

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 January 25, 2018 - * -

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...

Contents

41	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 Liquid argon detector development	20
61	3.4 The Booster Neutrino Beam	20
62	3.4.1 Creating the Booster Neutrino Beam	21
63	3.5 Event Reconstruction	22

64	4 Neutrino Identification: Finding MicroBooNE's first Neutrinos	25
65	4.1 Flash Finding	25
66	4.1.1 Flash Reconstruction	26
67	4.1.2 Beam Timing	27
68	4.1.3 Event Rates	28
69	4.2 TPC Topology Selection	28
70	4.2.1 Cosmic Tagging	29
71	4.2.2 2D Cluster Selection	29
72	4.2.3 3D Tracks and vertices Selection	31
73	4.2.4 TPC Updates	33
74	4.3 Conclusion	34
75	5 CC-Inclusive Cross Section Selection Filter	35
76	5.1 Data and MC Processing Chain	36
77	5.2 Normalization of data and MC	37
78	5.3 Optical Software Trigger and Reconstruction	38
79	5.3.1 Software Trigger	38
80	5.3.2 Flash Reconstruction	39
81	5.3.3 Beam Window	40
82	5.4 TPC Reconstruction	42
83	5.4.1 Hit Reconstruction	43
84	5.4.2 Clustering	43
85	5.4.3 Pandora	44
86	5.4.4 Trackkalmanhit	44
87	5.4.5 Cosmic Hit Removal	44
88	5.4.6 Projection Matching Algorithm	45
89	5.5 Event Selection	45
90	6 Background on Convolutional Neural Networks	49
91	6.1 Image Classification	49
92	6.2 CNN Structure	50
93	6.2.1 Backpropagation	53
94	6.3 Choosing Hyperparameters	54
95	7 Training Convolutional Neural Networks on particles WORKING TITLE	57
96	7.1 Hardware Frameworks used for Training	58
97	7.1.1 Syracuse CPU Machine setup	58

98	7.1.2 Syracuse University GPU Cluster Setup	58
99	7.2 Convolutional Neural Network Training	58
100	7.2.1 Image Making Scheme	58
101	7.2.2 Training CNN1075	59
102	7.2.3 Training CNN10000	60
103	7.2.4 Training CNN100000	61
104	8 Results of Convolutional Neural Networks on particles WORKING TITLE	75
105	8.1 Classification using CNN10000	75
106	8.1.1 Classification of MC data using Selection I Original CC-Inclusive	
107	Filter	75
108	8.1.2 Classification of MC data using Selection I Modified CC-Inclusive	
109	Filter	79
110	8.1.3 Conclusions and Future Work	87
111	8.2 Classification using CNN100000	89
112	8.2.1 Classification of MC data using Selection I Modified CC-Inclusive	
113	Filter	89
114	8.2.2 Classification of MicroBooNE data using Selection I Modified	
115	CC-Inclusive Filter	89
116	8.2.3 Comparing two CC-Inclusive Cross Section Selection Filters . .	89
117	9 Conclusion	91
118	Bibliography	93

¹¹⁹ List of figures

120	2.1	The Standard Solar Model	5
121	2.2	Solar Neutrino Experiments	7
122	2.2a	Ray Davis's Homestake Experiment	7
123	2.2b	Kamiokande Experiment	7
124	2.2c	SNO Experiment	7
125	2.3	Cosmic Ray Shower	8
126	2.4	Measurements of the double ratio for various atmospheric neutrino experiments	9
128	2.5	The flavor eigenstates are rotated by an angle θ with respect to the mass eigenstates	10
130	3.1	Low Energy excess seen in MiniBooNE	19
131	3.2	Arial view of the Main Injector and the Booster Neutrino Beam at Fermilab	21
132	3.3	Arial view of the Main Injector and the Booster Neutrino Beam at Fermilab	23
133	3.4	Energy spectrum of the Booster Neutrino Beam at Fermi National Laboratories	24
135	4.1	Efficiency for selecting beam events as a function of minimum total PE cut for all PE cuts as well as zoomed into interesting region.	26
137	4.2a	Predicted distribution of flash times with respect to trigger time for 1 day of data taking at nominal rate and intensity	31

139	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.	31
143	4.3	Percent of good clusters remaining for neutrinos and cosmics after the primary cuts were applied. This is relative to total number of initial clusters.	32
146	4.4	Percent matched cluster pairs remaining for neutrinos and cosmics after secondary cuts applied. This is relative to the number of events that contain clusters which pass the primary cuts.	32
149	5.1	5.1a Track range distribution for selection I. The track range is defined as the 3D distance between the start and end of the muon candidate track. No data is shown below 75 cm due to the track length cut described previously. 5.1b Efficiency of the selected events by process quasi-elastic (QE), resonant (RES), and deep-inelastic (DIS). Statistical uncertainty is shown in the bands and the distributions are a function of true muon momentum. The rise of the efficiency between 0 GeV and 0.5 GeV is due to the minimum track length cut and the decreasing efficiency for higher momentum tracks is caused by the containment requirement. .	38
158	5.1a	Track range distribution of selection I	38
159	5.1b	Selection efficiency as a function of the true muon momentum .	38
160	5.2	Time distribution of reconstructed optical flashes with a PE value of 50 or more for a sample of BNB unbiased triggered events.	40
162	5.3	Flash time distribution for all flashes (left plot) and flashes > 20PE (right plot). The different curves are as follows: on-beam data (black), off-beam data (red), CORSIKA inTime MC (light blue), BNB only MC (green), and BNB+Cosmic MC (purple). The dashed vertical lines mark the time window that was chosen for each sample	41
167	5.4	Reconstruction chain run on both data and MC. The red stars mean that the algorithm returns reconstructed 3D vertices.	42

169	5.5 Event selection diagram for selection I and selection II. This analysis focused on optimizing selection I. Boxes with the same color across the two selections symbolize similar cuts.	47
170		
171		
172	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	51
173		
174		
175	6.2 Visualization of filters found in first layer of a CNN.	51
176		
177		
178		
179		
180	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.	52
181		
182		
183	6.4 Pictorial Representation of Convolutional Neural Networks as well as a visual representation on CNN's complexity of layer feature extraction	52
184		
185	6.5 Pictoral representation of the AlexNet model. The AlexNet model consists of 5 convolution layers and 3 fully connected layers.	55
186		
187		
188		
189		
190		
191	6.6 Pictoral representation of the GoogleNet model. The GoogleNet model consists of 22 layers. The model implements 9 Inception modules which performs covolution and pooling in parallel and strays away from the basis that CNN layers need to be stacked up sequentially. The GoogleNet model also doesn't use fully connected layers, instead it uses average pooling which greatly reduces the amount of parameters. GoogleNet has 12x fewer parameters than AlexNet.	55
192		
193		
194	7.1 Accuracy vs. Loss of ImageNet 2-output μ/π sample consisting of 10000 images each.	61
195		
196		
197		
198	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.	62
199	7.2a Confusion Matrix showing Accuracy of CNN using training data	62

200	7.2b Confusion Matrix showing Accuracy of CNN using testing data	62
201	7.2c Probability plot of muons and pions from testing set. Images	
202	surrounding histogram are a random event from lowest bin and	
203	highest bin for each particle.	62
204	7.3 Confusion Matrix of all five particles	63
205	7.4 Training and testing accuracy of CNN trained on 100,000 images of	
206	$\mu/\pi/p/\gamma/e$ with 20,000 images of each particle. Each image was a size	
207	of 576x576 and the images per particle were split 90% use for training	
208	and 10% used for testing the network	64
209	7.5 Training and testing loss of CNN trained on 100,000 images of $\mu/\pi/p/\gamma/e$	64
210	7.6 t-SNE of CNN	65
211	7.7 Muon Prob	66
212	7.8 Pion Prob	66
213	7.9 Proton Prob	67
214	7.10 Electron Prob	67
215	7.11 Gamma Prob	68
216	7.12 Prob	68
217	7.13 mupi	69
218	7.14 mupi	69
219	7.15 mupi	70
220	7.16 mup	70
221	7.17 mup	71
222	7.18 mup	71
223	7.19 mue	72
224	7.20 mue	72
225	7.21 mue	73

226	7.22 mug	73
227	7.23 mug	74
228	7.24 mug	74
229	8.1 Snapshot of passing rates of Selection I from CC-Inclusive Filter	76
230	8.2 Results of CNN10000 classification of track candidate images output	
231	from cc-inclusive filter.	77
232	8.2a Confusion Matrix showing Accuracy of CNN using data with	
233	wrong normilazion	77
234	8.2b Probability plot showing μ/π separation of CNN using wrong	
235	normalization	77
236	8.2c Confusion Matrix showing Accuracy of CNN using data with	
237	correct normilazion	77
238	8.2d Probability plot showing μ/π separation of CNN using correct	
239	normalization	77
240	8.3 CNN10000 distributions of track candidate images output from Selec-	
241	tion I Original cc-inclusive filter with different image data normalizations	78
242	8.3a Track range distribution of events from Selection I Original	
243	passing CNN with 70% accuracy using image data with wrong	
244	normilazion	78
245	8.3b Track range distribution of events from Selection I Original pass-	
246	ing CNN with 70% accuracy using image data with correct	
247	normilazion	78
248	8.3c Momentum distribution of events from Selection I Original pass-	
249	ing CNN with 70% accuracy using image data with wrong	
250	normilazion	78
251	8.3d Momentum distribution of events from Selection I Original pass-	
252	ing CNN with 70% accuracy using image data with correct	
253	normilazion	78

254	8.4 Snapshot of passing rates of all cuts from Selection I Modified cc-inclusive filter	79
255		
256	8.5 Confusion matrix and probability plot of events passing selection I modified cc-inclusive cuts right before 75cm track length cut	80
257		
258	8.5a Confusion Matrix for CNN10000 classified events from selection I modified	80
259		
260	8.5b Probability plot for CNN10000 classified events from selection I modified	80
261		
262	8.6 CNN10000 distributions of track candidate images output from Selection I Modified cc-inclusive filter	81
263		
264	8.6a Track range distribution of events from Selection I Modified passing CNN with 70% accuracy	81
265		
266	8.6b Stacked signal and background track range distributions from Selection I Modified passing CNN with 70% accuracy	81
267		
268	8.6c Stacked signal and background track range distributions from Selection I Modified passing 75 cm track length cut	81
269		
270	8.6d Stacked signal muons and background muons/pions of track range distributions from Selection I Modified passing CNN with 70% accuracy	81
271		
272	8.6e Stacked signal muons and background muons/pions of track range distributions from Selection I Modified passing 75 cm track length cut	81
273		
274		
275		
276	8.7 CNN10000 stacked signal/background track range distributions of track candidate images output from Selection I Modified cc-inclusive filter	82
277		
278	8.7a Stacked signal muons and background muons/pions of track range distributions from Selection I Modified passing CNN with 75% accuracy	82
279		
280	8.7b Stacked signal muons and background muons/pions of track range distributions from Selection I Modified passing CNN with 80% accuracy	82
281		
282		
283		

