

1

2 μ/π 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

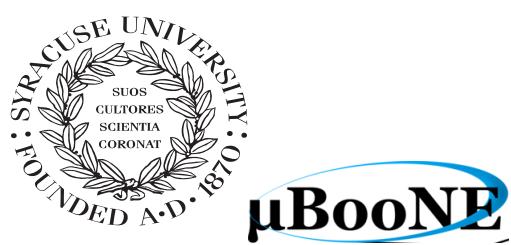
- * - DRAFT December 18, 2017 - * -

December, 2017
Syracuse University
Syracuse, New York

20
21
22
23

Copyright 2017
Jessica Nicole Esquivel
All Rights Reserved
Syracuse University

24



μ BooNE



26

Abstract

27

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

33

Dedication

34

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

37

Jessica Nicole Esquivel

38

Acknowledgements

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

41 Contents

42	List of figures	xv
43	List of tables	xxiii
44	1 Introduction	1
45	2 Neutrinos	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	3 The MicroBooNE Experiment	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 Astroparticle physics	19
61	3.3.4 Liquid argon detector development	19
62	3.4 The Booster Neutrino Beam	20
63	3.4.1 Creating the Booster Neutrino Beam	20
64	3.5 Event Reconstruction	21

65	4 Neutrino Identification: Finding MicroBooNE's first Neutrinos	23
66	4.1 Flash Finding	23
67	4.1.1 Flash Reconstruction	24
68	4.1.2 Beam Timing	25
69	4.1.3 Event Rates	26
70	4.2 TPC Topology Selection	26
71	4.2.1 Cosmic Tagging	27
72	4.2.2 2D Cluster Selection	27
73	4.2.3 3D Tracks and vertices Selection	29
74	4.2.4 TPC Updates	31
75	4.3 Conclusion	32
76	5 CC-Inclusive Cross Section Selection Filter	33
77	5.1 The importance of μ/π separation	34
78	6 Background on Convolutional Neural Networks	37
79	6.1 Image Classification	37
80	6.1.1 CNN Structure	38
81	6.1.2 Backpropagation	41
82	6.2 Choosing Hyperparameters	42
83	7 Training Convolutional Neural Networks on particles WORKING TITLE	43
84	7.1 Hardware Frameworks used for Training	44
85	7.1.1 Syracuse CPU Machine setup	44
86	7.1.2 Syracuse University GPU Cluster Setup	44
87	7.2 Convolutional Neural Network Training	44
88	7.2.1 Image Making Scheme	44
89	7.2.2 Training CNN1075	45
90	7.2.3 Training CNN10000	45
91	7.2.4 Training CNN100000	47
92	8 Results of Convolutional Neural Networks on particles WORKING TITLE	61
93	8.1 Classification using CNN10000	61
94	8.1.1 Classification of MC data using Selection I Original CC-Inclusive Filter	61
95	8.1.2 Classification of MC data using Selection I Modified CC-Inclusive Filter	65

98	8.1.3 Conclusions and Future Work	73
99	8.2 Classification using CNN100000	75
100	8.2.1 Classification of MC data using Selection I Modified CC-Inclusive	
101	Filter	75
102	8.2.2 Classification of MicroBooNE data using Selection I Modified	
103	CC-Inclusive Filter	75
104	8.2.3 Comparing two CC-Inclusive Cross Section Selection Filters . .	75
105	9 Conclusion	77
106	Bibliography	79

¹⁰⁷ List of figures

¹⁰⁸	2.1	The Standard Solar Model	5
¹⁰⁹	2.2	Solar Neutrino Experiments	7
¹¹⁰	2.2a	Ray Davis's Homestake Experiment	7
¹¹¹	2.2b	Kamiokande Experiment	7
¹¹²	2.2c	SNO Experiment	7
¹¹³	2.3	Cosmic Ray Shower	8
¹¹⁴	2.4	Measurements of the double ratio for various atmospheric neutrino experiments	9
¹¹⁶	2.5	The flavor eigenstates are rotated by an angle θ with respect to the mass eigenstates	10
¹¹⁸	3.1	3.1b Flux of BNB at FNAL	21
¹¹⁹	3.1a	Low Energy excess seen in MiniBooNE	21
¹²⁰	3.1b	Energy spectrum of the Booster Neutrino Beam at Fermi National Laboratories	21
¹²²	4.1	Efficiency for selecting beam events as a function of minimum total PE cut for all PE cuts as well as zoomed into interesting region.	24
¹²⁴	4.2a	Predicted distribution of flash times with respect to trigger time for 1 day of data taking at nominal rate and intensity	29

126	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 μ s after the trigger time.	29
130	4.3 Percent of good clusters remaining for neutrinos and cosmics after the primary cuts were applied. This is relative to total number of initial clusters.	30
133	4.4 Percent matched cluster pairs remaining for neutrinos and cosmics after secondary cuts applied. This is relative to the number of events that contain clusters which pass the primary cuts.	30
136	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. .	35
145	5.1a Track range distribution of selection I	35
146	5.1b Selection efficiency as a function of the true muon momentum .	35
147	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	39
150	6.2 Visualization of filters found in first layer of a CNN.	39
151	6.3 Applying a feature mask over a set of fashion items to extract necessary information for auto-encoding. Unnecessary information for example color or brand emblems are not saved. This feature map is an edge detec- tion mask that leaves only shape information which helps to distinguish between different types of clothes.	40

156	6.4 Pictorial Representation of Convolutional Neural Networks as well as a visual representation on CNN's complexity of layer feature extraction	40
157		
158	7.1 Accuracy vs. Loss of ImageNet 2-output μ/π sample consisting of 10000 images each.	47
159		
160	7.2 Description of confusion matrix variables: False pion rate = $false\pi/total\pi$	
161	True pion rate = $true\pi/total\pi$ Accuracy = $(true\pi + true\mu)/2$	
162	Pion prediction value = $true\pi/(true\pi + false\pi)$ Muon prediction value = $true\mu/(true\mu + false\mu)$	
163	7.2c The probability plot includes muons and pions that are classified as primary particles.	48
164		
165	7.2a Confusion Matrix showing Accuracy of CNN using training data	48
166	7.2b Confusion Matrix showing Accuracy of CNN using testing data	48
167	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.	48
168		
169		
170	7.3 Confusion Matrix of all five particles	49
171		
172	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	50
173		
174		
175	7.5 Training and testing loss of CNN trained on 100,000 images of $\mu/\pi/p/\gamma/e$	50
176	7.6 t-SNE of CNN	51
177	7.7 Muon Prob	52
178	7.8 Pion Prob	52
179	7.9 Proton Prob	53
180	7.10 Electron Prob	53
181	7.11 Gamma Prob	54
182	7.12 Prob	54
183	7.13 mupi	55

184	7.14 mupi	55
185	7.15 mupi	56
186	7.16 mup	56
187	7.17 mup	57
188	7.18 mup	57
189	7.19 mue	58
190	7.20 mue	58
191	7.21 mue	59
192	7.22 mug	59
193	7.23 mug	60
194	7.24 mug	60
195	8.1 Snapshot of passing rates of Selection I from CC-Inclusive Filter	62
196	8.2 Results of CNN10000 classification of track candidate images output from cc-inclusive filter.	63
198	8.2a Confusion Matrix showing Accuracy of CNN using data with wrong normilazion	63
200	8.2b Probability plot showing μ/π separation of CNN using wrong normalization	63
202	8.2c Confusion Matrix showing Accuracy of CNN using data with correct normilazion	63
204	8.2d Probability plot showing μ/π separation of CNN using correct normalization	63
206	8.3 CNN10000 distributions of track candidate images output from Selec- tion I Original cc-inclusive filter with different image data normalizations	64
208	8.3a Track range distribution of events from Selection I Original passing CNN with 70% accuracy using image data with wrong normilazion	64

211	8.3b	Track range distribution of events from Selection I Original passing CNN with 70% accuracy using image data with correct normilazion	64
212			
213			
214	8.3c	Momentum distribution of events from Selection I Original passing CNN with 70% accuracy using image data with wrong normilazion	64
215			
216			
217	8.3d	Momentum distribution of events from Selection I Original passing CNN with 70% accuracy using image data with correct normilazion	64
218			
219			
220	8.4	Snapshot of passing rates of all cuts from Selection I Modified cc-inclusive filter	65
221			
222	8.5	Confusion matrix and probability plot of events passing selection I modified cc-inclusive cuts right before 75cm track length cut	66
223			
224	8.5a	Confusion Matrix for CNN10000 classified events from selection I modified	66
225			
226	8.5b	Probability plot for CNN10000 classified events from selection I modified	66
227			
228	8.6	CNN10000 distributions of track candidate images output from Selection I Modified cc-inclusive filter	67
229			
230	8.6a	Track range distribution of events from Selection I Modified passing CNN with 70% accuracy	67
231			
232	8.6b	Stacked signal and background track range distributions from Selection I Modified passing CNN with 70% accuracy	67
233			
234	8.6c	Stacked signal and background track range distributions from Selection I Modified passing 75 cm track length cut	67
235			
236	8.6d	Stacked signal muons and background muons/pions of track range distributions from Selection I Modified passing CNN with 70% accuracy	67
237			
238			

239	8.6e	Stacked signal muons and background muons/pions of track range distributions from Selection I Modified passing 75 cm track length cut	67
240			
241			
242	8.7	CNN10000 stacked signal/background track range distributions of track candidate images output from Selection I Modified cc-inclusive filter	68
243			
244	8.7a	Stacked signal muons and background muons/pions of track range distributions from Selection I Modified passing CNN with 75% accuracy	68
245			
246			
247	8.7b	Stacked signal muons and background muons/pions of track range distributions from Selection I Modified passing CNN with 80% accuracy	68
248			
249			
250	8.7c	Stacked signal muons and background muons/pions of track range distributions from Selection I Modified passing CNN with 85% accuracy	68
251			
252			
253	8.7d	Stacked signal muons and background muons/pions of track range distributions from Selection I Modified passing CNN with 90% accuracy	68
254			
255			
256	8.8	CNN10000 momentum distributions of track candidate images output from Selection I Modified cc-inclusive filter	69
257			
258	8.8a	Momentum distribution of events from Selection I Modified passing CNN with 70% accuracy	69
259			
260	8.8b	Stacked signal and background momentum distributions from Selection I Modified passing CNN with 70% accuracy	69
261			
262	8.8c	Stacked signal and background momentum distributions from Selection I Modified passing 75 cm track length cut	69
263			
264	8.9	Track distribution comparisons of true CC muons plotted vs true CC muons and pions plotted	71
265			
266	8.9a	Stacked signal μ /background μ and π track range distribution of CNN @ 70%	71
267			

268	8.9b Stacked signal $\mu\&\pi$ /background $\mu\&\pi$ track range distribution of CNN @ 70%	71
270	8.10 Images of true CC events where the pion was the tagged track candidate	71
271	8.10a Pion reconstructed track range is less than 75 cm and longer than muon track due to dead wires	71
273	8.10b Pion reconstructed track range is less than 75 cm and larger than muon reconstructed track	71
275	8.10c Pion reconstructed track range is greater than 75 cm and larger than muon reconstructed track	71
277	8.11 Snapshot of signal and background event numbers of Selection I modified from cc-inclusive note [?]	74
279	8.12 CNN performance of classified muons and pions compared to the already implemented 75 cm track length cut	74

²⁸¹ List of tables

²⁸² 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.	²⁸³ 29
²⁸⁸ 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 ν_μ 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 ν_μ 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. . .	²⁹⁹ 72

304	8.2 Signal and background event numbers at modified selection level with	
305	CNN cut estimated from a BNB+Cosmic sample and Cosmic only sam-	
306	ple normalized to $5 * 10^{19}$ PoT. The last column gives the fraction of	
307	this signal or background type to the total selected events per CNN	
308	probability.	72

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

— Shirley Chisholm

³¹¹ **Chapter 1**

³¹² **Introduction**

³¹³ This thesis will be a description of work done to further increase efficiency and purity
³¹⁴ of the charged current inclusive cross section measurement using the MicroBooNE
³¹⁵ detector. It will also describe the MicroBooNE detector, what neutrinos are, the charged
³¹⁶ current inclusive cross section measurement and its importance as well as convolutional
³¹⁷ neural networks and how they can be used in μ/π separation. Chapter 2 will talk
³¹⁸ about the background of neutrinos and the people and detectors that discovered
³¹⁹ neutrinos as well as an in depth history of neutrino oscillation and the discovery that
³²⁰ neutrinos have mass. Chapter 3 will discuss the MicroBooNE experiment, specifically,
³²¹ how Liquid Argon Time Projection Chambers work, the Light Collection System and
³²² the Electronic and Readout Trigger systems. Chapter ?? will describe the Booster
³²³ Neutrino Beam sationed at Fermi National Accelerator Lab. Chapter 4 will discuss the
³²⁴ work that was done to detect the first neutrinos seen in the MicroBooNE detector and
³²⁵ the software reconstruction efforts required to create an automated neutrino ID filter
³²⁶ that was used to find the first neutrinos and then was later expanded on to create the
³²⁷ charged current inclusive filter that will be discussed in chapter 5 Chapter 6 will give a
³²⁸ brief description of what Convolutional Neural Networks are and how it will be used
³²⁹ for μ/π separation in this selection. Chapter 7 will discuss the hardware frameworks
³³⁰ and training methods used to train multiple Convolutional Neural Networks for use
³³¹ in the charged current inclusive cross section measurement. Chapters 8, ?? and ?? will
³³² discuss the results of using Convolutional Neural Networks on monte-carlo and data
³³³ to sift out charged current inclusive neutrino events.

³³⁴ Chapter 2

³³⁵ Neutrinos

³³⁶ 2.1 What are Neutrinos

³³⁷ Neutrinos are one of the fundamental particles which make up the universe. They are
³³⁸ also one of the least understood. Neutrinos are not affected by the electromagnetic
³³⁹ forces because they do not have electric charge. Neutrinos are affected by a "weak"
³⁴⁰ sub-atomic force of much shorter range than electromagnetism, and are therefore able
³⁴¹ to pass through great distances in matter without being affected by it. Until the late
³⁴² 90's, neutrinos were thought to have no mass. Due to their mass, neutrinos are also
³⁴³ affected by gravity. Neutrinos are created by radioactive decay or nuclear reactions
³⁴⁴ such as the ones that happen in the sun, in nuclear reactors or when cosmic rays hit
³⁴⁵ atoms. There are three types of neutrinos, ν_e , ν_μ and ν_τ which correspond to their
³⁴⁶ charged lepton pairs.

³⁴⁷ As previously stated, neutrinos are very weakly interacting; in fact, neutrinos can
³⁴⁸ pass unscathed through a wall of lead several hundred light-years thick. Because
³⁴⁹ neutrinos interact so rarely, studying neutrinos requires a massive detector and a
³⁵⁰ powerful neutrino source. With that being said, we can only infer their existence when
³⁵¹ they interact in a detector. In a collision, distinct charged particles are produced with
³⁵² each type of neutrino. An electron neutrino will create an electron, a muon neutrino
³⁵³ will create a muon, and a tau neutrino will create a tau. The track the particle leaves
³⁵⁴ in the detector is how one figures out what type of neutrino interaction was "seen".
³⁵⁵ Liquid Argon Time Projection Chambers are the newest type of detectors being used to
³⁵⁶ study neutrinos due to their excellent imaging and particle identification capabilities.

³⁵⁷ 2.2 History of Neutrinos

³⁵⁸ The neutrino was first postulated by Wolfgang Pauli in 1931 to explain how beta
³⁵⁹ decay could resolve the conservation of energy, momentum and angular momentum
³⁶⁰ problem. Pauli suggested that this missing energy might be carried off, unseen, by a
³⁶¹ neutral particle (he called neutron) which was escaping detection. James Chadwick
³⁶² discovered a much heavier nuclear particle in 1932 that he also named neutron, leaving
³⁶³ two particles with the same name. Enrico Fermi was the first person to coin the
³⁶⁴ term neutrino (which means little neutral one in latin) in 1933 to fix this confusion.
³⁶⁵ Fermi's paper, which was published in 1934, unified Pauli's neutrino with Paul Dirac's
³⁶⁶ positron and Werner Heisenberg's neutron-proton model and his theory accurately
³⁶⁷ explained many experimentally observed results. Wang Ganchang first proposed the
³⁶⁸ use of beta capture to experimentally detect neutrinos and in 1959 Clyde Cowan and
³⁶⁹ Frederick Reines published their work stating that they had detected the neutrino.
³⁷⁰ The experiment called for antineutrinos created in a nuclear reactor by beta decay that
³⁷¹ reacted with protons producing neutrons and positrons: $\nu_e + p^+ \rightarrow n^0 + e^+$. Once
³⁷² this happens, the positron finds an electron and they annihilate each other and the
³⁷³ resulting gamma rays are detectable. The neutron is detected by neutron capture and
³⁷⁴ the releasing of another gamma ray. In 1962 Leon M. Lenderman, Melvin Schwartz
³⁷⁵ and Jack Steinberger were the first to detect interactions of the muon neutrino. The
³⁷⁶ first detection of the tau neutrino was announced in the summer of 2000 by the
³⁷⁷ DONUT collaboration at Fermilab. In the late 1960s, many experiments found that the
³⁷⁸ number of electron neutrinos arriving from the sun was around 1/3 to 1/2 the number
³⁷⁹ predicted by the Standard Solar Model. This became known as the solar neutrino
³⁸⁰ problem and remained unresolved for around thirty years. This problem was resolved
³⁸¹ by the discovery of neutrino oscillation and mass. [1]

³⁸² 2.3 Neutrino Oscillations

³⁸³ Neutrino oscillation was first predicted by Bruno Pontecorvo. It describes the phe-
³⁸⁴ nomenon of a neutrino created with a specific lepton flavor (electron, muon or tau)
³⁸⁵ that is later measured to have a different flavor. Neutrino oscillation is important
³⁸⁶ theoretically and experimentally due to the fact that this observation implies that the
³⁸⁷ neutrino has a non-zero mass, which is not part of the original Standard Model of
³⁸⁸ particle physics. [2]

389 2.3.1 Solar Oscillations and the Solar Neutrino Problem

390 The solar neutrino flux derived from Bahcall's Standard Solar Model is shown in figure
391 2.1. Nuclear fusion and decay processes produce an abundant amount of neutrinos.
392 The standard solar model predicts that these reactions produce several groups of
393 neutrinos, each with differing fluxes and energy spectra. The figure also shows the
394 ranges of detection of existing solar neutrino experiments in different shades of blue
395 to illustrate that they sample different portions of the solar neutrino energy spectrum.
396 Three of these experiments, plus a new one, are discussed below.

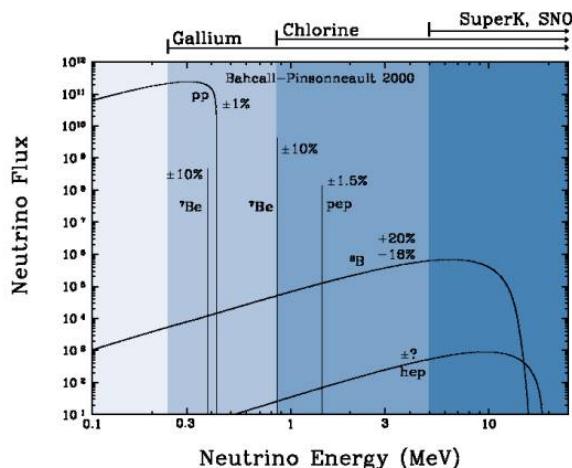


Figure 2.1: The Standard Solar Model

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

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

- ⁴⁰⁹ The unexplained difference between the measured solar neutrino flux and model
⁴¹⁰ predictions lead to the Solar Neutrino Problem. [4]



⁴¹¹ While it is now known that the Homestake Experiment detected neutrinos, some
⁴¹² physicist were weary of the results. Conclusive evidence of the Solar Neutrino Problem
⁴¹³ was provided by the Kamiokande-II experiment, a water Cherenkov detector with
⁴¹⁴ a low enough energy threshold to detect neutrinos through neutrino-electron elastic
⁴¹⁵ scattering. In the elastic scattering interaction the electrons coming out of the point of
⁴¹⁶ reaction strongly point in the direction that the neutrino was traveling, away from the
⁴¹⁷ sun. While the neutrinos observed in Kamiokande-II were clearly from the sun, there
⁴¹⁸ was still a discrepancy between Kamiokande-II and Homestake; The Kamiokande-
⁴¹⁹ II experiment measured about 1/2 the predicted flux, rather than the 1/3 that the
⁴²⁰ Homestake Experiment saw.

