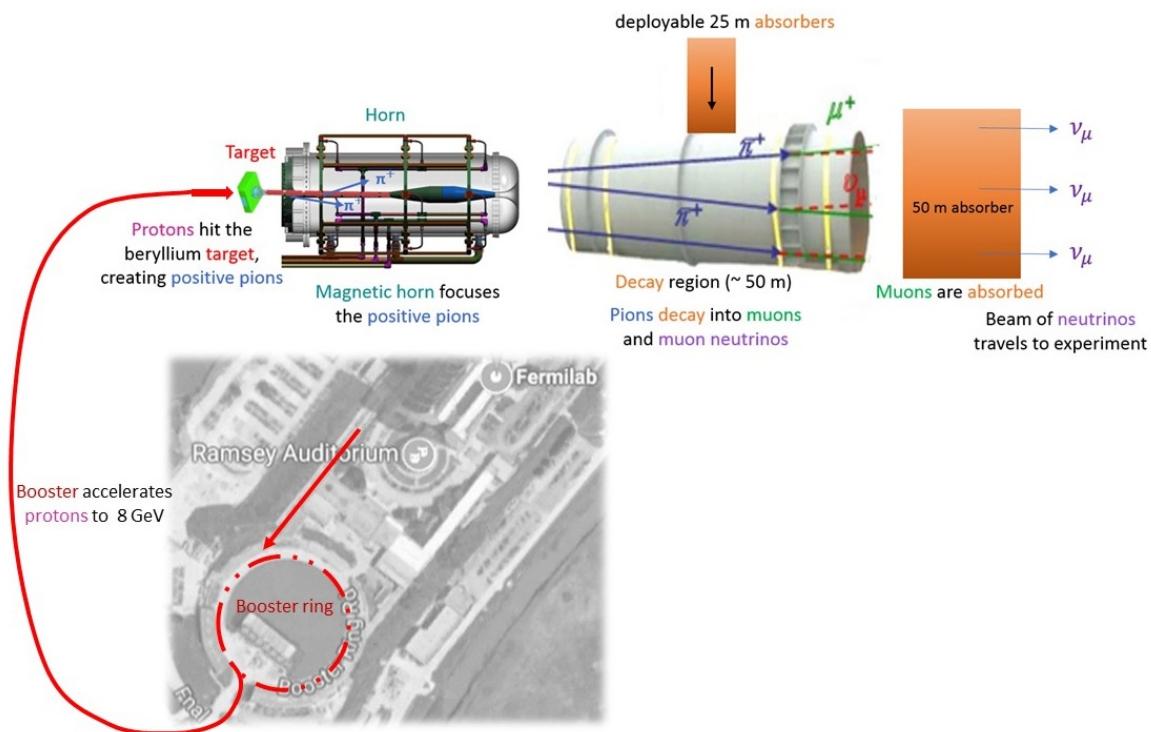
### <sup>742</sup> 3.4.1 Creating the Booster Neutrino Beam

<sup>743</sup> The BNB is a very pure  $\nu_\mu$  beam, with only 0.6% contamination from  $\nu_e$ s. The energy  
<sup>744</sup> also peaks around 700 MeV which is desired based on the probability of oscillation  
<sup>745</sup> equation which depends on the the value of  $L/E$ , where  $L$  is the distance of the  
<sup>746</sup> detector from the neutrino beam and  $E$  is the energy of the neutrino beam.  $L/E$  was  
<sup>747</sup> chosen to increase the probability of seeing neutrino oscillations in the MiniBooNE  
<sup>748</sup> Low Energy Excess (LEE) range based on the probability of oscillation equation, which  
<sup>749</sup> is  $P_{\nu_\mu \rightarrow \nu_e}(L, E) = \sin^2 2\theta \sin^2 \left( 1.27 \Delta m^2 \frac{L}{E} \right)$ . The BNB collides 8.9 GeV/c momentum  
<sup>750</sup> protons from the FNAL booster synchrotron into a beryllium target which produces a  
<sup>751</sup> high flux of neutrinos. The protons originate from  $H^2$  gas molecules that are turned  
<sup>752</sup> into  $H^-$  ions by a Cockcroft-Walton generator shown in figure ???. The  $H^-$  initially are  
<sup>753</sup> accelerated to 1MeV kinetic energy and are then passed to a linear accelerator using  
<sup>754</sup> alternating electromagnetic fields to increase their energy to 400MeV. The ions are  
<sup>755</sup> stripped of electrons by passing them through a carbon foil. The protons are bunched

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

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

777 **3.5 Event Reconstruction**



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

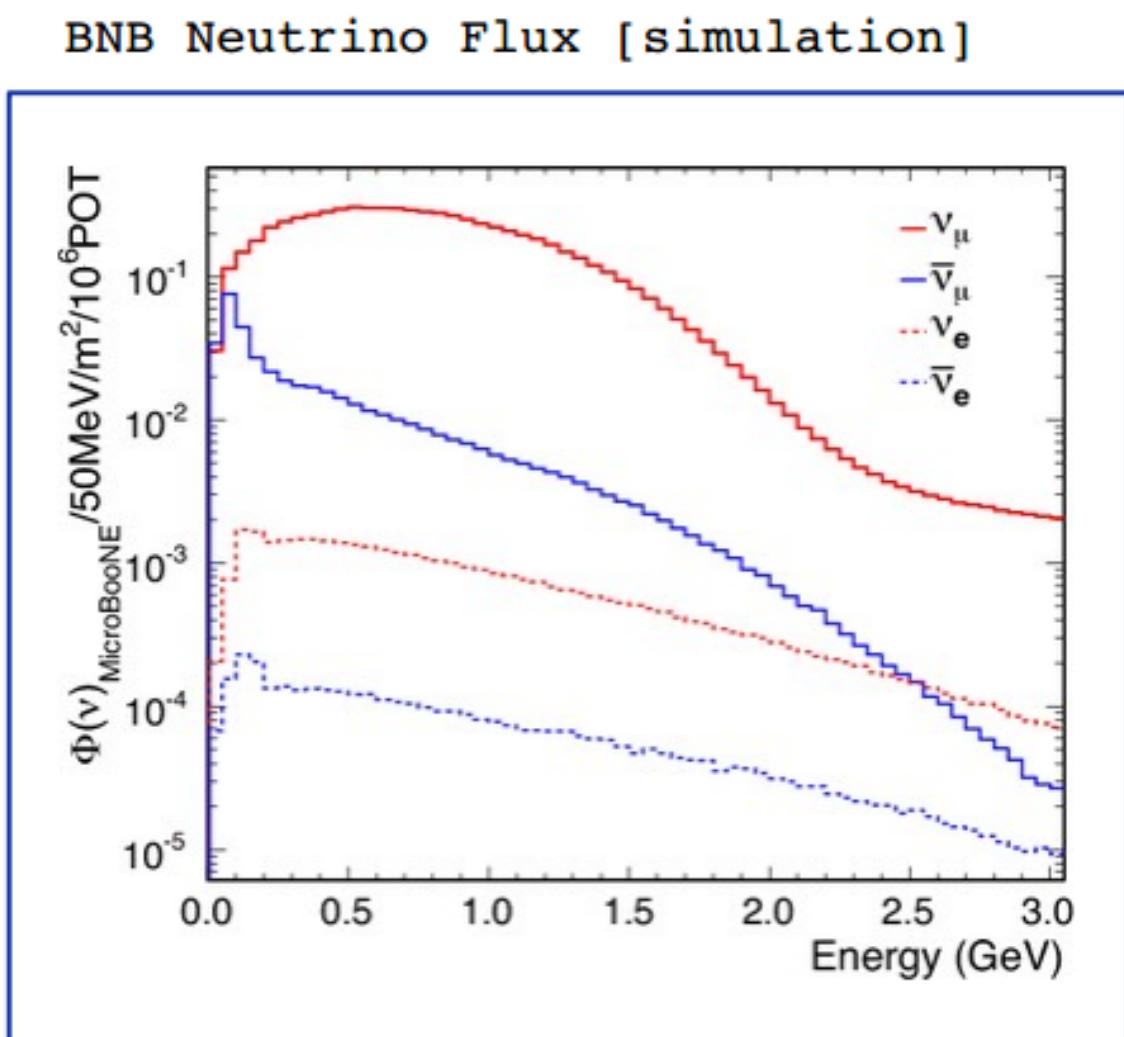


Figure 3.4: Energy spectrum of the Booster Neutrino Beam at Fermi National Laboratories

# <sup>778</sup> Chapter 4

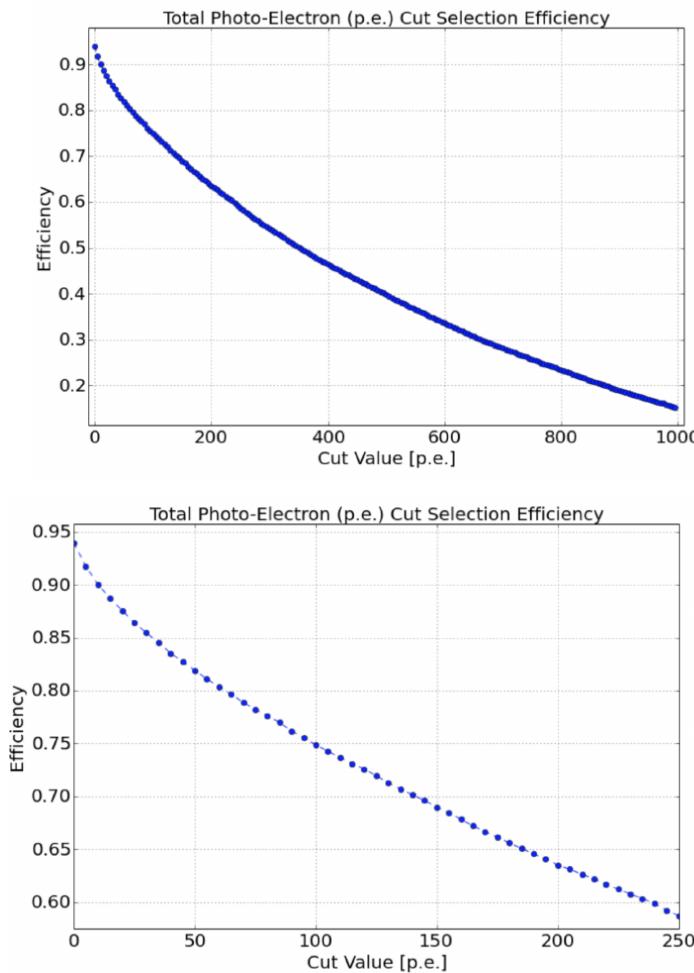
## <sup>779</sup> Neutrino Identification: Finding <sup>780</sup> MicroBooNE's first Neutrinos

<sup>781</sup> The goal of the Neutrino Identification analysis was to positively identify BNB neutrino  
<sup>782</sup> interactions in the MicroBooNE detector collected during the first days of running.  
<sup>783</sup> Neutrino event candidates were identified in part by using a cut on detected flash of  
<sup>784</sup> scintillation light during the  $1.6 \mu\text{s}$  beam-spill length of the BNB as well as identifying  
<sup>785</sup> reconstructed object from the TPC that are neutrino like. After this selection, 2D  
<sup>786</sup> and 3D event displays were used for verification of the selection performance. This  
<sup>787</sup> selection was targeted to reduce the ratio of neutrino events to cosmic-only events from  
<sup>788</sup> the initial 1 neutrino to 675 cosmics to a ratio of 1 to 0.5 or better which is equivalent to  
<sup>789</sup> a background reduction by a factor of 1000 or more. These selected events were used  
<sup>790</sup> for MicroBooNE's public displays of neutrino interactions. A clearly visible neutrino  
<sup>791</sup> interaction with an identifiable vertex and at least 2 tracks originating from the vertex  
<sup>792</sup> was what the analysis focused on. This analysis wasn't optimized for high purity  
<sup>793</sup> or efficiency, but rather for very distinguishable neutrino interactions that could be  
<sup>794</sup> identified by the public.

### <sup>795</sup> 4.1 Flash Finding

<sup>796</sup> Flash finding is the first step used in finding neutrino interactions. This section will  
<sup>797</sup> detail how optical information is reconstructed as well as analysis scripts and event  
<sup>798</sup> filters were used.

799 **4.1.1 Flash Reconstruction**

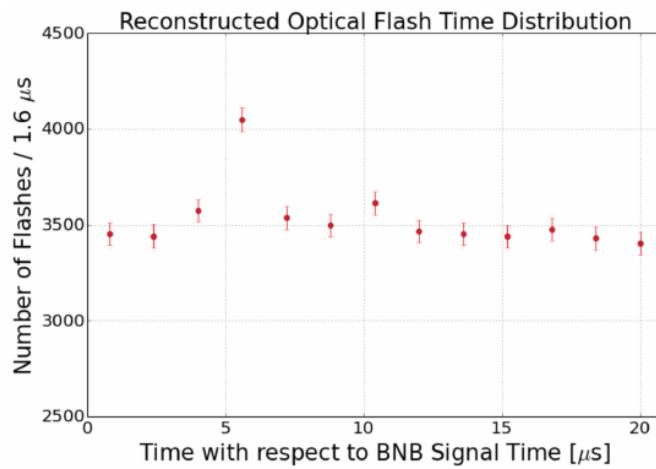


**Figure 4.1:** Efficiency for selecting beam events as a function of minimum total PE cut for all PE cuts as well as zoomed into interesting region.

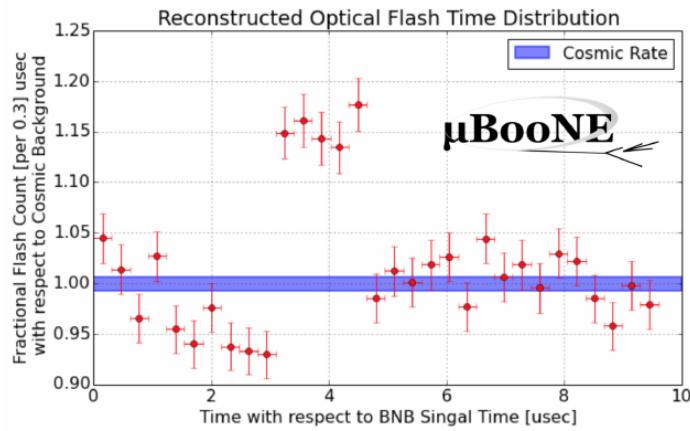
800 A flash is described as a collection of light seen at the same time within the detector.  
801 They are then reconstructed by identifying signal from the PMTs above a specific  
802 photoelectron (PE) threshold. These signals are called optical hits. Optical hits from  
803 all the PMTs are then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

<sup>810</sup> cut of 50 PE was deemed sufficient for this analysis. Figure 4.1 show the total PE  
<sup>811</sup> versus the selection efficency of selecting neutrino beam events.

### <sup>812</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>813</sup> It is necessary to get the specific time from flashes if one uses flashes to filter out  
<sup>814</sup> neutrino interactions coincident with the neutrino beam spill period and background.  
<sup>815</sup> Before a filter can be applied, an understanding of the timing of the trigger and PMT

816 readout with respect to the arrival of neutrinos from the BNB. To do this, a  $1.6 \mu\text{s}$   
 817 window near the expected beamtime was created and verified by finding that the  
 818 number of flashes was significantly above the cosmic-ray background flashes. Beam  
 819 data during the first week of running, October 16th 2016 through October 22nd 2016  
 820 and were used for a timing measurement. The total POT uses corresponds to roughly  
 821 24 hours of data taking at nominal intensity ( $4 \times 10^{12} \text{ ppp}$ ) and a 5 Hz repetition rate.  
 822 Figure 4.2a shows size of the expected neutrino signal in time using Monte Carlo  
 823 predictions and figure 4.2b shows the neutrino signal in data. The intensity in data is  
 824 lower, however there can still be seen a significant excess above data.

825 **4.1.3 Event Rates**

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

831 **4.2 TPC Topology Selection**

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

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

842 **4.2.1 Cosmic Tagging**

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

- 847 • 1: The track is tagged as entering or exiting the TPC
- 848 • 0.95: The track is a delta ray associated with a tagged track
- 849 • 0.5: The track is either entering or exiting, but not both
- 850 • 0.4: The track is entering or exiting through the Z boundary
- 851 • 0: The track isn't tagged

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

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

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

867 **4.2.2 2D Cluster Selection**

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

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

877 **Primary Cuts**

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

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

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

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

897 **Secondary Cuts**

898 The secondary cuts look to match long, low-angle clusters with short, high-charge  
899 clusters. Only clusters that have passed previous cuts are used. First clusters with  
900 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.

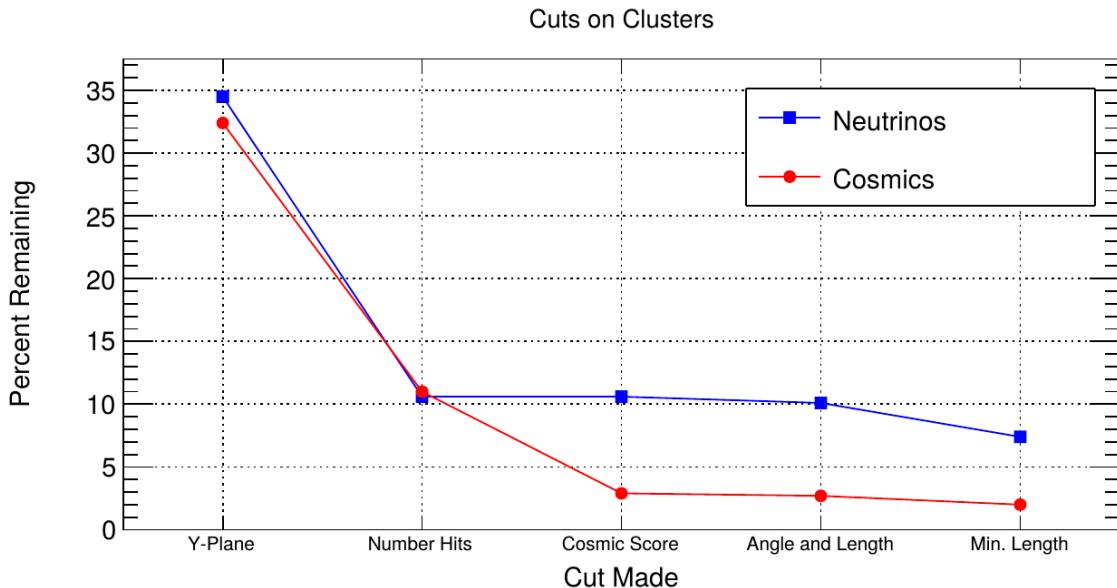
901 direction. Then we search for any cluster that is within approximately 3 cm ( 10 wires  
 902 and 30 time ticks) away from the low-z end of the long cluster. This cluster must also  
 903 be shorter than the first. In our reconstruction, the start and end point of a cluster can  
 904 be swapped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

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

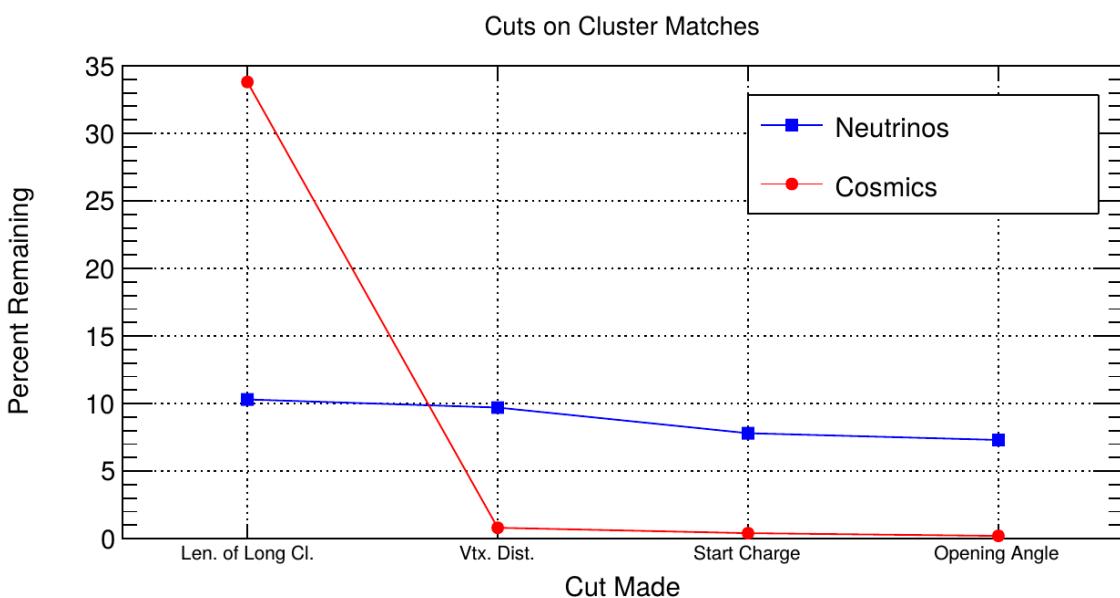
### 913 4.2.3 3D Tracks and vertices Selection

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

- 917 •  $d$ : distance between the start points of the two tracks.
- 918 •  $d_1$ : distance between vertex and start of track 1.
- 919 •  $d_2$ : 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.