284	8.7c	Stacked signal muons and background muons/pions of track range distributions from Selection I Modified passing CNN with 85% accuracy	82
285			
286			
287	8.7d	Stacked signal muons and background muons/pions of track range distributions from Selection I Modified passing CNN with 90% accuracy	82
288			
289			
290	8.8	CNN10000 momentum distributions of track candidate images output from Selection I Modified cc-inclusive filter	83
291			
292	8.8a	Momentum distribution of events from Selection I Modified passing CNN with 70% accuracy	83
293			
294	8.8b	Stacked signal and background momentum distributions from Selection I Modified passing CNN with 70% accuracy	83
295			
296	8.8c	Stacked signal and background momentum distributions from Selection I Modified passing 75 cm track length cut	83
297			
298	8.9	Track distribution comparisons of true CC muons plotted vs true CC muons and pions plotted	85
299			
300	8.9a	Stacked signal μ /background μ and π track range distribution of CNN @ 70%	85
301			
302	8.9b	Stacked signal $\mu\&\pi$ /background $\mu\&\pi$ track range distribution of CNN @ 70%	85
303			
304	8.10	Images of true CC events where the pion was the tagged track candidate	85
305	8.10a	Pion reconstructed track range is less than 75 cm and longer than muon track due to dead wires	85
306			
307	8.10b	Pion reconstructed track range is less than 75 cm and larger than muon reconstructed track	85
308			
309	8.10c	Pion reconstructed track range is greater than 75 cm and larger than muon reconstructed track	85
310			
311	8.11	Snapshot of signal and background event numbers of Selection I modified from cc-inclusive note [?]	88
312			

313	8.12 CNN performance of classified muons and pions compared to the al-	
314	ready implemented 75 cm track length cut	88

List of tables

315	4.1	Passing rates for 2D cluster cuts for neutrino on MC set and a cosmic only MC set. First column shows event rates with no cuts applied to both sets. Columns two and three show event rates after primary and secondary cuts are applied. Line three shows the second line scaled with the flash finding factor of 0.008. All events are normalized to per day assuming we are running at 5 Hz.	31
321	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. . .	86

338	8.2 Signal and background event numbers at modified selection level with	
339	CNN cut estimated from a BNB+Cosmic sample and Cosmic only sam-	
340	ple normalized to $5 * 10^{19}$ PoT. The last column gives the fraction of	
341	this signal or background type to the total selected events per CNN	
342	probability.	86

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

— Shirley Chisholm

³⁴⁵ **Chapter 1**

³⁴⁶ **Introduction**

³⁴⁷ This thesis will be a description of work done to further increase efficiency and purity
³⁴⁸ of the charged current inclusive cross section measurement using the MicroBooNE
³⁴⁹ detector. It will also describe the MicroBooNE detector, what neutrinos are, the
³⁵⁰ charged current inclusive cross section measurement and its importance as well as
³⁵¹ convolutional neural networks and how they can be used in μ/π separation. Chapter
³⁵² 2 will talk about the background of neutrinos and the people and detectors that
³⁵³ discovered neutrinos as well as an in depth history of neutrino oscillation and the
³⁵⁴ discovery that neutrinos have mass.

³⁵⁵ Chapter 3 will discuss the MicroBooNE experiment, specifically, how Liquid
³⁵⁶ Argon Time Projection Chambers work, the Light Collection System and the Electronic
³⁵⁷ and Readout Trigger systems. This chapter will also describe the Booster Neutrino
³⁵⁸ Beam sationed at Fermilab.

³⁵⁹ Chapter 4 will discuss the work that was done to detect the first neutrinos seen in
³⁶⁰ the MicroBooNE detector and the software reconstruction efforts required to create an
³⁶¹ automated neutrino ID filter that was used to find the first neutrinos and then was
³⁶² later expanded on to create the charged current inclusive filter that will be discussed
³⁶³ in chapter 5

³⁶⁴ Chapter 6 will give a brief description of what Convolutional Neural Networks are
³⁶⁵ and how it will be used for μ/π separation in this selection. Chapter 7 will discuss
³⁶⁶ the hardware frameworks and training methods used to train multiple Convolutional
³⁶⁷ Neural Networks for use in the charged current inclusive cross section measurement.
³⁶⁸ Chapters 8 and ?? will discuss the results of using Convolutional Neural Networks on
³⁶⁹ monte-carlo and data to sift out charged current inclusive neutrino events.

³⁷⁰ **Chapter 2**

³⁷¹ **Neutrinos**

³⁷² **2.1 What are Neutrinos**

³⁷³ Neutrinos are one of the fundamental particles which make up the universe. They are
³⁷⁴ also one of the least understood. Neutrinos are not affected by the electromagnetic
³⁷⁵ forces because they do not have electric charge. Neutrinos are affected by a "weak"
³⁷⁶ sub-atomic force of much shorter range than electromagnetism, and are therefore able
³⁷⁷ to pass through great distances in matter without being affected by it. Until the late
³⁷⁸ 90's, neutrinos were thought to have no mass. Due to their mass, neutrinos are also
³⁷⁹ affected by gravity. Neutrinos are created by radioactive decay or nuclear reactions
³⁸⁰ such as the ones that happen in the sun, in nuclear reactors or when cosmic rays hit
³⁸¹ atoms. There are three types of neutrinos, ν_e , ν_μ and ν_τ which correspond to their
³⁸² charged lepton pairs.

³⁸³ As previously stated, neutrinos are very weakly interacting; in fact, neutrinos can
³⁸⁴ pass unscathed through a wall of lead several hundred light-years thick. Because
³⁸⁵ neutrinos interact so rarely, studying neutrinos requires a massive detector and a
³⁸⁶ powerful neutrino source. With that being said, we can only infer their existence when
³⁸⁷ they interact in a detector. In a collision, distinct charged particles are produced with
³⁸⁸ each type of neutrino. An electron neutrino will create an electron, a muon neutrino
³⁸⁹ will create a muon, and a tau neutrino will create a tau. The track the particle leaves
³⁹⁰ in the detector is how one figures out what type of neutrino interaction was "seen".
³⁹¹ Liquid Argon Time Projection Chambers are the newest type of detectors being used to
³⁹² study neutrinos due to their excellent imaging and particle identification capabilities.

³⁹³ 2.2 History of Neutrinos

³⁹⁴ The neutrino was first postulated by Wolfgang Pauli in 1931 to explain how beta
³⁹⁵ decay could resolve the conservation of energy, momentum and angular momentum
³⁹⁶ problem. Pauli suggested that this missing energy might be carried off, unseen, by a
³⁹⁷ neutral particle (he called neutron) which was escaping detection. James Chadwick
³⁹⁸ discovered a much heavier nuclear particle in 1932 that he also named neutron, leaving
³⁹⁹ two particles with the same name. Enrico Fermi was the first person to coin the
⁴⁰⁰ term neutrino (which means little neutral one in latin) in 1933 to fix this confusion.
⁴⁰¹ Fermi's paper, which was published in 1934, unified Pauli's neutrino with Paul Dirac's
⁴⁰² positron and Werner Heisenberg's neutron-proton model and his theory accurately
⁴⁰³ explained many experimentally observed results. Wang Ganchang first proposed the
⁴⁰⁴ use of beta capture to experimentally detect neutrinos and in 1959 Clyde Cowan and
⁴⁰⁵ Frederick Reines published their work stating that they had detected the neutrino.
⁴⁰⁶ The experiment called for antineutrinos created in a nuclear reactor by beta decay that
⁴⁰⁷ reacted with protons producing neutrons and positrons: $\nu_e + p^+ \rightarrow n^0 + e^+$. Once
⁴⁰⁸ this happens, the positron finds an electron and they annihilate each other and the
⁴⁰⁹ resulting gamma rays are detectable. The neutron is detected by neutron capture and
⁴¹⁰ the releasing of another gamma ray. In 1962 Leon M. Lenderman, Melvin Schwartz
⁴¹¹ and Jack Steinberger were the first to detect interactions of the muon neutrino. The
⁴¹² first detection of the tau neutrino was announced in the summer of 2000 by the
⁴¹³ DONUT collaboration at Fermilab. In the late 1960s, many experiments found that the
⁴¹⁴ number of electron neutrinos arriving from the sun was around 1/3 to 1/2 the number
⁴¹⁵ predicted by the Standard Solar Model. This became known as the solar neutrino
⁴¹⁶ problem and remained unresolved for around thirty years. This problem was resolved
⁴¹⁷ by the discovery of neutrino oscillation and mass. [1]

⁴¹⁸ 2.3 Neutrino Oscillations

⁴¹⁹ Neutrino oscillation was first predicted by Bruno Pontecorvo. It describes the phe-
⁴²⁰ nomenon of a neutrino created with a specific lepton flavor (electron, muon or tau)
⁴²¹ that is later measured to have a different flavor. Neutrino oscillation is important
⁴²² theoretically and experimentally due to the fact that this observation implies that the
⁴²³ neutrino has a non-zero mass, which is not part of the original Standard Model of
⁴²⁴ particle physics. [2]

425 2.3.1 Solar Oscillations and the Solar Neutrino Problem

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

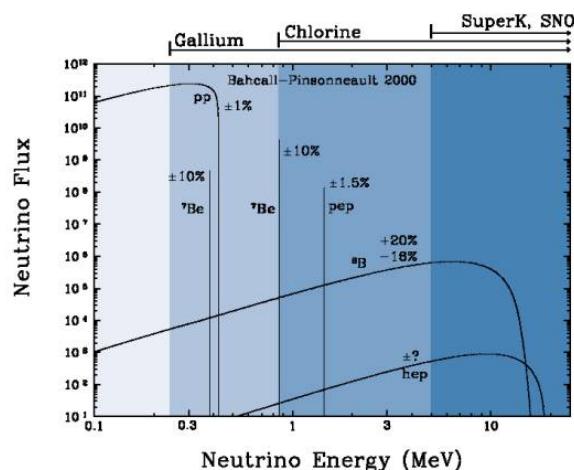


Figure 2.1: The Standard Solar Model

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

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

⁴⁴⁵ The unexplained difference between the measured solar neutrino flux and model
⁴⁴⁶ predictions lead to the Solar Neutrino Problem. [4]



⁴⁴⁷ While it is now known that the Homestake Experiment detected neutrinos, some
⁴⁴⁸ physicist were weary of the results. Conclusive evidence of the Solar Neutrino Problem
⁴⁴⁹ was provided by the Kamiokande-II experiment, a water Cherenkov detector with
⁴⁵⁰ a low enough energy threshold to detect neutrinos through neutrino-electron elastic
⁴⁵¹ scattering. In the elastic scattering interaction the electrons coming out of the point of
⁴⁵² reaction strongly point in the direction that the neutrino was traveling, away from the
⁴⁵³ sun. While the neutrinos observed in Kamiokande-II were clearly from the sun, there
⁴⁵⁴ was still a discrepancy between Kamiokande-II and Homestake; The Kamiokande-
⁴⁵⁵ II experiment measured about 1/2 the predicted flux, rather than the 1/3 that the
⁴⁵⁶ Homestake Experiment saw.