⁴²¹ The solution to the solar neutrino problem was finally experimentally determined
⁴²² by the Sudbury Neutrino Observatory(SNO). The Ray Davis's Homestake Experiment
⁴²³ was only sensitive to electron neutrinos, and the Kamiokande-II Experiment was
⁴²⁴ dominated by the electron neutrino signal. The SNO experiment had the capability to
⁴²⁵ see all three neutrino flavors. Because of this, it was possible to measure the electron
⁴²⁶ neutrinos and total neutrino flux. The experiment demonstrated that the deficit was
⁴²⁷ due to the MSW effect, the conversion of electron neutrinos from their pure flavor
⁴²⁸ state into the second neutrino mass eigenstate as they passed through a resonance
⁴²⁹ due to the changing density of the sun. The resonance is energy dependent, and is
⁴³⁰ visible near 2MeV. The water cherenkov detectors only detect neutrinos above about
⁴³¹ 5MeV, while the radiochemical experiments were sensitive to lower energy (0.8MeV
⁴³² for chlorine, 0.2MeV for gallium), and this turned out to be the source of the difference
⁴³³ in the observed neutrino rates at the two types of experiments. Figure 2.2 shows
⁴³⁴ Homestake, Kamiokande-II and SNO experiments.

⁴³⁵ MSW Effect

⁴³⁶ The Mikheyev-Smirnov-Wolfenstein effect is a process which acts to modify neu-
⁴³⁷ trino oscillations in matter. The presence of electrons in matter changes the energy

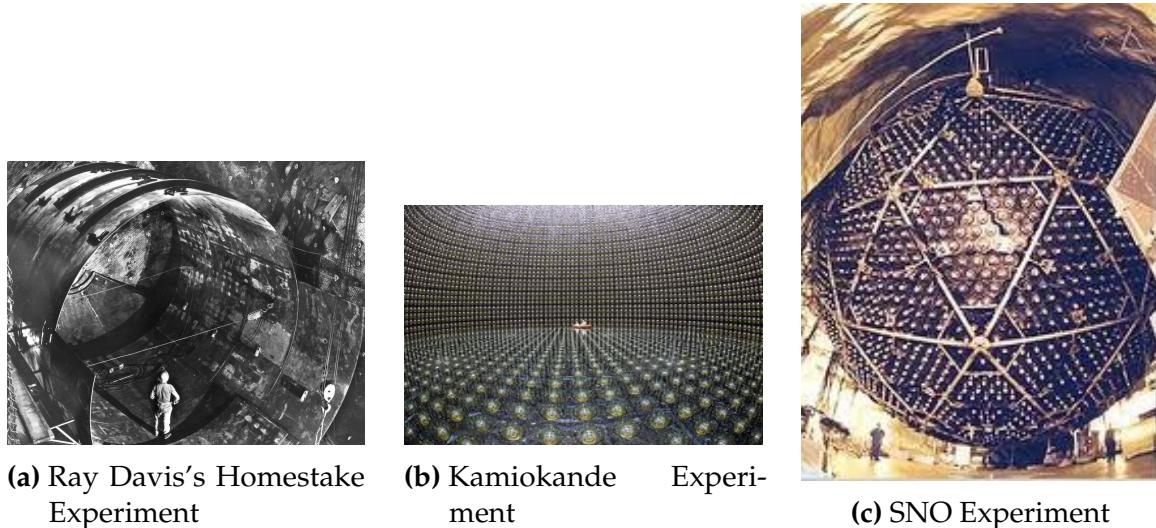


Figure 2.2: Solar Neutrino Experiments

438 levels of the mass eigenstates of neutrinos due to charged current coherent forward
 439 scattering of the electron neutrinos. This coherent forward scattering is similar to
 440 the electromagnetic process with respect to the refractive index of light in a medium.
 441 Because of this MSW Effect, neutrinos in vacuum have a different effective mass than
 442 neutrinos in matter and because neutrino oscillations depend on the squared mass
 443 difference of the neutrinos, the neutrino oscillations are different in matter than in
 444 vacuum. This effect is important at the sun where electron neutrinos are produced.
 445 The neutrinos of high energy leaving the sun are in a vacuum propagation eigenstate
 446 ν_2 that has a very small overlap with the electron neutrino $\nu_e = \nu_1\cos(\theta) + \nu_2\sin(\theta)$
 447 seen by the charged current reactions in Kamiokande-II and SNO. The discrepancy of
 448 the deficit between SNO, Kamiokande-II and Homestake is due to the energy of the
 449 solar neutrinos. The MSW effect "turns on" at about 2MeV and at lower energies, this
 450 MSW effect is negligible. [5]

451 2.3.2 Atmospheric Oscillations and the Atmospheric Neutrino 452 Anomaly

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

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

456

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



Figure 2.3: Cosmic Ray Shower

457 Figure 2.3 shows the cosmic ray shower. In general, these neutrinos have energies
 458 from 1GeV to 100s of GeV and the ratio of ν_μ s to ν_e s equals to 2 (see equation 2.4)

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

459 There have been two types of detectors used to study atmospheric neutrinos: Water
 460 Cherenkov detectors and tracking calorimeters. Super-Kamiokande is the detector we
 461 will focus on. These atmospheric detector experiments measure the ratio of ν_μ to ν_e .
 462 They also measure the zenith angle distribution of the neutrinos. These experiments
 463 report a double ratio (shown in equation 2.5). This double ratio is the ratio measured
 464 in the detector to the ratio that's expected which is 2. If the double ratio equals to 1, the
 465 data agrees with the prediction. Various measurements from multiple experiments
 466 are shown in figure 2.4. Except for Frejus, all R measurements are less than 1. This
 467 discrepancy between the predicted R and the measured R became known as the
 468 Atmospheric Neutrino Anomaly.

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

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

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

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

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

485 2.3.3 Two Flavor Neutrino Oscillation Formulation

486 The flavor eigenstates can oscillate between each other because they are composed
 487 of an add-mixture of mass eigenstates(ν_1, ν_2). Figure 2.5 shows the mass and flavor
 488 eigenstates rotated by an angle θ which is the mixing angle.

489 In matrix form the wavefunctions are:

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

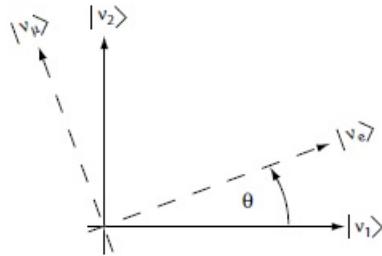


Figure 2.5: The flavor eigenstates are rotated by an angle θ with respect to the mass eigenstates

490 Applying the time evolution operator to ν_μ :

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

491 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
492 being, let us assume $\hbar = c = 1$. With this assumption: $E_1 = \sqrt{p^2 + m_1^2}$ and $E_2 =$
493 $\sqrt{p^2 + m_2^2}$. The next modifications is to assume neutrinos are relativistic:

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

494 because of this,

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

495

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

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

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

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

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

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

⁴⁹⁸ 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:

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

⁴⁹⁹ Finding the Probability for a $\nu_\mu \rightarrow \nu_e$:

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

⁵⁰⁰ Remembering that $\langle \nu_i | \nu_j \rangle = \delta_{ij}$

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

⁵⁰¹ Taking the absolute value squared gives us:

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

⁵⁰² Since the neutrino is relativistic we can set $p = E_\nu$ and change $x = L$. Also
⁵⁰³ recognizing the trigonometric relation $(1 - \cos 2\theta)/2 = \sin^2\theta$ the above equation
⁵⁰⁴ becomes:

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

⁵⁰⁵ All that's left to do now is re-introduce \hbar and c doing this we get:

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

⁵⁰⁶ This equations has three important variables.

- The angle θ : 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, Δm^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 Δm^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 δ_{ij} which is the CP violating term and has not been measured yet, and θ_{13} which has just been measured. CP violation is a violation

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

531 **2.3.5 Reactor Oscillation**

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

539 **Chapter 3**

540 **The MicroBooNE Experiment**

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

546 **3.1 Liquid argon time projection chambers**

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

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

587 3.2 The MicroBooNE Time Projection Chamber

588 MicroBooNE (Micro Booster Neutrino Experiment) is a 89 ton active volume (180 ton
 589 total mass) LArTPC which is then inserted into a cylindrical cryostat on axis of the
 590 Booster Neutrino Beam (BNB) stationed at Fermilab in Batavia, Illinois. Understanding
 591 LArTPC technology and detector physics is necessary to build a LArTPC the size of
 592 DUNE, and MicroBooNE has made many advances in developing this technology [7]
 593 [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° 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.

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

631 Both the light collection system and the TPC create analog signal that is read out and
632 digitized by the electronics system. The process requires amplification and shaping of
633 the signal which then goes to the data acquisition (DAQ) software for writing of the
634 digitized data to disk. The anode plane wires are connected to detector specific circuit
635 boards (ASICS) that are submerged and operate inside the liquid argon volume. These
636 ASICS send amplified signal to 11 feed-throughs where further amplification of the
637 signal happens outside the cryostat. The signal is received by custom LArTPC readout
638 modules distributed over nine readout crates which do the digitization. The TPC wires
639 are digitized at 16 MHz then downsampled to 2 MHz. The TPC system reads out 4
640 frames of wire signal data per event, 1 frame before a trigger and 2 frames after the
641 triggered frame. The four frames allows for identification of a neutrino interaction as
642 well as cosmic background rejection. The process of digitization is similar for the light
643 collection system. Each PMT signal undergoes a shaping with a 60 ns peaking time
644 for digitization of multiple samples. The digitization occurs at 64 MHz but are not
645 read out continuously during the TPC readout time. Only shaped PMT signal samples
646 above a small threshold are read out and saved. Both the TPC and PMT readouts are
647 initiated via triggers on a separate trigger board located in a warm electronics crate.
648 The timing trigger is created by a timing signal from the BNB accelerator which is
649 shaped and sent to the trigger board. The PMT trigger is generated when the PMT
650 signal multiplicity is greater than 1 and the summed PMT pulse-height is more than 2
651 photo-electrons summed up over all PMT channels. When the trigger board gets both
652 a timing trigger and a PMT trigger in coincidence, at BNB trigger is then generated by
653 the board. This signal is then passed to all readout crates initiating the readout of data.
654 The data is then sent to the DAQ software which then saves the data to disk into one
655 event memory.

656 3.3 MicroBooNE's Physics Goals

657 3.3.1 The low-energy excess

658 The primary goal of the MicroBooNE experiment is to study and investigate the low-
659 energy excess seen in MiniBooNE. 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.

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.

3.3.3 Astroparticle physics

99% of energy leaving a supernova leaves in the form of neutrinos which can be seen by detectors on earth's surface. MicroBooNE will have a dedicated supernova data stream and a connection to the SuperNova Early Warning System [?] so when a nearby supernova explosion occurs, continuous data will be written for several hours that can be later analyzed for supernova neutrino events. This information coupled with information from other experiments can be used to better understand supernovae.[add infograph of supernova here](#)

3.3.4 Liquid argon detector development

The last physics goal for the MicroBooNE collaboration is to provide important information regarding LArTPC technology. Being the first in large scale LArTPCs in the US,

688 MicroBooNE will be able to provide improvements to High Voltage (HV) distribution,
689 Noise Characterization [?], and Michel Electron Reconstruction [8].

690 **3.4 The Booster Neutrino Beam**

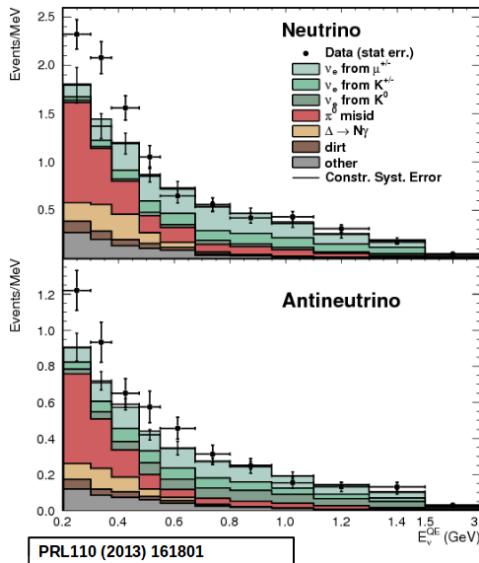
691 The MicroBooNE detector is stationed at Fermi National Accelerator Laboratory
692 (FNAL) where it receives neutrinos from both the Booster Neutrino Beam (BNB)
693 and Neutrinos from the Main Injector (NuMI) beams. MicroBooNE is on-axis for the
694 BNB and off-axis by 135 mrad for NuMI. For the purpose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

699 **3.4.1 Creating the Booster Neutrino Beam**

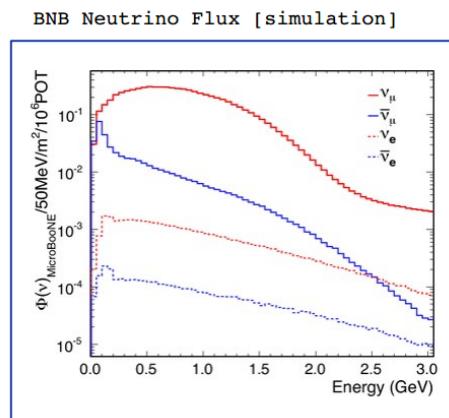
700 The BNB is a very pure ν_μ beam, with only 0.6% contamination from ν_e s. The energy
701 also peaks around 700 MeV which is desired based on the probability of oscillation
702 equation which depends on the the value of L/E , where L is the distance of the
703 detector from the neutrino beam and E is the energy of the neutrino beam. L/E was
704 chosen to increase the probability of seeing neutrino oscillations in the MiniBooNE
705 Low Energy Excess (LEE) range based on the probability of oscillation equation, which
706 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
707 protons from the FNAL booster synchrotron into a beryllium target which produces a
708 high flux of neutrinos. The protons originate from H^2 gas molecules that are turned
709 into H^- ions by a Cockcroft-Walton generator shown in figure ???. The H^- initially are
710 accelerated to 1MeV kinetic energy and are then passed to a linear accelerator using
711 alternating electromagnetic fields to increase their energy to 400MeV. The ions are
712 stripped of electrons by passing them through a carbon foil. The protons are bunched
713 into beam spills which contain $4 * 10^{12}$ protons in a 1.6 μ s time window per spill. It's
714 at this point that the protons are directed towards the beryllium target. The amount
715 of protons directed towards the target (POT) is measured by two toroids upstream of
716 the target with an error of 2%. Beam intensity, timing, width, position, and direction
717 are monitored by beam position monitors, multi-wire chamber and resistive monitors.

The beryllium target is 71.1 cm long, 1.7 proton interaction lengths, and is 0.51 cm in radius. The target is located inside a larger focusing electromagnet called the horn. THe horn is an aluminum alloy pulsed toroidal electromagnet. The pulsed current peaks at 170 kA with a time-width of 143 μ s which coincides with the protons arriving on the target. The current flows from the inner conductor to the outer conductor with a maximum magnetic field of 1.5 Tesla. The magnetic field focuses the charged secondary particles produced by the p-Be interactions. The direction of current can be switched to changed to polarity of the secondary particles being focused creating a beam of either primarily neutrinos, with positively charged secondary particles, or antineutrinos.

Further down the beamline is a concrete collimator which absorbs particles not necessary to the neutrino flux. The collimator is 214 cm long and 30 cm in radius. After the collimator comes a 45 meter long, 1 meter raduis, air-filled cylindrical decay region which then ends in a beam-stop made of steam and concrete. The beam-stop contains an array of gas proportional counters to detect muons. **add beam diagram here**



(a) Low Energy excess seen in Mini-BooNE



(b) Energy spectrum of the Booster Neutrino Beam at Fermi National Laboratories

Figure 3.1: 3.1b Flux of BNB at FNAL.

3.5 Event Reconstruction

⁷³⁵ Chapter 4

⁷³⁶ Neutrino Identification: Finding ⁷³⁷ MicroBooNE's first Neutrinos

⁷³⁸ The goal of the Neutrino Identification analysis was to positively identify BNB neutrino
⁷³⁹ interactions in the MicroBooNE detector collected during the first days of running.
⁷⁴⁰ Neutrino event candidates were identified in part by using a cut on detected flash of
⁷⁴¹ scintillation light during the $1.6 \mu\text{s}$ beam-spill length of the BNB as well as identifying
⁷⁴² reconstructed object from the TPC that are neutrino like. After this selection, 2D
⁷⁴³ and 3D event displays were used for verification of the selection performance. This
⁷⁴⁴ selection was targeted to reduce the ratio of neutrino events to cosmic-only events from
⁷⁴⁵ the initial 1 neutrino to 675 cosmics to a ratio of 1 to 0.5 or better which is equivalent to
⁷⁴⁶ a background reduction by a factor of 1000 or more. These selected events were used
⁷⁴⁷ for MicroBooNE's public displays of neutrino interactions. A clearly visible neutrino
⁷⁴⁸ interaction with an identifiable vertex and at least 2 tracks originating from the vertex
⁷⁴⁹ was what the analysis focused on. This analysis wasn't optimized for high purity
⁷⁵⁰ or efficiency, but rather for very distinguishable neutrino interactions that could be
⁷⁵¹ identified by the public.

⁷⁵² 4.1 Flash Finding

⁷⁵³ Flash finding is the first step used in finding neutrino interactions. This section will
⁷⁵⁴ detail how optical information is reconstructed as well as analysis scripts and event
⁷⁵⁵ filters were used.