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

- 927 • trackkalmanhit with cccluster  $\min(\max_{d,i}) < 3$  cm.  
928 • trackkalmanhit with pandoraNu  $\min(\max_{d,i}) < 4.5$  cm.  
929 • pandoraNu with cccluster  $\min(\max_{d,i}) < 5$  cm.

#### 930 4.2.4 TPC Updates

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

#### 937 2D Filter Updates

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

#### 943 3D Filter Updates

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

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

949 **4.3 Conclusion**

950 After proccesing these filters in parallel, it was shown that the 3D filter had a higher  
951 purity than the 2D filter because of the higher cosmic rejection being used due to 3D  
952 reconstruction. The 2D filter is blind to track entering/exiting from the top or bottom  
953 of the TPC. Although the 3D filter had a higher purity, the 2D filter was still able to  
954 find identifiable events in data that were used as public event displays. A sample of  
955 event displays are shown in figures ?? and ??.

956 **Chapter 5**

957 **CC-Inclusive Cross Section Selection**  
958 **Filter**

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

964 MicroBooNE requires fully automated event reconstruction and selection algorithms for use in the many physics measurements being worked on to date due to  
965 the large data rate MicroBooNE receives. Being able to automatically pluck out the  
966 neutrino interaction among a sea of cosmics proved to be challenging but was accomplished.  
967 MicroBooNE has developed two complementary and preliminary selection  
968 algorithms to select charged-current  $\nu_\mu - Ar$  interactions. Both are fully automated  
969 and cut based. The results of this thesis will focus on selection I and selection I modified  
970 and will focus on further improving these algorithms using Convolutional Neural  
971 Network (CNN) implementations. These selections identify the muon from a neutrino  
972 interaction without biasing towards track multiplicity. To combat cosmic and neutral  
973 current background, the analysis is strongly biased towards forward-going long tracks  
974 which are contained. This limits phase space and reduces acceptance.  
975

## 5.1 Data and MC Processing Chain

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

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

- larsoft v04\_36\_00
- GEANT v04\_09\_06\_p04d
- GENIE v02\_08\_06d
- GENIE xsec v02\_08\_06a
- pandora v02\_03\_0a
- CORSIKA v07\_4003

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

- MC fcl files
  - reco\_uboone\_mcc7\_driver\_stage1.fcl
  - reco\_uboone\_mcc7\_driver\_stage2.fcl
- Data fcl files
  - reco\_uboone\_data\_Feb2016\_driver\_stage1.fcl
  - reco\_uboone\_data\_Feb2016\_driver\_stage2.fcl

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

- 1004     • **Detector conditions:** the detector has to be in a good operating condition. The  
1005        detector conditions are read from the slow monitoring database and are required  
1006        to be within the alarm thresholds. The variables of interest for events passing  
1007        the good run list criteria include DAQ, PMT, HV, Drift HV, wire bias, electron  
1008        lifetime and detector power. These conditions need to be met on a run-by-run  
1009        basis in order to pass the selection.
- 1010     • **Data quality:** normal and stable behavior for basic reconstruction quantities.  
1011        These reconstruction variables include average number of tracks, hits, and flashes  
1012        in each event, the average length of tracks, the average amplitude and area of  
1013        hits, the average PE and the average spread of each one of these quantities.
- 1014     • **Beam Conditions:** the BNB must be on and stable and the POT per spill needs  
1015        to above the intensity threshold. Beam quality conditions include checking the  
1016        fraction of proton beam interacting within the target, the horn current, and the  
1017        intensity of protons per spill. The final sample is  $5 * 10^{19}$  and a per-spill intensity  
1018        of  $4 * 10^{12}$
- 1019     • **Run processed:** the full run must be processed completely without missing  
1020        subruns or crashes in the data processing.

1021        The selection begins with a cut that requires an optical flash greater than 50 photo  
1022        electrons (PE) in the  $1.6 \mu\text{s}$  beam window. Next, two or more 3D reconstructed tracks  
1023        must be within 5 cm from a 3D reconstructed vertex. The most forward going track  
1024        vertex-track association is then selected for further cuts. The vertex from the chosen  
1025        association must be in the fiducial volume, and the longest track from this association  
1026        must be matched to a flash 80 cm in z. Lastly the longest track must be contained and  
1027        longer than 75 cm.

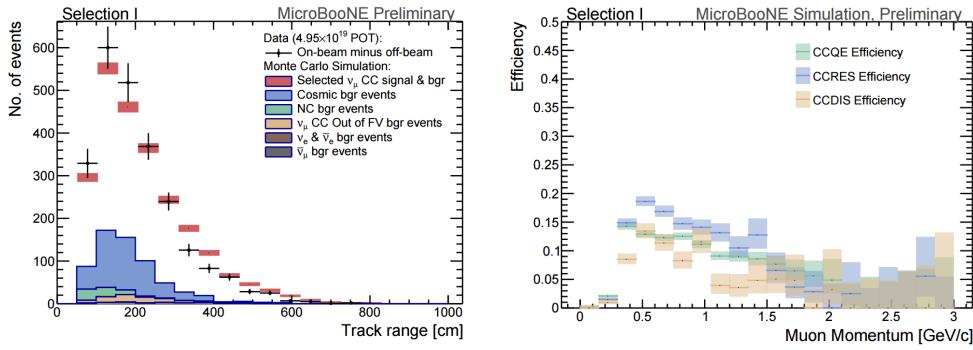
## 1028     5.2 Normalization of data and MC

1029        The off-beam sample is used to measure beam unrelated backgrounds. For normaliza-  
1030        tion, one needs the total number of BNB spills ( $N_{BNB}$ ) and the total number of external  
1031        triggers. The BNB spills used need to pass the beam quality cuts. The normalization  
1032        factor is then  $N_{BNB} / N_{EXT}$  which is 1.23.

1033        To normalize generated BNB MC events to POT, we used the following:

- 1034 •  $5 * 10^{19} POT = 41524.3$  generated events

1035 where this scaling factor only applies to mcc7 generated events. The inTime cosmic  
 1036 sample is normalized with respect to the open cosmic sample so an understanding  
 1037 of both is necessary. The POT per beam spill for mcc7 BNB samples is  $5 * 10^{12}$ . To  
 1038 calculate how many spills are necessary to produce a specific POT one would multiply  
 1039 the total POT by the average 1/POT per spill. For a total POT of  $5 * 10^{19}$  the amount  
 1040 of spills necessary is  $\frac{5*10^{19}}{5*10^{12}} = 1 * 10^7$ . This is only one in  $\sim 241$  events therefore each  
 1041 cosmic event needs to be scaled up by a factor of 240.8 when comparing to BNB  
 1042 MC. For inTime cosmics however, two filters are applied to reduce computing and  
 1043 processing time and only leave cosmics that will interact within the detector. The  
 1044 passing rate after these two filters is 0.02125, therefore the total inTime cosmic scaling  
 1045 factor to compare inTime cosmics to BNB is  $0.02125 * 240.8 = 5.12$ .



(a) Track range distribution of selection I

(b) Selection efficiency as a function of the true muon momentum

**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.

## 1046 5.3 Optical Software Trigger and Reconstruction

### 1047 5.3.1 Software Trigger

1048 Most of the BNB spills from the accelerator do not have a neutrino interaction in  
1049 MicroBooNE. To save computation resources and reduce data-rates, we require a  
1050 burst of light in the light collection system in coincidence with the 1.6  $\mu$ s beam spill.  
1051 Requiring light activity in coincidence with the beam spill eliminates the vast majority  
1052 of triggers with no neutrino interaction in the detector, however, it doesn't guarantee  
1053 the activity in the detector is a neutrino interaction since a cosmic ray can interact in  
1054 coincidence with the beam spill as well.

1055 To implement this, a software trigger was used on the PMT waveforms to decide  
1056 whether or not to keep that event. The software trigger is implemented after the event  
1057 builder combines data from the PMTs and triggers into a single event. The software  
1058 trigger uses the digitized output of the 32 PMT channels in the light collection system.  
1059 Only the waveform region in coincidence with the beam spill is used to search for  
1060 possible triggers. For each PMT, a waveform is found by taking the difference of ADC  
1061 values is calculated between  $t$  and  $t + s$ . This waveform is then scanned for ADC  
1062 values above a threshold  $X_0$ . Once an ADC is above this threshold, a discriminator  
1063 window is opened for a fixed number of time ticks ( $W_0$ ). If the ADC count within this  
1064 window  $W_0$  is greater than a second larger threshold  $X_3$ , a final window of width  $W_3$   
1065 is opened. The max ADC value within this final window is set as the peak amplitude  
1066 for the PMT and then summed across all 32 PMTs and set to the variable PHMAX. The  
1067 software trigger places a final cut on the PHMAX variable to decide whether or not  
1068 to keep the event. The thresholds were found by the Trigger task force using Monte  
1069 Carlo Studies and are as follows:

- 1070 •  $X_0 = 5$  ADC
- 1071 •  $X_3 = 10$  ADC
- 1072 •  $W_0 = 6$  Ticks
- 1073 •  $W_3 = 6$  Ticks
- 1074 • PHMAX cut = 130 ADC

**1075 5.3.2 Flash Reconstruction**

1076 MicroBooNE collects light from each of the 32 PMTs either in a continuous readout  
 1077 window of  $23.4 \mu\text{s}$  activated by a beam gate signal on the trigger board, or in discrimi-  
 1078 nated pulses of  $\sim 1 \mu\text{s}$  duration activated if the ADC count for any PMT goes above 80  
 1079 ADC count. These two formats are saved as output waveforms and put onto an event.  
 1080 Additionally, each PMT can provide two output streams, high-gain ( $\sim 20 \text{ ADC/PE}$ )  
 1081 and low-gain ( $\sim 2 \text{ ADC/PE}$ ) channels. The first step in the reconstruction is to merge  
 1082 both these channels into a “saturation corrected waveform” which uses information  
 1083 from the low-gain waveform to correct for saturating high-gain pulses.

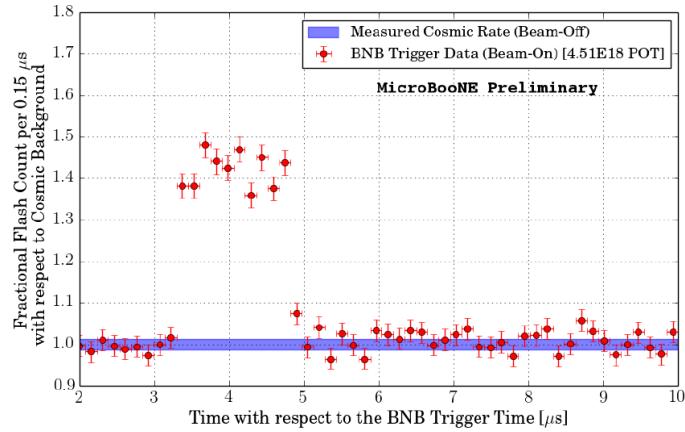
1084 The saturation corrected waveform in the continuous readout window is used to  
 1085 reconstruct optical hits. Each PMT’s waveform is scanned for hits then a threshold  
 1086 based hit reconstruction algorithm is applied which requires pulses of a minimum  
 1087 area in order to be reconstructed. Each reconstructed hit is associated to a PMT, a time  
 1088 in  $\mu\text{s}$ , and a PE count.

1089 Once hits are reconstructed for all 32 PMTs, all PMT information is then combined  
 1090 into optical flashes which represent optical information seen by the PMTs from interac-  
 1091 tions in the detector. Each flash has information on total light seen per interaction, the  
 1092 distribution of the light across all 32 PMTs, the flash time with respect to the trigger  
 1093 time of the flash, and lastly, the spacial information of the flash in Y-Z plane of the  
 1094 detector. These flashes are reconstructed by requiring that there is a  $\sim 1 \mu\text{s}$  coincidence  
 1095 between the reconstructed hits in all 32 PMTs. The total PE is summed up among  
 1096 all coincident hits across the PMTs and if the total PE is greater than 2 PE, a flash is  
 1097 reconstructed. There are also safe guards in place to take care of late scintillation light.

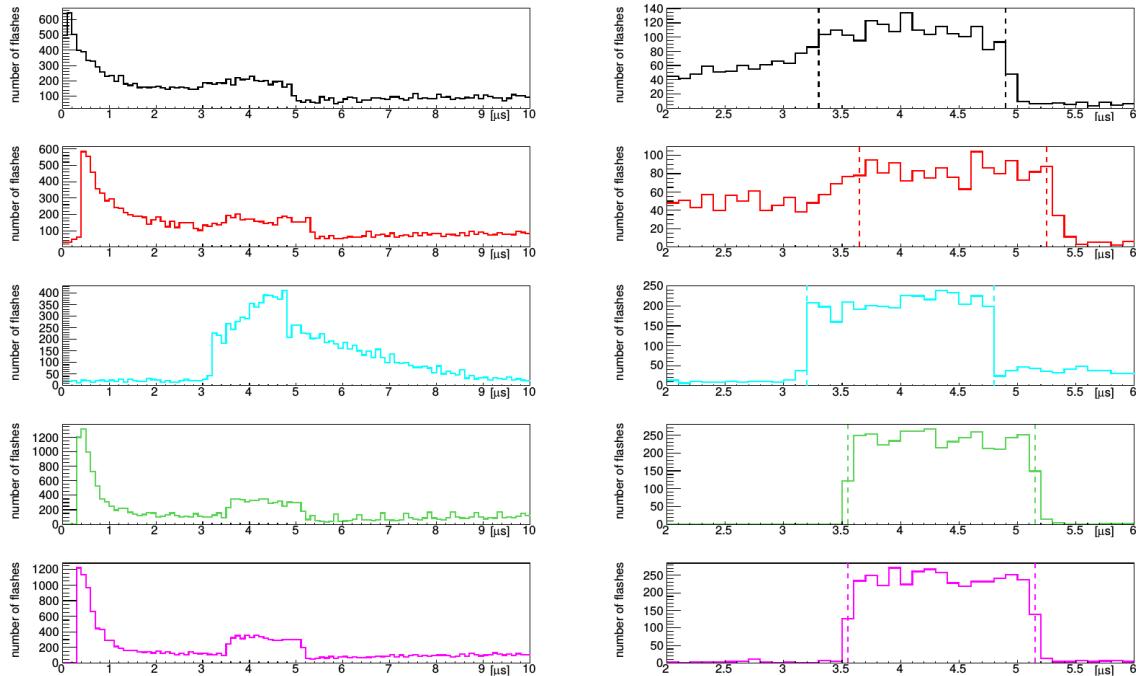
1098 Figure 5.2 shows the time distribution of reconstructed optical flashes using the  
 1099 BNB continuous stream. You can see a clear excess in coincidence with the expected  
 1100 arrival time of neutrinos. The same flash reconstruction that was used in the cc-  
 1101 inclusive filter detailed here was used to create this plot in data.

**1102 5.3.3 Beam Window**

1103 Figure 5.3 shows the distribution of flashes for on-beam, off-beam and various MC  
 1104 samples. The software trigger has been applied to these samples. The pile-up seen just  
 1105 after  $0 \mu\text{s}$  is a feature of the flash finding algorithm and consists of low PE flashes and



**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.

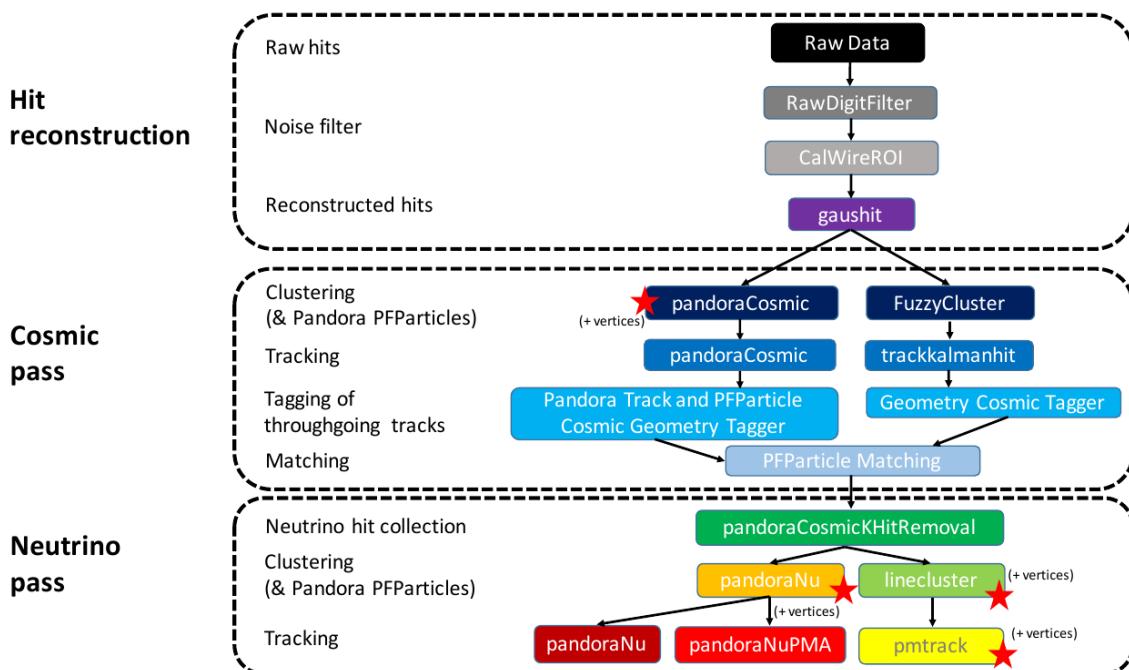


**Figure 5.3:** Flash time distribution for all flashes (left plot) and flashes  $> 20\text{PE}$  (right plot). The different curves are as follows: on-beam data (black), off-beam data (red), CORSIKA inTime MC (light blue), BNB only MC (green), and BNB+Cosmic MC (purple). The dashed vertical lines mark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

1106 is removed in the second column of distributions with a low 20 PE threshold cut. The  
1107 plots show that the time window for the distributions are shifted a small amount from  
1108 each-other. This is caused by different hardware configurations per sample. Using  
1109 these distributions, the windows chosen per sample are as follows:

- 1110 • On-Beam: 3.3 to 4.9  $\mu$ s
  - 1111 • Off-Beam: 3.65 to 5.25  $\mu$ s
  - 1112 • CORSIKA inTime: 3.2 to 4.8  $\mu$ s
  - 1113 • BNB only: 3.55 to 5.15  $\mu$ s
  - 1114 • BNB+Cosmic: 3.55 to 5.15  $\mu$ s
- 1115 Each window has a width of 1.6  $\mu$ s.