⁴⁵⁷ The solution to the solar neutrino problem was finally experimentally determined
⁴⁵⁸ by the Sudbury Neutrino Observatory(SNO). The Ray Davis's Homestake Experiment
⁴⁵⁹ was only sensitive to electron neutrinos, and the Kamiokande-II Experiment was
⁴⁶⁰ dominated by the electron neutrino signal. The SNO experiment had the capability to
⁴⁶¹ see all three neutrino flavors. Because of this, it was possible to measure the electron
⁴⁶² neutrinos and total neutrino flux. The experiment demonstrated that the deficit was
⁴⁶³ due to the MSW effect, the conversion of electron neutrinos from their pure flavor
⁴⁶⁴ state into the second neutrino mass eigenstate as they passed through a resonance
⁴⁶⁵ due to the changing density of the sun. The resonance is energy dependent, and is
⁴⁶⁶ visible near 2MeV. The water cherenkov detectors only detect neutrinos above about
⁴⁶⁷ 5MeV, while the radiochemical experiments were sensitive to lower energy (0.8MeV
⁴⁶⁸ for chlorine, 0.2MeV for gallium), and this turned out to be the source of the difference
⁴⁶⁹ in the observed neutrino rates at the two types of experiments. Figure 2.2 shows
⁴⁷⁰ Homestake, Kamiokande-II and SNO experiments.

⁴⁷¹ MSW Effect

⁴⁷² The Mikheyev-Smirnov-Wolfenstein effect is a process which acts to modify neu-
⁴⁷³ trino oscillations in matter. The presence of electrons in matter changes the energy

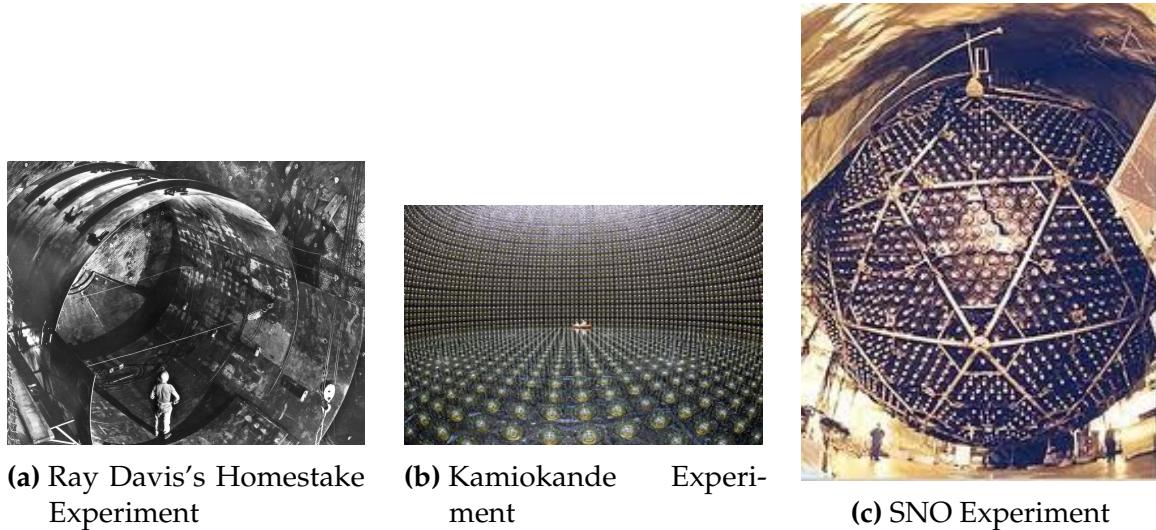


Figure 2.2: Solar Neutrino Experiments

474 levels of the mass eigenstates of neutrinos due to charged current coherent forward
 475 scattering of the electron neutrinos. This coherent forward scattering is similar to
 476 the electromagnetic process with respect to the refractive index of light in a medium.
 477 Because of this MSW Effect, neutrinos in vacuum have a different effective mass than
 478 neutrinos in matter and because neutrino oscillations depend on the squared mass
 479 difference of the neutrinos, the neutrino oscillations are different in matter than in
 480 vacuum. This effect is important at the sun where electron neutrinos are produced.
 481 The neutrinos of high energy leaving the sun are in a vacuum propagation eigenstate
 482 ν_2 that has a very small overlap with the electron neutrino $\nu_e = \nu_1\cos(\theta) + \nu_2\sin(\theta)$
 483 seen by the charged current reactions in Kamiokande-II and SNO. The discrepancy of
 484 the deficit between SNO, Kamiokande-II and Homestake is due to the energy of the
 485 solar neutrinos. The MSW effect "turns on" at about 2MeV and at lower energies, this
 486 MSW effect is negligible. [5]

487 **2.3.2 Atmospheric Oscillations and the Atmospheric Neutrino 488 Anomaly**

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

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

492

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



Figure 2.3: Cosmic Ray Shower

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

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

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

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

505 Kamiokande-II has the the capability of measuring the direction of the incoming
 506 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

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

2.3.3 Two Flavor Neutrino Oscillation Formulation

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

In matrix form the wavefunctions are:

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

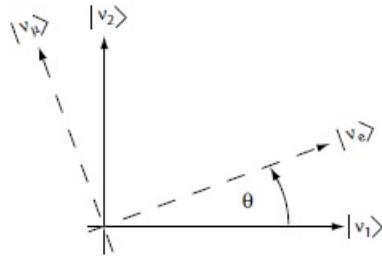


Figure 2.5: The flavor eigenstates are rotated by an angle θ with respect to the mass eigenstates

526 Applying the time evolution operator to ν_μ :

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

527 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
528 being, let us assume $\hbar = c = 1$. With this assumption: $E_1 = \sqrt{p^2 + m_1^2}$ and $E_2 =$
529 $\sqrt{p^2 + m_2^2}$. The next modifications is to assume neutrinos are relativistic:

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

530 because of this,

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

531

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

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

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

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

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

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

⁵³⁴ 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:

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

⁵³⁵ Finding the Probability for a $\nu_\mu \rightarrow \nu_e$:

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

⁵³⁶ Remembering that $\langle \nu_i | \nu_j \rangle = \delta_{ij}$

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

⁵³⁷ Taking the absolute value squared gives us:

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

⁵³⁸ Since the neutrino is relativistic we can set $p = E_\nu$ and change $x = L$. Also
⁵³⁹ recognizing the trigonometric relation $(1 - \cos 2\theta)/2 = \sin^2\theta$ the above equation
⁵⁴⁰ becomes:

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

⁵⁴¹ All that's left to do now is re-introduce \hbar and c doing this we get:

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

⁵⁴² 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

⁵⁶⁴ of the postulated CP-symmetry. CP-symmetry states that the laws of physics should
⁵⁶⁵ be the same if a particle were to be exchanged with its antiparticle and then if the left
⁵⁶⁶ hand side of a decay were switched with the right hand side.

⁵⁶⁷ **2.3.5 Reactor Oscillation**

⁵⁶⁸ Many experiments have searched for oscillation of electron anti-neutrinos produced at
⁵⁶⁹ nuclear reactors. Such oscillations give the value of the parameter θ_{13} . The KamLAND
⁵⁷⁰ experiment, started in 2002, has made a high precision observation of reactor neutrino
⁵⁷¹ oscillation. Neutrinos produced in nuclear reactors have energies similar to solar
⁵⁷² neutrinos, a few MeV. The baselines of these experiments have ranged from tens
⁵⁷³ of meters to over 100 km. On 8 March 2012, the Daya Bay team announced a 5.2σ
⁵⁷⁴ discovery that $\theta_{13} \neq 0$.

575 **Chapter 3**

576 **The MicroBooNE Experiment**

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

582 **3.1 Liquid argon time projection chambers**

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

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

623 3.2 The MicroBooNE Time Projection Chamber

624 MicroBooNE (Micro Booster Neutrino Experiment) is a 89 ton active volume (180 ton
 625 total mass) LArTPC which is then inserted into a cylindrical cryostat on axis of the
 626 Booster Neutrino Beam (BNB) stationed at Fermilab in Batavia, Illinois. Understanding
 627 LArTPC technology and detector physics is necessary to build a LArTPC the size of
 628 DUNE, and MicroBooNE has made many advances in developing this technology [7]
 629 [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.

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

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

692 3.3 MicroBooNE's Physics Goals

693 3.3.1 The low-energy excess

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

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

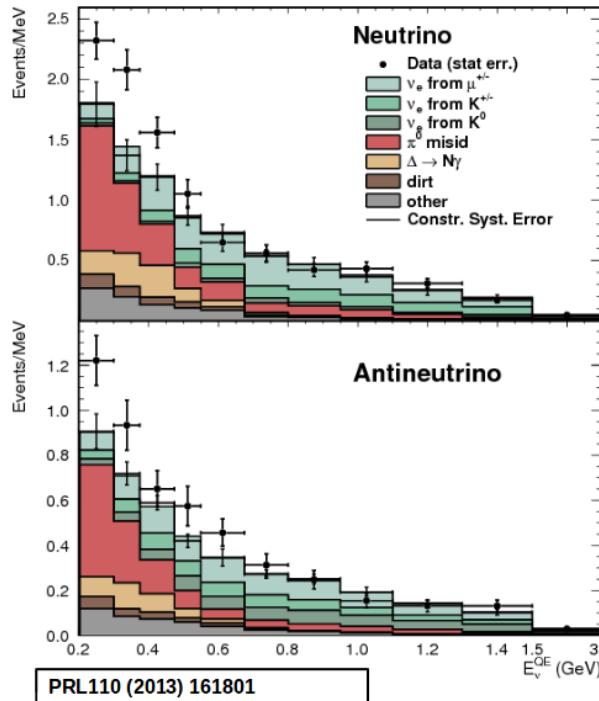


Figure 3.1: Low Energy excess seen in MiniBooNE

3.3.2 Cross sections

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

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

729 3.3.3 Liquid argon detector development

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

734 3.4 The Booster Neutrino Beam

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

⁷⁴² introduces to a measurement. An aerial view of fermilab as well as the BNB is shown
⁷⁴³ in figure 3.2

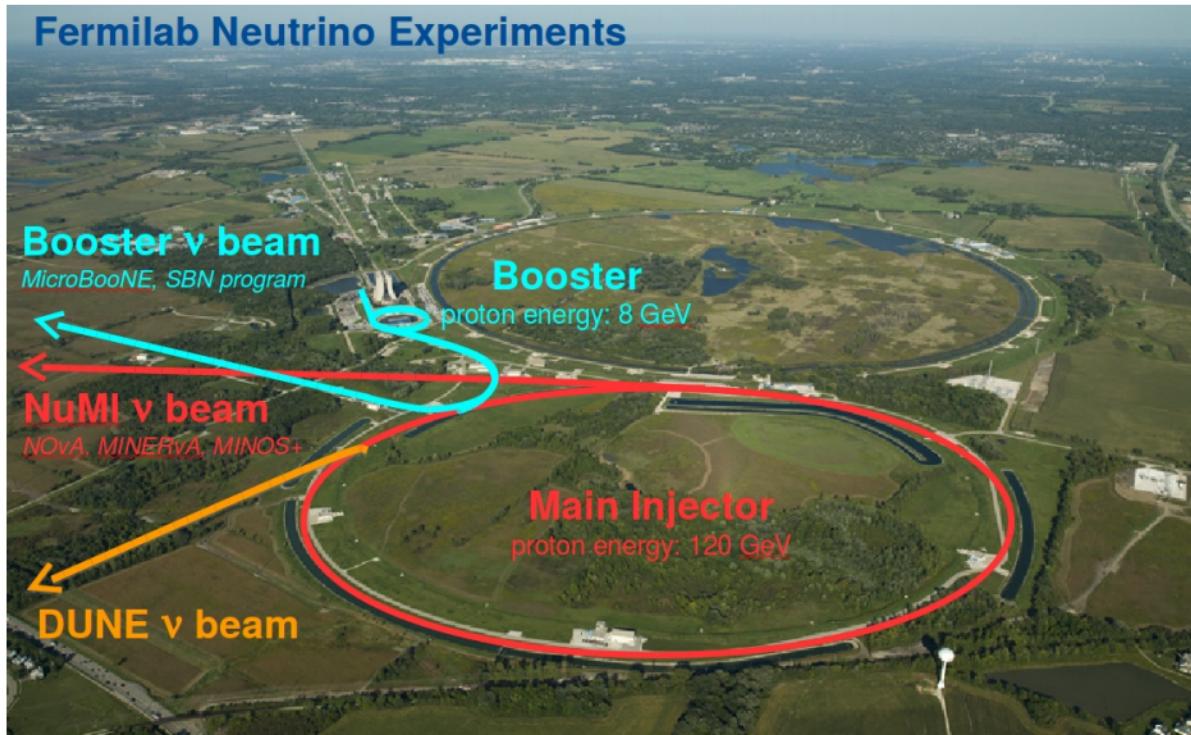


Figure 3.2: Arial view of the Main Injector and the Booster Neutrino Beam at Fermilab