756 **4.1.1 Flash Reconstruction**

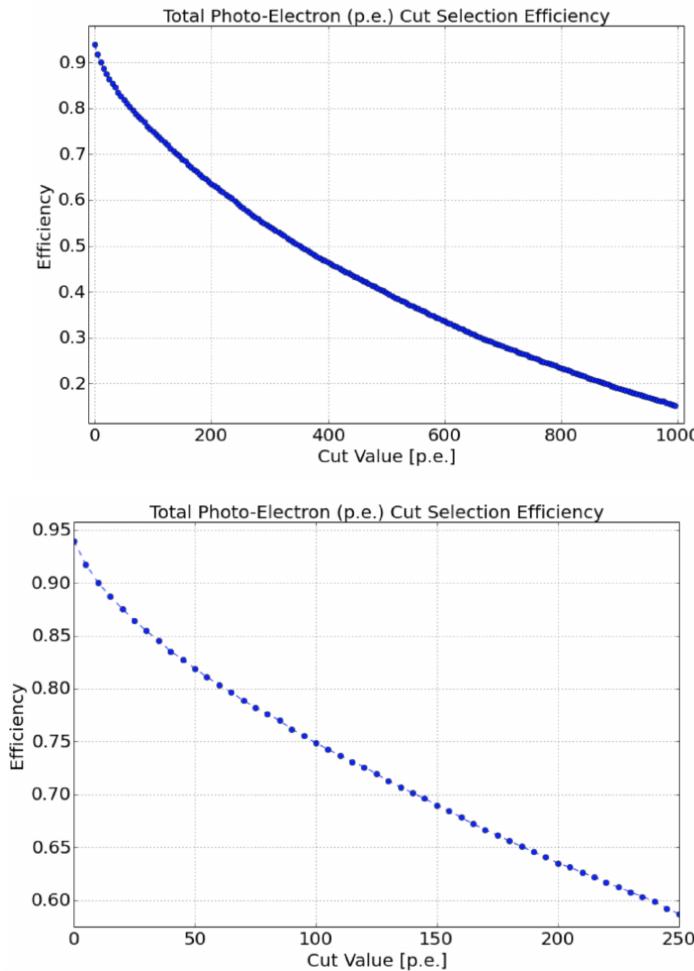
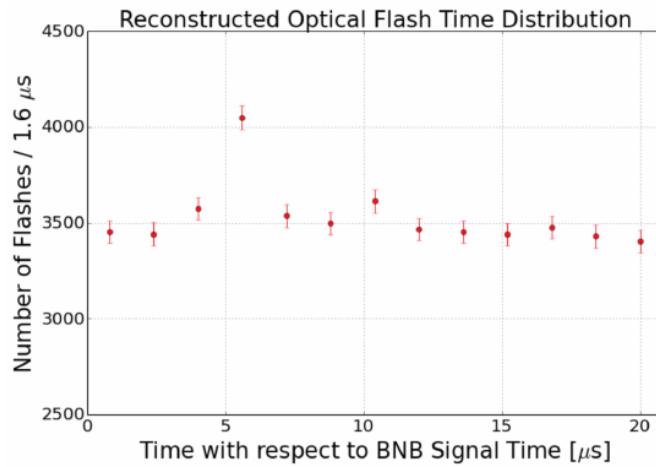


Figure 4.1: Efficiency for selecting beam events as a function of minimum total PE cut for all PE cuts as well as zoomed into interesting region.

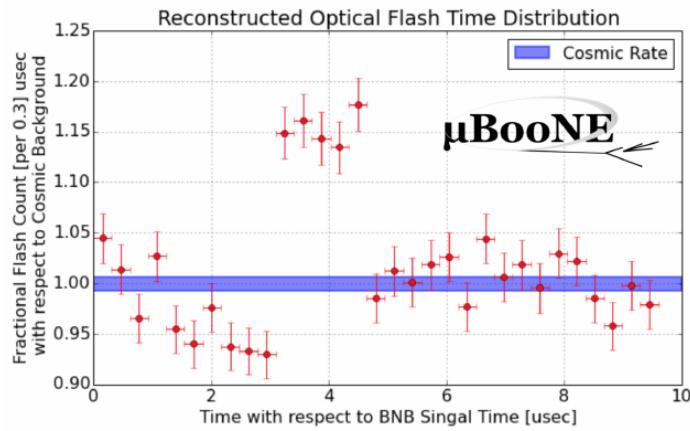
757 A flash is described as a collection of light seen at the same time within the detector.
758 They are then reconstructed by identifying signal from the PMTs above a specific
759 photoelectron (PE) threshold. These signals are called optical hits. Optical hits from
760 all the PMTs are then accumulated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

⁷⁶⁷ cut of 50 PE was deemed sufficient for this analysis. Figure 4.1 show the total PE
⁷⁶⁸ versus the selection efficency of selecting neutrino beam events.

⁷⁶⁹ 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 μ s after the trigger time.

⁷⁷⁰ It is necessary to get the specific time from flashes if one uses flashes to filter out
⁷⁷¹ neutrino interactions coincident with the neutrino beam spill period and background.
⁷⁷² Before a filter can be applied, an understanding of the timing of the trigger and PMT

773 readout with respect to the arrival of neutrinos from the BNB. To do this, a $1.6 \mu\text{s}$
774 window near the expected beamtime was created and verified by finding that the
775 number of flashes was significantly above the cosmic-ray background flashes. Beam
776 data during the first week of running, October 16th 2016 through October 22nd 2016
777 and were used for a timing measurement. The total POT uses corresponds to roughly
778 24 hours of data taking at nominal intensity ($4 \times 10^{12} \text{ ppp}$) and a 5 Hz repetition rate.
779 Figure 4.2a shows size of the expected neutrino signal in time using Monte Carlo
780 predictions and figure 4.2b shows the neutrino signal in data. The intensity in data is
781 lower, however there can still be seen a significant excess above data.

782 **4.1.3 Event Rates**

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

788 **4.2 TPC Topology Selection**

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

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

⁷⁹⁹ 4.2.1 Cosmic Tagging

⁸⁰⁰ The first step in TPC selection was based on the geometry of cosmic tracks in an event.
⁸⁰¹ The cosmic ray muon geometry tagger runs on 3D tracks and assigns a score to each
⁸⁰² reconstructed track on the likeliness of the track originating from a cosmic. The cosmic
⁸⁰³ scores are detailed below:

- ⁸⁰⁴ • 1: The track is tagged as entering or exiting the TPC
- ⁸⁰⁵ • 0.95: The track is a delta ray associated with a tagged track
- ⁸⁰⁶ • 0.5: The track is either entering or exiting, but not both
- ⁸⁰⁷ • 0.4: The track is entering or exiting through the Z boundary
- ⁸⁰⁸ • 0: The track isn't tagged

⁸⁰⁹ Clusters are assigned either a 0 or 1, 1 being a cosmic. In simulation, 90% of cosmics
⁸¹⁰ are tagged as cosmics. These tracks are no longer considered when looking for a
⁸¹¹ neutrino topology. Requiring that the tracks be contained in turn affects the neutrino
⁸¹² efficiency by 20%. The algorithm checks that each track is contained within a boundary
⁸¹³ region of 10 cm from all sides of the TPC. This boundary region was optimized via
⁸¹⁴ handscanning of experimental data.

⁸¹⁵ As can be expected, cosmic tagging is more efficient in the 3D channel (tracks) than
⁸¹⁶ the 2D channel (clusters) because the reconstructed tracks can use the full 3D position
⁸¹⁷ information of the entering and exiting points while the 2D channel mainly use the
⁸¹⁸ reconstructed x position of the cluster which is associated to timing.

⁸¹⁹ Cosmic tagging uses timing information to reject tracks and clusters that are outside
⁸²⁰ of drift window. The drift window for 128 kV is $1.6 \mu\text{s}$ while for 70 kV, the actual
⁸²¹ voltage MicroBooNE is running at, is $2.3 \mu\text{s}$. Due to this variation between simulation
⁸²² and data, we expect to see $2.3/1.6 = 1.44$ times more cosmic induced tracks or clusters
⁸²³ in the drift window.

⁸²⁴ 4.2.2 2D Cluster Selection

⁸²⁵ This selection was spearheaded by myself and Katherine Woodruff. After looking at
⁸²⁶ experimental cosmics data, 2D clustering performs well, while 3D track reconstruction
⁸²⁷ is affected by more variations in simulation, for example noise filters. This was the

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

834 **Primary Cuts**

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

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

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

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

854 **Secondary Cuts**

855 The secondary cuts look to match long, low-angle clusters with short, high-charge
 856 clusters. Only clusters that have passed previous cuts are used. First clusters with
 857 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.

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

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

870 4.2.3 3D Tracks and vertices Selection

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

- 874 • d : distance between the start points of the two tracks.
- 875 • d_1 : distance between vertex and start of track 1.
- 876 • d_2 : distance between vertex and start of track 2.

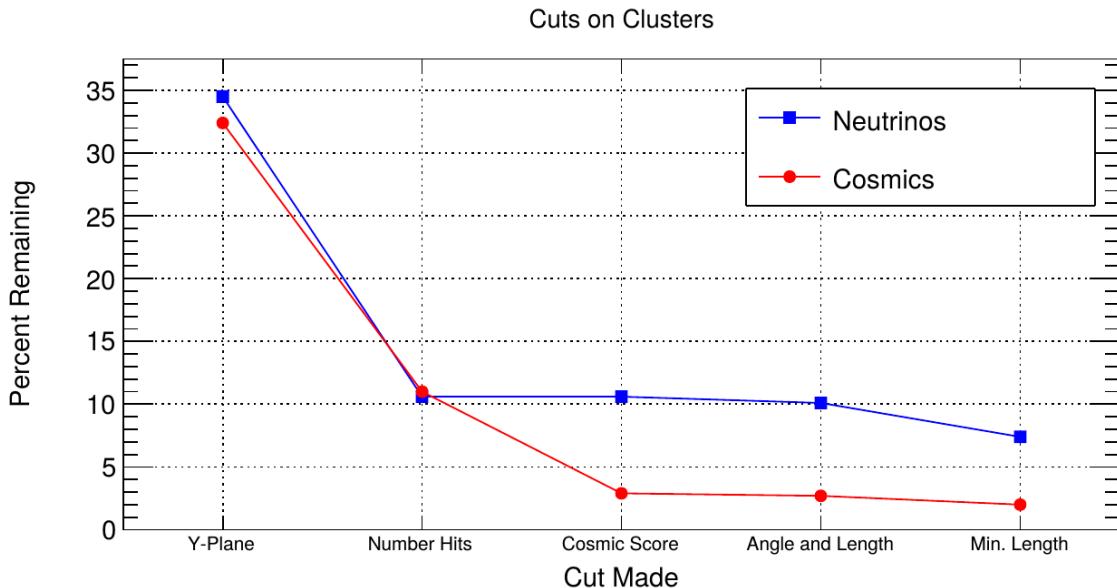


Figure 4.3: Percent of good clusters remaining for neutrinos and cosmics after the primary cuts were applied. This is relative to total number of initial clusters.

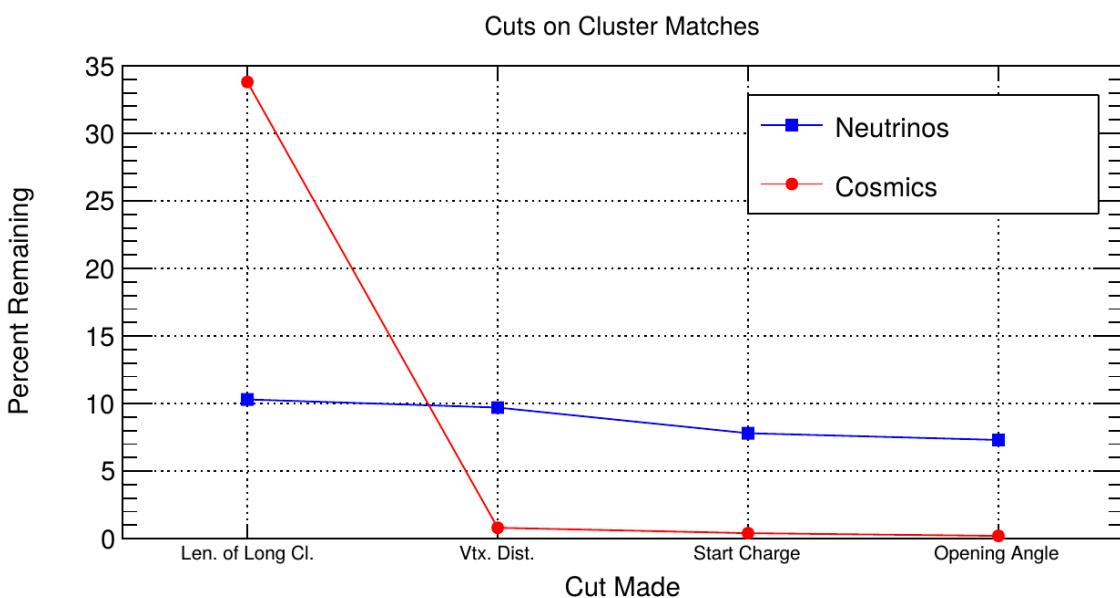


Figure 4.4: Percent matched cluster pairs remaining for neutrinos and cosmics after secondary cuts applied. This is relative to the number of events that contain clusters which pass the primary cuts.

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

- 884 • trackkalmanhit with cccluster $\min(\max_{d,i}) < 3$ cm.
885 • trackkalmanhit with pandoraNu $\min(\max_{d,i}) < 4.5$ cm.
886 • pandoraNu with cccluster $\min(\max_{d,i}) < 5$ cm.

887 4.2.4 TPC Updates

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

894 2D Filter Updates

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

900 3D Filter Updates

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

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

906 **4.3 Conclusion**

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

913 **Chapter 5**

914 **CC-Inclusive Cross Section Selection
915 Filter**

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

932 MicroBooNE requires fully automated event reconstruction and selection algo-
933 rithms for use in the many physics measurements being worked on to date due to
934 the large data rate MicroBooNE receives. Being able to automatically pluck out the
935 neutrino interaction among a sea of cosmics proved to be challenging but was accom-
936 plished. MicroBooNE has developed two complementary and preliminary selection
937 algorithms to select charged-current $\nu_\mu - Ar$ interactions. Both are fully automated

and cut based. The results below focus on the first selection and the “In-Progress” plots presented on the poster associated with this proceeding will focus on further improving this algorithm using Convolutional Neural Network (CNN) implementations. The full details can be found in MicroBooNE public note [?] and for more information of CNN implementation on MicroBooNE data refer to [?]. Selection I is based on cuts developed in a MC performance study described in [?]. It identifies the muon from a neutrino interaction without biasing towards track multiplicity. To combat cosmic and neutral current background, the analysis is strongly biased towards forward-going long tracks which are contained. This limits phase space and reduces acceptance.

The efficiency and purity are used as performance values of selection I. Efficiency is described as the number of selected true ν_μ CC events divided by the number of expected true ν_μ CC events. The purity is described as the number of selected true ν_μ CC events divided by the sum of itself and all the backgrounds. The efficiency of selection I is 12% and the purity is 39.7%. The poster related to this proceedings will focus on the last cut which requires the longest track to be longer than 75 cm. This cut has a passing rate of 30% w.r.t the previous cut and is implemented in part to separate charged-current events from neutral-current events that mimic our signal. Implementing a CNN for $\mu - \pi$ separation picks out differences in these two particles that are track range independent therefore eliminating the need for the 75 cm track length cut and increase efficiency and passing rate at low muon momentum. Figure 5.1a shows the track distribution of selection I and the lack of data below the 75 cm track length cut. Figure 5.1b shows the efficiency of selection I as a function of muon momentum. The selection begins with a cut that requires an optical flash greater than 50 photo electrons (PE) in the 1.6 μs beam window. Next, two or more 3D reconstructed tracks must be within 5 cm from a 3D reconstructed vertex. The most forward going track vertex-track association is then selected for further cuts. The vertex from the chosen association must be in the fiducial volume, and the longest track from this association must be matched to a flash 80 cm in z. Lastly the longest track must be contained and longer than 75 cm.

5.1 The importance of μ/π separation

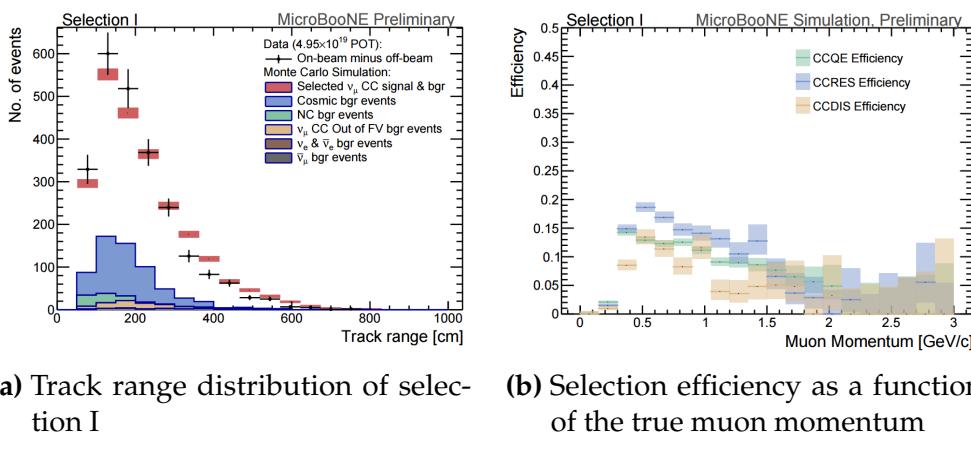


Figure 5.1: **5.1a** Track range distribution for selection I. The track range is defined as the 3D distance between the start and end of the muon candidate track. No data is shown below 75 cm due to the track length cut described previously. **5.1b** Efficiency of the selected events by process quasi-elastic (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.

968 **Chapter 6**

969 **Background on Convolutional Neural
970 Networks**

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

979 **6.1 Image Classification**

980 Image classification is the process of inputting an image into the CNN and receiving a
981 probability of classes that best describes what is happening in the image. As humans,
982 image classification is something that is learned at a very young age and is easy to
983 do without much effort. This is also apparent when hand-scanning LArTPC images.
984 After learning what a neutrino event looks like in MicroBooNE, it is relatively easy
985 to recognize simple neutrino events from cosmic ray background as well as highly
986 ionizing particles like protons from minimum ionizing particles like muons. The very
987 detailed images LArTPC detectors output are prime candidates for input images into
988 a CNN. CNNs mimic a human's ability to classify objects by creating an architecture
989 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.

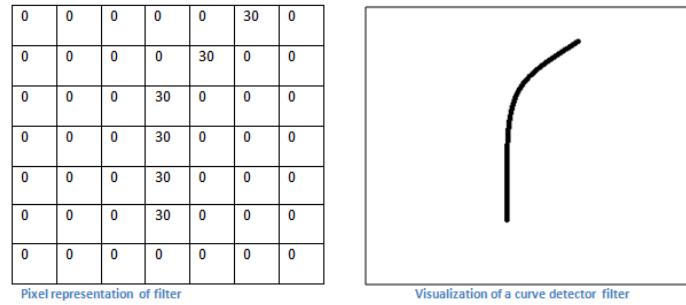


Figure 6.1: Pixel representation and visualization of a curve detector filter. As you can see, in the pixel representation, the weights of this filter are greater along a curve we are trying to find in the input image

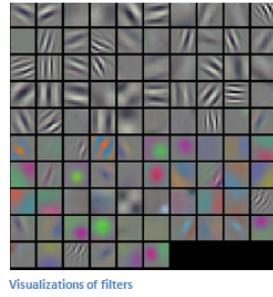


Figure 6.2: Visualization of filters found in first layer of a CNN.

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



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

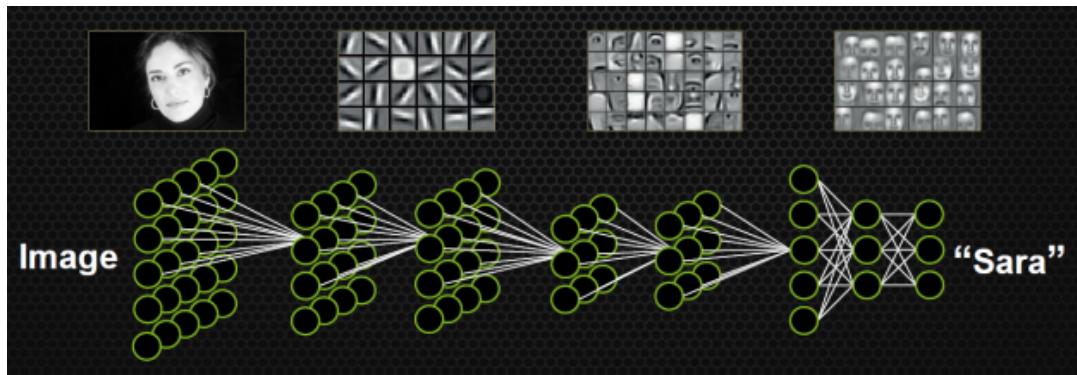


Figure 6.4: Pictorial Representation of Convolutional Neural Networks as well as a visual representation on CNN's complexity of layer feature extraction

1039 There are other layers in a CNN architecture that will not be covered in the scope
 1040 of this thesis but in a general sense, these layers are interspersed between convolution
 1041 layers to preserve dimensionality and control overfitting of the network. The last layer
 1042 is called a fully connected layer and it's job is to output an N dimensional vector where
 1043 N is the number of classes the network has been trained on. Each number in this vector
 1044 represents the probability that the input image is a certain class. Fully connected layers
 1045 use the feature maps of the high level features to compute the products between the
 1046 weights of the previous layer to get the probabilities of each class. These weights are
 1047 then adjusted through the training process using backpropagation.