## 1116 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.

1117     Figure 5.4 summarizes the reconstruction chain applied to both MC and data for  
1118    this analysis. After the hit reconstruction, a cosmic pass is applied which removes all  
1119    hits associated to through-going tracks. A description of these TPC reconstruction  
1120    algorithms will be detailed below.

#### 1121    **5.4.1 Hit Reconstruction**

1122    The waveforms used for hit reconstruction consist of charge deposited on the sense  
1123    wire in drift time. The first step in hit reconstruction is to pass the waveforms through  
1124    a filtering algorithm to filter out the noise introduced from the electronics. The input  
1125    waveforms are also truncated from 9600 time ticks to 6400 time ticks in this first step  
1126    to reduce the data footprint of these waveforms.

1127    Once noise filtering is complete, a deconvolution algorithm is applied to the wave-  
1128    forms to remove the drift field and electronics response, therefore leaving only the  
1129    ionized electrons kicked off the argon atoms by an incident track. During this process,  
1130    Region of Interests (ROI) are identified and cut out of the waveforms to further reduce  
1131    the data volume.

1132    The hit finding algorithm then finds candidate peaks in these ROI's and fits the  
1133    peaks to Gaussian curves. These Gaussian shaped peaks are now called hits and  
1134    represent the charge deposition on a wire by the incoming track. These hit objects  
1135    have a peak time and width and are the basic object input to further algorithms down  
1136    the reconstruction chain.

#### 1137    **5.4.2 Clustering**

1138    There are multiple clustering algorithms used in this analysis. The main purpose of all  
1139    the clustering algorithms is to associate hits together in 2D space to create objects like  
1140    tracks, vertices and showers. For the fuzzy cluster algorithm, three steps are used to  
1141    achieve this. The first step is to associate hits to each-other using a fuzzy clustering  
1142    algorithm which gives each hit a degree of belonging to the cluster. Second, a Hough  
1143    transform is used to find hits associated to candidate tracks and showers within each  
1144    of the clusters found in the first step. The last step merges smaller candidate tracks  
1145    and showers into large clusters. The last step also associates unclustered hits into

<sub>1146</sub> nearby objects which helps shower reconstruction. The result is a set of clusters made  
<sub>1147</sub> up of associate hits that represent tracks or showers per plane.

<sub>1148</sub> The pandora algorithm utilizes it's own clustering algorithm and will be detailed  
<sub>1149</sub> in the next section. The last clustering algorithm is called linecluster. The linecluster  
<sub>1150</sub> algorithm reconstructs 2D linear clusters per plane by fitting a line onto nearby hits  
<sub>1151</sub> which is then extrapolated to neighboring wires. 2D vertices are found per plane by  
<sub>1152</sub> using the intersection points of the ends of nearby clusters. These 2D vertices are then  
<sub>1153</sub> matched in time across all three planes to get a 3D vertex in space.

<sub>1154</sub> **5.4.3 Pandora**

<sub>1155</sub> **5.4.4 Trackkalmanhit**

<sub>1156</sub> The trackkalmanhit algorithm takes 2D clusters returned from the fuzzy cluster algo-  
<sub>1157</sub> rithm and outputs track objects. There are no hierarchy structure as it is in pandora,  
<sub>1158</sub> each track is independent. There also is no vertex reconstruction with this algorithm  
<sub>1159</sub> as well.

<sub>1160</sub> **5.4.5 Cosmic Hit Removal**

<sub>1161</sub> **5.4.6 Projection Matching Algorithm**

<sub>1162</sub> **5.4.7 Calorimetry**

<sub>1163</sub> **5.5 Event Selection**

<sub>1164</sub> **Chapter 6**

<sub>1165</sub> **Background on Convolutional Neural  
Networks**

<sub>1167</sub> Convolutional neural networks (CNNs) have been one of the most influential inno-  
<sub>1168</sub> vations in the field of computer vision. Neural networks became popular in 2012  
<sub>1169</sub> when Alex Krizhevsky used them to win that year's ImageNet competition [?] by  
<sub>1170</sub> dropping the error from 26% to 15%. Since then, many companies are using deep  
<sub>1171</sub> learning including Facebook's tagging algorithms, Google for their photo search and  
<sub>1172</sub> Amazon for product recommendations. For the purpose of this thesis CNNs were  
<sub>1173</sub> used for image classification, specifically, images of varying particles created using  
<sub>1174</sub> LArTPC data.

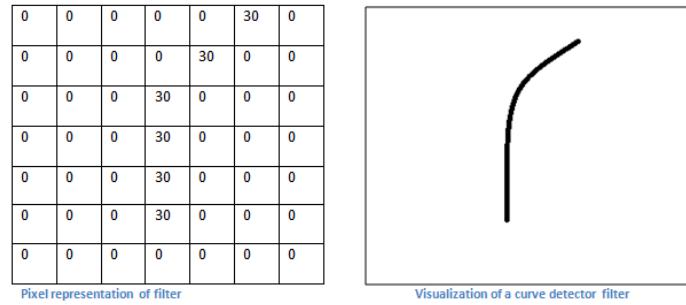
<sub>1175</sub> **6.1 Image Classification**

<sub>1176</sub> Image classification is the process of inputting an image into the CNN and receiving a  
<sub>1177</sub> probability of classes that best describes what is happening in the image. As humans,  
<sub>1178</sub> image classification is something that is learned at a very young age and is easy to  
<sub>1179</sub> do without much effort. This is also apparent when hand-scanning LArTPC images.  
<sub>1180</sub> After learning what a neutrino event looks like in MicroBooNE, it is relatively easy  
<sub>1181</sub> to recognize simple neutrino events from cosmic ray background as well as highly  
<sub>1182</sub> ionizing particles like protons from minimum ionizing particles like muons. The very  
<sub>1183</sub> detailed images LArTPC detectors output are prime candidates for input images into  
<sub>1184</sub> a CNN. CNNs mimic a human's ability to classify objects by creating an architecture  
<sub>1185</sub> that can learn differences between all the images it's given as well as figure out the

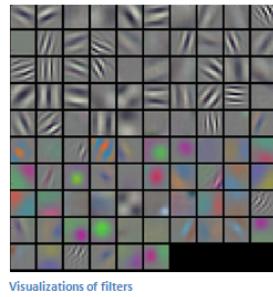
unique features that make up each object. CNNs are modeled after the visual cortex. Hubel and Wiesel [?] found that there are small regions of neuronal cells in the brain that respond to specific regions of the visual field. They saw that some neurons fired when exposed to vertical edges while others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

## 6.2 CNN Structure

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

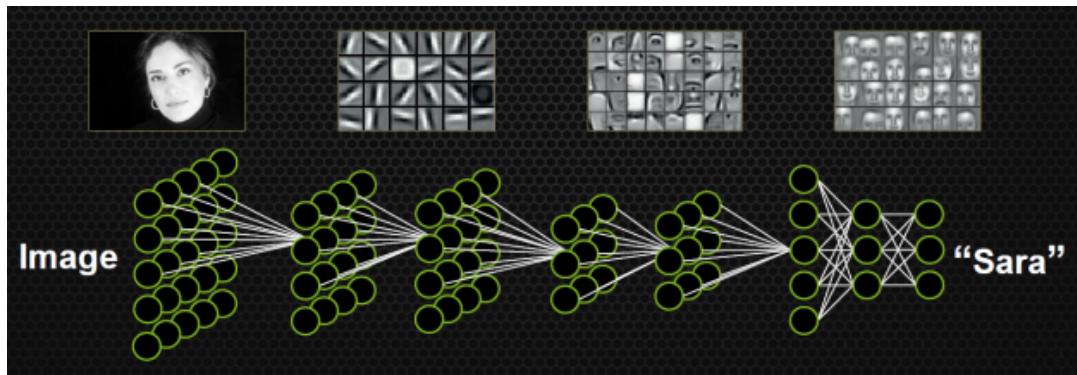


**Figure 6.2:** Visualization of filters found in first layer of a CNN.

1219 In figure 6.3 you can see how an edge detection filter is used to save only necessary  
 1220 information for recognizing different types of clothes. You can also see by having  
 1221 multiple filters you can get more detail or less detail from an image which can then  
 1222 simplify or complicate the object recognition task. Being able to distinguish between a  
 1223 shirt or a leg garment is as much information you want, having a filter that extracts  
 1224 outline edge or shape information would be all that you need. But if instead you  
 1225 wanted to distinguish between a formal cocktail dress or a summer dress, more  
 1226 information would need to be saved equating to many more filters for one image.  
 1227 Rather than trying to come up with how many filters and what features are important  
 1228 for detection, CNNs do this automatically. CNNs take input parameters, called  
 1229 hyperparameters, for example number of layers, number of filters per layers, number  
 1230 of weights per filter, and uses these to create the output feature maps. The layers build  
 1231 upon each-other, for example if we were creating a CNN for facial recognition the  
 1232 convolutional layers will start learning feature combinations off of the previous layers.  
 1233 The low level features like edges, gradients, and corners of the first layers become high  
 1234 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

1235 There are other layers in a CNN architecture that will not be covered in the scope  
 1236 of this thesis but in a general sense, these layers are interspersed between convolution  
 1237 layers to preserve dimensionality and control overfitting of the network. The last layer  
 1238 is called a fully connected layer and it's job is to output an  $N$  dimensional vector where  
 1239  $N$  is the number of classes the network has been trained on. Each number in this vector  
 1240 represents the probability that the input image is a certain class. Fully connected layers  
 1241 use the feature maps of the high level features to compute the products between the  
 1242 weights of the previous layer to get the probabilities of each class. These weights are  
 1243 then adjusted through the training process using backpropagation.

---

### 1244 6.2.1 Backpropagation

1245 A CNN at it's onset has weights that are randomized. The filters themselves don't  
 1246 know how to pull out identifying information per class. For a neural network to learn,  
 1247 it must be trained on a training set that is labeled. Backpropagation has four seperate  
 1248 steps: foward pass, loss function, backward pass and updating weights. In the forward  
 1249 pass, a training image is passed through the whole network. All of our weights at this  
 1250 time are randomly initialized so the output for the first image will have no preference  
 1251 to a specific class. A common loss function is mean squared error (MSE):

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

1252 If we assume that the MSE is the loss of our CNN, the goal would be that our  
 1253 predicted label (output of CNN) is the same as our training label. To do this, we need  
 1254 to minimize the loss function. To do this, it is necessary to find out which weights most  
 1255 directly affect the loss of the network i.e  $\frac{dL}{dW}$  where L is our loss function and W are  
 1256 the weights of a specific layer. The next step is the backward pass which determines  
 1257 which weights contribute the most to the loss and finds ways to adjust these weights  
 1258 so that the loss decreases. After the derivative is computed, the last step updates the  
 1259 weights in the opposite direction of the gradient.

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

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

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

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

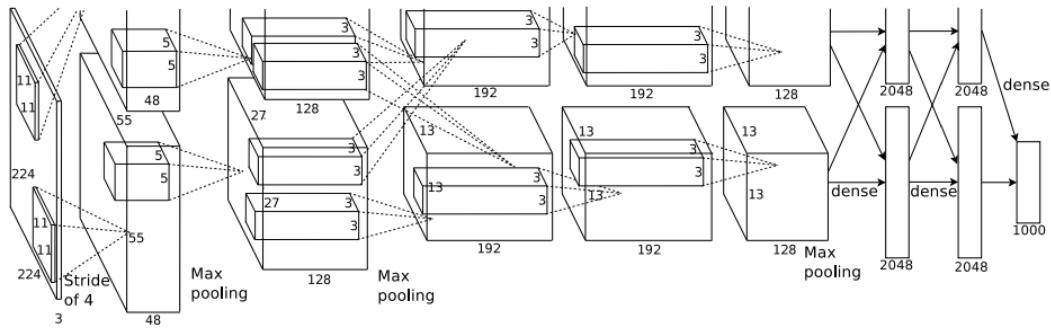
1260 The learning rate is a parameter given to the CNN and it describes the steps the  
 1261 network takes to update the weights. Higher learning rate equals large steps and a  
 1262 lower training time, but a learning rate that is too large can mean the CNN never  
 1263 converges.

1264 Going through backpropagation consists of one training iteration. Once the net-  
 1265 work completes a specific number of iterations, another parameter given, and runs  
 1266 over all training images that are split up into batches, the process is considered com-  
 1267 plete. User input parameters, called hyperparameters, help the network converge to

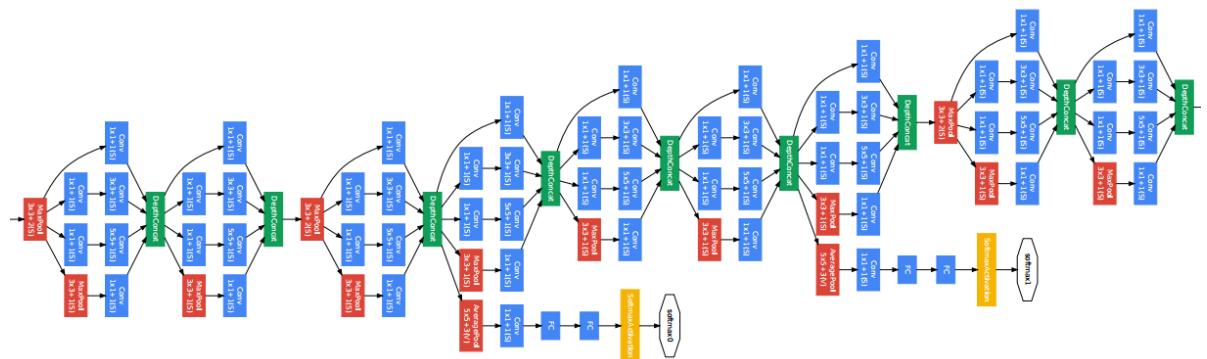
1268 optimal weights for each layer. Batch size, learning rate, and training iteration are just  
1269 some of the user input hyperparameters that help. Lastly, to check if the network has  
1270 learned, a different set of labeled images are fed to the CNN iteratively through the  
1271 training process to see how well it's learning. This process is especially important to  
1272 make sure the network architecture isn't being affected by overfitting (memorizing  
1273 training input rather than learning).

## 1274 6.3 Choosing Hyperparameters

1275 Convolutional neural networks are a relatively new tools in computer vision. Choosing  
1276 hyperparameters for your specific dataset is a non-trivial task. Hyperparameters can  
1277 range from the amount of layers and filters per layer in an CNN architecture to the  
1278 stride the receptive field of a filter takes, not to mention training hyperparameters  
1279 such as learning rate and batch size described above. They're ways to optimize these  
1280 hyperparameters via hyperparameter optimization using Bayesian Optimization [?]  
1281 but as you can imagine, optimizing an CNN architecture from scratch can be very  
1282 computationally intensive. For the purpose of this thesis, two well known CNN  
1283 architectures were used, AlexNet [?], which won the ImageNet Large-Scale Visual  
1284 Recognition Challenge (ILSVRC) in 2012 and therefore bringing awareness of CNNs,  
1285 and GoogleNet [?], which won the ILSVRC in 2014, giving rise to deep networks. Both  
1286 AlexNet and GoogleNet architectures were used to train on LArTPC images and their  
1287 low level filter weights. Higher level filter weights were randomly initialized before  
1288 training so the network can learn high level features of LArTPC image classes. The  
1289 AlexNet architecture is shown in figure 6.5 and the GoogleNet architecture is shown  
1290 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.



# 1291 Chapter 7

## 1292 Training Convolutional Neural 1293 Networks on particles **WORKING** 1294 **TITLE**

1295 Three Convolutional Neural Networks CNNs were trained throughout this analysis.  
1296 There are differences to each CNN and will be described fully in the next sections but  
1297 the main difference are the amount of particle images used for training and validation.  
1298 CNN1075 used 1,075 muons and 10,75 pions for training and the same amount of each  
1299 particle for validation. CNN10000 used 10,000 muons and 10,000 pions split in half  
1300 for testing and training. Lastly CNN100000 had muons, pions, protons, electrons,  
1301 and gammas in its training and validation set. Each particle had 20,000 images and  
1302 training and validation was split 90% training, 10% validation. This chapter will also  
1303 describe the different hardware frameworks used for training beginning on a CPU  
1304 and ending on a GPU cluster.

---

1305 **7.1 Hardware Frameworks used for Training**

1306 **7.1.1 Syracuse CPU Machine setup**

1307 **7.1.2 Syracuse University GPU Cluster Setup**

1308 **7.2 Convolutional Neural Network Training**

1309 **7.2.1 Image Making Scheme**

1310 **Images used for Traing/Validation of Convolutional Neural Networks**

1311 **add image making for CNN1075** The  $\mu/\pi$  image dataset used to train and validate  
1312 the CNN10000 was created using single generated isotropic muons and pions from  
1313 0-2 GeV energ range. 10,000 muons and 10,000 pions were used for training and  
1314 testing split 50%. The images were created based on wire number and time tick in the  
1315 collection plane. Uboonecode v06\_23\_00 was used instead of v05\_08\_00 which was  
1316 used previously. The wire signal was the raw ADC value after noise filtering. Each  
1317 collection plane grayscale image was 3456x1280x1 where 5 time ticks were pooled into  
1318 1 bin which is different than the previous dataset and was implemented due to the fact  
1319 that the time ticks of an event went from 9400 to 6400 with the change of uboonecode  
1320 version. The grayscale color standard is 8bit therefore the ADC value of wire and time  
1321 tick was also downsampled due to the 12bit ADC value MicroBooNE has. To do this,  
1322 the highest ADC pixel in the image was found and then this was divided by the rest  
1323 placing all pixel values between 0-1. From there, all pixel values are then multiplied  
1324 by 255. All images were made using a LArSoft module. Once the images were created,  
1325 using and image manipulation framework called OpenCV images were read into a  
1326 numpy array and cropped to the region of interest by only keeping rows and columns  
1327 where all ADC values are higher than 0 and then resized it to 224x224 using OpenCV's  
1328 resize function. This downsampling of ADC values creates a problem of information  
1329 loss for example, a proton which is highly ionizing will have the same brightness as a  
1330 minimum ionizing muon by virtue of how the images are created. Issues that arose  
1331 in CNN1075 that were fixed in CNN10000 include zero-padding images in X and Y  
1332 that are smaller than 224X224 to eliminate over-zooming effect and fixing a bug that  
1333 shifted pixels separated by a dead-wire region.