⁷⁴⁴ 3.4.1 Creating the Booster Neutrino Beam

⁷⁴⁵ The BNB is a very pure ν_μ beam, with only 0.6% contamination from ν_e s. The energy
⁷⁴⁶ also peaks around 700 MeV which is desired based on the probability of oscillation
⁷⁴⁷ equation which depends on the the value of L/E , where L is the distance of the
⁷⁴⁸ detector from the neutrino beam and E is the energy of the neutrino beam. L/E was
⁷⁴⁹ chosen to increase the probability of seeing neutrino oscillations in the MiniBooNE
⁷⁵⁰ Low Energy Excess (LEE) range based on the probability of oscillation equation, which
⁷⁵¹ 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
⁷⁵² protons from the FNAL booster synchrotron into a beryllium target which produces a
⁷⁵³ high flux of neutrinos. The protons originate from H^2 gas molecules that are turned
⁷⁵⁴ into H^- ions by a Cockcroft-Walton generator shown in figure ???. The H^- initially are
⁷⁵⁵ accelerated to 1MeV kinetic energy and are then passed to a linear accelerator using
⁷⁵⁶ alternating electromagnetic fields to increase their energy to 400MeV. The ions are
⁷⁵⁷ stripped of electrons by passing them through a carbon foil. The protons are bunched

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

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

779 **3.5 Event Reconstruction**

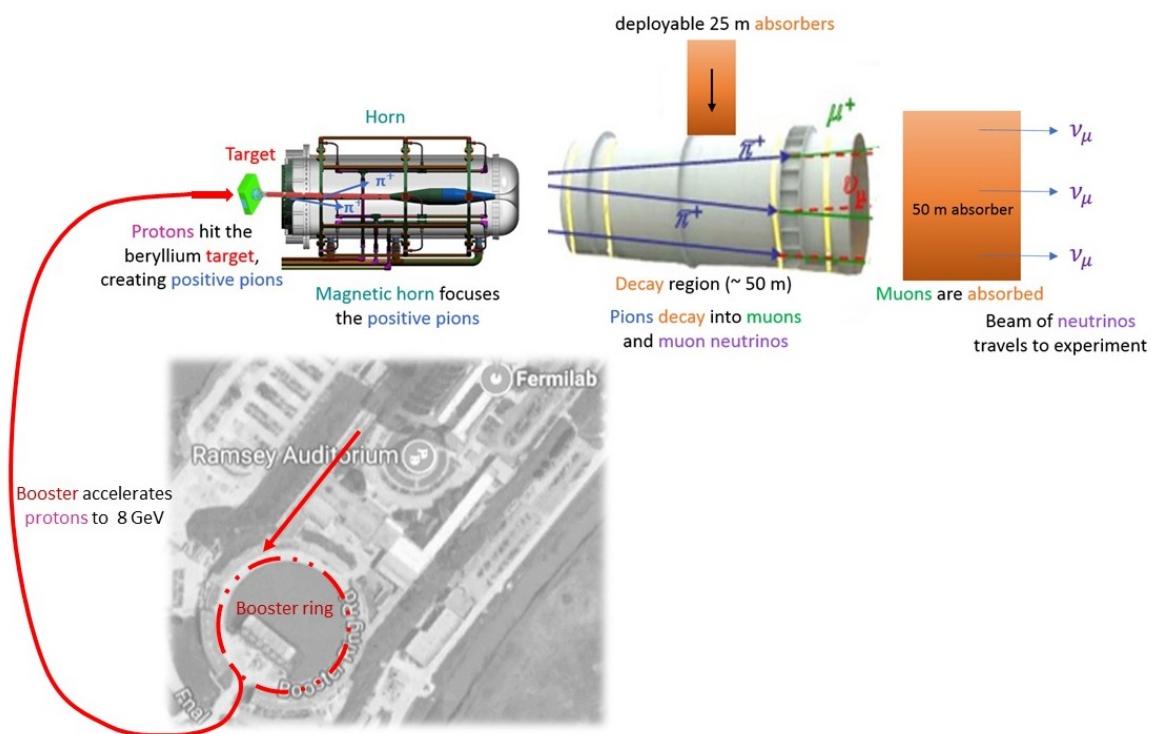


Figure 3.3: Aerial view of the Main Injector and the Booster Neutrino Beam at Fermilab

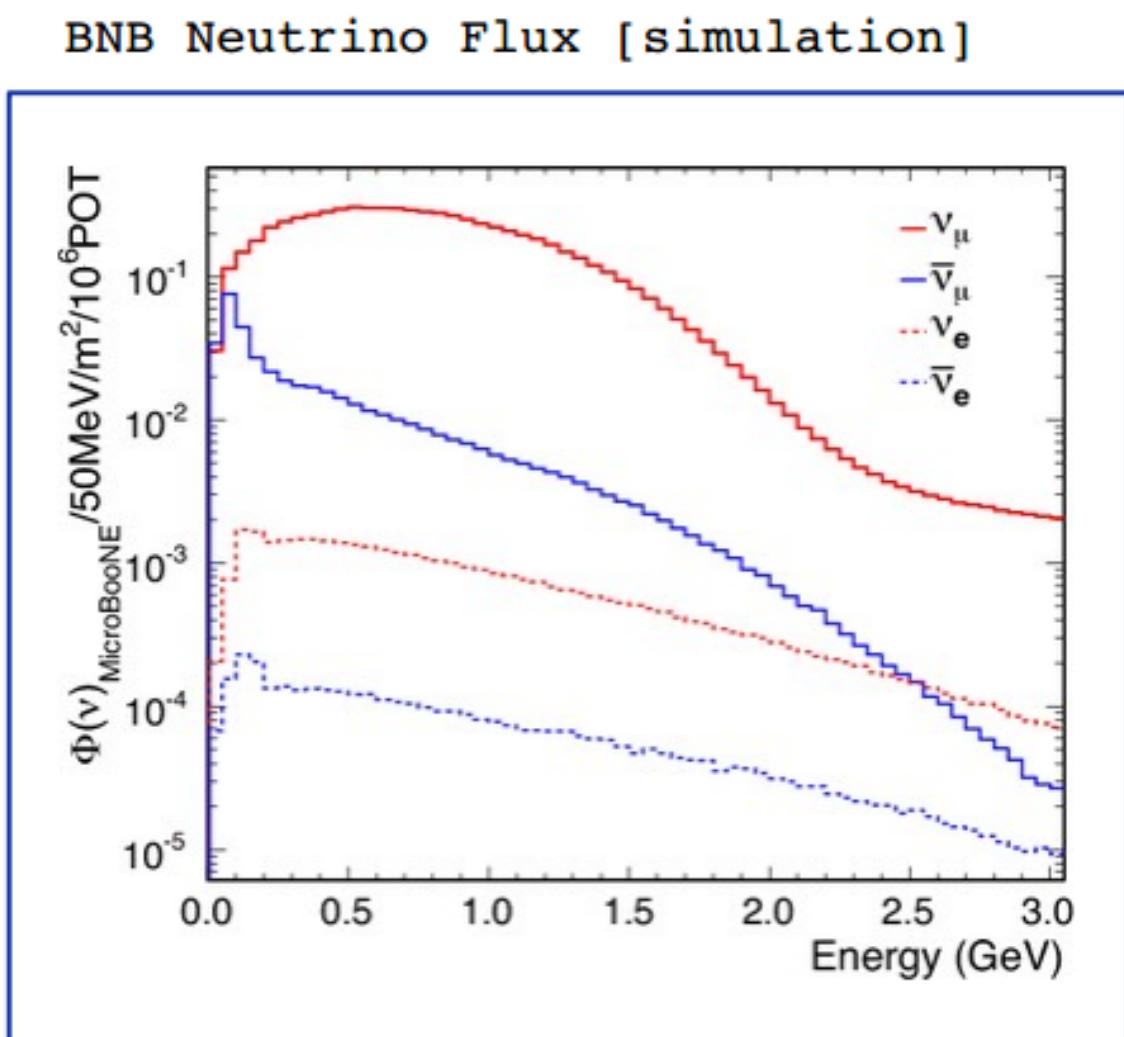


Figure 3.4: Energy spectrum of the Booster Neutrino Beam at Fermi National Laboratories

780 **Chapter 4**

781 **Neutrino Identification: Finding**
782 **MicroBooNE's first Neutrinos**

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

797 **4.1 Flash Finding**

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

801 **4.1.1 Flash Reconstruction**

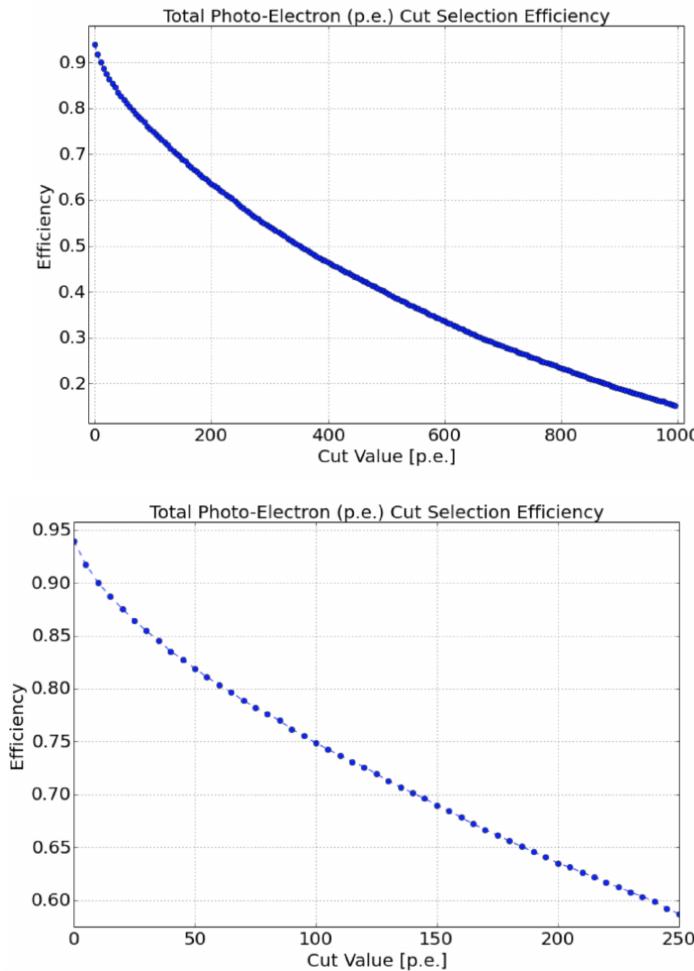
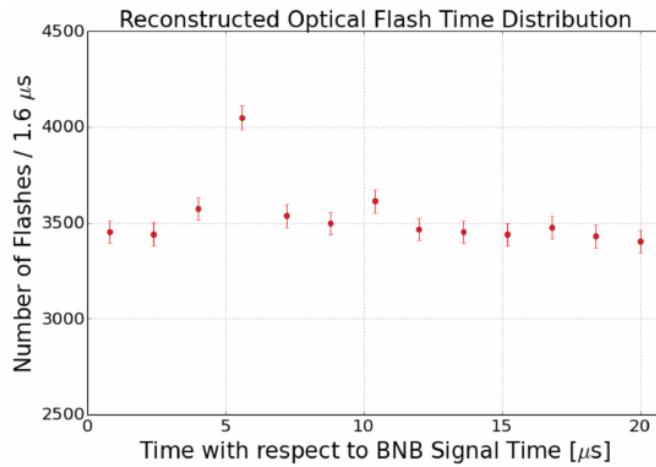


Figure 4.1: Efficiency for selecting beam events as a function of minimum total PE cut for all PE cuts as well as zoomed into interesting region.