1048 6.2.1 Backpropagation

1049 A CNN at it's onset has weights that are randomized. The filters themselves don't
 1050 know how to pull out identifying information per class. For a neural network to learn,
 1051 it must be trained on a training set that is labeled. Backpropagation has four seperate
 1052 steps: foward pass, loss function, backward pass and updating weights. In the forward
 1053 pass, a training image is passed through the whole network. All of our weights at this
 1054 time are randomly initialized so the output for the first image will have no preference
 1055 to a specific class. A common loss function is mean squared error (MSE):

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

1056 If we assume that the MSE is the loss of our CNN, the goal would be that our
 1057 predicted label (output of CNN) is the same as our training label. To do this, we need
 1058 to minimize the loss function. To do this, it is necessary to find out which weights most
 1059 directly affect the loss of the network i.e $\frac{dL}{dW}$ where L is our loss function and W are
 1060 the weights of a specific layer. The next step is the backward pass which determines
 1061 which weights contribute the most to the loss and finds ways to adjust these weights
 1062 so that the loss decreases. After the derivative is computed, the last step updates the
 1063 weights in the opposite direction of the gradient.

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

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

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

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

1064 The learning rate is a parameter given to the CNN and it describes the steps the
 1065 network takes to update the weights. Higher learning rate equals large steps and a
 1066 lower training time, but a learning rate that is too large can mean the CNN never
 1067 converges.

1068 Going through backpropagation consists of one training iteration. Once the net-
 1069 work completes a specific number of iterations, another parameter given, and runs
 1070 over all training images that are split up into batches, the process is considered com-
 1071 plete. User input parameters, called hyperparameters, help the network converge to

1072 optimal weights for each layer. Batch size, learning rate, and training iteration are just
1073 some of the user input hyperparameters that help. Lastly, to check if the network has
1074 learned, a different set of labeled images are fed to the CNN iteratively through the
1075 training process to see how well it's learning. This process is especially important to
1076 make sure the network architecture isn't being affected by overfitting (memorizing
1077 training input rather than learning).

1078 6.3 Choosing Hyperparameters

1079 Convolutional neural networks are a relatively new tools in computer vision. Choosing
1080 hyperparameters for your specific dataset is a non-trivial task. Hyperparameters can
1081 range from the amount of layers and filters per layer in an CNN architecture to the
1082 stride the receptive field of a filter takes, not to mention training hyperparameters
1083 such as learning rate and batch size described above. They're ways to optimize these
1084 hyperparameters via hyperparameter optimization using Bayesian Optimization [?]
1085 but as you can imagine, optimizing an CNN architecture from scratch can be very
1086 computationally intensive. For the purpose of this thesis, two well known CNN
1087 architectures were used, AlexNet [?], which won the ImageNet Large-Scale Visual
1088 Recognition Challenge (ILSVRC) in 2012 and therefore bringing awareness of CNNs,
1089 and GoogleNet [?], which won the ILSVRC in 2014, giving rise to deep networks. Both
1090 AlexNet and GoogleNet architectures were used to train on LArTPC images and their
1091 low level filter weights. Higher level filter weights were randomly initialized before
1092 training so the network can learn high level features of LArTPC image classes. The
1093 AlexNet architecture is shown in figure ?? and the GoogleNet architecture is shown in
1094 figure ??

1095 **Chapter 7**

1096 **Training Convolutional Neural
1097 Networks on particles WORKING
1098 TITLE**

1099 Three Convolutional Neural Networks CNNs were trained throughout this analysis.
1100 There are differences to each CNN and will be described fully in the next sections but
1101 the main difference are the amount of particle images used for training and validation.
1102 CNN1075 used 1,075 muons and 10,75 pions for training and the same amount of each
1103 particle for validation. CNN10000 used 10,000 muons and 10,000 pions split in half
1104 for testing and training. Lastly CNN100000 had muons, pions, protons, electrons,
1105 and gammas in its training and validation set. Each particle had 20,000 images and
1106 training and validation was split 90% training, 10% validation. This chapter will also
1107 describe the different hardware frameworks used for training beginning on a CPU
1108 and ending on a GPU cluster.

₁₁₀₉ **7.1 Hardware Frameworks used for Training**

₁₁₁₀ **7.1.1 Syracuse CPU Machine setup**

₁₁₁₁ **7.1.2 Syracuse University GPU Cluster Setup**

₁₁₁₂ **7.2 Convolutional Neural Network Training**

₁₁₁₃ **7.2.1 Image Making Scheme**

₁₁₁₄ **Images used for Traing/Validation of Convolutional Neural Networks**

₁₁₁₅ **add image making for CNN1075** The μ/π image dataset used to train and validate
₁₁₁₆ the CNN10000 was created using single generated isotropic muons and pions from
₁₁₁₇ 0-2 GeV energ range. 10,000 muons and 10,000 pions were used for training and
₁₁₁₈ testing split 50%. The images were created based on wire number and time tick in the
₁₁₁₉ collection plane. Uboonecode v06_23_00 was used instead of v05_08_00 which was
₁₁₂₀ used previously. The wire signal was the raw ADC value after noise filtering. Each
₁₁₂₁ collection plane grayscale image was 3456x1280x1 where 5 time ticks were pooled into
₁₁₂₂ 1 bin which is different than the previous dataset and was implemented due to the fact
₁₁₂₃ that the time ticks of an event went from 9400 to 6400 with the change of uboonecode
₁₁₂₄ version. The grayscale color standard is 8bit therefore the ADC value of wire and time
₁₁₂₅ tick was also downsampled due to the 12bit ADC value MicroBooNE has. To do this,
₁₁₂₆ the highest ADC pixel in the image was found and then this was divided by the rest
₁₁₂₇ placing all pixel values between 0-1. From there, all pixel values are then multiplied
₁₁₂₈ by 255. All images were made using a LArSoft module. Once the images were created,
₁₁₂₉ using and image manipulation framework called OpenCV images were read into a
₁₁₃₀ numpy array and cropped to the region of interest by only keeping rows and columns
₁₁₃₁ where all ADC values are higher than 0 and then resized it to 224x224 using OpenCV's
₁₁₃₂ resize function. This downsampling of ADC values creates a problem of information
₁₁₃₃ loss for example, a proton which is highly ionizing will have the same brightness as a
₁₁₃₄ minimum ionizing muon by virtue of how the images are created. Issues that arose
₁₁₃₅ in CNN1075 that were fixed in CNN10000 include zero-padding images in X and Y
₁₁₃₆ that are smaller than 224X224 to eliminate over-zooming effect and fixing a bug that
₁₁₃₇ shifted pixels separated by a dead-wire region.

1138 Images were also made from events that passed the cc-inclusive selection 1 filter
1139 right before the 75 cm track length cut and were classified using the CNN10000. The
1140 dataset used to create these images is the same one used in [?], prodgenie_bnb_nu_cosmic_uboone_mc
1141 These images were created using information from the track candidate that passed
1142 the filter. Only wire number and time ticks associated to the track candidate were
1143 drawn on the image to mimic a single particle generated image. These images were
1144 then classified using CNN10000. Two approaches were taken in making these images.
1145 The first was using the image normalization above where the maximum pixel in each
1146 image is used as a normalization constant to get all pixels between 0-1 then multiply
1147 all pixels by 255. As described above, this is the incorrect way to normalize; it should
1148 be normalized by dataset not by event, which is the second way the images were
1149 created. The results of CNN10000 performance are shown in section 7.2.

1150 7.2.2 Training CNN1075

1151 The work shown in these next sections are based on the previous work done described
1152 in [?]. That CNN (now referred to as CNN1075) was trained using single generated
1153 isotropic muons and pions from 0-2 GeV energy range. 1,075 muons and pions were
1154 used to train the network and 1,075 μ/π were used as a validation set. The accuracy is
1155 how well CNN1075 is doing by epoch and was 74.5%. The loss is gradient descent
1156 or minimization of the error of the weights and biases used in each neuron of each
1157 layer of CNN1075 and was 58% with a trend sloping downwards on the loss curve
1158 as well as a trend sloping upward in the accuracy curve. Due to the depth of the
1159 neural network framework, it was necessary to train with a larger dataset and for
1160 more epochs, however, the downward slope of the loss curve is an indication that once
1161 trained for longer with a higher training sample, neural networks can be used for μ/π
1162 separation. Updates in the image making and downsampling algorithm were made to
1163 fix issues that arose in CNN1075.

1164 7.2.3 Training CNN10000

1165 The hyperparameters used for CNN10000 are shown. The batch size for the training
1166 and testing as well as the test iter were chosen to encompass the whole training/testing
1167 image set when doing accuracy/loss calculations. To do this, multiplying the test

1168 iter by the test batch size give you the amount of images used when calculating
1169 accuracy/loss curves. For reference, the accuracy and loss are defined as well.

```
1170     • train_batch_size: 100
1171     • test_batch_size: 100
1172     • test_iter: 100
1173     • test_interval: 100
1174     • base_lr: 0.001
1175     • lr_policy: "step"
1176     • gamma: 0.1
1177     • stepsize: 1000
1178     • display: 100
1179     • max_iter: 10000
1180     • momentum: 0.99
1181     • weight_decay: 0.0005
1182     • snapshot: 100
1183     • Accuracy: How often the CNN predicts the truth over total number of images
1184     • Loss: Error between truth and prediction. Minimize loss by gradient descent to
1185       update weights and biases of CNN
```

1186 The same architecure that was used to train CNN1075 was employed on CNN10000,
1187 Imagenet. Caffe [?] was the software package used for both CNNs. The differences
1188 include batch size and test_iter and momentum to account for the larger dataset. Both
1189 CNNs were trained on a CPU machine, Syracuse01. Further training will be done
1190 on a GPU cluster stationed at Syracuse University. Figure 7.1 shows the loss and
1191 accuracy of CNN10000. There is around a 10% increase in accuracy from CNN1075 to
1192 CNN10000, 85%, and around a 20% decrease in loss, 36%.

1193 Figure 7.2 show a breakdown of μ/π separation for CNN10000. It also shows
1194 the network is not being overtrained due to the Accuracy of both the training and
1195 testing datasets being within .01% of eachother. The CNN is doing a very good job of

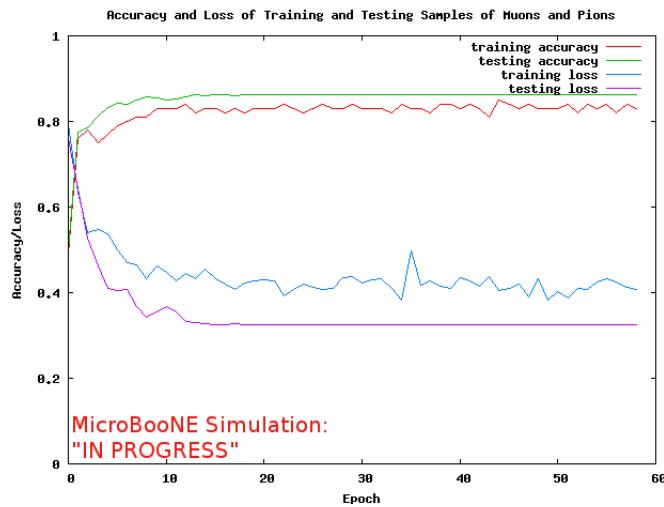
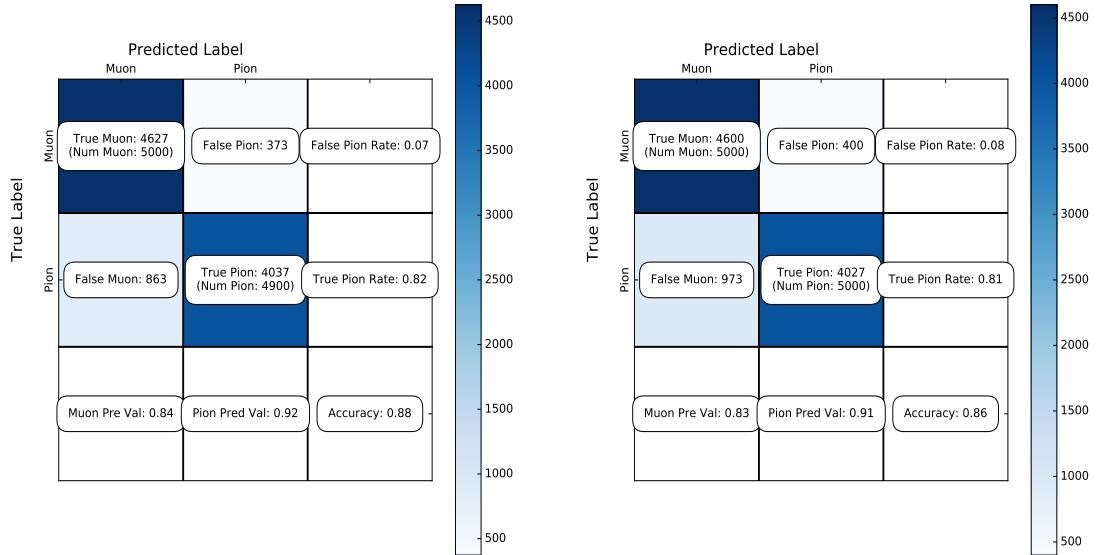


Figure 7.1: Accuracy vs. Loss of ImageNet 2-output μ/π sample consisting of 10000 images each.

₁₁₉₆ classifying true muons as muons, and our loss increase from CNN1075 is due to the
₁₁₉₇ increase in accurately classifying pions as pions.

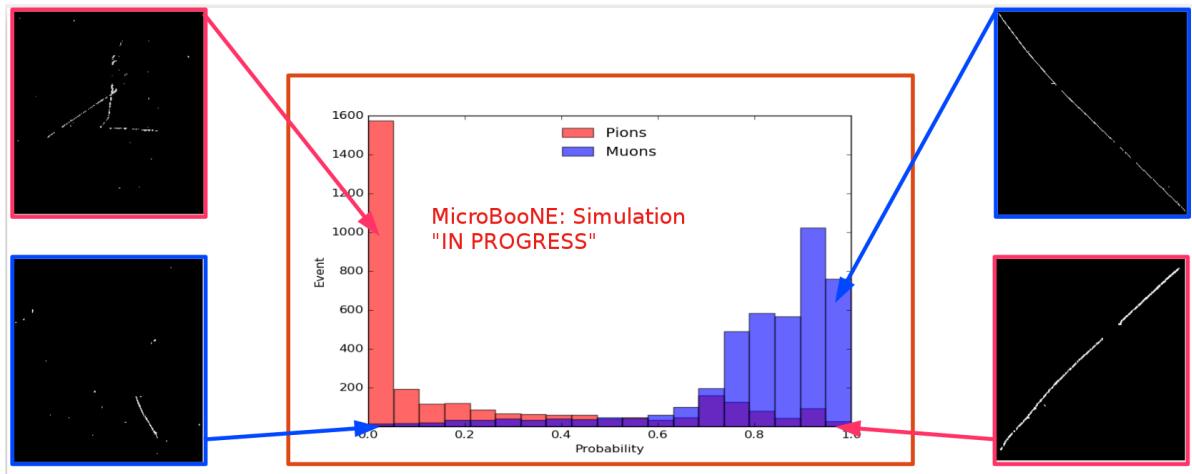
₁₁₉₈ 7.2.4 Training CNN100000

₁₁₉₉ 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.

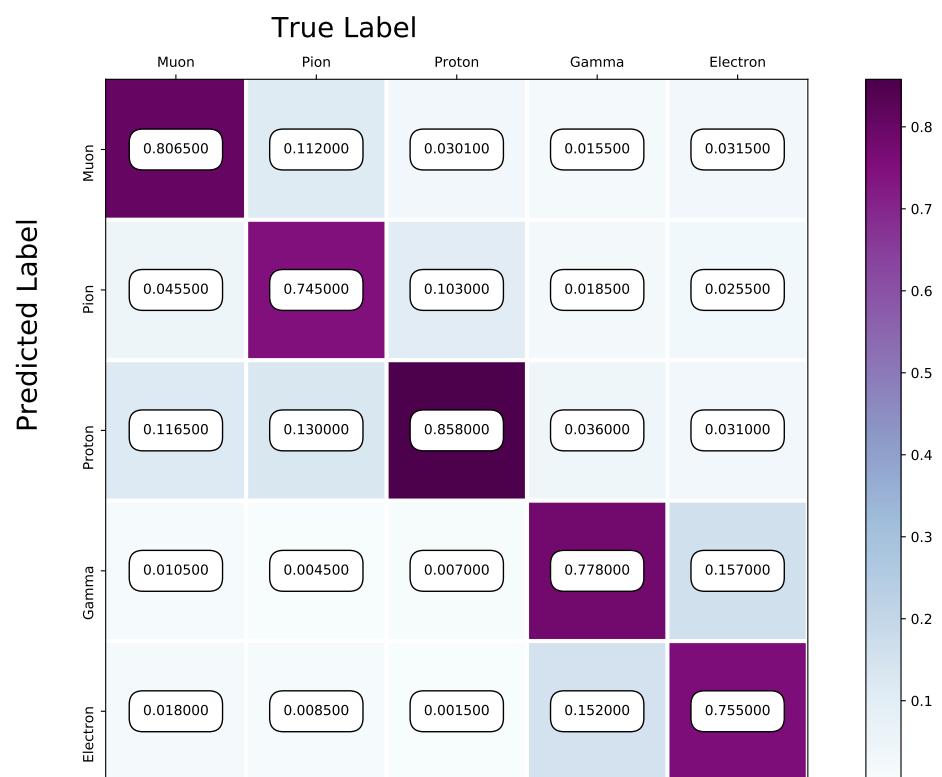


Figure 7.3: Confusion Matrix of all five particles

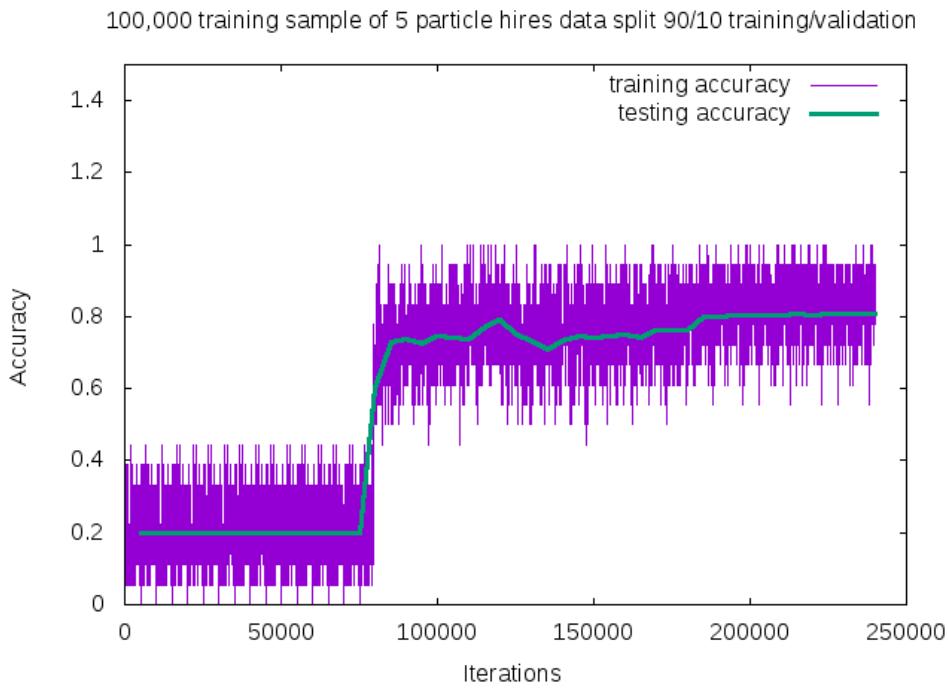


Figure 7.4: Training and testing accuracy of CNN trained on 100,000 images of $\mu/\pi/p/\gamma/e$ with 20,000 images of each particle. Each image was a size of 576x576 and the images per particle were split 90% use for training and 10% used for testing the network