1334 Images were also made from events that passed the cc-inclusive selection 1 filter  
1335 right before the 75 cm track length cut and were classified using the CNN10000. The  
1336 dataset used to create these images is the same one used in [?], prodgenie\_bnb\_nu\_cosmic\_uboone\_mc  
1337 These images were created using information from the track candidate that passed  
1338 the filter. Only wire number and time ticks associated to the track candidate were  
1339 drawn on the image to mimic a single particle generated image. These images were  
1340 then classified using CNN10000. Two approaches were taken in making these images.  
1341 The first was using the image normalization above where the maximum pixel in each  
1342 image is used as a normalization constant to get all pixels between 0-1 then multiply  
1343 all pixels by 255. As described above, this is the incorrect way to normalize; it should  
1344 be normalized by dataset not by event, which is the second way the images were  
1345 created. The results of CNN10000 performance are shown in section 7.2.

### 1346 7.2.2 Training CNN1075

1347 The work shown in these next sections are based on the previous work done described  
1348 in [?]. That CNN (now referred to as CNN1075) was trained using single generated  
1349 isotropic muons and pions from 0-2 GeV energy range. 1,075 muons and pions were  
1350 used to train the network and 1,075  $\mu/\pi$  were used as a validation set. The accuracy is  
1351 how well CNN1075 is doing by epoch and was 74.5%. The loss is gradient descent  
1352 or minimization of the error of the weights and biases used in each neuron of each  
1353 layer of CNN1075 and was 58% with a trend sloping downwards on the loss curve  
1354 as well as a trend sloping upward in the accuracy curve. Due to the depth of the  
1355 neural network framework, it was necessary to train with a larger dataset and for  
1356 more epochs, however, the downward slope of the loss curve is an indication that once  
1357 trained for longer with a higher training sample, neural networks can be used for  $\mu/\pi$   
1358 separation. Updates in the image making and downsampling algorithm were made to  
1359 fix issues that arose in CNN1075.

### 1360 7.2.3 Training CNN10000

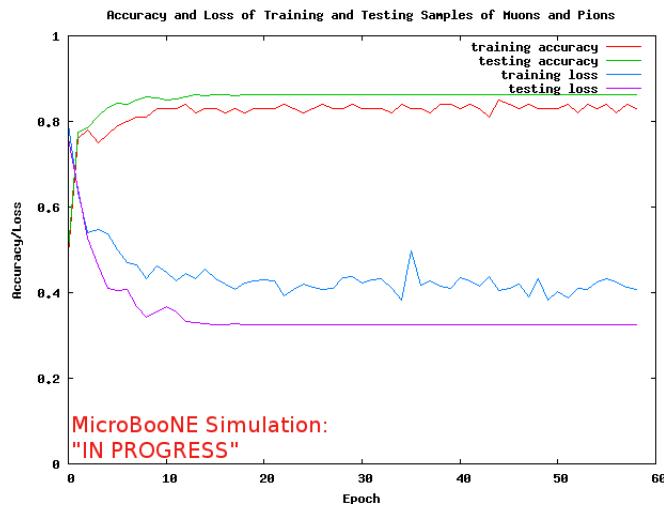
1361 The hyperparameters used for CNN10000 are shown. The batch size for the training  
1362 and testing as well as the test iter were chosen to encompass the whole training/testing  
1363 image set when doing accuracy/loss calculations. To do this, multiplying the test

1364 iter by the test batch size give you the amount of images used when calculating  
1365 accuracy/loss curves. For reference, the accuracy and loss are defined as well.

```
1366     • train_batch_size: 100  
1367     • test_batch_size: 100  
1368     • test_iter: 100  
1369     • test_interval: 100  
1370     • base_lr: 0.001  
1371     • lr_policy: "step"  
1372     • gamma: 0.1  
1373     • stepsize: 1000  
1374     • display: 100  
1375     • max_iter: 10000  
1376     • momentum: 0.99  
1377     • weight_decay: 0.0005  
1378     • snapshot: 100  
1379     • Accuracy: How often the CNN predicts the truth over total number of images  
1380     • Loss: Error between truth and prediction. Minimize loss by gradient descent to  
1381         update weights and biases of CNN
```

1382 The same architecure that was used to train CNN1075 was employed on CNN10000,  
1383 Imagenet. Caffe [?] was the software package used for both CNNs. The differences  
1384 include batch size and test\_iter and momentum to account for the larger dataset. Both  
1385 CNNs were trained on a CPU machine, Syracuse01. Further training will be done  
1386 on a GPU cluster stationed at Syracuse University. Figure 7.1 shows the loss and  
1387 accuracy of CNN10000. There is around a 10% increase in accuracy from CNN1075 to  
1388 CNN10000, 85%, and around a 20% decrease in loss, 36%.

1389 Figure 7.2 show a breakdown of  $\mu/\pi$  separation for CNN10000. It also shows  
1390 the network is not being overtrained due to the Accuracy of both the training and  
1391 testing datasets being within .01% of eachother. The CNN is doing a very good job of

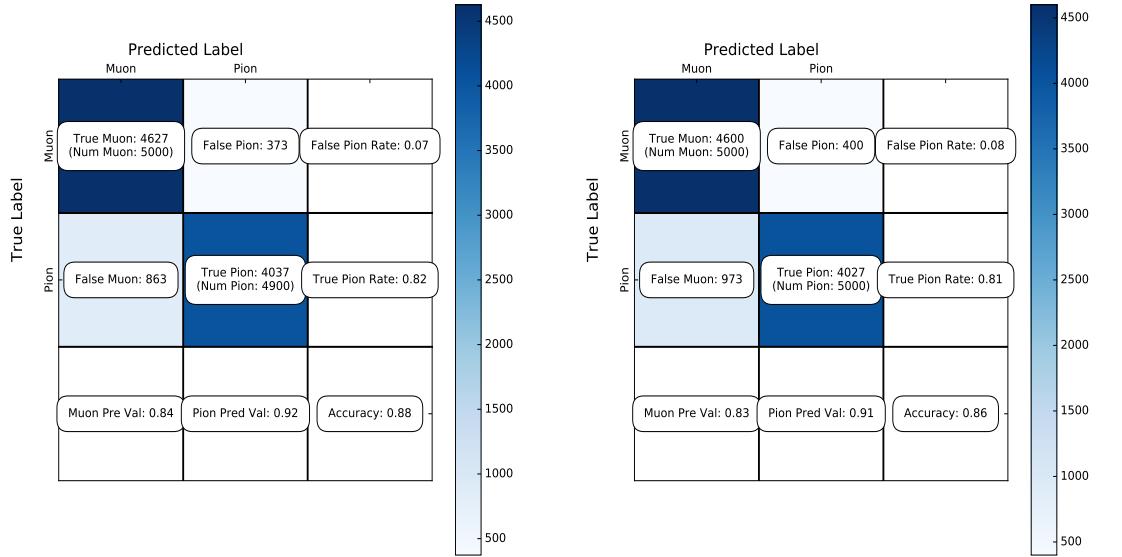


**Figure 7.1:** Accuracy vs. Loss of ImageNet 2-output  $\mu/\pi$  sample consisting of 10000 images each.

1392 classifying true muons as muons, and our loss increase from CNN1075 is due to the  
1393 increase in accurately classifying pions as pions.

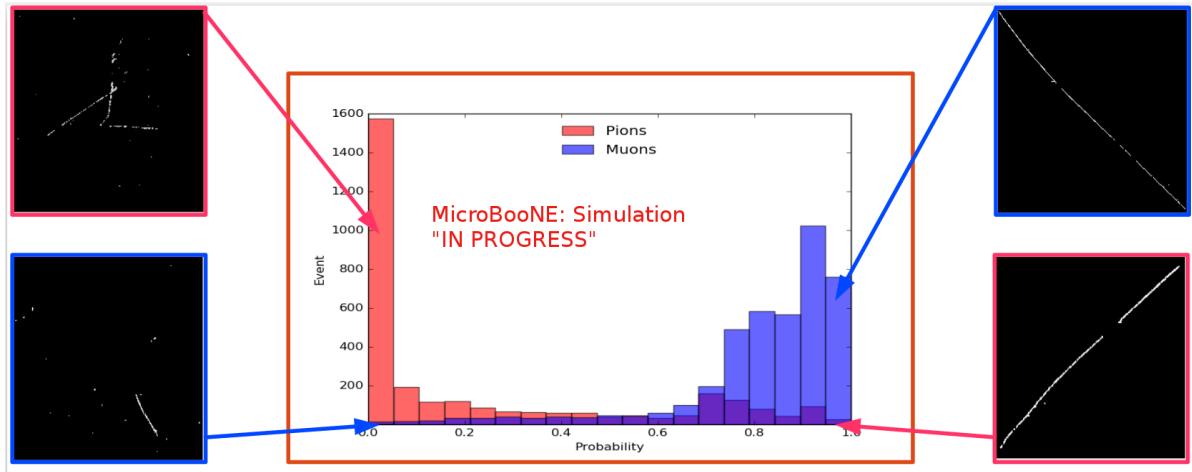
#### 1394 7.2.4 Training CNN100000

1395 Results of training using 100,000 images, 20,000 images per  $\mu/\pi/p/\gamma/e$ .



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

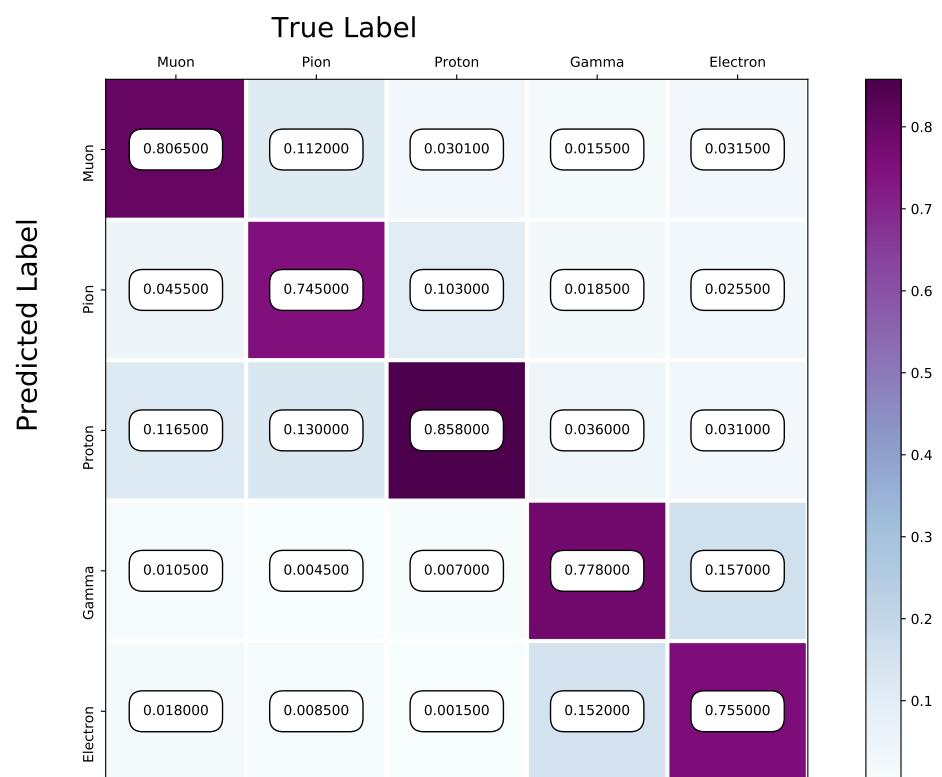
**(b)** Confusion Matrix showing Accuracy of CNN using testing data



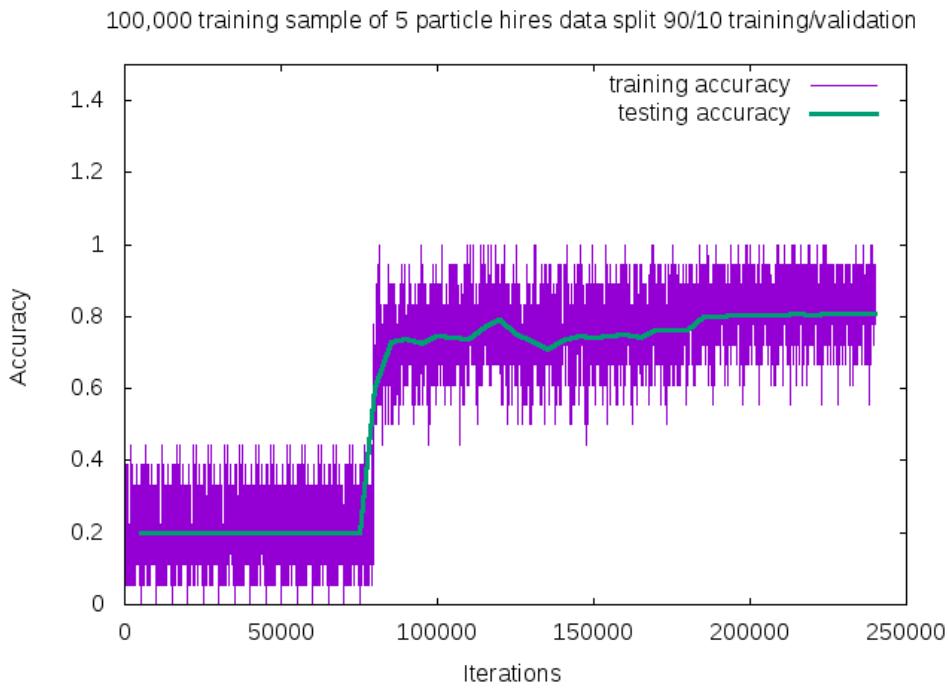
**(c)** Probability plot of muons and pions from testing set. Images surrounding histogram are a random event from lowest bin and highest bin for each particle.

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

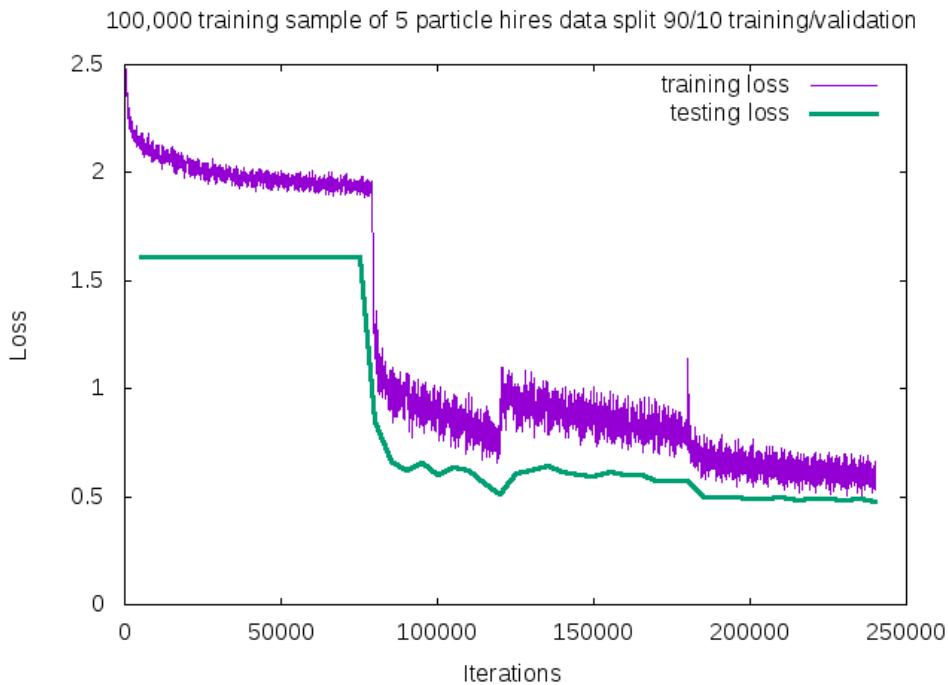
**7.2c** The probability plot includes muons and pions that are classified as primary particles.



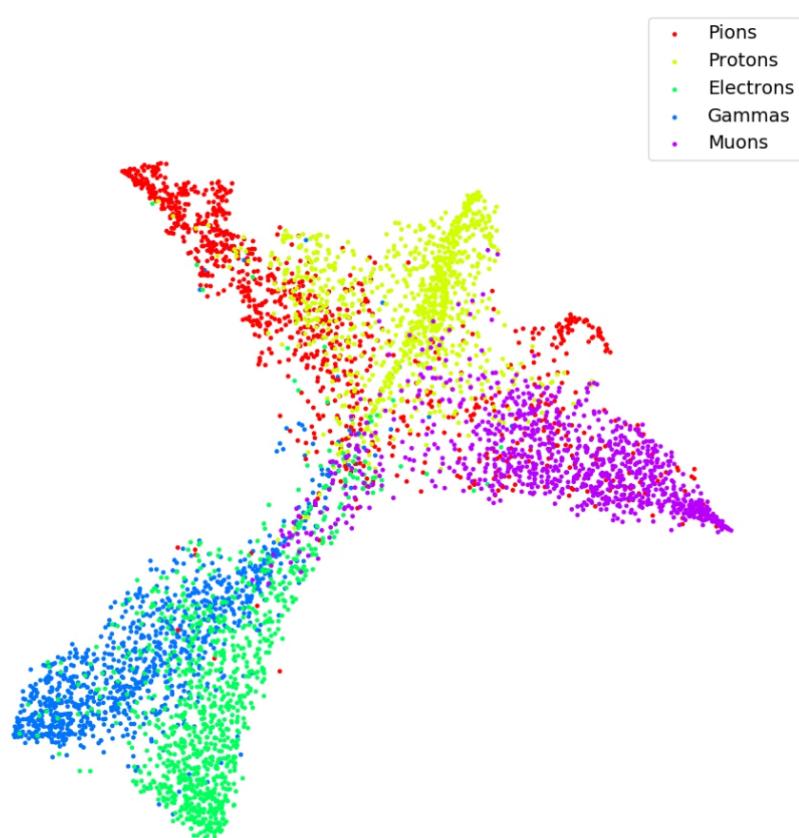
**Figure 7.3:** Confusion Matrix of all five particles



**Figure 7.4:** 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 7.5:** Training and testing loss of CNN trained on 100,000 images of  $\mu/\pi/p/\gamma/e$