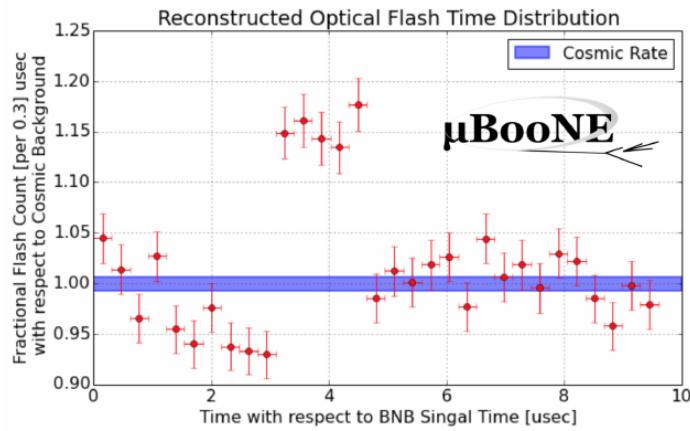
802 A flash is described as a collection of light seen at the same time within the detector.
803 They are then reconstructed by identifying signal from the PMTs above a specific
804 photoelectron (PE) threshold. These signals are called optical hits. Optical hits from
805 all the PMTs are then accumulated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

⁸¹² cut of 50 PE was deemed sufficient for this analysis. Figure 4.1 show the total PE
⁸¹³ versus the selection efficency of selecting neutrino beam events.

⁸¹⁴ 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.

⁸¹⁵ It is necessary to get the specific time from flashes if one uses flashes to filter out
⁸¹⁶ neutrino interactions coincident with the neutrino beam spill period and background.
⁸¹⁷ Before a filter can be applied, an understanding of the timing of the trigger and PMT

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

827 **4.1.3 Event Rates**

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

833 **4.2 TPC Topology Selection**

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

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

⁸⁴⁴ **4.2.1 Cosmic Tagging**

⁸⁴⁵ The first step in TPC selection was based on the geometry of cosmic tracks in an event.
⁸⁴⁶ The cosmic ray muon geometry tagger runs on 3D tracks and assigns a score to each
⁸⁴⁷ reconstructed track on the likeliness of the track originating from a cosmic. The cosmic
⁸⁴⁸ scores are detailed below:

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

⁸⁵⁴ Clusters are assigned either a 0 or 1, 1 being a cosmic. In simulation, 90% of cosmics
⁸⁵⁵ are tagged as cosmics. These tracks are no longer considered when looking for a
⁸⁵⁶ neutrino topology. Requiring that the tracks be contained in turn affects the neutrino
⁸⁵⁷ efficiency by 20%. The algorithm checks that each track is contained within a boundary
⁸⁵⁸ region of 10 cm from all sides of the TPC. This boundary region was optimized via
⁸⁵⁹ handscanning of experimental data.

⁸⁶⁰ As can be expected, cosmic tagging is more efficient in the 3D channel (tracks) than
⁸⁶¹ the 2D channel (clusters) because the reconstructed tracks can use the full 3D position
⁸⁶² information of the entering and exiting points while the 2D channel mainly use the
⁸⁶³ reconstructed x position of the cluster which is associated to timing.

⁸⁶⁴ Cosmic tagging uses timing information to reject tracks and clusters that are outside
⁸⁶⁵ of drift window. The drift window for 128 kV is 1.6 μ s while for 70 kV, the actual
⁸⁶⁶ voltage MicroBooNE is running at, is 2.3 μ s. Due to this variation between simulation
⁸⁶⁷ and data, we expect to see $2.3/1.6 = 1.44$ times more cosmic induced tracks or clusters
⁸⁶⁸ in the drift window.

⁸⁶⁹ **4.2.2 2D Cluster Selection**

⁸⁷⁰ This selection was spearheaded by myself and Katherine Woodruff. After looking at
⁸⁷¹ experimental cosmics data, 2D clustering performs well, while 3D track reconstruction
⁸⁷² is affected by more variations in simulation, for example noise filters. This was the

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

879 **Primary Cuts**

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

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

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

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

899 **Secondary Cuts**

900 The secondary cuts look to match long, low-angle clusters with short, high-charge
 901 clusters. Only clusters that have passed previous cuts are used. First clusters with
 902 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.

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

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

915 4.2.3 3D Tracks and vertices Selection

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

- 919 • d : distance between the start points of the two tracks.
- 920 • d_1 : distance between vertex and start of track 1.
- 921 • d_2 : distance between vertex and start of track 2.

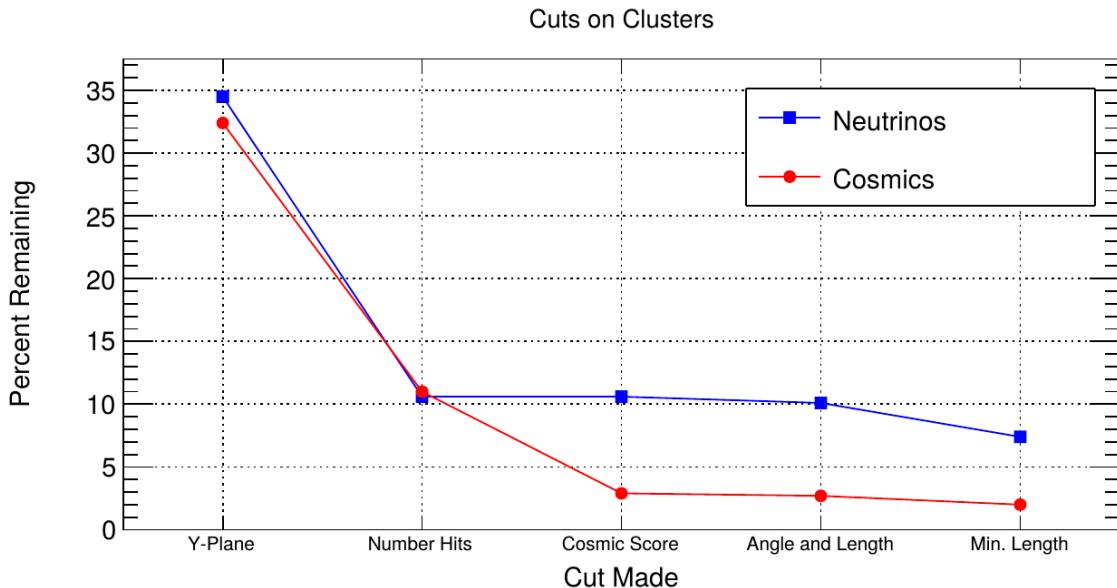


Figure 4.3: Percent of good clusters remaining for neutrinos and cosmics after the primary cuts were applied. This is relative to total number of initial clusters.

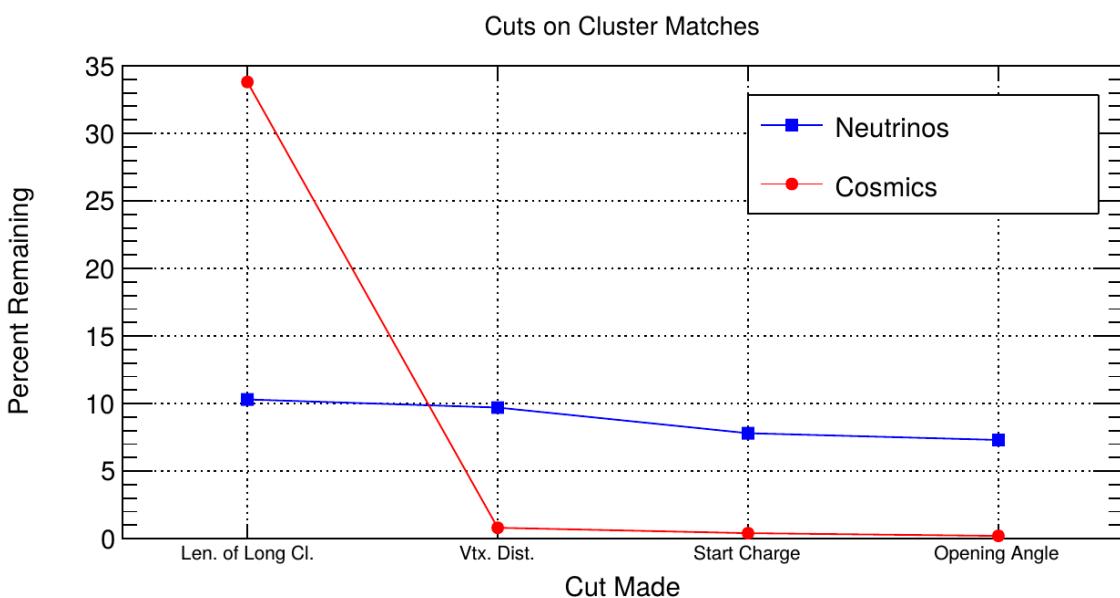


Figure 4.4: Percent matched cluster pairs remaining for neutrinos and cosmics after secondary cuts applied. This is relative to the number of events that contain clusters which pass the primary cuts.

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

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

932 4.2.4 TPC Updates

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

939 2D Filter Updates

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

945 3D Filter Updates

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

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

951 **4.3 Conclusion**

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

958 **Chapter 5**

959 **CC-Inclusive Cross Section Selection**

960 **Filter**

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

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

⁹⁷⁸ **5.1 Data and MC Processing Chain**

⁹⁷⁹ The data used for this analysis were based on hardware and software triggers. Events
⁹⁸⁰ used came from the *BNB_INCLUSIVE* and *EXT_BNB_INCLUSIVE* streams and were
⁹⁸¹ used for signal and background. The *BNB_INCLUSIVE* stream is chosen by requiring
⁹⁸² that the hardware trigger bit is fired and that the event passed an optical software
⁹⁸³ trigger within a BNB spill timing window. The *EXT_BNB_INCLUSIVE* stream requires
⁹⁸⁴ the EXT hardware trigger to fire as well as pass the same optical software trigger
⁹⁸⁵ within a BNB spill size timing window similar to the *BNB_INCLUSIVE*.