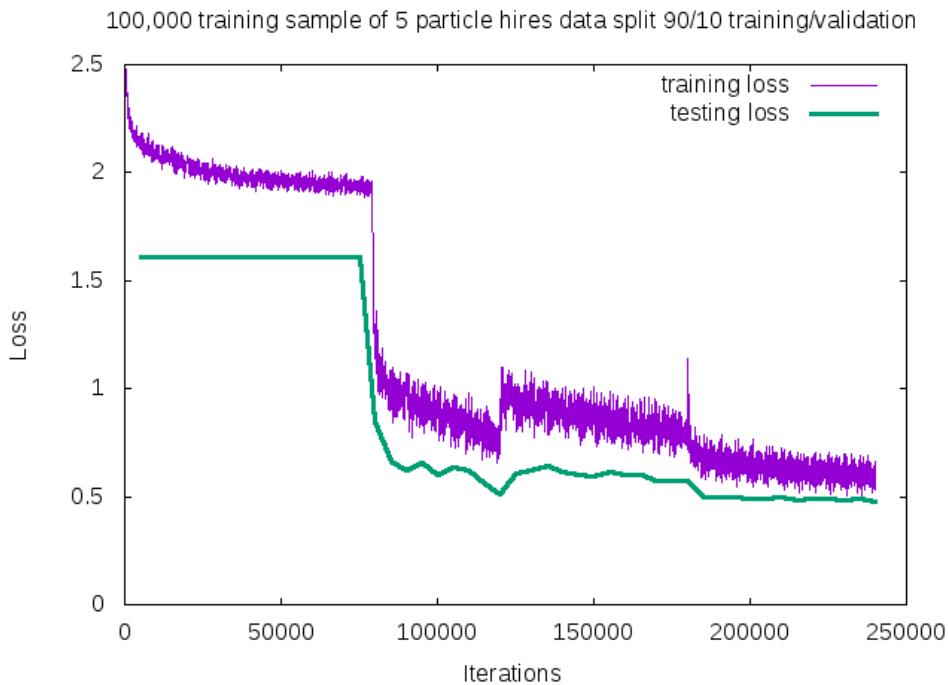


Figure 7.5: Training and testing loss of CNN trained on 100,000 images of $\mu/\pi/p/\gamma/e$