**Figure 7.6:** t-SNE of CNN

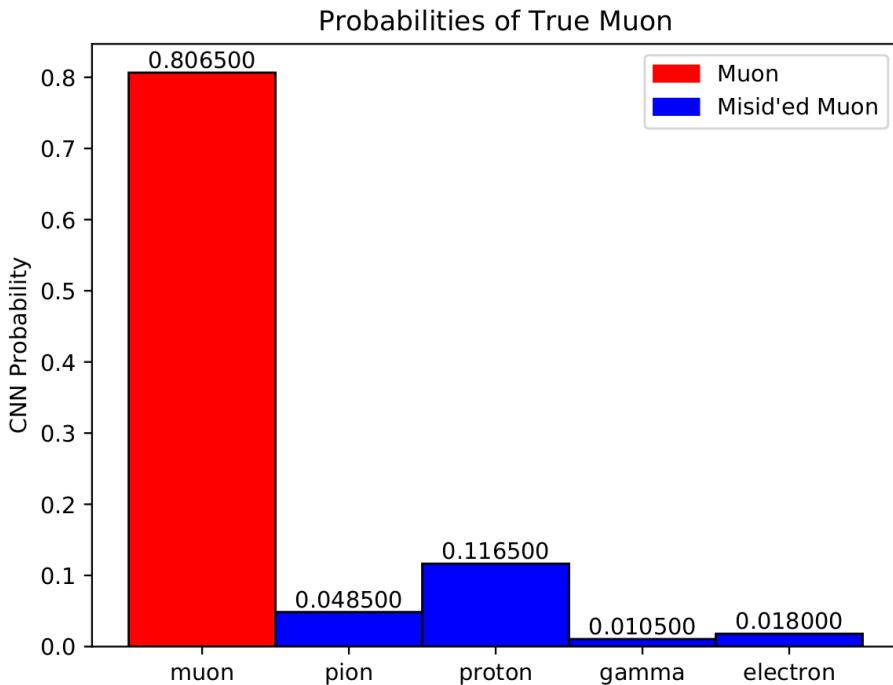


Figure 7.7: Muon Prob

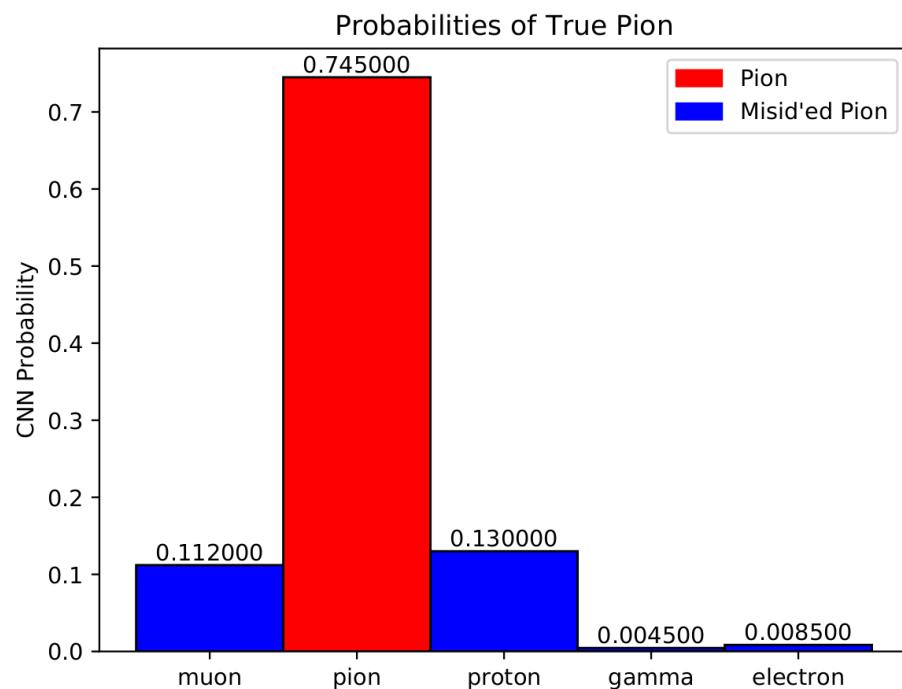
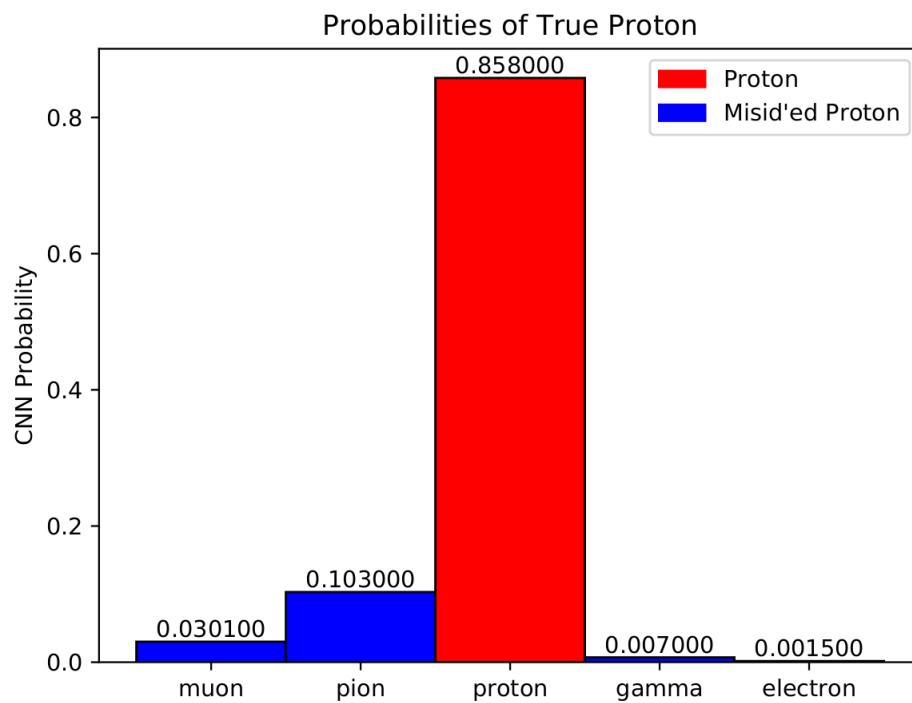
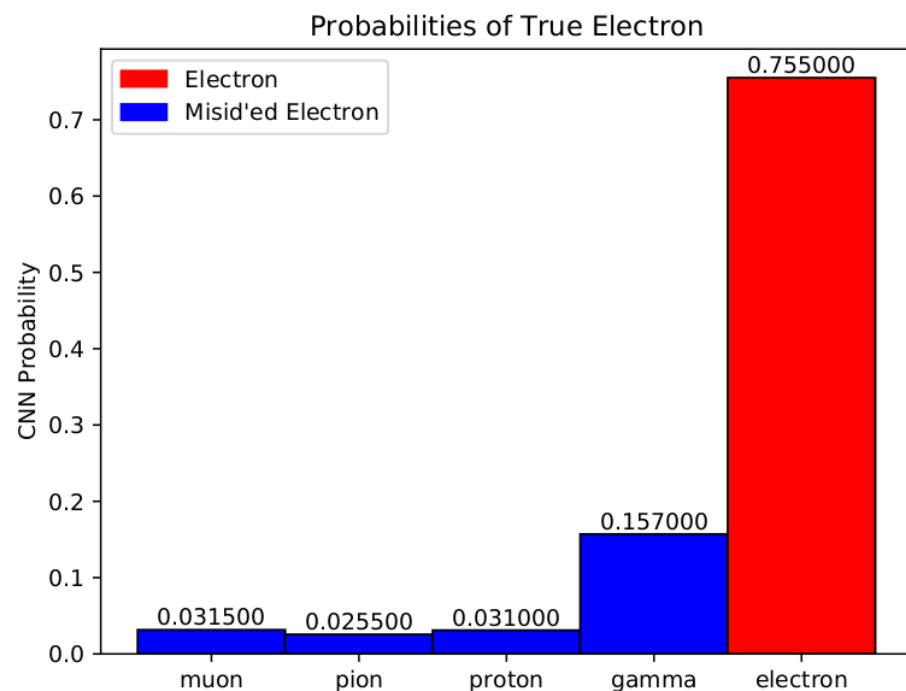


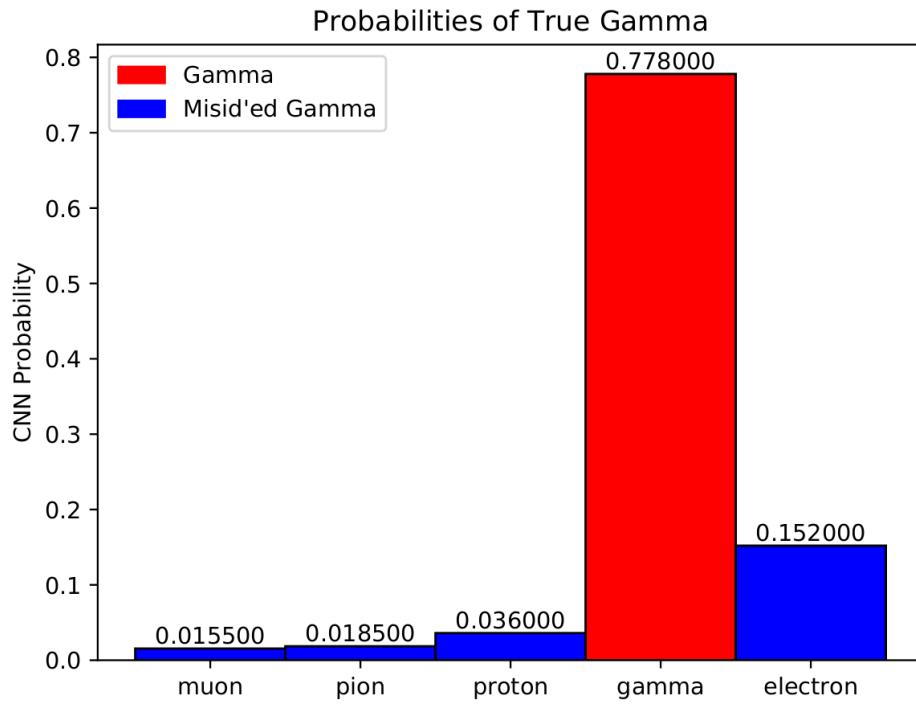
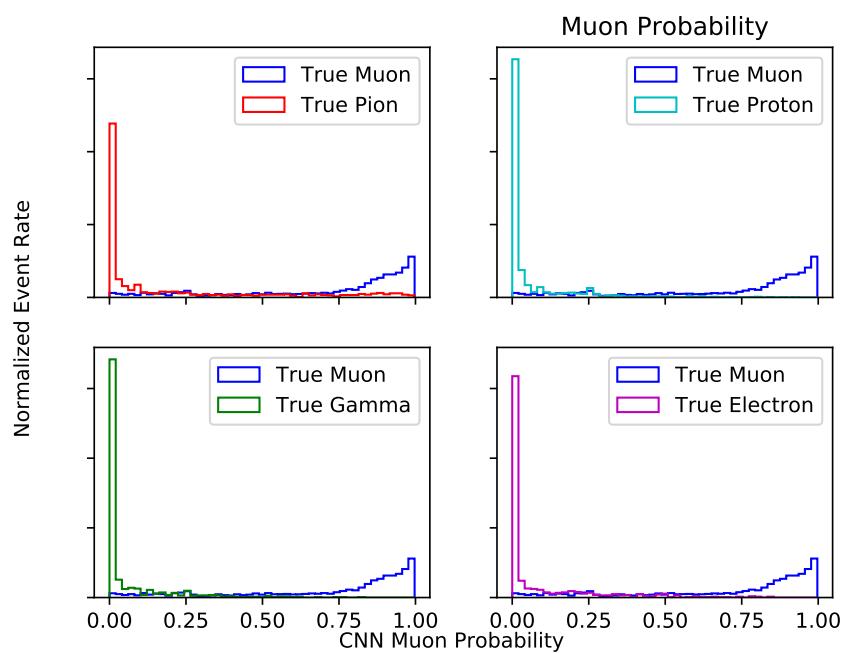
Figure 7.8: Pion Prob



**Figure 7.9:** Proton Prob



**Figure 7.10:** Electron Prob

**Figure 7.11:** Gamma Prob**Figure 7.12:** Prob

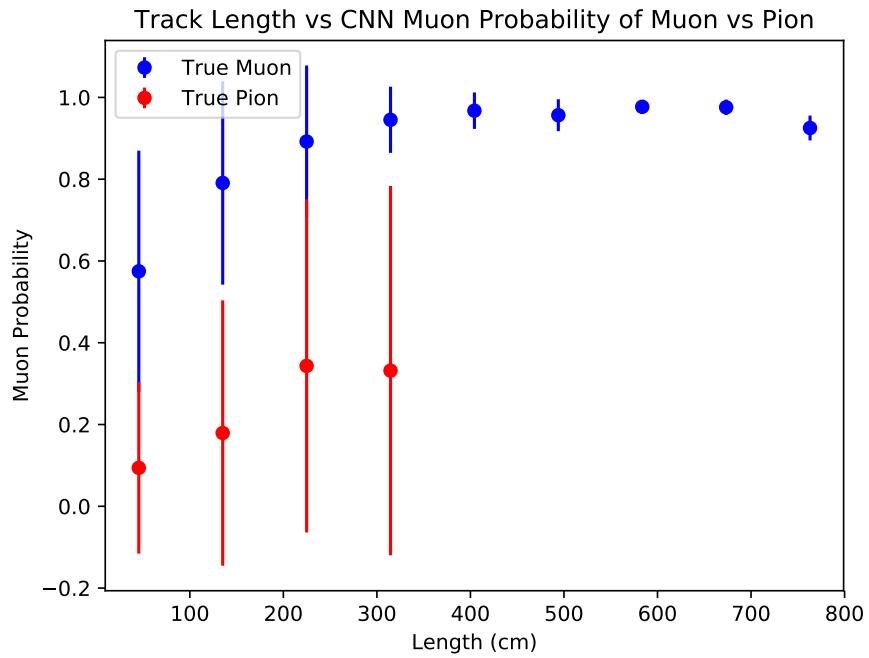


Figure 7.13: mupi

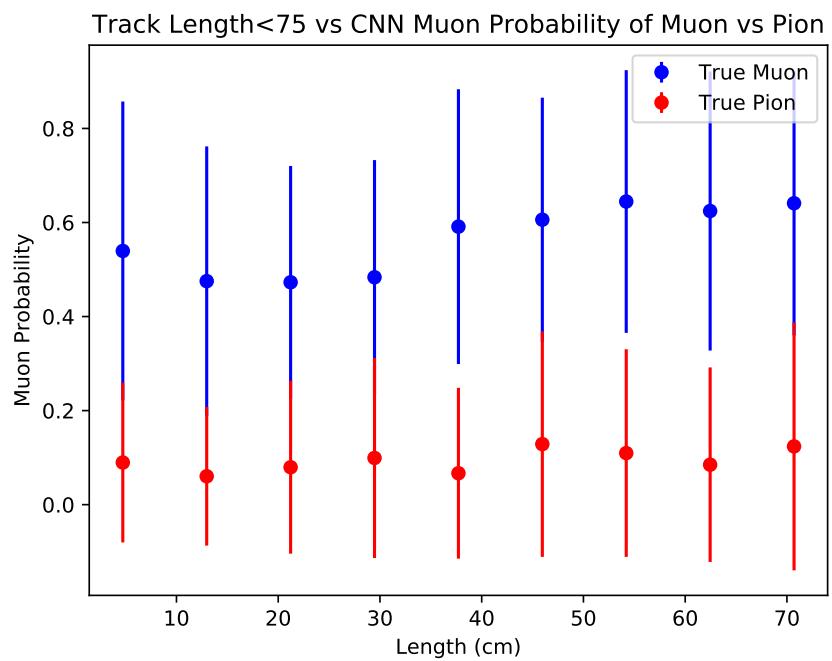
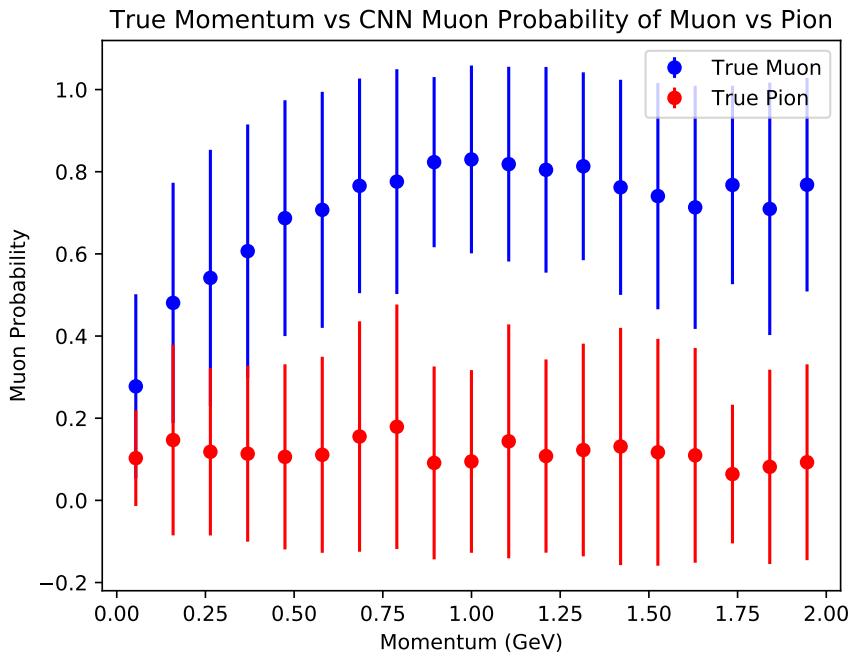
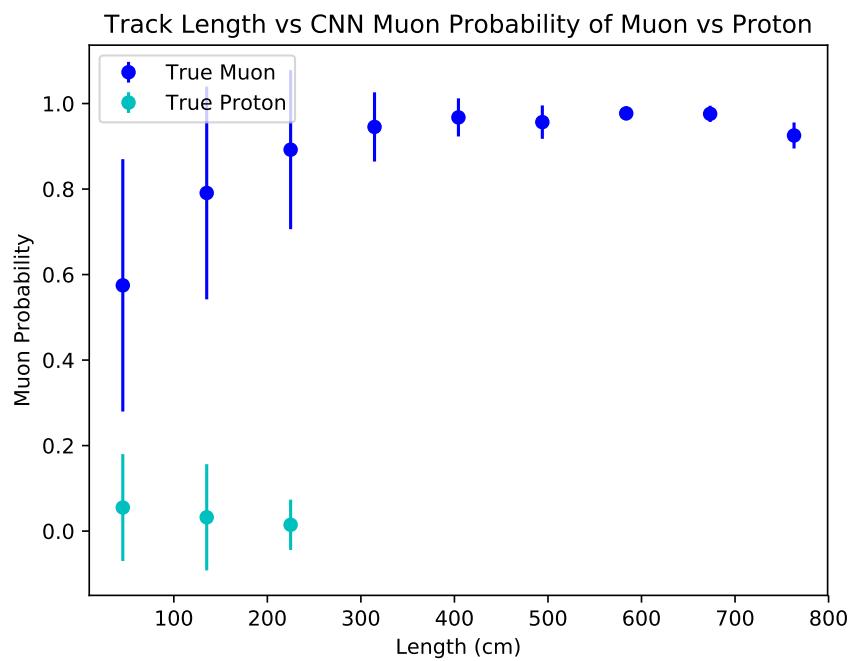
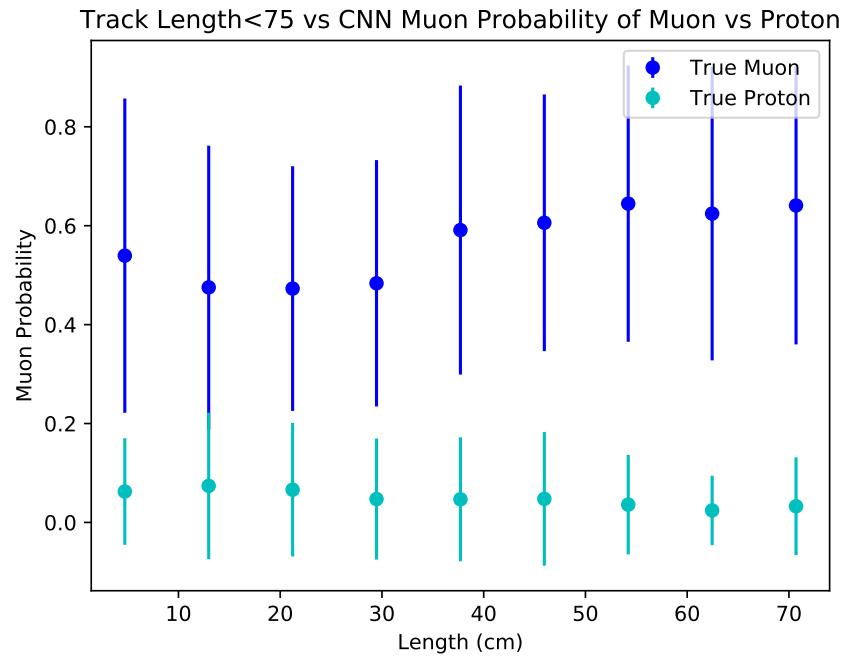
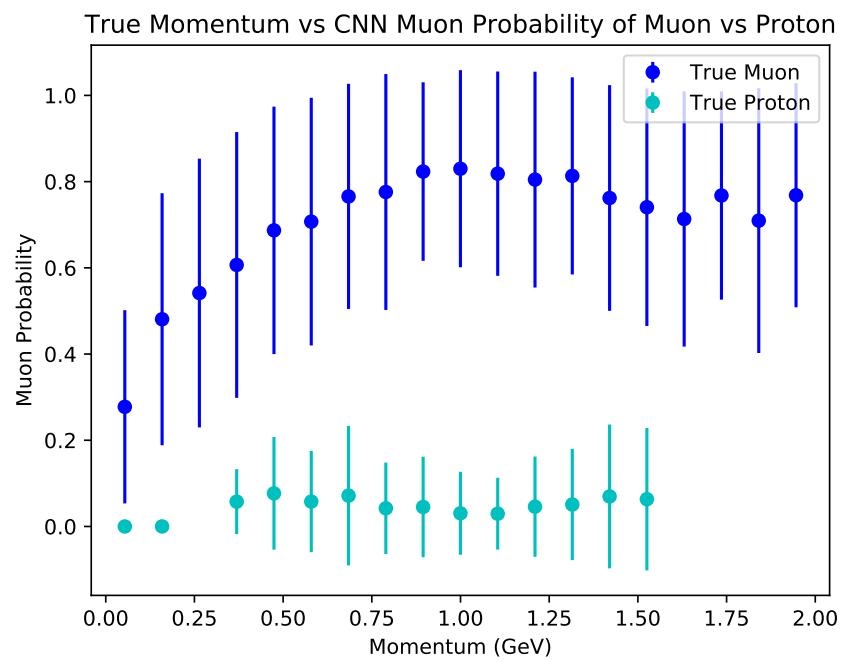