⁹⁸⁶ The two MC samples used in this analysis and for determining selection efficiencies
⁹⁸⁷ and purities were GENIE BNB neutrino interactions with CORSIKA cosmic ray overlay
⁹⁸⁸ within the readout window and inTime CORSIKA cosmic rays. The MC samples
⁹⁸⁹ generated used *uboonecode v04_36_00* and are based on the following packages:

- ⁹⁹⁰ • larsoft v04_36_00
- ⁹⁹¹ • GEANT v04_09_06_p04d
- ⁹⁹² • GENIE v02_08_06d
- ⁹⁹³ • GENIE xsec v02_08_06a
- ⁹⁹⁴ • pandora v02_03_0a
- ⁹⁹⁵ • CORSIKA v07_4003

⁹⁹⁶ Both data and MC samples were processed using the same reconstruction release,
⁹⁹⁷ *uboonecode v05_08_00* and the fcl files used for reconstruction are listed below:

- ⁹⁹⁸ • MC fcl files
 - ⁹⁹⁹ – reco_uboone_mcc7_driver_stage1.fcl
 - ¹⁰⁰⁰ – reco_uboone_mcc7_driver_stage2.fcl
- ¹⁰⁰¹ • Data fcl files
 - ¹⁰⁰² • reco_uboone_data_Feb2016_driver_stage1.fcl
 - ¹⁰⁰³ • reco_uboone_data_Feb2016_driver_stage2.fcl

¹⁰⁰⁴ On top of the hardware and software triggers, the data also had to pass more
¹⁰⁰⁵ criteria to be identified as part of the good run list. The criteria is detailed below.

- 1006 • **Detector conditions:** the detector has to be in a good operating condition. The
1007 detector conditions are read from the slow monitoring database and are required
1008 to be within the alarm thresholds. The variables of interest for events passing
1009 the good run list criteria include DAQ, PMT, HV, Drift HV, wire bias, electron
1010 lifetime and detector power. These conditions need to be met on a run-by-run
1011 basis in order to pass the selection.
- 1012 • **Data quality:** normal and stable behavior for basic reconstruction quantities.
1013 These reconstruction variables include average number of tracks, hits, and flashes
1014 in each event, the average length of tracks, the average amplitude and area of
1015 hits, the average PE and the average spread of each one of these quantities.
- 1016 • **Beam Conditions:** the BNB must be on and stable and the POT per spill needs
1017 to above the intensity threshold. Beam quality conditions include checking the
1018 fraction of proton beam interacting within the target, the horn current, and the
1019 intensity of protons per spill. The final sample is $5 * 10^{19}$ and a per-spill intensity
1020 of $4 * 10^{12}$
- 1021 • **Run processed:** the full run must be processed completely without missing
1022 subruns or crashes in the data processing.

1023 5.2 Normalization of data and MC

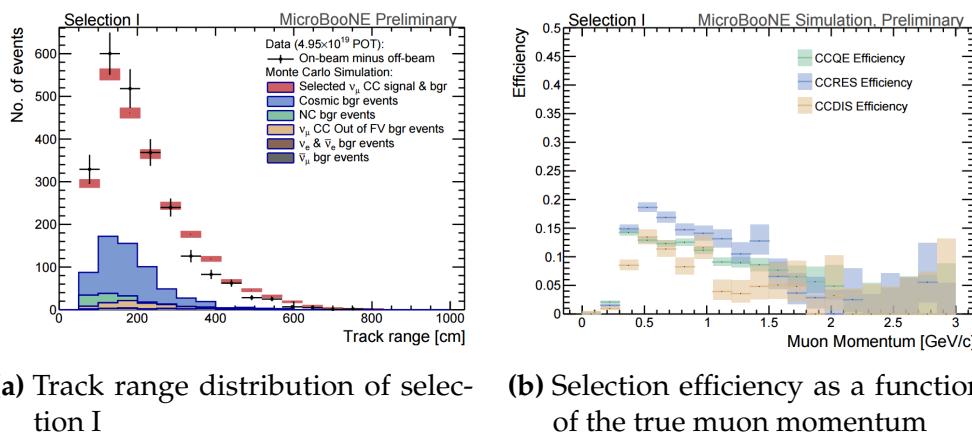
1024 The off-beam sample is used to measure beam unrelated backgrounds. For normalization
1025 one needs the total number of BNB spills (N_{BNB}) and the total number of external
1026 triggers. The BNB spills used need to pass the beam quality cuts. The normalization
1027 factor is then N_{BNB}/N_{EXT} which is 1.23.

1028 To normalize generated BNB MC events to POT, we used the following:

- 1029 • $5 * 10^{19} POT = 41524.3$ generated events

1030 where this scaling factor only applies to mcc7 generated events. The inTime cosmic
1031 sample is normalized with respect to the open cosmic sample so an understanding
1032 of both is necessary. The POT per beam spill for mcc7 BNB samples is $5 * 10^{12}$. To
1033 calculate how many spills are necessary to produce a specific POT one would multiply
1034 the total POT by the average 1/POT per spill. For a total POT of $5 * 10^{19}$ the amount
1035 of spills necessary is $\frac{5*10^{19}}{5*10^{12}} = 1 * 10^7$. This is only one in ~ 241 events therefore each

1036 cosmic event needs to be scaled up by a factor of 240.8 when comparing to BNB
 1037 MC. For inTime cosmics however, two filters are applied to reduce computing and
 1038 processing time and only leave cosmics that will interact within the detector. The
 1039 passing rate after these two filters is 0.02125, therefore the total inTime cosmic scaling
 1040 factor to compare inTime cosmics to BNB is $0.02125 * 240.8 = 5.12$.



(a) Track range distribution of selection I

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

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

1041 5.3 Optical Software Trigger and Reconstruction

1042 5.3.1 Software Trigger

1043 Most of the BNB spills from the accelerator do not have a neutrino interaction in
 1044 MicroBooNE. To save computation resources and reduce data-rates, we require a
 1045 burst of light in the light collection system in coincidence with the 1.6 μ s beam spill.
 1046 Requiring light activity in coincidence with the beam spill eliminates the vast majority
 1047 of triggers with no neutrino interaction in the detector, however, it doesn't guarantee
 1048 the activity in the detector is a neutrino interaction since a cosmic ray can interact in
 1049 coincidence with the beam spill as well.

1050 To implement this, a software trigger was used on the PMT waveforms to decide
1051 whether or not to keep that event. The software trigger is implemented after the event
1052 builder combines data from the PMTs and triggers into a single event. The software
1053 trigger uses the digitized output of the 32 PMT channels in the light collection system.
1054 Only the waveform region in coincidence with the beam spill is used to search for
1055 possible triggers. For each PMT, a waveform is found by taking the difference of ADC
1056 values is calculated between t and $t + s$. This waveform is then scanned for ADC
1057 values above a threshold X_0 . Once an ADC is above this threshold, a discriminator
1058 window is opened for a fixed number of time ticks (W_0). If the ADC count within this
1059 window W_0 is greater than a second larger threshold X_3 , a final window of width W_3
1060 is opened. The max ADC value within this final window is set as the peak amplitude
1061 for the PMT and then summed across all 32 PMTs and set to the variable PHMAX. The
1062 software trigger places a final cut on the PHMAX variable to decide whether or not
1063 to keep the event. The thresholds were found by the Trigger task force using Monte
1064 Carlo Studies and are as follows:

- 1065 • $X_0 = 5$ ADC
- 1066 • $X_3 = 10$ ADC
- 1067 • $W_0 = 6$ Ticks
- 1068 • $W_3 = 6$ Ticks
- 1069 • PHMAX cut = 130 ADC

1070 5.3.2 Flash Reconstruction

1071 MicroBooNE collects light from each of the 32 PMTs either in a continuous readout
1072 window of $23.4 \mu\text{s}$ activated by a beam gate signal on the trigger board, or in discrimi-
1073 nated pulses of $\sim 1 \mu\text{s}$ duration activated if the ADC count for any PMT goes above 80
1074 ADC count. These two formats are saved as output waveforms and put onto an event.
1075 Additionally, each PMT can provide two output streams, high-gain (~ 20 ADC/PE)
1076 and low-gain (~ 2 ADC/PE) channels. The first step in the reconstruction is to merge
1077 both these channels into a “saturation corrected waveform” which uses information
1078 from the low-gain waveform to correct for saturating high-gain pulses.

1079 The saturation corrected waveform in the continuous readout window is used to
1080 reconstruct optical hits. Each PMT’s waveform is scanned for hits then a threshold

1081 based hit reconstruction algorithm is applied which requires pulses of a minimum
 1082 area in order to be reconstructed. Each reconstructed hit is associated to a PMT, a time
 1083 in μs , and a PE count.

1084 Once hits are reconstructed for all 32 PMTs, all PMT information is then combined
 1085 into optical flashes which represent optical information seen by the PMTs from interac-
 1086 tions in the detector. Each flash has information on total light seen per interaction, the
 1087 distribution of the light across all 32 PMTs, the flash time with respect to the trigger
 1088 time of the flash, and lastly, the spacial information of the flash in Y-Z plane of the
 1089 detector. These flashes are reconstructed by requiring that there is a $\sim 1 \mu\text{s}$ coincidence
 1090 between the reconstructed hits in all 32 PMTs. The total PE is summed up among
 1091 all coincident hits across the PMTs and if the total PE is greater than 2 PE, a flash is
 1092 reconstructed. There are also safe guards in place to take care of late scintillation light.

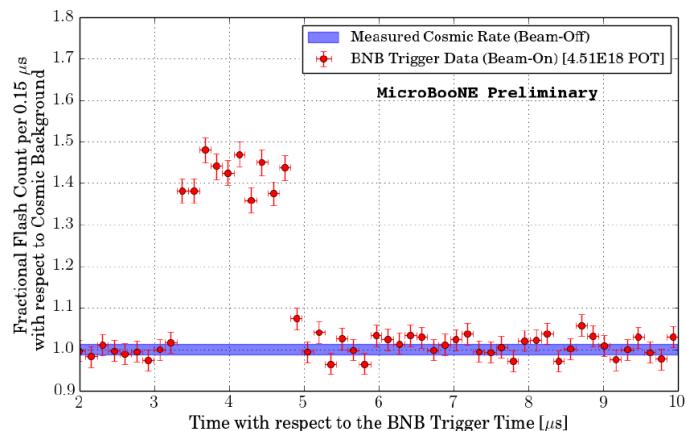


Figure 5.2: Time distribution of reconstructed optical flashes with a PE value of 50 or more for a sample of BNB unbiased triggered events.

1093 Figure 5.2 shows the time distribution of reconstructed optical flashes using the
 1094 BNB continuous stream. You can see a clear excess in coincidence with the expected
 1095 arrival time of neutrinos. The same flash reconstruction that was used in the cc-
 1096 inclusive filter detailed here was used to create this plot in data.