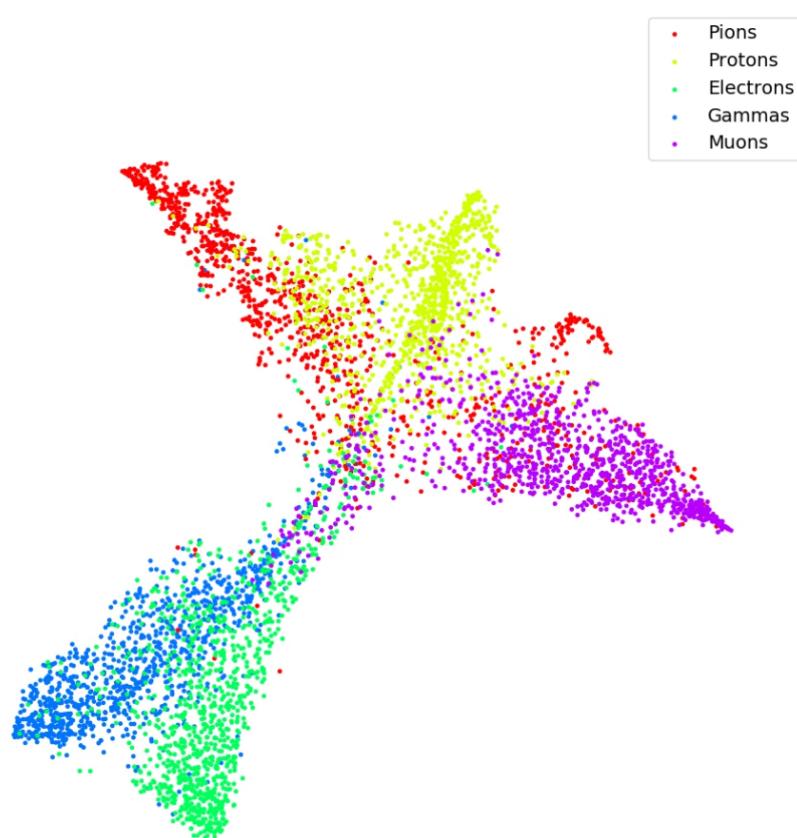


Figure 7.6: t-SNE of CNN

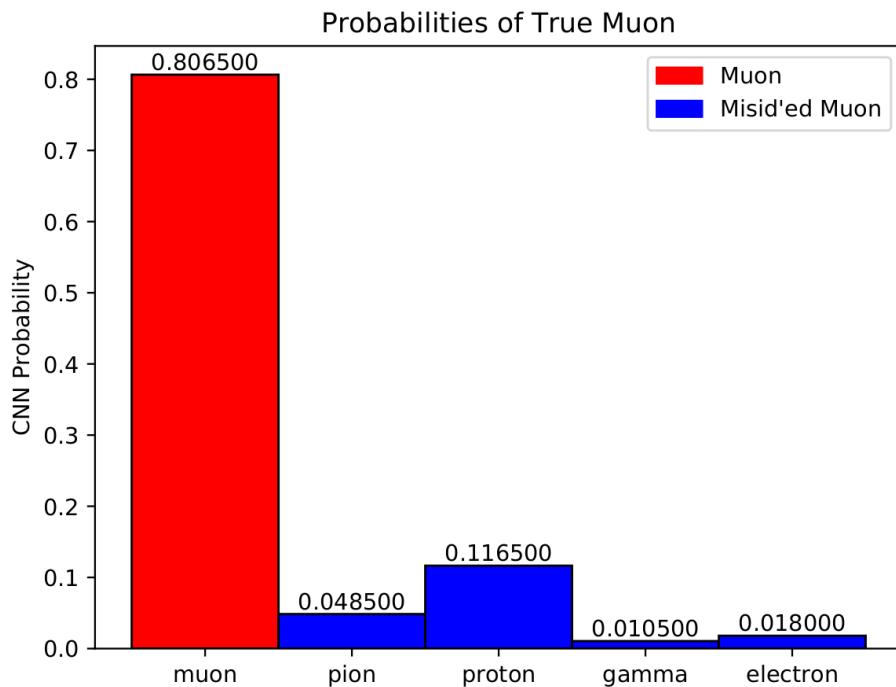


Figure 7.7: Muon Prob

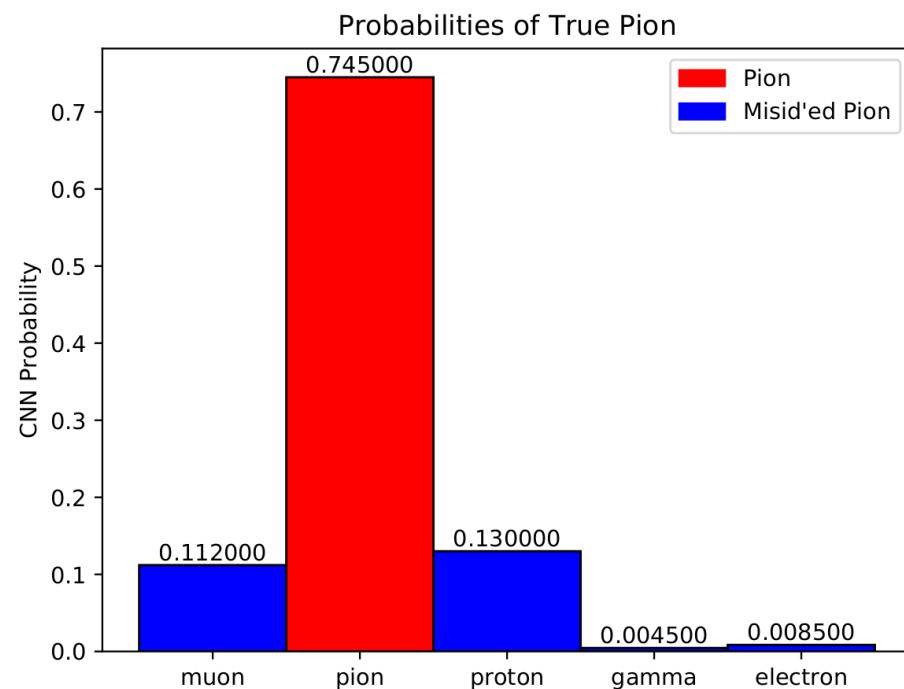


Figure 7.8: Pion Prob

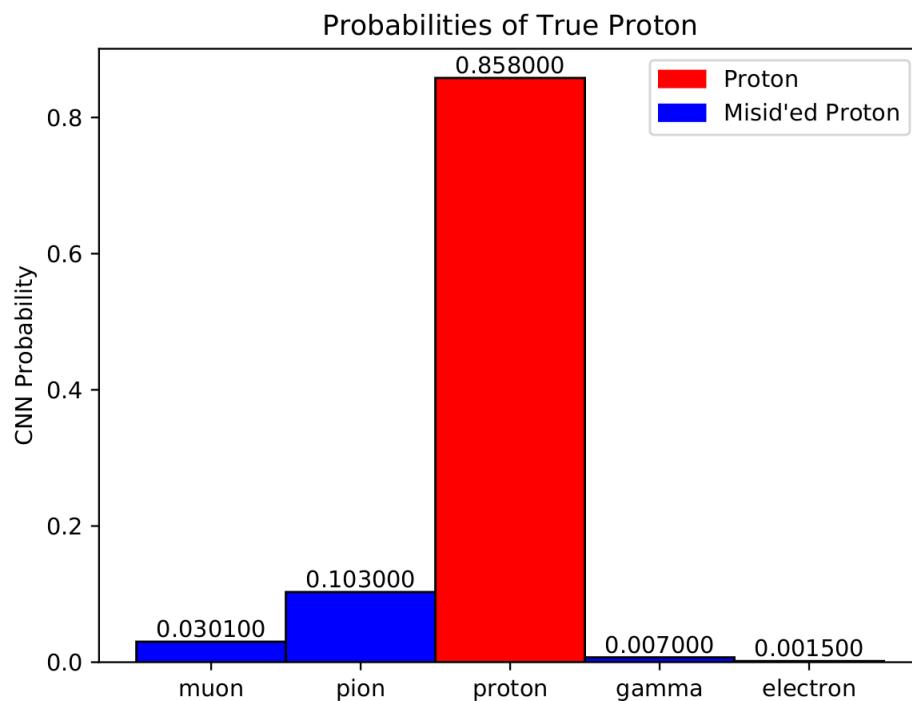


Figure 7.9: Proton Prob

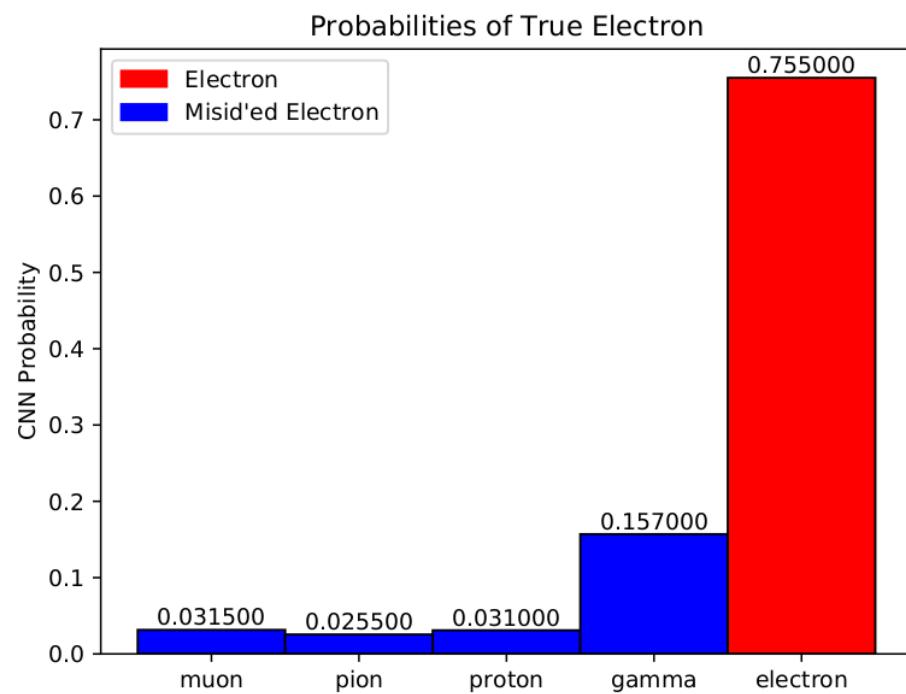


Figure 7.10: Electron Prob

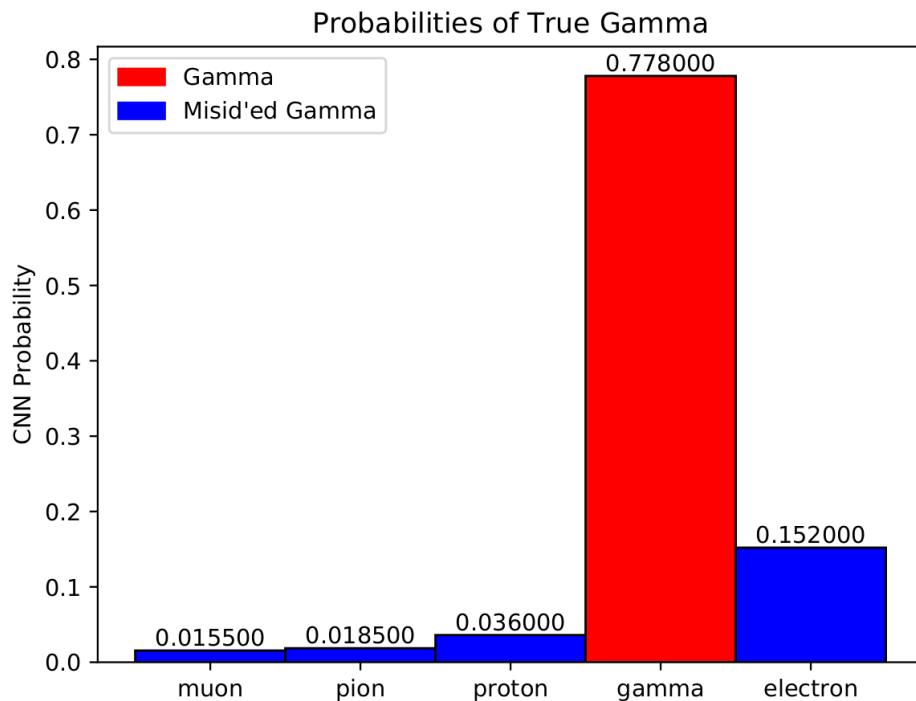


Figure 7.11: Gamma Prob

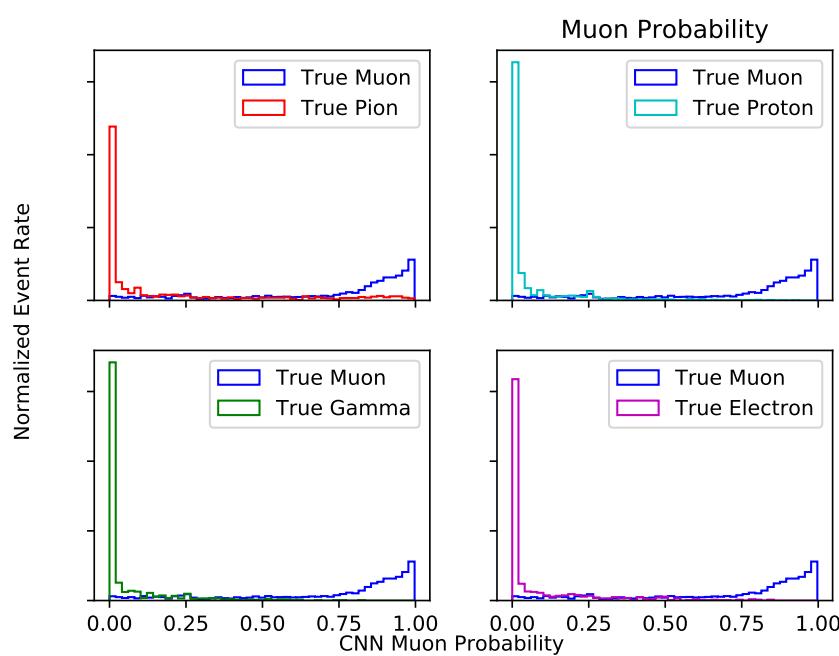


Figure 7.12: Prob

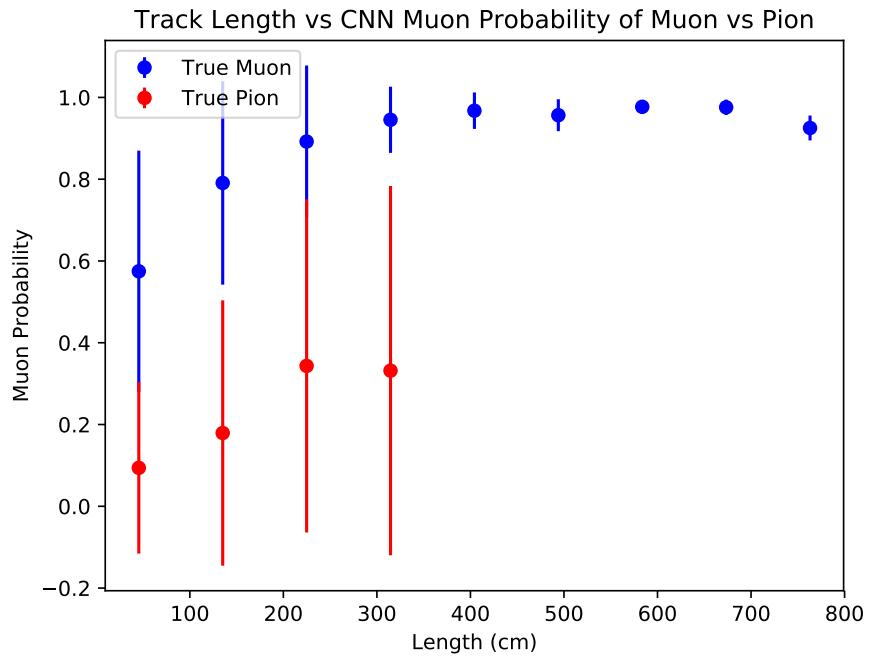


Figure 7.13: mupi

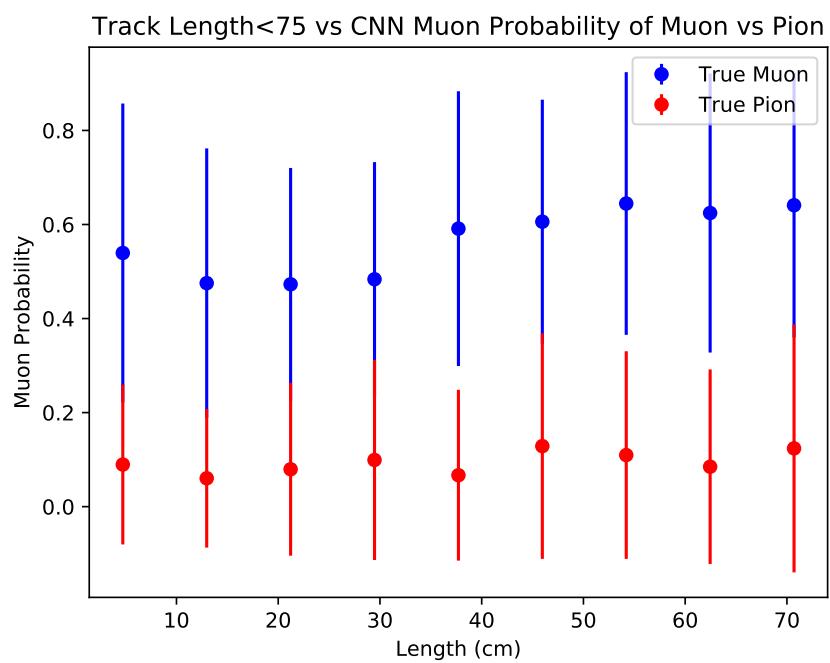
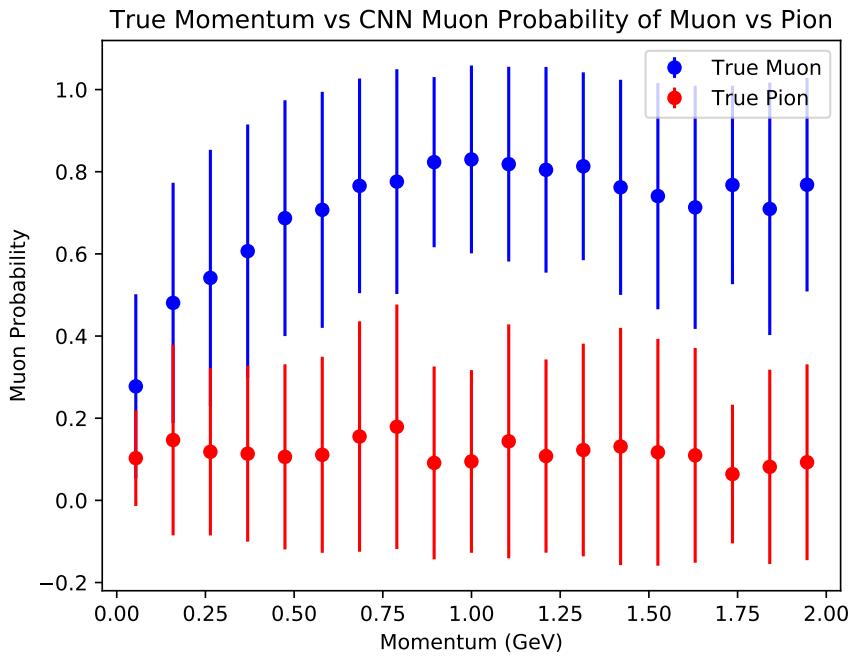
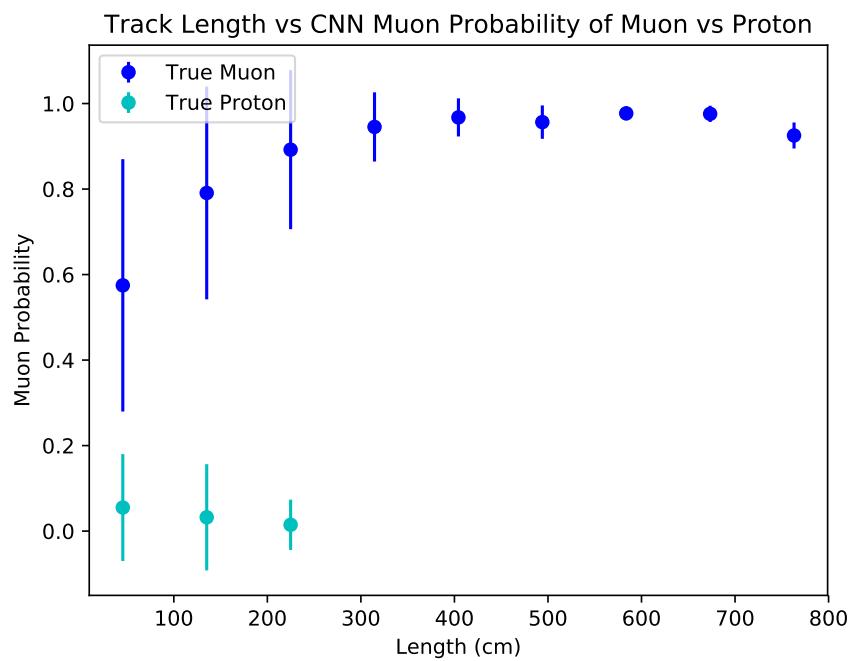


Figure 7.14: mupi

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

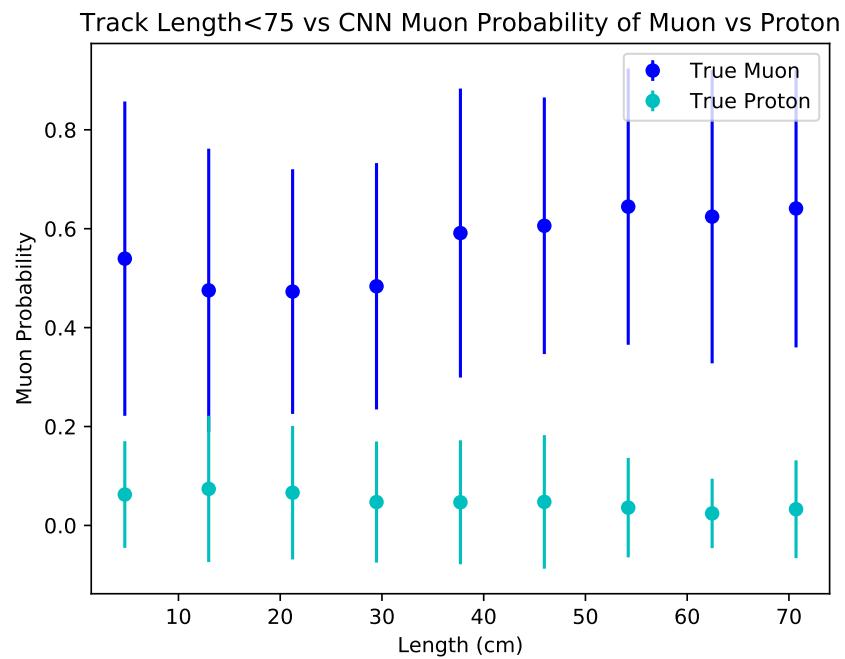


Figure 7.17: mup

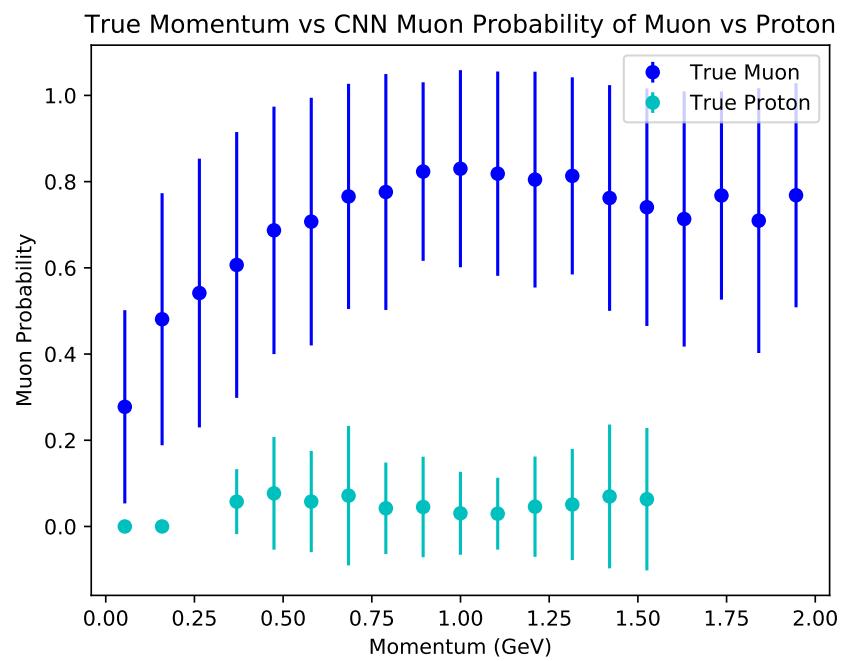
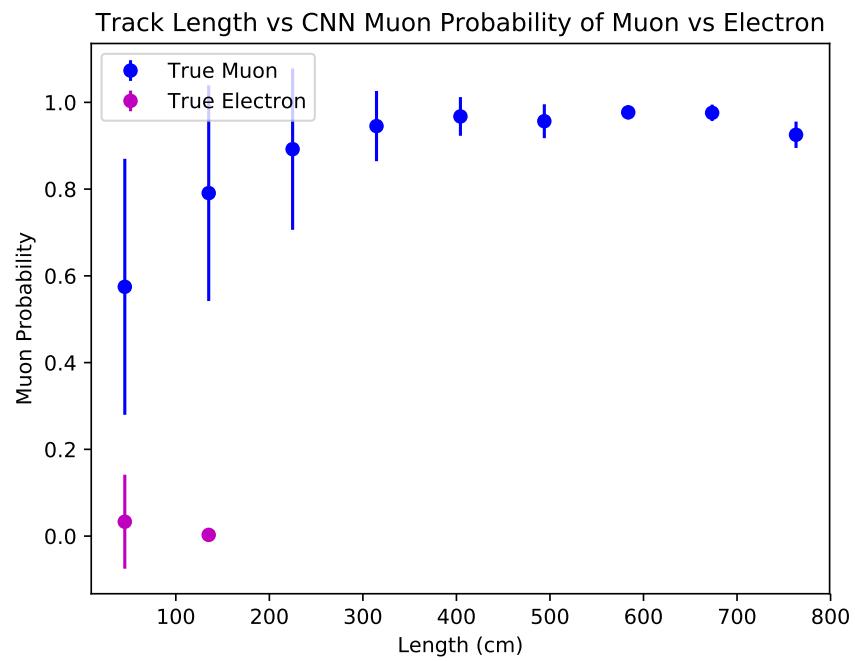
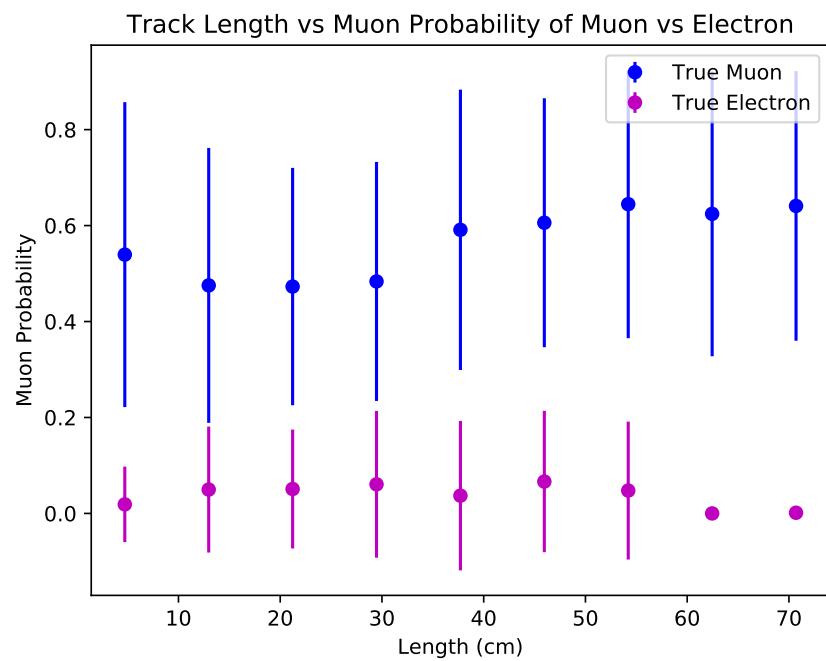


Figure 7.18: mup

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

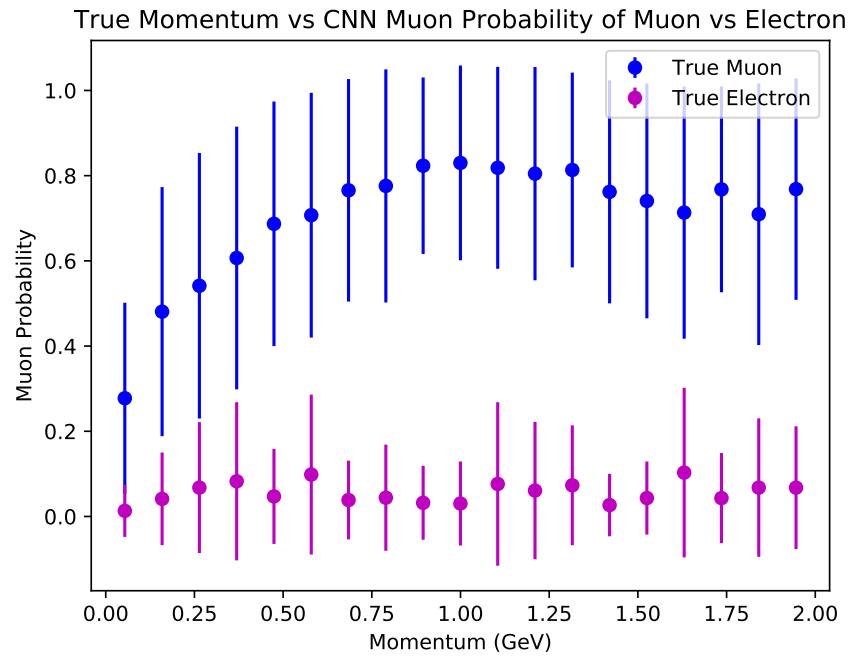


Figure 7.21: mue

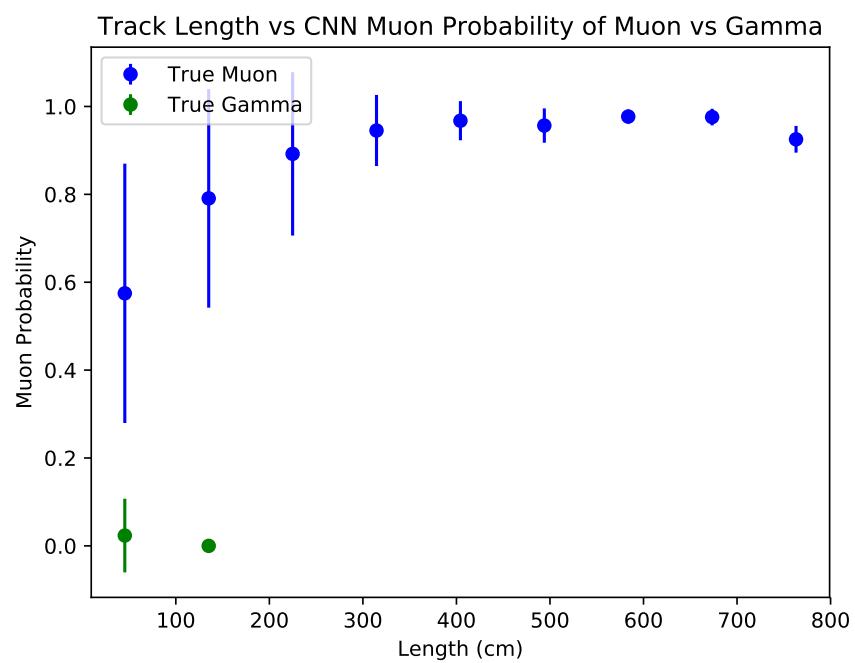


Figure 7.22: mug

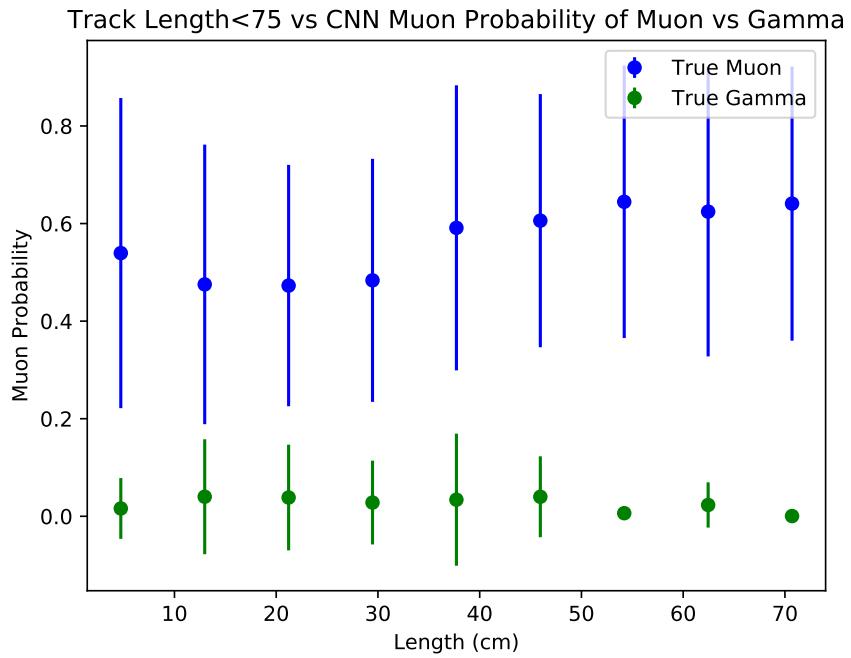


Figure 7.23: mug

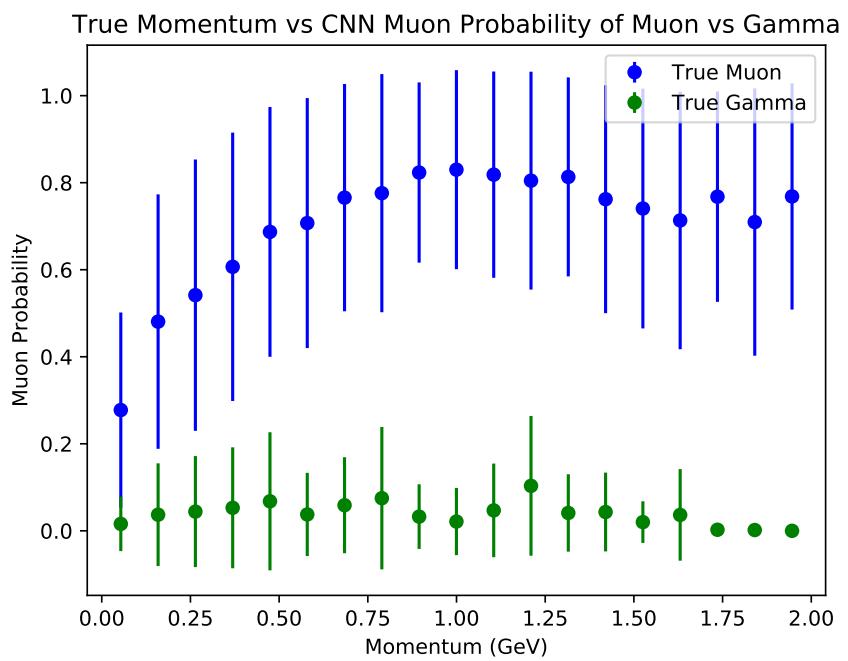


Figure 7.24: mug

1200 Chapter 8

1201 Results of Convolutional Neural 1202 Networks on particles **WORKING** 1203 **TITLE**

1204 8.1 Classification using CNN10000

1205 8.1.1 Classification of MC data using Selection I Original 1206 CC-Inclusive Filter