Figure 7.14: mupi

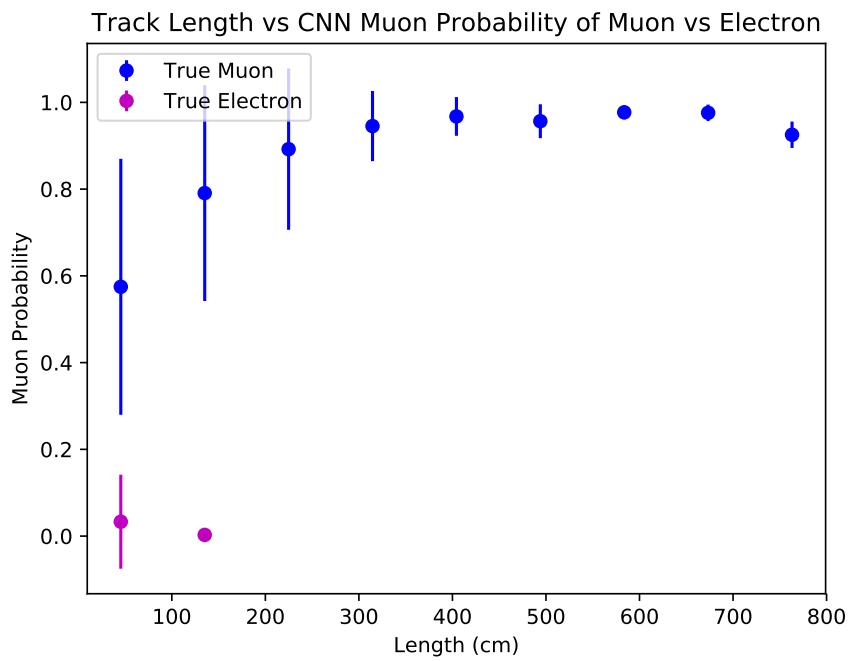
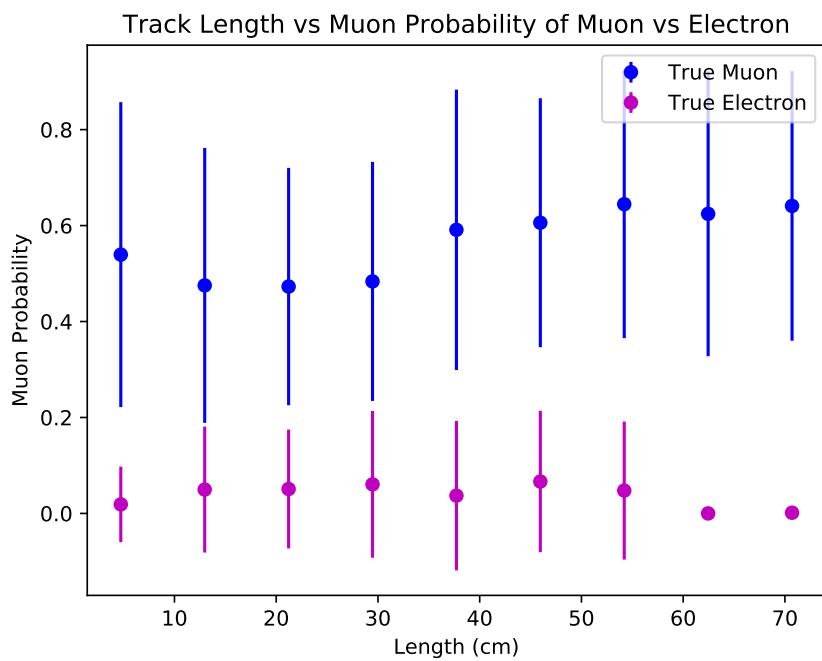
**Figure 7.15:** mupi**Figure 7.16:** mup



**Figure 7.17:** mup



**Figure 7.18:** mup

**Figure 7.19:** mue**Figure 7.20:** mue

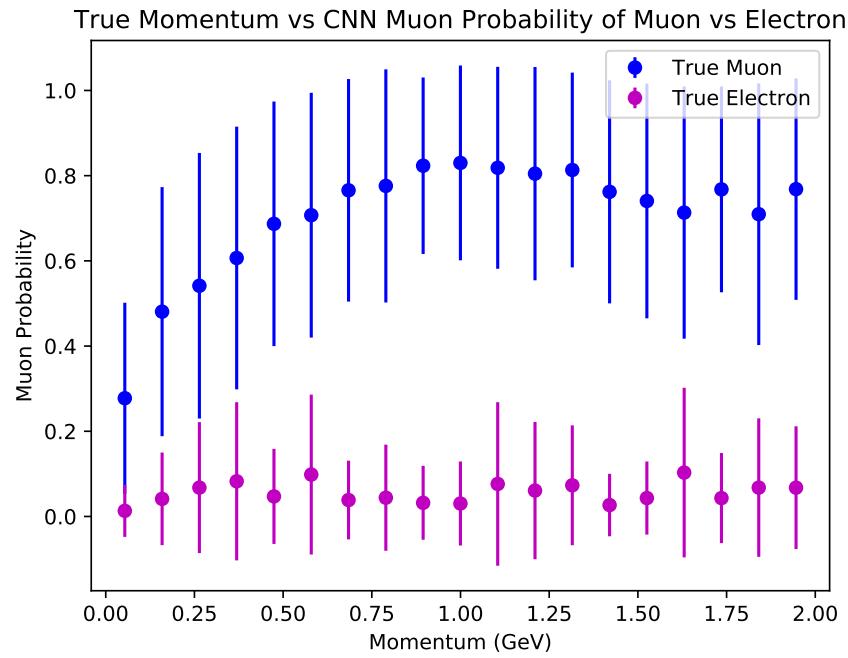


Figure 7.21: mue

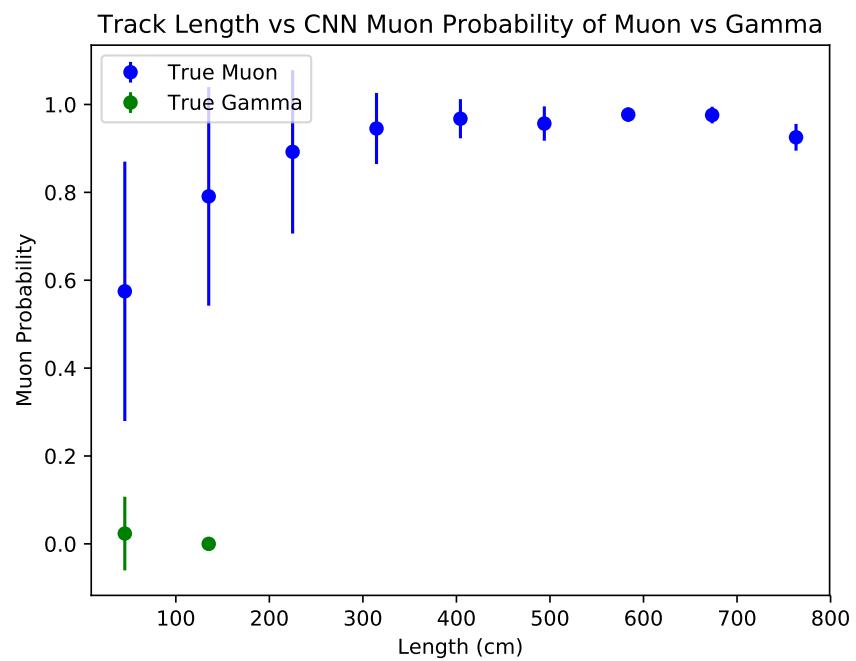


Figure 7.22: mug

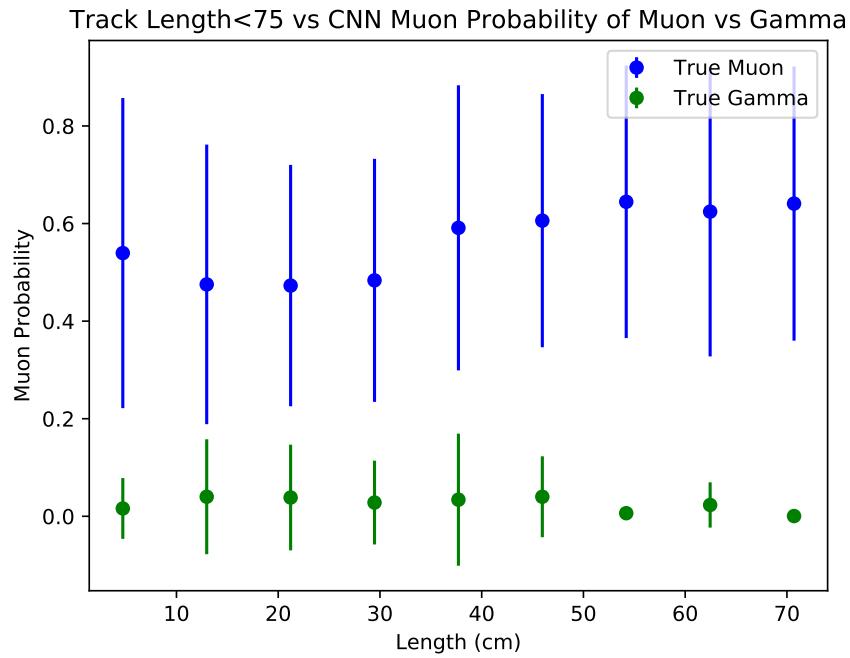


Figure 7.23: mug

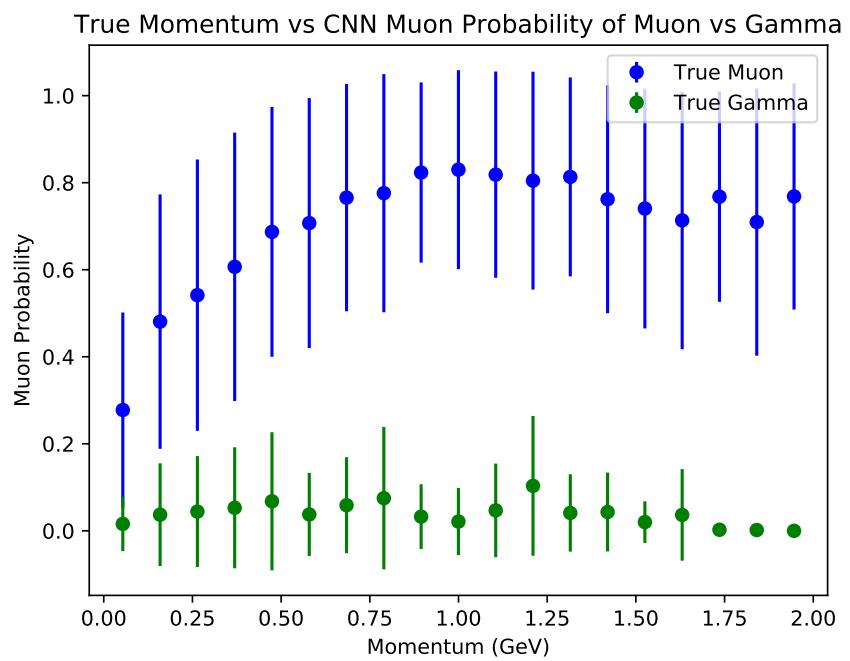


Figure 7.24: mug

# <sup>1396</sup> Chapter 8

## <sup>1397</sup> Results of Convolutional Neural Networks on particles **WORKING** <sup>1398</sup> **TITLE** <sup>1399</sup>

### <sup>1400</sup> 8.1 Classification using CNN10000

#### <sup>1401</sup> 8.1.1 Classification of MC data using Selection I Original <sup>1402</sup> CC-Inclusive Filter

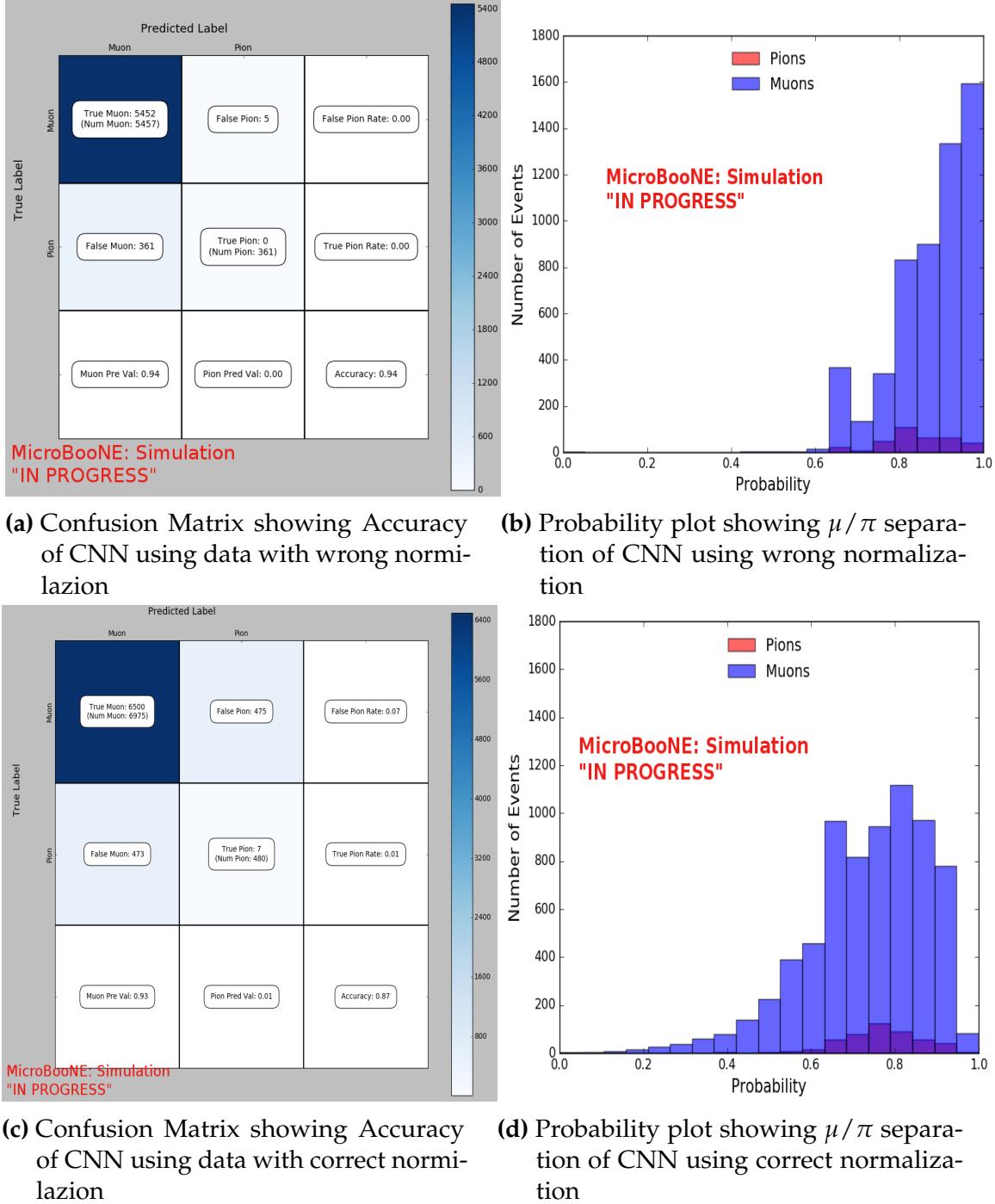
<sup>1403</sup> The next step that was taken was to use CNN10000 to classify track candidate images  
<sup>1404</sup> that were identified by the selection I original cc-inclusive filter described in [?].  
<sup>1405</sup> Passing rates for each cut in cc-inclusive filter are show in figure 8.1. For the incorrect  
<sup>1406</sup> image making normalization dataset, out of 188,880 events, 7438 passed the cut right  
<sup>1407</sup> before 75 cm track length cut which is 3.9% of total data. Discrepancies in passing rates  
<sup>1408</sup> are due to grid submission issues, however, this dataset is used to check if changes  
<sup>1409</sup> in image making normalization affects  $\mu/\pi$  separation probability due to CNN10000  
<sup>1410</sup> being trained with incorrectly image making normalized data. For the second dataset  
<sup>1411</sup> with correct image making normalization, out of 188,880 events, 9552 events passed the  
<sup>1412</sup> cut right before the 75 cm track length cut which is 5.1% passing rate and is comparable  
<sup>1413</sup> to figure 8.1. In time cosmics were also run over for efficiency and purity calculations.  
<sup>1414</sup> Out of 14395 in time cosmic events, 175 passed the cut right before the 75 cm track  
<sup>1415</sup> length cut which is a passing rate of 1.2% compared to 1.3% shown in figure 8.1.