1097 5.3.3 Beam Window

1098 Figure 5.3 shows the distribution of flashes for on-beam, off-beam and various MC
 1099 samples. The software trigger has been applied to these samples. The pile-up seen just
 1100 after 0 μs is a feature of the flash finding algorithm and consists of low PE flashes and

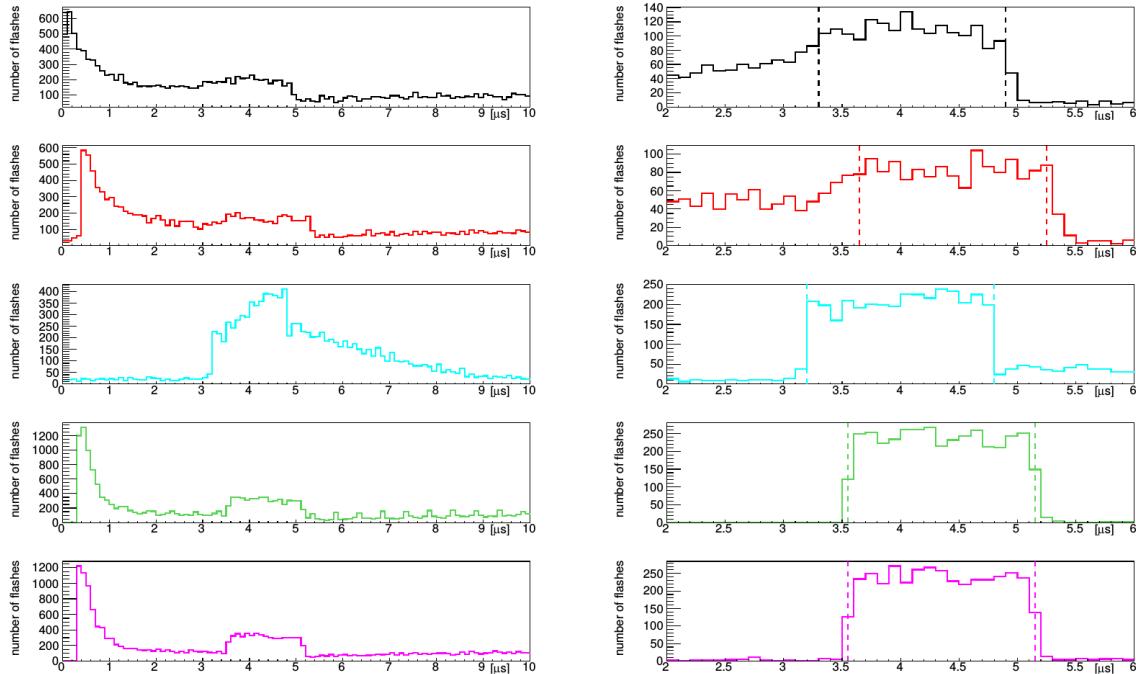


Figure 5.3: Flash time distribution for all flashes (left plot) and flashes $> 20\text{PE}$ (right plot). The different curves are as follows: on-beam data (black), off-beam data (red), CORSIKA inTime MC (light blue), BNB only MC (green), and BNB+Cosmic MC (purple). The dashed vertical lines mark the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

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

- On-Beam: 3.3 to 4.9 μ s
 - Off-Beam: 3.65 to 5.25 μ s
 - CORSIKA inTime: 3.2 to 4.8 μ s
 - BNB only: 3.55 to 5.15 μ s
 - BNB+Cosmic: 3.55 to 5.15 μ s
- Each window has a width of 1.6 μ s.

5.4 TPC Reconstruction

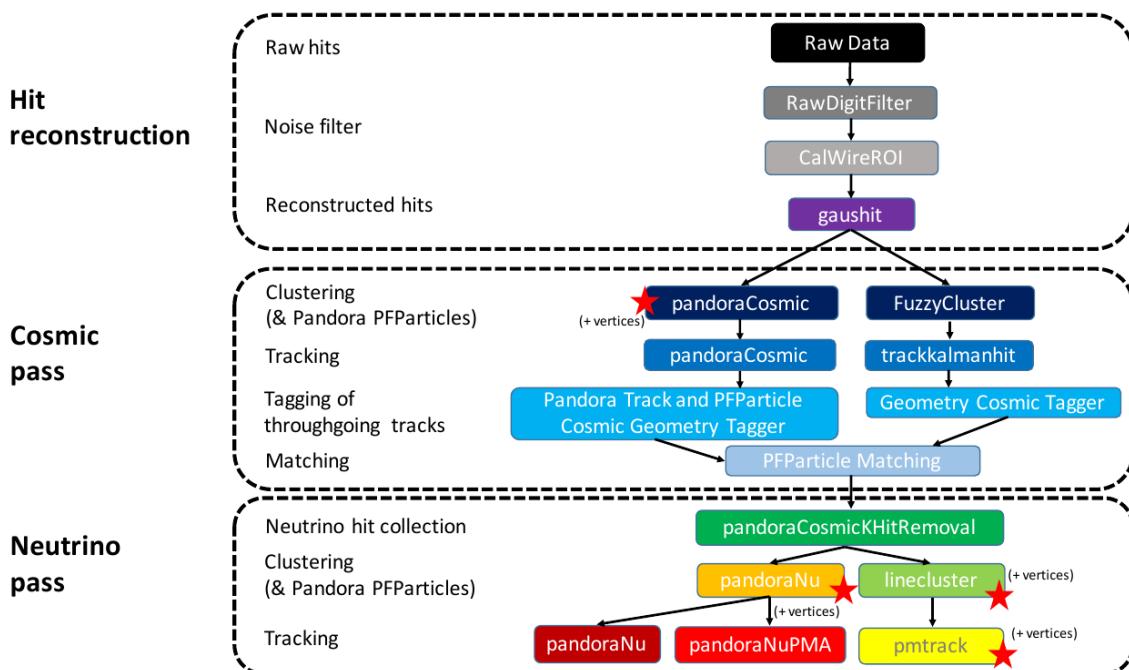


Figure 5.4: Reconstruction chain run on both data and MC. The red stars mean that the algorithm returns reconstructed 3D vertices.

1112 Figure 5.4 summarizes the reconstruction chain applied to both MC and data for
1113 this analysis. After the hit reconstruction, a cosmic pass is applied which removes all
1114 hits associated to through-going tracks. A description of these TPC reconstruction
1115 algorithms will be detailed below.

1116 **5.4.1 Hit Reconstruction**

1117 The waveforms used for hit reconstruction consist of charge deposited on the sense
1118 wire in drift time. The first step in hit reconstruction is to pass the waveforms through
1119 a filtering algorithm to filter out the noise introduced from the electronics. The input
1120 waveforms are also truncated from 9600 time ticks to 6400 time ticks in this first step
1121 to reduce the data footprint of these waveforms.

1122 Once noise filtering is complete, a deconvolution algorithm is applied to the wave-
1123 forms to remove the drift field and electronics response, therefore leaving only the
1124 ionized electrons kicked off the argon atoms by an incident track. During this process,
1125 Region of Interests (ROI) are identified and cut out of the waveforms to further reduce
1126 the data volume.

1127 The hit finding algorithm then finds candidate peaks in these ROI's and fits the
1128 peaks to Gaussian curves. These Gaussian shaped peaks are now called hits and
1129 represent the charge deposition on a wire by the incoming track. These hit objects
1130 have a peak time and width and are the basic object input to further algorithms down
1131 the reconstruction chain.

1132 **5.4.2 Clustering**

1133 There are multiple clustering algorithms used in this analysis. The main purpose of all
1134 the clustering algorithms is to associate hits together in 2D space to create objects like
1135 tracks, vertices and showers. For the fuzzy cluster algorithm, three steps are used to
1136 achieve this. The first step is to associate hits to each-other using a fuzzy clustering
1137 algorithm which gives each hit a degree of belonging to the cluster. Second, a Hough
1138 transform is used to find hits associated to candidate tracks and showers within each
1139 of the clusters found in the first step. The last step merges smaller candidate tracks
1140 and showers into large clusters. The last step also associates unclustered hits into

1141 nearby objects which helps shower reconstruction. The result is a set of clusters made
1142 up of associate hits that represent tracks or showers per plane.

1143 The pandora algorithm utilizes it's own clustering algorithm and will be detailed
1144 in the next section. The last clustering algorithm is called linecluster. The linecluster
1145 algorithm reconstructs 2D linear clusters per plane by fitting a line onto nearby hits
1146 which is then extrapolated to neighboring wires. 2D vertices are found per plane by
1147 using the intersection points of the ends of nearby clusters. These 2D vertices are then
1148 matched in time across all three planes to get a 3D vertex in space.

1149 5.4.3 Pandora

1150 5.4.4 Trackkalmanhit

1151 The trackkalmanhit algorithm takes 2D clusters returned from the fuzzy cluster algo-
1152 rithm and outputs track objects. There are no hierarchy structure as it is in pandora,
1153 each track is independent. There also is no vertex reconstruction with this algorithm
1154 as well.

1155 5.4.5 Cosmic Hit Removal

1156 The Pandora algorithm is applied to the events twice, the first to remove downward
1157 going tracks primarily from cosmic ray muon like particles. The second pass only runs
1158 on a subset of hits that aren't associated with cosmic ray muon tracks.

1159 After the first pass, the output of PFParticle hierarchy is then passed to a cosmic
1160 ray tagger to look through all hits to determine start and end points. If the start or
1161 end point trajectories are consistent with entering or exiting the TPC, then these hits
1162 are removed from the second pass. Hits are considered entering or exiting the TPC
1163 if the drift time are outside of the neutrino drift window or outside of the fiducial
1164 volume of the TPC. The fiducial volume was based on a montecarlo study and is 20
1165 cm from the top or bottom of the TPC and 10 cm from the TPC ends. Hits associated
1166 with candidate cosmic ray tracks are removed from the input hit collection and the
1167 remaining hits are passed to the neutrino optimized pass of Pandora.

1168 5.4.6 Projection Matching Algorithm

1169 The projection matching algorithm (PMA) was inherited from ICARUS and has been
1170 implemented in LArSoft. PMA differs from traditional LArSoft 3D reconstruction
1171 algorithms. Most 3D reconstruction attempts to match 2D objects from all three planes
1172 by drift time, while the PMA algorithm projects a track hypothesis on each plane
1173 then the distance between this projection and the hits on each plane is minimized
1174 simultaneously. More information can be found in [?].

1175 5.5 Event Selection

1176 The first requirement for selecting ν_μ CC events is that the event has at least one
1177 scintillation light flash in the beam trigger window with more than 50 PE on all PMTs
1178 combined. From the flashes that pass, the most intense is chosen and considered to be
1179 originating from a neutrino interaction and will be the only flash used in further cuts.

1180 Vertices are then required to have at least one reconstructed track start or endpoint
1181 within a 5 cm radius. Showers associated with a vertex do not pass this cut. All
1182 tracks associated with a vertex are then used to calculate a track length weighted
1183 average of the θ -angle. Of all the vertices that do pass, only the vertex with the most
1184 forward going θ -angle average of all associated tracks is considered the neutrino vertex
1185 candidate. The most forward going θ -angle average is chosen by picking the largest
1186 track range weighted average of $|\cos(\theta)|$, seeing as $\cos(\theta) = 1$ is the beam direction.
1187 Next, it is required that the reconstructed neutrino vertex candidate be within the
1188 fiducial volume as well as within the drift time starting at t_0 . The fiducial volume
1189 boundaries chosen are 10 cm from the edges of the TPC in x and z which is the drift
1190 direction and beam direction respectively, and 20 cm from the edges of the TPC in y
1191 which is the vertical direction. For all further cuts, only the longest track associated
1192 with the neutrino vertex candidate and this track is assumed to be the muon candidate
1193 of the neutrino event.