1207 The next step that was taken was to use CNN10000 to classify track candidate images
1208 that were identified by the selection I original cc-inclusive filter described in [?].
1209 Passing rates for each cut in cc-inclusive filter are show in figure 8.1. For the incorrect
1210 image making normalization dataset, out of 188,880 events, 7438 passed the cut right
1211 before 75 cm track length cut which is 3.9% of total data. Discrepancies in passing rates
1212 are due to grid submission issues, however, this dataset is used to check if changes
1213 in image making normalization affects μ/π separation probability due to CNN10000
1214 being trained with incorrectly image making normalized data. For the second dataset
1215 with correct image making normalization, out of 188,880 events, 9552 events passed the
1216 cut right before the 75 cm track length cut which is 5.1% passing rate and is comparable
1217 to figure 8.1. In time cosmics were also run over for efficiency and purity calculations.
1218 Out of 14395 in time cosmic events, 175 passed the cut right before the 75 cm track
1219 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 ν_μ 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 ν_μ 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 ν_μ 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
≥ 1 flash with ≥ 50 PE	136219 (71%/71%)	44002 (97%/97%)	2979 (62%/62%)	1:14
≥ 1 vertex in FV	131170 (96%/69%)	43794 (99%/97%)	2805 (94%/58%)	1:13
≥ 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 ≥ 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 μ/π separation of both the correct and incorrect normalized images. The confusion matrices are only composed of μ/π 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 μ/π 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 μ or π 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

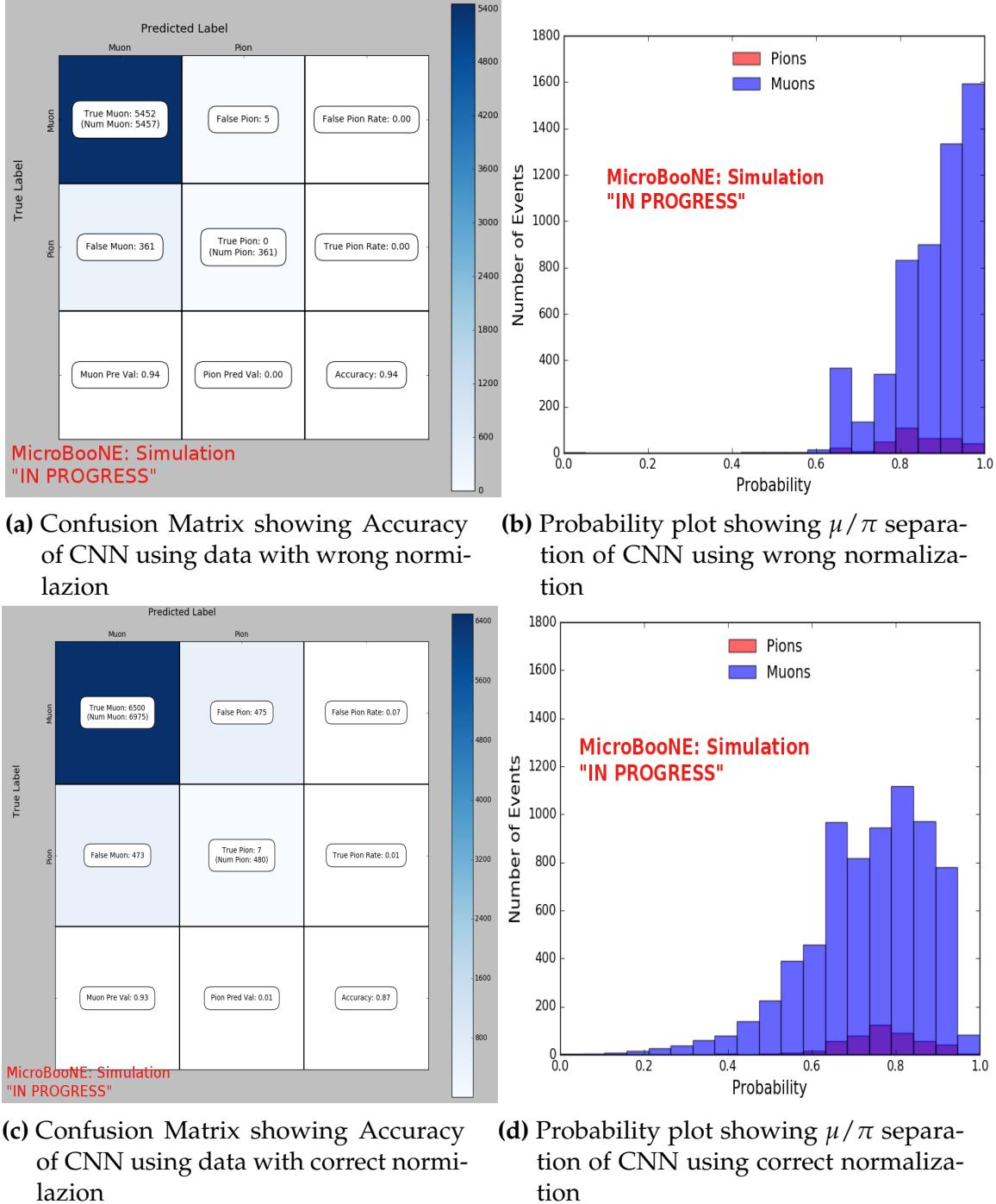


Figure 8.2: Results of CNN10000 classification of track candidate images output from cc-inclusive filter.

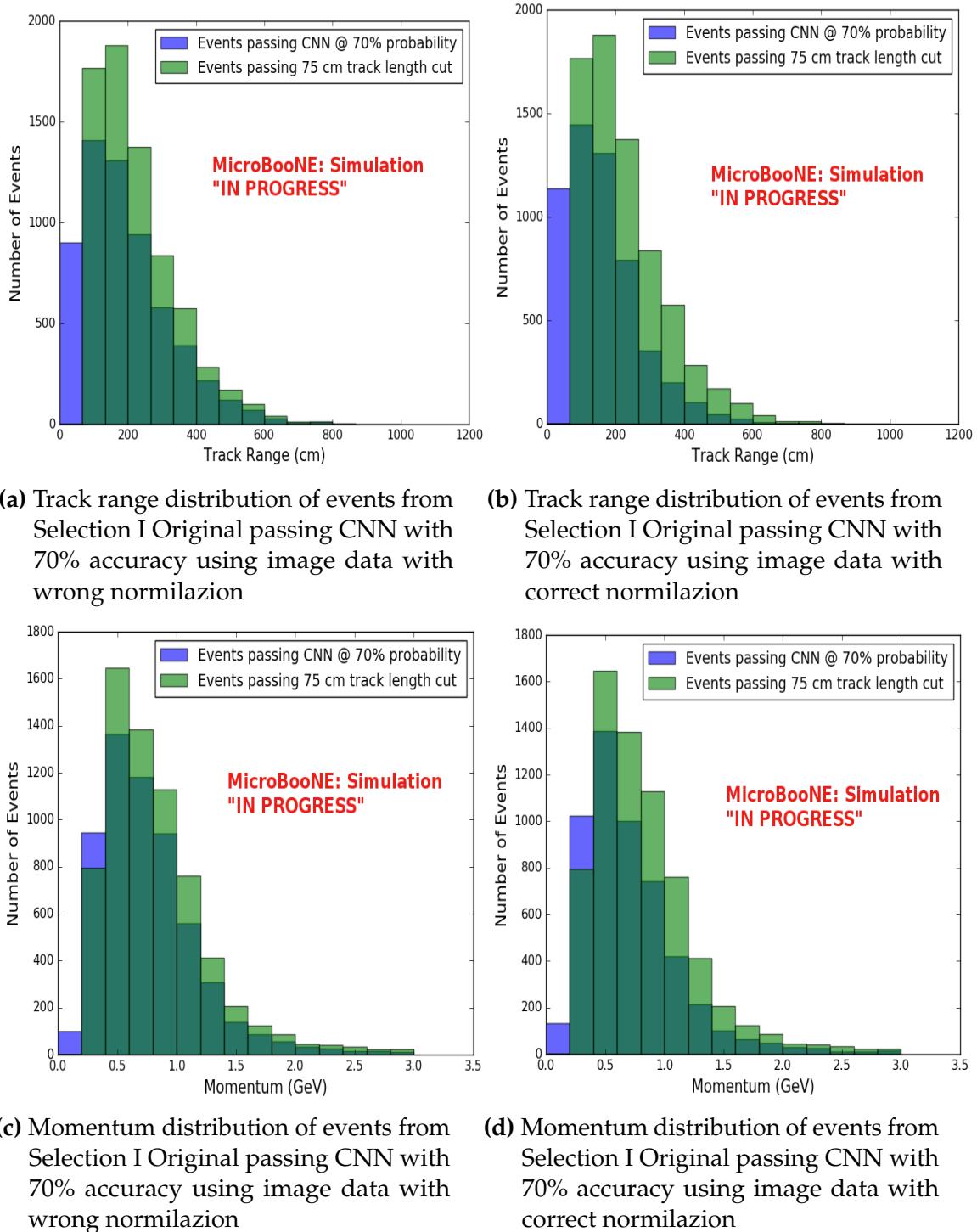


Figure 8.3: CNN10000 distributions of track candidate images output from Selection I Original cc-inclusive filter with different image data normalizations

1240 and electrons. Although μ/π separation is lacking, CNN10000 does an excellent job of
1241 classifying muons and using higher CNN probability can increase purity. Figures 8.3a,
1242 8.3b, 8.3c and 8.3d show the track and momentum distributions for these two datasets.
1243 In both sets you have an increase in data in the bin below 75 cm and at bins below 0.5
1244 GeV. These distributions were made with events classified with 70% probability of
1245 being a muon regardless of true particle type.

1246 **8.1.2 Classification of MC data using Selection I Modified**
1247 **CC-Inclusive Filter**

1248 **Table 5: Selection I: Modified** The table shows passing rates for the above
1249 described event selection. Numbers are absolute event counts and Cosmic back-
1250 ground is not scaled appropriately. The BNB+Cosmic sample contains all
1251 events, not just ν_μ CC inclusive. The selected events are further broken up
1252 in the following subsection. The numbers in brackets give the passing rate wrt
1253 the step before (first percentage) and wrt the generated events (second per-
1254 centage). In the BNB+Cosmic Mc Truth column shows how many true ν_μ CC
1255 inclusive events (in FV) are left in the sample. This number includes possible
1256 mis-identifications where a cosmic track is picked by the selection instead of the
1257 neutrino interaction in the same event. The cosmic only sample has low statis-
1258 tics, but please note that it is not used in any plots, it is just for illustrating the
1259 cut efficiency. The last column Signal:Cosmic only gives an estimate of the ν_μ
1260 CC events wrt the cosmic only background at each step. For this number, the
1261 cosmic background has been scaled as described in chapter 2. Note that this
1262 numbers is not a purity, since other backgrounds can't be determined at this
1263 step.

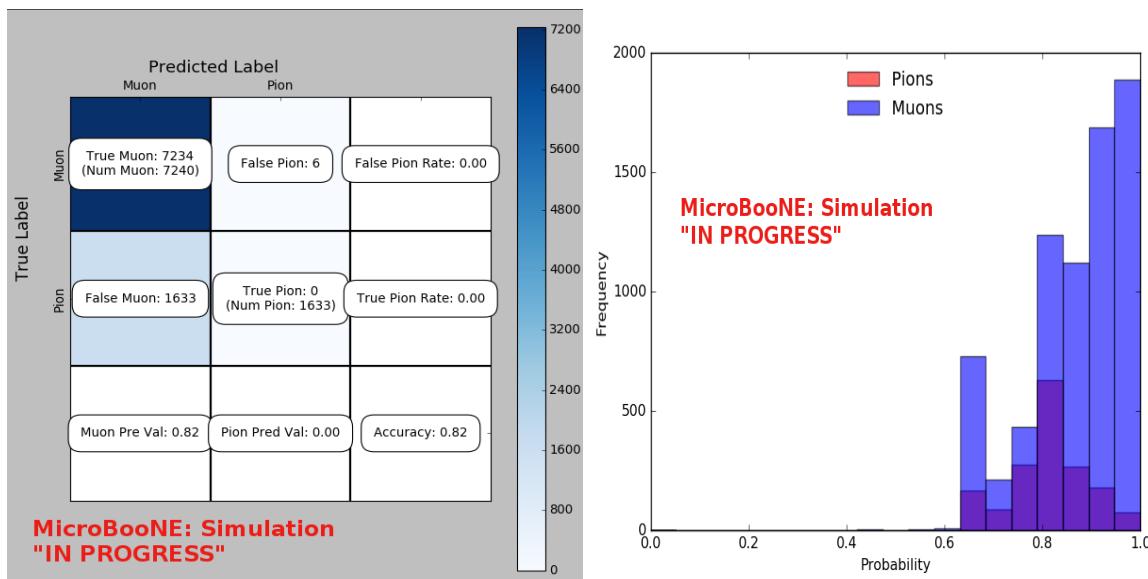
	BNB + Cosmic		Cosmic only	Signal: Cosmic only
	Selection	MC-Truth		
Generated events	191362	45273	4804	1:22
≥ 1 flash with ≥ 50 PE	136219 (71%/71%)	44002 (97%/97%)	2979 (62%/62%)	1:14
≥ 1 track within 5 cm of vertex	135830 (99%/71%)	43974 (99%/97%)	2975 (99%/62%)	1:14
vertex candidate in FV	79112 (58%/41%)	34891 (79%/77%)	1482 (50%/31%)	1:8.9
flash matching of longest track	40267 (51%/21%)	25891 (74%/57%)	340 (23%/7.1%)	1:2.8
track containment	19391 (48%/10%)	11693 (45%/26%)	129 (38%/2.7%)	1:2.3
track ≥ 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

1248 CNN10000 was also used to classify track candidate images that were identified by
1249 the selection I modified cc-inclusive filter described in [?]. Passing rates for each cut in
1250 this filter are shown in figure 8.4. As seen in section 8.1.1, wrong image normalization
1251 had a higher muon classification probability so all work done using selection I modified
1252 cc-inclusive filter was done using this normalization. Out of 188,880 events, 19,112
1253 passed the cut right before the 75 cm track length cut which is a 10.1% passing rate and
1254 comparable to the 10% passing rate shown in figure 8.4. In time cosmics were also run
1255 over, out of 14,606 in time cosmics events, 302 passed the cut right before the 75 cm
1256 track length cut which is a 2.1% passing rate comparable to the 2.7% passing rate in the
1257 cc-inclusive tech-note. Figures 8.5a and 8.5b show the accuracy and μ/π separation.
1258 Both plots are only composed of muons and pions and like selection I original data,

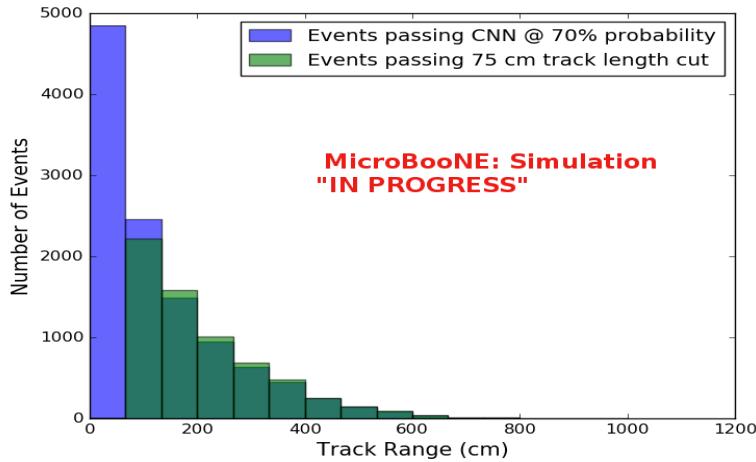
1259 all other particles were id'ed as muons. Also like selection I original data, muons are
 1260 being identified at a very high rate. Figure 8.6a shows the track range distributions
 1261 of all events from selection I modified being classified by the CNN as a muon with a
 1262 probability of 70% regardless of true particle type. We get entries for the CNN curve
 1263 in the lowest bin and none for the 75 cm curve. To see how many true CC events
 1264 were identified by CNN10000 breaking down figure 8.6a by event type was necessary.
 1265 Figures 8.6b and 8.6c show track range distributions separated by signal and various
 1266 backgrounds. Particle type was not taken into consideration in these plots so true CC
 1267 event images can be any track candidate particle passing selection I modified cut right
 1268 before track length cut including pions and protons.

1269 To gain an even deeper understanding on how CNN10000 is performing, plotting
 1270 these distributions with only muons and pions was done due to the fact that CNN10000
 1271 was trained with only those particles for μ/π separation. Figures 8.6d-8.7d show the
 1272 stacked histograms of signal and background of the track range distributions with
 1273 varying CNN probabilities starting from 70% and ending at 90% probability. With
 1274 higher probabilities we get a purer sample in the lower bin but we end up losing
 1275 events as well. Momentum distributions for all signal/background events are shown
 1276 in figure 8.8.