Table 3: **Selection I: Original** The table shows passing rates for the above described event selection. Numbers are absolute event counts and Cosmic background is not scaled appropriately. The BNB+Cosmic sample contains all events, not just  $\nu_\mu$  CC inclusive. The selected events are further broken up in the following subsection. 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 cosmic only sample has low statistics, but please note that it is not used in any plots, it is just for illustrating 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 chapter 2. Note that this numbers is not a purity, since other backgrounds can't be determined at this step.

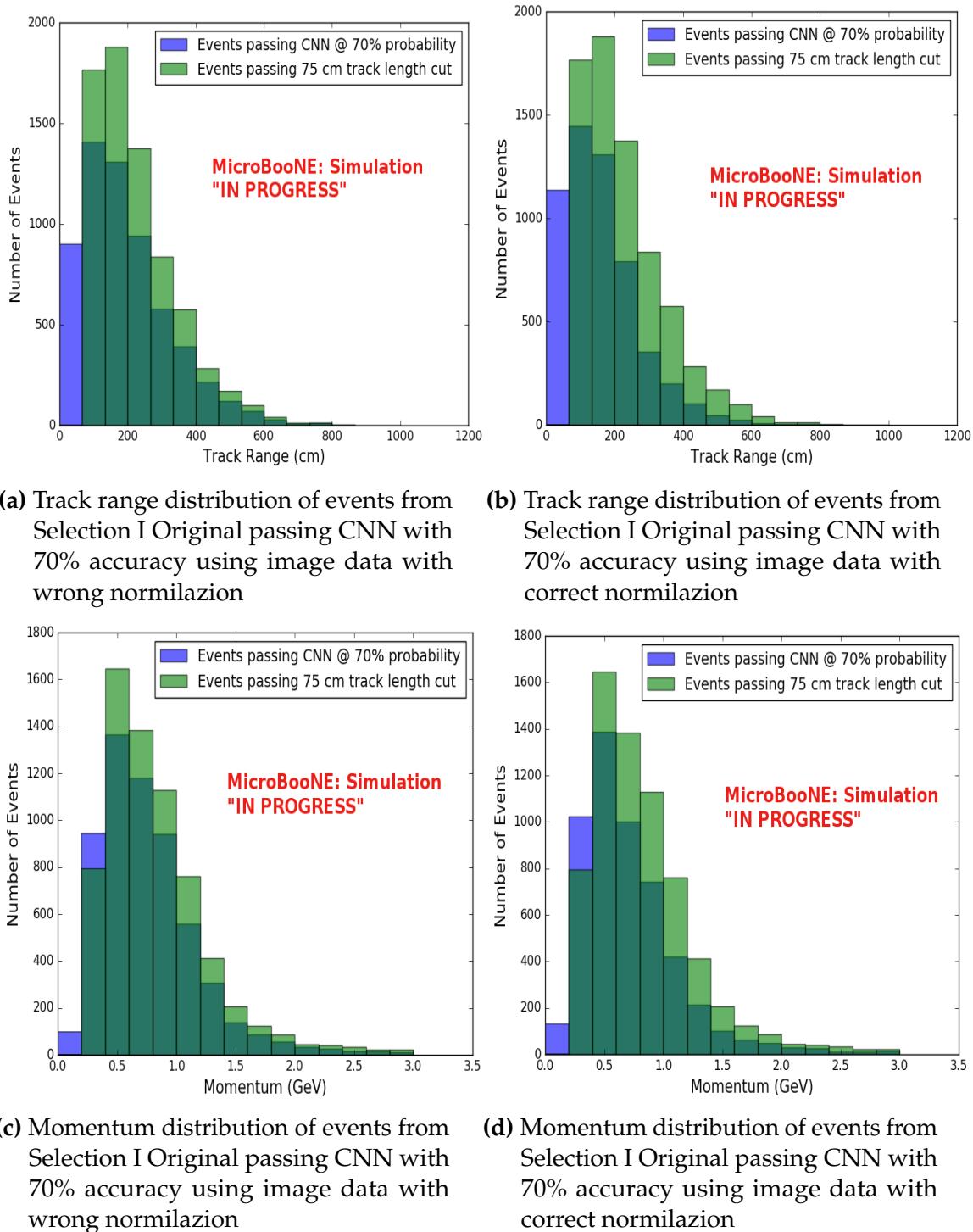
	BNB + Cosmic Selection	MC-Truth	Cosmic only	Signal: Cosmic only
Generated events	191362	45273	4804	1:22
$\geq 1$ flash with $\geq 50$ PE	136219 (71%/71%)	44002 (97%/97%)	2979 (62%/62%)	1:14
$\geq 1$ vertex in FV	131170 (96%/69%)	43794 (99%/97%)	2805 (94%/58%)	1:13
$\geq 1$ track within 5 cm of vertex	129784 (99%/68%)	43689 (99%/97%)	2756 (98%/58%)	1:13
flash matching of longest track	44775 (34%/23%)	23647 (54%/52%)	647 (23%/13%)	1:5.7
track containment	10114 (23%/5.3%)	6882 (29%/15%)	61 (9.4%/1.3%)	1:1.9
track $\geq 75$ cm	7358 (73%/3.8%)	5801 (84%/13%)	31 (51%/0.6%)	1:1.1

**Figure 8.1:** Snapshot of passing rates of Selection I from CC-Inclusive Filter

Figures 8.2a, 8.2b, 8.2c and 8.2d show the accuracy and  $\mu/\pi$  separation of both the correct and incorrect normalized images. The confusion matrices are only composed of  $\mu/\pi$  data. Other particles passed the cc-inclusive filter before the 75 cm track length cut and were all mis-id'ed as muons. Since CNN10000 has not seen any particles other than muons and pions, it makes sense that those get mis-id'ed. Figures 8.2b and 8.2d don't have  $\mu/\pi$  separation comparable to 7.2c, but 8.2b does skew to higher probabilities compared to 8.2d. This is to be expected and further work on quantifying the performance of CNN10000 should use the incorrect image making normalization. It is also expected that the separation isn't as defined as the testing dataset for CNN10000. CNN10000 was trained and tested using single particle muons and pions and the track candidate dataset come from BNB+Cosmic events, not to mentions all track candidates have passed the cc-inclusive filter that tags "muon-like" tracks therefore the pions in this sample look much closer in muon topology than the network has seen. Also, these images were made from wire and time ticks associated to hits from the track candidate that passed the cc-inclusive filter. This is different from the training images where a bounding box was drawn over the total  $\mu$  or  $\pi$  interaction. Spurious energy deposition from a  $\pi - Ar$  interaction is most likely not included in the BNB+Cosmic images due to the tracking algorithm. To remedy this, the neural network needs to see more "muon-like" pions and muons and pions from a neutrino interaction passing the cc-inclusive filter as well as a larger particle variety including protons, photons



**Figure 8.2:** Results of CNN10000 classification of track candidate images output from cc-inclusive filter.



**Figure 8.3:** CNN10000 distributions of track candidate images output from Selection I Original cc-inclusive filter with different image data normalizations

<sup>1436</sup> and electrons. Although  $\mu/\pi$  separation is lacking, CNN10000 does an excellent job of  
<sup>1437</sup> classifying muons and using higher CNN probability can increase purity. Figures 8.3a,  
<sup>1438</sup> 8.3b, 8.3c and 8.3d show the track and momentum distributions for these two datasets.  
<sup>1439</sup> In both sets you have an increase in data in the bin below 75 cm and at bins below 0.5  
<sup>1440</sup> GeV. These distributions were made with events classified with 70% probability of  
<sup>1441</sup> being a muon regardless of true particle type.

<sup>1442</sup> **8.1.2 Classification of MC data using Selection I Modified**  
<sup>1443</sup> **CC-Inclusive Filter**

Table 5: **Selection I: Modified** The table shows passing rates for the above described event selection. Numbers are absolute event counts and Cosmic background is not scaled appropriately. The BNB+Cosmic sample contains all events, not just  $\nu_\mu$  CC inclusive. The selected events are further broken up in the following subsection. 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 cosmic only sample has low statistics, but please note that it is not used in any plots, it is just for illustrating 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 chapter 2. Note that this numbers is not a purity, since other backgrounds can't be determined at this step.

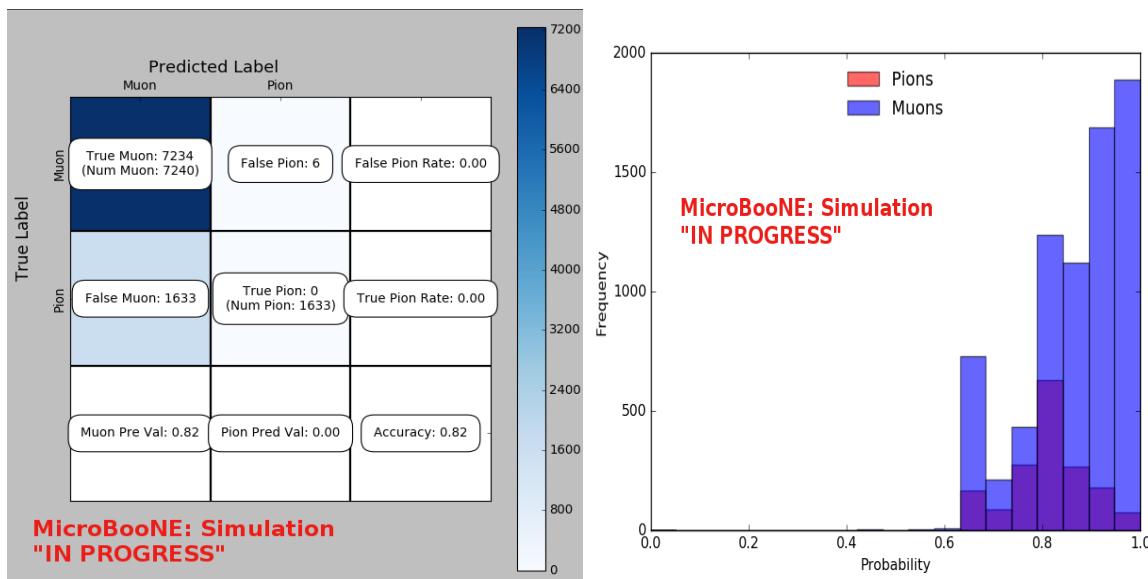
	BNB + Cosmic		Cosmic only	Signal: Cosmic only
	Selection	MC-Truth		
Generated events	191362	45273	4804	1:22
$\geq 1$ flash with $\geq 50$ PE	136219 (71%/71%)	44002 (97%/97%)	2979 (62%/62%)	1:14
$\geq 1$ track within 5cm of vertex	135830 (99%/71%)	43974 (99%/97%)	2975 (99%/62%)	1:14
vertex candidate in 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

**Figure 8.4:** Snapshot of passing rates of all cuts from Selection I Modified cc-inclusive filter

<sup>1444</sup> CNN10000 was also used to classify track candidate images that were identified by  
<sup>1445</sup> the selection I modified cc-inclusive filter described in [?]. Passing rates for each cut in  
<sup>1446</sup> this filter are shown in figure 8.4. As seen in section 8.1.1, wrong image normalization  
<sup>1447</sup> had a higher muon classification probability so all work done using selection I modified  
<sup>1448</sup> cc-inclusive filter was done using this normalization. Out of 188,880 events, 19,112  
<sup>1449</sup> passed the cut right before the 75 cm track length cut which is a 10.1% passing rate and  
<sup>1450</sup> comparable to the 10% passing rate shown in figure 8.4. In time cosmics were also run  
<sup>1451</sup> over, out of 14,606 in time cosmics events, 302 passed the cut right before the 75 cm  
<sup>1452</sup> track length cut which is a 2.1% passing rate comparable to the 2.7% passing rate in the  
<sup>1453</sup> cc-inclusive tech-note. Figures 8.5a and 8.5b show the accuracy and  $\mu/\pi$  separation.  
<sup>1454</sup> Both plots are only composed of muons and pions and like selection I original data,

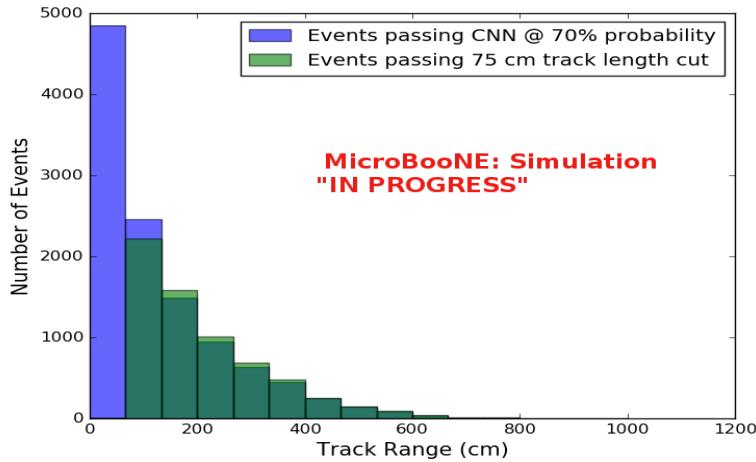
1455 all other particles were id'ed as muons. Also like selection I original data, muons are  
 1456 being identified at a very high rate. Figure 8.6a shows the track range distributions  
 1457 of all events from selection I modified being classified by the CNN as a muon with a  
 1458 probability of 70% regardless of true particle type. We get entries for the CNN curve  
 1459 in the lowest bin and none for the 75 cm curve. To see how many true CC events  
 1460 were identified by CNN10000 breaking down figure 8.6a by event type was necessary.  
 1461 Figures 8.6b and 8.6c show track range distributions separated by signal and various  
 1462 backgrounds. Particle type was not taken into consideration in these plots so true CC  
 1463 event images can be any track candidate particle passing selection I modified cut right  
 1464 before track length cut including pions and protons.

1465 To gain an even deeper understanding on how CNN10000 is performing, plotting  
 1466 these distributions with only muons and pions was done due to the fact that CNN10000  
 1467 was trained with only those particles for  $\mu/\pi$  separation. Figures 8.6d-8.7d show the  
 1468 stacked histograms of signal and background of the track range distributions with  
 1469 varying CNN probabilities starting from 70% and ending at 90% probability. With  
 1470 higher probabilities we get a purer sample in the lower bin but we end up losing  
 1471 events as well. Momentum distributions for all signal/background events are shown  
 1472 in figure 8.8.

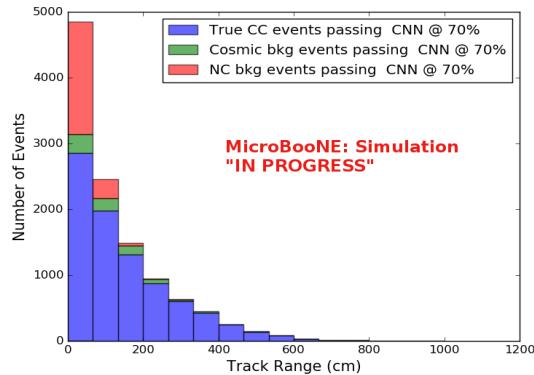


(a) Confusion Matrix for CNN10000 classified events from selection I modified      (b) Probability plot for CNN10000 classified events from selection I modified

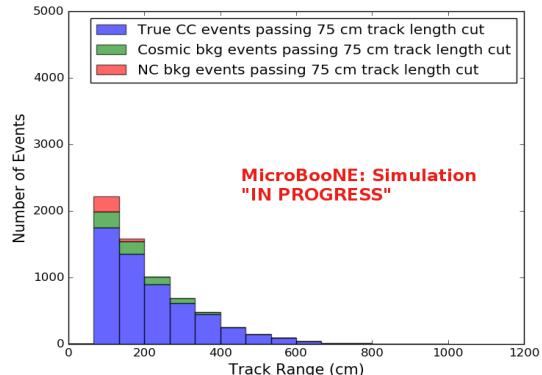
**Figure 8.5:** Confusion matrix and probability plot of events passing selection I modified cc-inclusive cuts right before 75cm track length cut



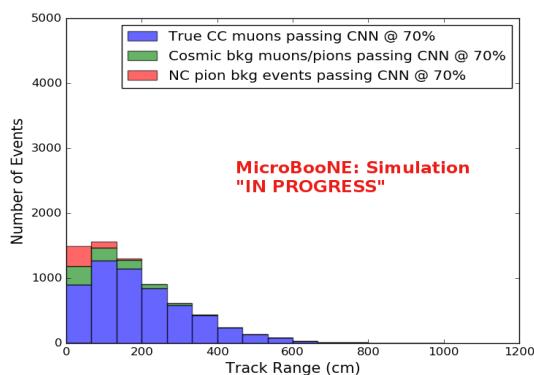
**(a)** Track range distribution of events from Selection I Modified passing CNN with 70% accuracy



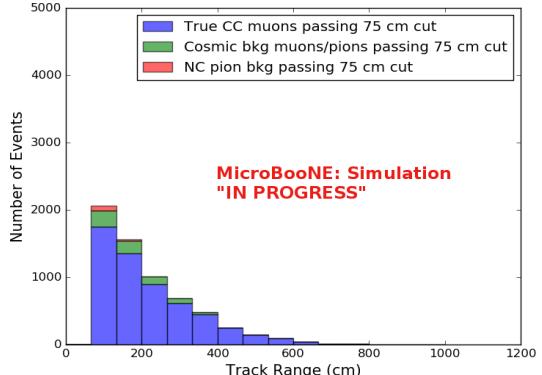
**(b)** Stacked signal and background track range distributions from Selection I Modified passing CNN with 70% accuracy



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

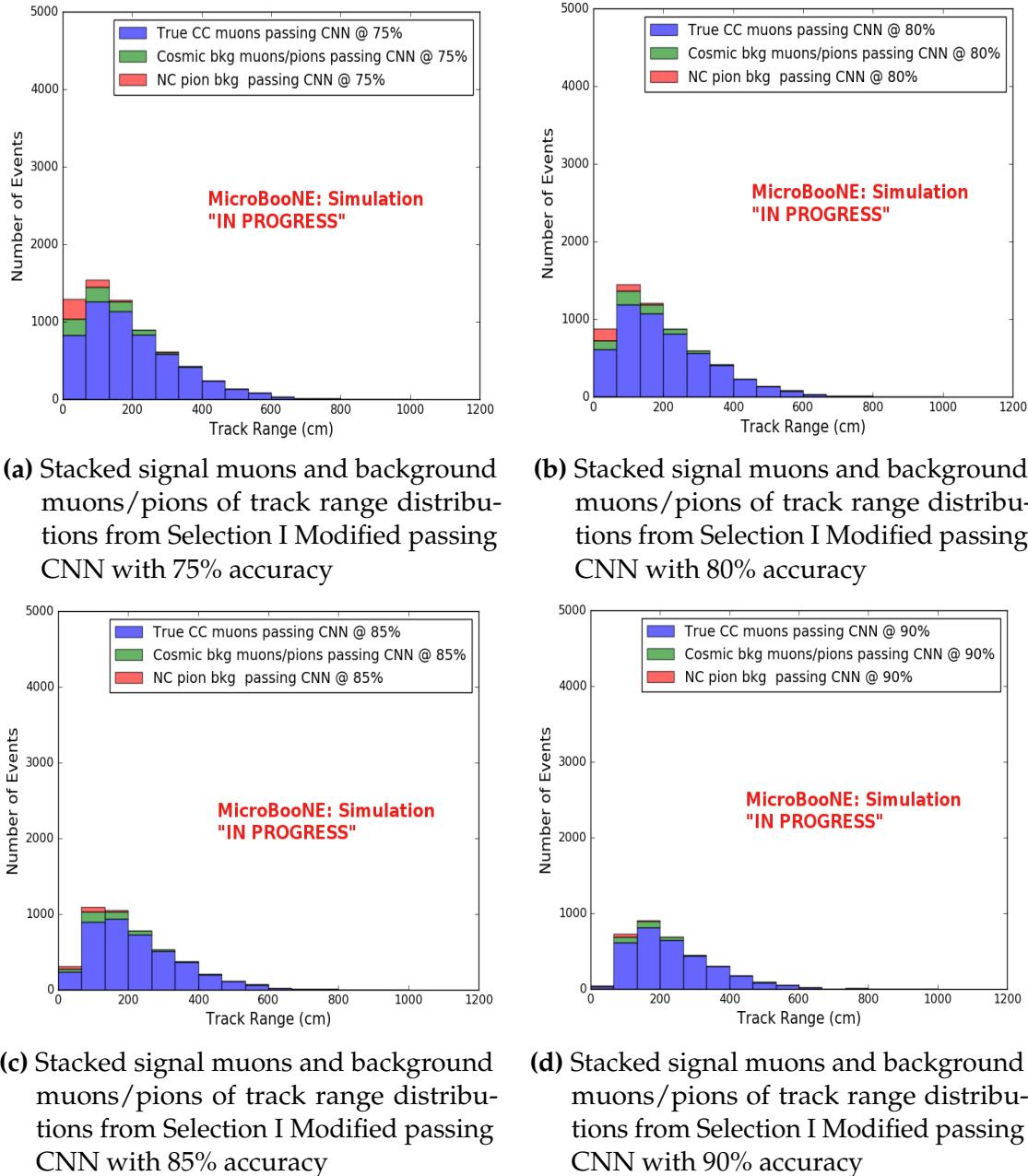


**(d)** Stacked signal muons and background muons/pions of track range distributions from Selection I Modified passing CNN with 70% accuracy

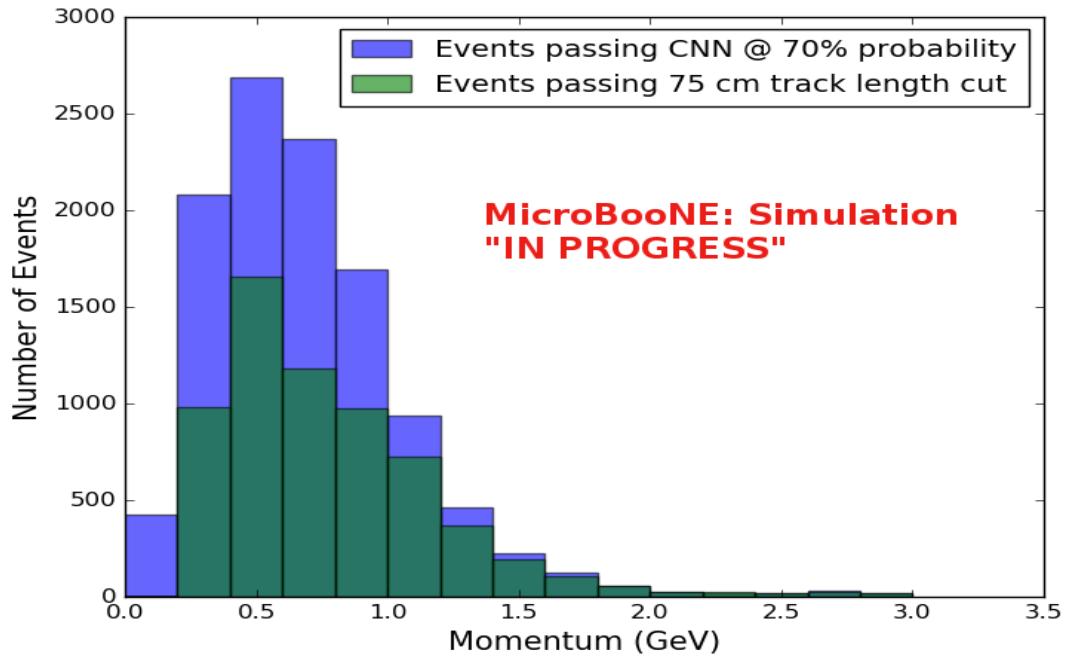


**(e)** Stacked signal muons and background muons/pions of track range distributions from Selection I Modified passing 75 cm track length cut

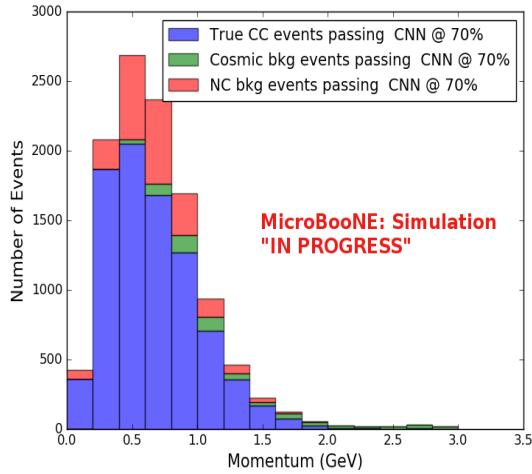
**Figure 8.6:** CNN10000 distributions of track candidate images output from Selection I Modified cc-inclusive filter



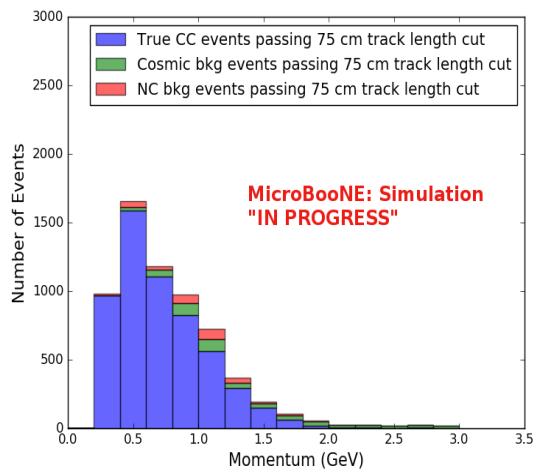
**Figure 8.7:** CNN10000 stacked signal/background track range distributions of track candidate images output from Selection I Modified cc-inclusive filter



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



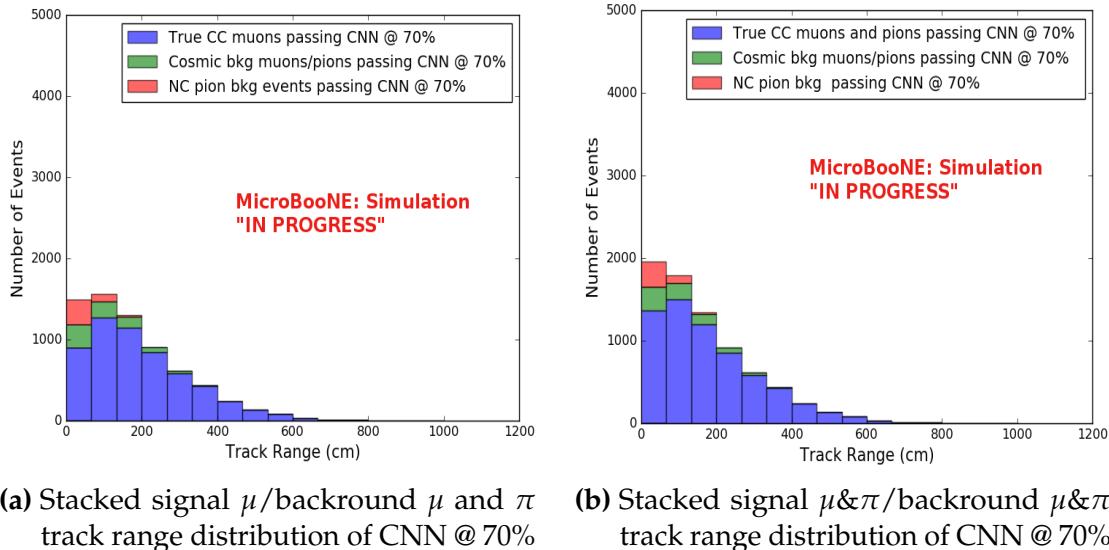
**(b)** Stacked signal and background momentum distributions from Selection I Modified passing CNN with 70% accuracy



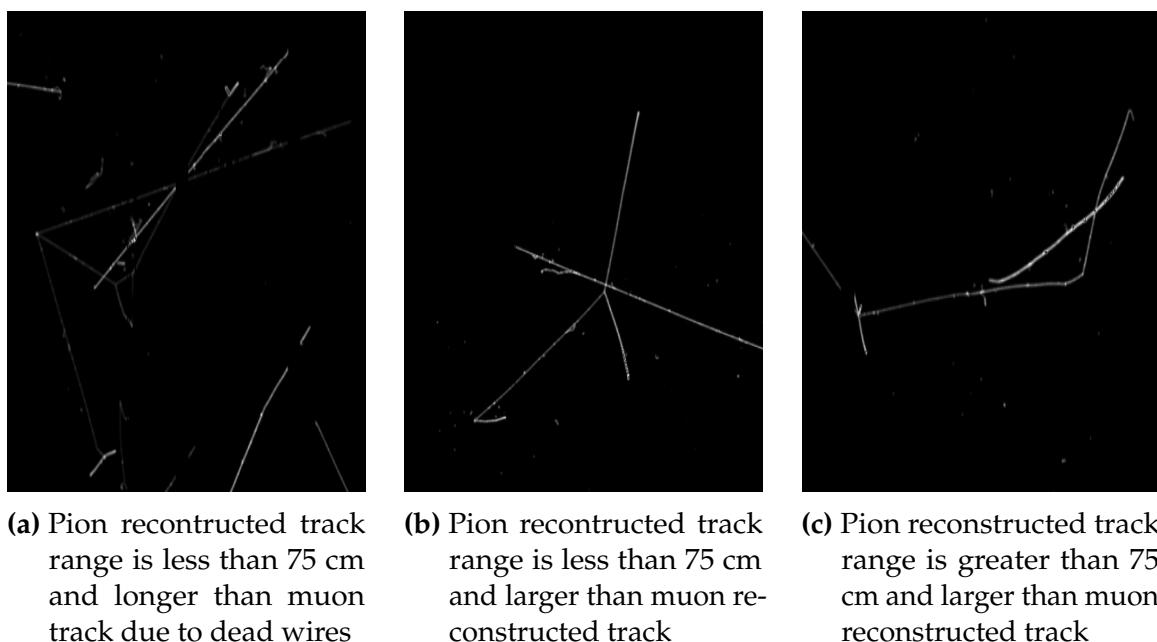
**(c)** Stacked signal and background momentum distributions from Selection I Modified passing 75 cm track length cut

**Figure 8.8:** CNN10000 momentum distributions of track candidate images output from Selection I Modified 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 8.9 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 8.10, 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 8.10a 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 8.10b 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 8.10c shows a pion with a reconstructed track length greater than 75 cm and the muon. These three cases show that a broader question must be asked when training the network other than is it a muon or pion. There are different routes to recover interesting events like these. One route is to ask the network “Is it a CC event or is it an NC event?” and obtain an image dataset consisting of whole CC/NC events that will train the network to answer this question. The other route is to ask the network “Is this a  $\mu/\pi/p/$  from a CC event or NC event and obtain an image dataset consisting of primary particles from a CC/NC event. Both these paths will be explored in future work.



**Figure 8.9:** Track distribution comparisons of true CC muons plotted vs true CC muons and pions plotted



**Figure 8.10:** 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 CNN @ 70% Probability	passing CNN @ 83% Probability
		10605(35%)	6142(61%)
Backgrounds	Cosmics Only Events	13573(45%)	2582(26%)
	Cosmics in BNB Events	2249(7.4%)	492(4.9%)
	NC Events	3412(11%)	778(7.7%)
	$\nu_e$ and $\bar{\nu}_e$ Events	139(0.5%)	32(0.3%)
	$\bar{\nu}_\mu$ Events	97(0.3%)	67(0.7%)

**Table 8.2:** Signal and background event numbers at modified selection level 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.

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 modified with 75 cm cut and selection I modified with CNN cut. As it stands, a CNN cut at 83% probability has

1506 a MC true CC event passing rate of 14% compared to the 13% passing rate of the 75  
 1507 cm track length cut. The Signal:Cosmic Only background is also reduced from 1:0.6  
 1508 to 1:0.4 The total passing rate is also higher than the 75 cm cut, 3.6% vs 4.0%. Table  
 1509 8.2 shows the breakdown of signal and backgrounds for the CNN at the different  
 1510 probabilities. We have a 61% signal passing rate with the CNN cut @ 83% versus the  
 1511 53.8% signal passing rate of the 75 cm cut.

1512 Based on these numbers, the following performance values of the modified selec-  
 1513 tion with 75 cm cut versus modified selection with CNN @ 83% probability cut were  
 1514 calculated:

- 1515 • Efficiency: Number of selected true  $\nu_\mu$  CC events divided by the number of  
 1516 expected true  $\nu_\mu$  CC events with interaction in the FV.
  - 1517 – Selection I modified: 13%
  - 1518 – Selection I modified with CNN cut @ 83% probability: 14%
- 1519 • Purity: Number of selected true  $\nu_\mu$  CC events divided by sum of itself and the  
 1520 number of all backgrounds.
  - 1521 – Selection I modified: 53.8%
  - 1522 – Selection I modified with CNN cut @ 83% probability: 61%

1523 Lastly, figure 8.12 shows a more representative performance of the CNN. Due to  
 1524 the fact that the CNN was trained on muons and pions, showing the performance  
 1525 of CC muon events versus NC pion events with respect to CNN probability gives a  
 1526 better picture of how the network is performing. Figure 8.12 shows that at 83% we  
 1527 are below the 75 cm cut NC pion threshold and still above the CC muon threshold.  
 1528 Using 83% probability not only reduced the NC pion background, it also dramatically  
 1529 reduced the in time cosmics and cosmics in the BNB.

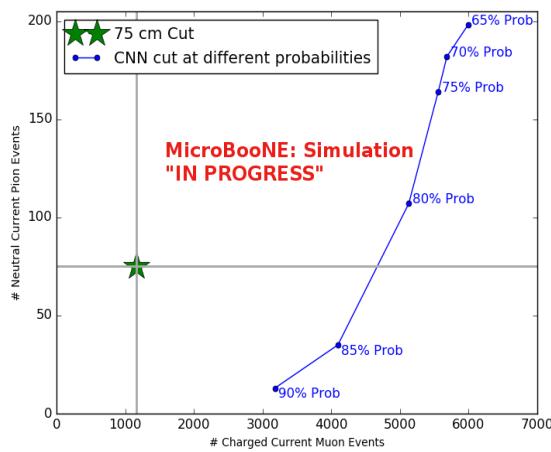
### 1530 8.1.3 Conclusions and Future Work

1531 It was shown that even though CNN10000 was trained with single particle generated  
 1532 muons and pions, it performs fairly well at classifying track candidate images from  
 1533 BNB+Cosmic events. Events have been regained below the 75 cm track length cut and  
 1534 the momentum and track range distributions have similar shapes to the distributions of  
 1535 Selection I original and modified. Efficiencies and purities were calculated for selection

Table 8: **Selection I: Modified** Signal and background event numbers at modified selection level estimated from a BNB+Cosmic sample and Cosmic only sample normalized to  $5 \times 10^{19}$  PoT. The last column gives the fraction of this signal or background type to the total 2189 selected events.

Signal	#Events	
$\nu_\mu$ CC events with true vertex in FV	1168	53.8%
Backgrounds		
Cosmics only events	725	33.4%
Cosmics in BNB events	144	6.6%
NC events	75	3.5%
$\nu_e$ and $\bar{\nu}_e$ events	4	0.2%
$\bar{\nu}_\mu$ events	15	0.7%
$\nu_\mu$ CC events with true vertex outside FV	40	1.8%

**Figure 8.11:** Snapshot of signal and background event numbers of Selection I modified from cc-inclusive note [?]



**Figure 8.12:** CNN performance of classified muons and pions compared to the already implemented 75 cm track length cut

1536 I modified events before 75 cm track length cut with the CNN at 83% probability and  
1537 are 14% and 62% respectively. Although the CNN doesn't have separation between  
1538 muons and pions and although all particles passing CNN are classified as muon,  
1539 increasing CNN probability allows us to increase the purity as well as maintain an  
1540 efficiency comparable to the 75 cm track length cut all while recovering events below  
1541 that 75 cm cut. Out of the 6142 events that passed the CNN @ 83% 1470 events were  
1542 below the 75 cm cut, a recovery of 3.3% of data with a purity of 15%. Although  
1543 these numbers are low, it is an improvement from the selection I modified in both total  
1544 efficiency and purity and an increase in phase space by recovering these events.

1545 **8.2 Classification using CNN100000**

1546 All future classifications will be done using Selection I Modified CC-Inclusive Filter  
1547 because it has a higher efficiency and purity than Selection I Original CC-Inclusive  
1548 Filter. To reiterate, CNN100000 was trained using 20,000 images of each  $\mu/\pi/p/\gamma/e$ .  
1549 The results of using CNN100000 to classify BNB+Cosmics will be outlined below.

1550 **8.2.1 Classification of MC data using Selection I Modified  
1551 CC-Inclusive Filter**

1552 **8.2.2 Classification of MicroBooNE data using Selection I Modified  
1553 CC-Inclusive Filter**

1554 **8.2.3 Comparing two CC-Inclusive Cross Section Selection Filters**



1555 **Chapter 9**

1556 **Conclusion**

1557 Your Conclusions here.

1558



<sub>1559</sub> **Bibliography**

- <sub>1560</sub> [1] Wikipedia, Neutrino, <http://wikipedia.org/wiki/Neutrino>, 2013.
- <sub>1561</sub> [2] Wikipedia, Neutrino oscillation, [http://en.wikipedia.org/wiki/Neutrino\\_oscillation](http://en.wikipedia.org/wiki/Neutrino_oscillation), 2013.
- <sub>1563</sub> [3] B. N. Laboratory, Neutrinos and nuclear chemistry, <http://www.chemistry.bnl.gov/sciandtech/sn/default.htm>, 2010.
- <sub>1565</sub> [4] K. Heeger, Big world of small neutrinos, <http://conferences.fnal.gov/lp2003/forthepublic/neutrinos/>.
- <sub>1567</sub> [5] A. Y. Smirnov, p. 23 (2003), hep-ph/0305106.
- <sub>1568</sub> [6] The MicroBooNE Collaboration, R. e. a. Acciari, (2015), arXiv:1512.06148.
- <sub>1569</sub> [7] The MicroBooNE Collaboration, R. e. a. Acciari, (2016), arXiv:1705.07341.
- <sub>1570</sub> [8] The MicroBooNE Collaboration, R. e. a. Acciari, (2017), arXiv:1704.02927.
- <sub>1571</sub> [9] The MicroBooNE Collaboration, R. e. a. Acciari, JINST **12** (2017).