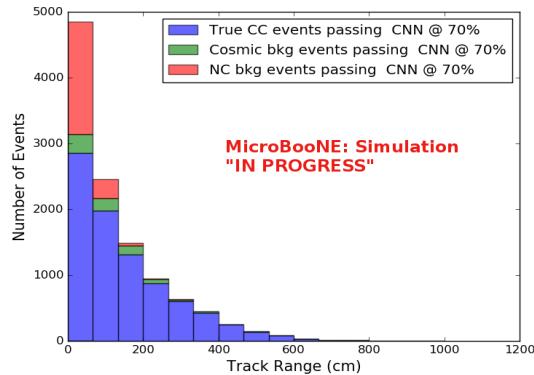


(a) Confusion Matrix for CNN10000 classified events from selection I modified (b) Probability plot for CNN10000 classified events from selection I modified

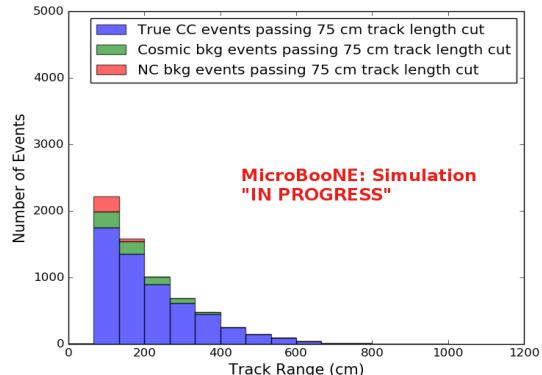
Figure 8.5: Confusion matrix and probability plot of events passing selection I modified cc-inclusive cuts right before 75cm track length cut



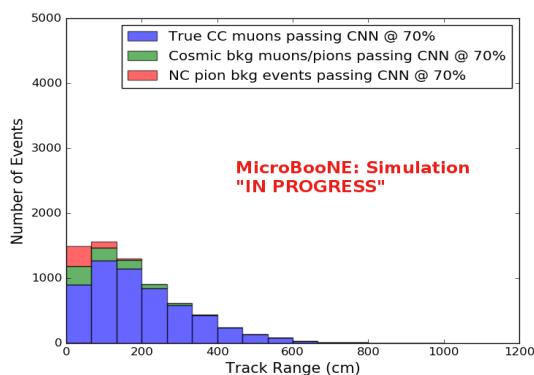
(a) Track range distribution of events from Selection I Modified passing CNN with 70% accuracy



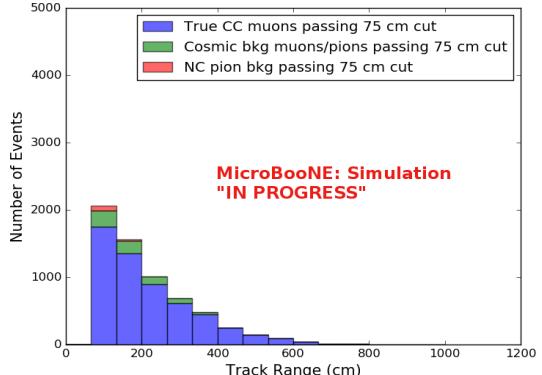
(b) Stacked signal and background track range distributions from Selection I Modified passing CNN with 70% accuracy



(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

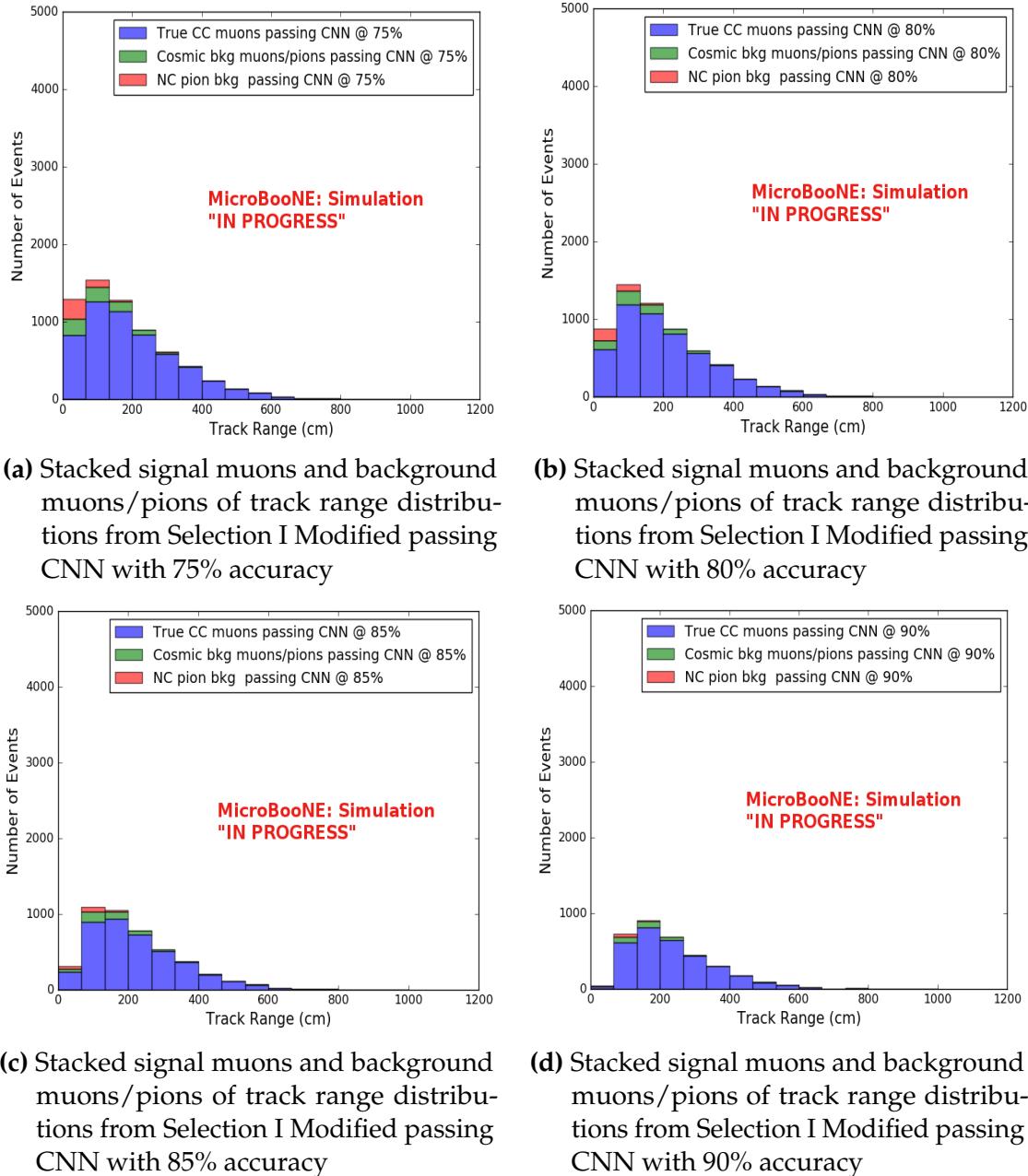
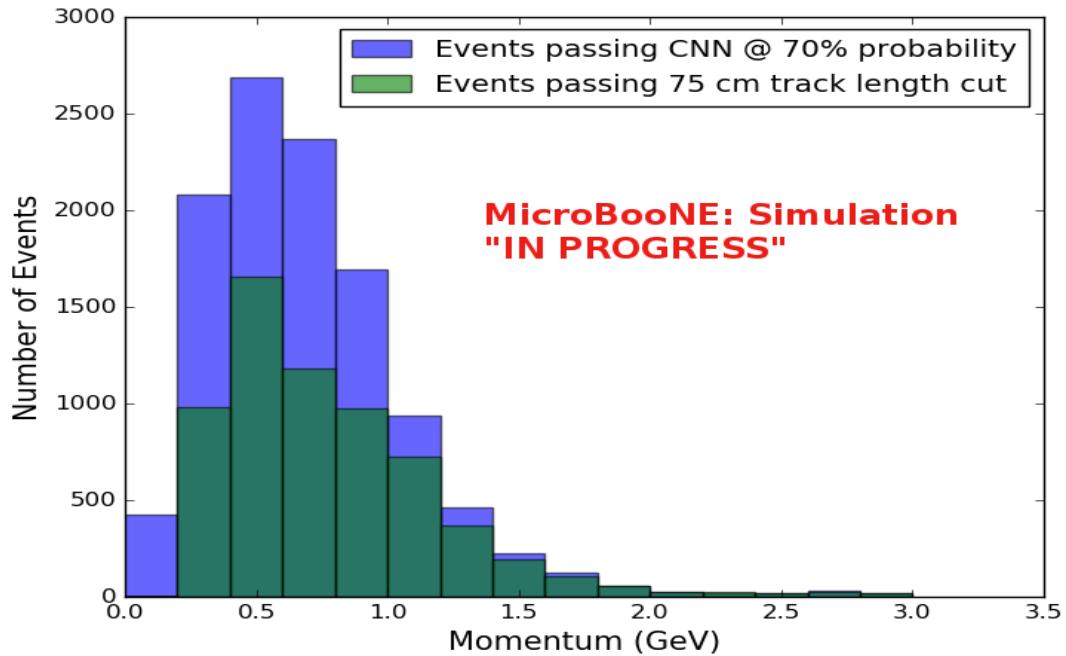
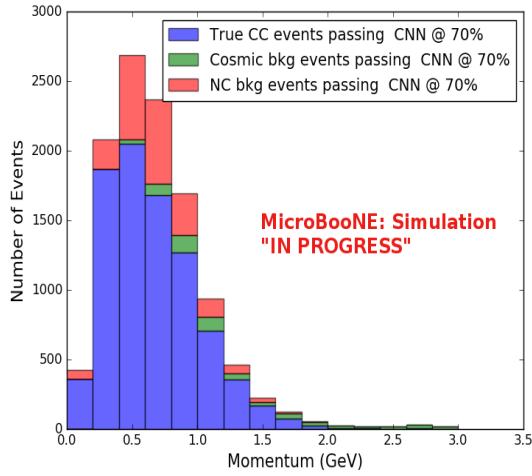


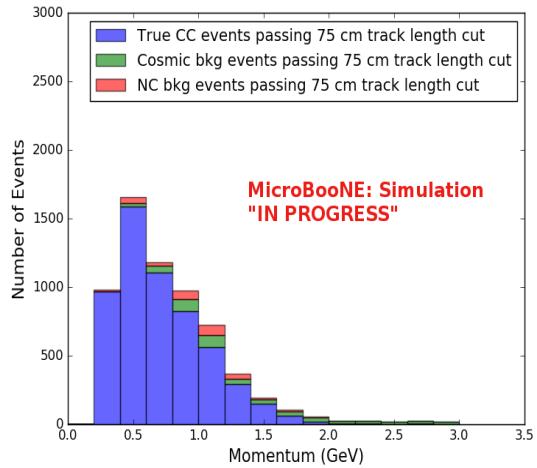
Figure 8.7: CNN10000 stacked signal/background track range distributions of track candidate images output from Selection I Modified cc-inclusive filter



(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.

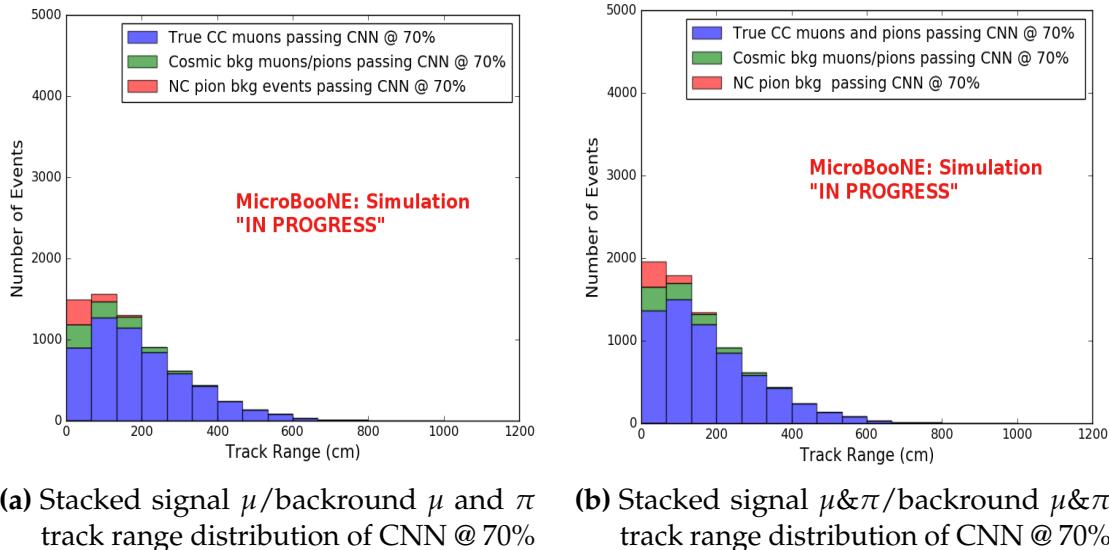


Figure 8.9: Track distribution comparisons of true CC muons plotted vs true CC muons and pions plotted

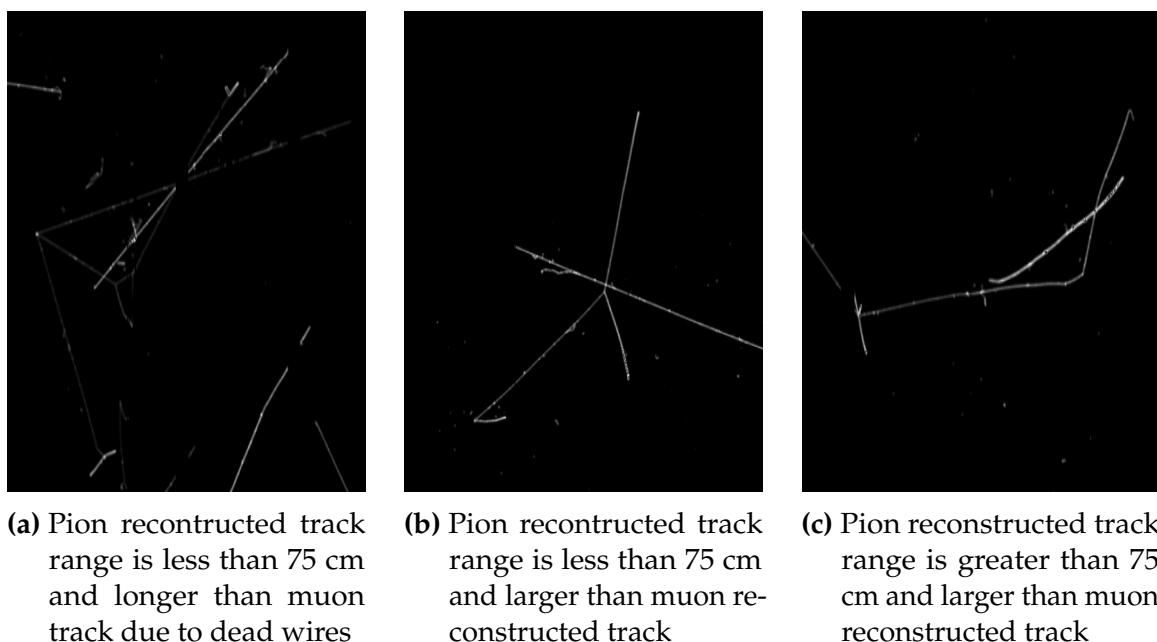


Figure 8.10: Images of true CC events where the pion was the tagged track candidate

		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 ν_μ 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 ν_μ 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	ν_μ CC events with true vertex in FV	#Events(Fraction)	#Events(Fraction)
		passing CNN @ 70% 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%)
	ν_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

1310 a MC true CC event passing rate of 14% compared to the 13% passing rate of the 75
 1311 cm track length cut. The Signal:Cosmic Only background is also reduced from 1:0.6
 1312 to 1:0.4 The total passing rate is also higher than the 75 cm cut, 3.6% vs 4.0%. Table
 1313 8.2 shows the breakdown of signal and backgrounds for the CNN at the different
 1314 probabilities. We have a 61% signal passing rate with the CNN cut @ 83% versus the
 1315 53.8% signal passing rate of the 75 cm cut.

1316 Based on these numbers, the following performance values of the modified selec-
 1317 tion with 75 cm cut versus modified selection with CNN @ 83% probability cut were
 1318 calculated:

- 1319 • Efficiency: Number of selected true ν_μ CC events divided by the number of
 1320 expected true ν_μ CC events with interaction in the FV.
 - 1321 – Selection I modified: 13%
 - 1322 – Selection I modified with CNN cut @ 83% probability: 14%
- 1323 • Purity: Number of selected true ν_μ CC events divided by sum of itself and the
 1324 number of all backgrounds.
 - 1325 – Selection I modified: 53.8%
 - 1326 – Selection I modified with CNN cut @ 83% probability: 61%

1327 Lastly, figure 8.12 shows a more representative performance of the CNN. Due to
 1328 the fact that the CNN was trained on muons and pions, showing the performance
 1329 of CC muon events versus NC pion events with respect to CNN probability gives a
 1330 better picture of how the network is performing. Figure 8.12 shows that at 83% we
 1331 are below the 75 cm cut NC pion threshold and still above the CC muon threshold.
 1332 Using 83% probability not only reduced the NC pion background, it also dramatically
 1333 reduced the in time cosmics and cosmics in the BNB.

1334 8.1.3 Conclusions and Future Work

1335 It was shown that even though CNN10000 was trained with single particle generated
 1336 muons and pions, it performs fairly well at classifying track candidate images from
 1337 BNB+Cosmic events. Events have been regained below the 75 cm track length cut and
 1338 the momentum and track range distributions have similar shapes to the distributions of
 1339 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 5×10^{19} PoT. The last column gives the fraction of this signal or background type to the total 2189 selected events.

Signal	#Events	
ν_μ 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%
ν_e and $\bar{\nu}_e$ events	4	0.2%
$\bar{\nu}_\mu$ events	15	0.7%
ν_μ 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 [?]

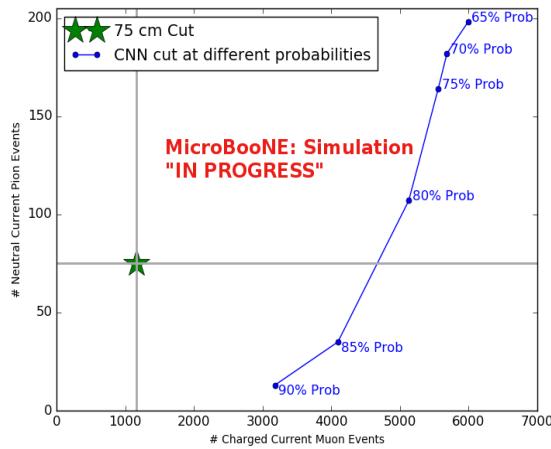


Figure 8.12: CNN performance of classified muons and pions compared to the already implemented 75 cm track length cut

1340 I modified events before 75 cm track length cut with the CNN at 83% probability and
1341 are 14% and 62% respectively. Although the CNN doesn't have separation between
1342 muons and pions and although all particles passing CNN are classified as muon,
1343 increasing CNN probability allows us to increase the purity as well as maintain an
1344 efficiency comparable to the 75 cm track length cut all while recovering events below
1345 that 75 cm cut. Out of the 6142 events that passed the CNN @ 83% 1470 events were
1346 below the 75 cm cut, a recovery of 3.3% of data with a purity of 15%. Although
1347 these numbers are low, it is an improvement from the selection I modified in both total
1348 efficiency and purity and an increase in phase space by recovering these events.

1349 **8.2 Classification using CNN100000**

1350 All future classifications will be done using Selection I Modified CC-Inclusive Filter
1351 because it has a higher efficiency and purity than Selection I Original CC-Inclusive
1352 Filter. To reiterate, CNN100000 was trained using 20,000 images of each $\mu/\pi/p/\gamma/e$.
1353 The results of using CNN100000 to classify BNB+Cosmics will be outlined below.

1354 **8.2.1 Classification of MC data using Selection I Modified
1355 CC-Inclusive Filter**

1356 **8.2.2 Classification of MicroBooNE data using Selection I Modified
1357 CC-Inclusive Filter**

1358 **8.2.3 Comparing two CC-Inclusive Cross Section Selection Filters**

¹³⁵⁹ **Chapter 9**

¹³⁶⁰ **Conclusion**

¹³⁶¹ Your Conclusions here.

¹³⁶²

¹³⁶³ Bibliography

- ¹³⁶⁴ [1] Wikipedia, Neutrino, <http://wikipedia.org/wiki/Neutrino>, 2013.
- ¹³⁶⁵ [2] Wikipedia, Neutrino oscillation, http://en.wikipedia.org/wiki/Neutrino_oscillation, 2013.
- ¹³⁶⁶
- ¹³⁶⁷ [3] B. N. Laboratory, Neutrinos and nuclear chemistry, <http://www.chemistry.bnl.gov/sciandtech/sn/default.htm>, 2010.
- ¹³⁶⁸
- ¹³⁶⁹ [4] K. Heeger, Big world of small neutrinos, <http://conferences.fnal.gov/lp2003/forthepublic/neutrinos/>.
- ¹³⁷⁰
- ¹³⁷¹ [5] A. Y. Smirnov, p. 23 (2003), hep-ph/0305106.
- ¹³⁷² [6] The MicroBooNE Collaboration, R. e. a. Acciari, (2015), arXiv:1512.06148.
- ¹³⁷³ [7] The MicroBooNE Collaboration, R. e. a. Acciari, (2016), arXiv:1705.07341.
- ¹³⁷⁴ [8] The MicroBooNE Collaboration, R. e. a. Acciari, (2017), arXiv:1704.02927.
- ¹³⁷⁵ [9] The MicroBooNE Collaboration, R. e. a. Acciari, JINST **12** (2017).