1194 The next cut requires the position of the flash in the z-direction and the track z-
1195 projection to be compared. This basic flash matching algorithm is rudimentary and a
1196 placeholder for a more sophisticated algorithm. The z-position of the flash needs to be
1197 within 80 cm to the z-positions of track start or endpoints. If the flash is between the
1198 track start and endpoint, the distance of the flash to the track is considered to be 0 cm.

1199 Lastly, the track needs to be fully contained within the fiducial volume and have a
1200 track range greater than 75 cm. The range is the 3D distance between the track's start
1201 and endpoint. The length cut was optimized to remove NC background that contain
1202 a pion due to the pion interaction rate to be ~ 70 cm. A track that makes all the cuts
1203 is considered to be the muon of a ν_μ CC event. The list of cuts for this selection is
1204 described below:

- 1205 1. At least one flash > 50 PE within the beam gate.
- 1206 2. At least one track within 5 cm around a vertex.
- 1207 3. Vertex with flattest tracks is chosen to be vertex candidate.
- 1208 4. Vertex candidate in fiducial volume.
- 1209 5. Longest track associated with vertex candidate is chosen to be track candidate.
- 1210 6. Longest track is within 80 cm (z-axis only) of the flash.
- 1211 7. Longest track is fully contained.
- 1212 8. Longest track is greater than 75 cm.

1213 The event selection scheme can also be seen in figure 5.5. Table ?? lists the passing
1214 rates for MC events for the selection scheme described above. Table ?? lists the passing
1215 rates for on-beam and off-beam data for the selection scheme. The normalization
1216 factors applied between on-beam and off-beam data are described in section 5.2.

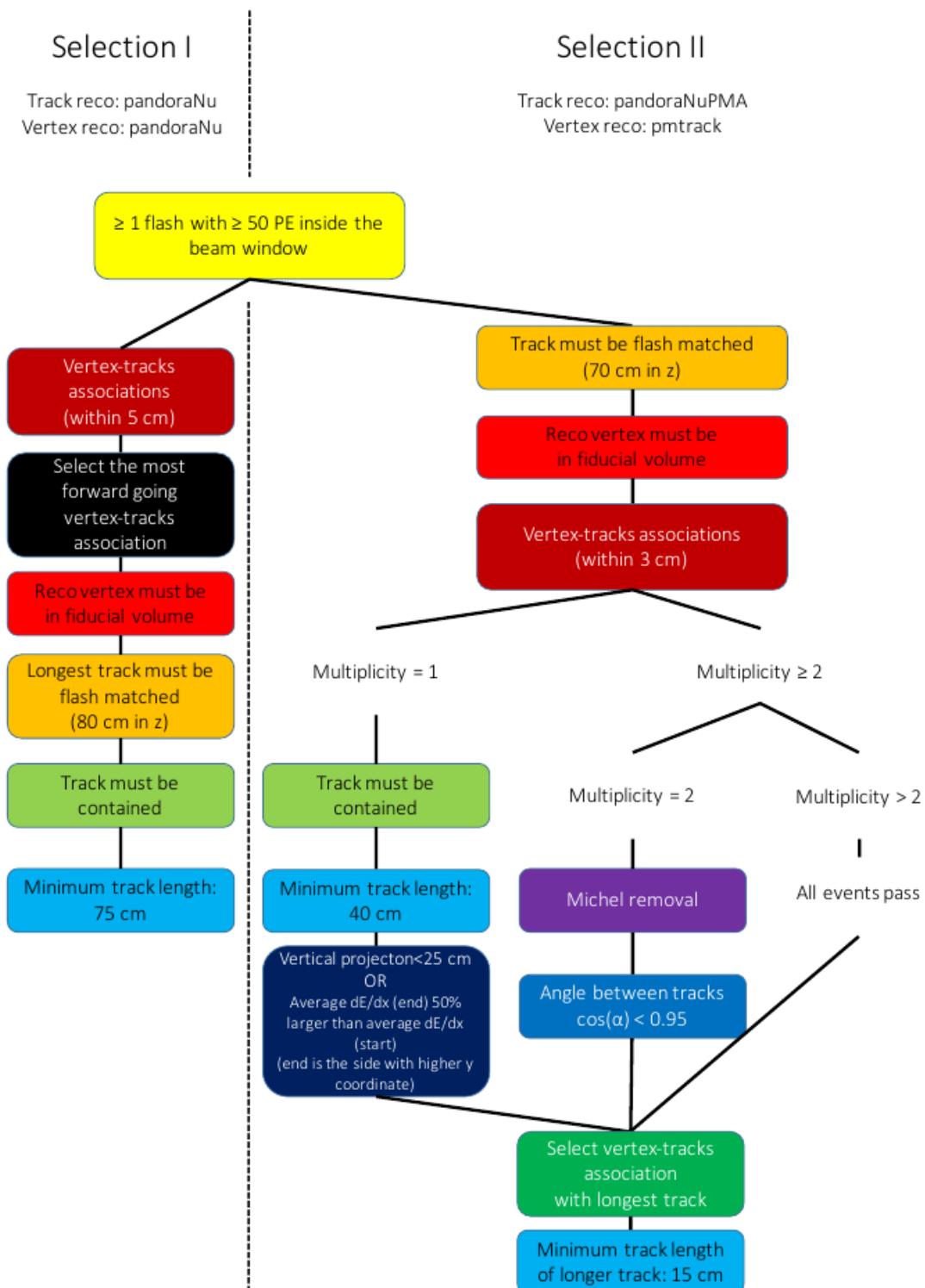


Figure 5.5: Event selection diagram for selection I and selection II. This analysis focused on optimizing selection I. Boxes with the same color across the two selections symbolize similar cuts.

₁₂₁₇ **Chapter 6**

₁₂₁₈ **Background on Convolutional Neural Networks**
₁₂₁₉

₁₂₂₀ Convolutional neural networks (CNNs) have been one of the most influential innovations in the field of computer vision. Neural networks became popular in 2012
₁₂₂₁ when Alex Krizhevsky used them to win that year's ImageNet competition [?] by
₁₂₂₂ dropping the error from 26% to 15%. Since then, many companies are using deep
₁₂₂₃ learning including Facebook's tagging algorithms, Google for their photo search and
₁₂₂₄ Amazon for product recommendations. For the purpose of this thesis CNNs were
₁₂₂₅ used for image classification, specifically, images of varying particles created using
₁₂₂₆ LArTPC data.
₁₂₂₇

₁₂₂₈ **6.1 Image Classification**

₁₂₂₉ Image classification is the process of inputting an image into the CNN and receiving a
₁₂₃₀ probability of classes that best describes what is happening in the image. As humans,
₁₂₃₁ image classification is something that is learned at a very young age and is easy to
₁₂₃₂ do without much effort. This is also apparent when hand-scanning LArTPC images.
₁₂₃₃ After learning what a neutrino event looks like in MicroBooNE, it is relatively easy
₁₂₃₄ to recognize simple neutrino events from cosmic ray background as well as highly
₁₂₃₅ ionizing particles like protons from minimum ionizing particles like muons. The very
₁₂₃₆ detailed images LArTPC detectors output are prime candidates for input images into
₁₂₃₇ a CNN. CNNs mimic a human's ability to classify objects by creating an architecture
₁₂₃₈ 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.

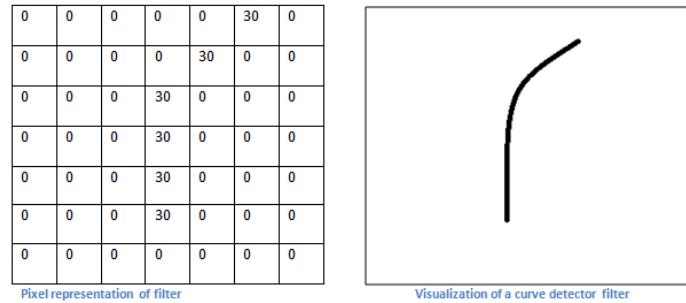


Figure 6.1: Pixel representation and visualization of a curve detector filter. As you can see, in the pixel representation, the weights of this filter are greater along a curve we are trying to find in the input image

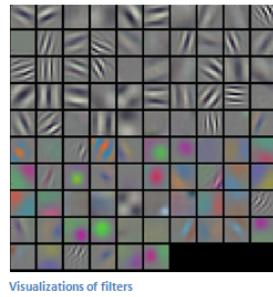


Figure 6.2: Visualization of filters found in first layer of a CNN.

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.

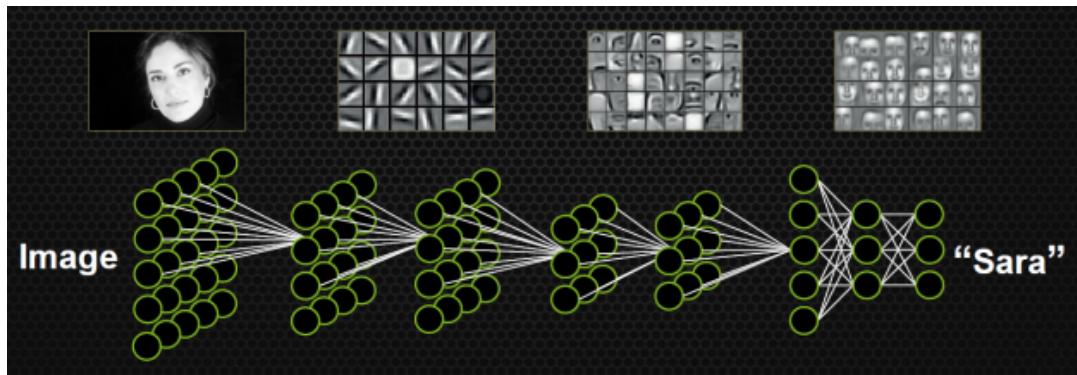


Figure 6.4: Pictorial Representation of Convolutional Neural Networks as well as a visual representation on CNN's complexity of layer feature extraction

1288 There are other layers in a CNN architecture that will not be covered in the scope
 1289 of this thesis but in a general sense, these layers are interspersed between convolution
 1290 layers to preserve dimensionality and control overfitting of the network. The last layer
 1291 is called a fully connected layer and it's job is to output an N dimensional vector where
 1292 N is the number of classes the network has been trained on. Each number in this vector
 1293 represents the probability that the input image is a certain class. Fully connected layers
 1294 use the feature maps of the high level features to compute the products between the
 1295 weights of the previous layer to get the probabilities of each class. These weights are
 1296 then adjusted through the training process using backpropagation.

¹²⁹⁷ 6.2.1 Backpropagation

¹²⁹⁸ A CNN at it's onset has weights that are randomized. The filters themselves don't
¹²⁹⁹ know how to pull out identifying information per class. For a neural network to learn,
¹³⁰⁰ it must be trained on a training set that is labeled. Backpropagation has four seperate
¹³⁰¹ steps: foward pass, loss function, backward pass and updating weights. In the forward
¹³⁰² pass, a training image is passed through the whole network. All of our weights at this
¹³⁰³ time are randomly initialized so the output for the first image will have no preference
¹³⁰⁴ to a specific class. A common loss function is mean squared error (MSE):

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

¹³⁰⁵ If we assume that the MSE is the loss of our CNN, the goal would be that our
¹³⁰⁶ predicted label (output of CNN) is the same as our training label. To do this, we need
¹³⁰⁷ to minimize the loss function. To do this, it is necessary to find out which weights most
¹³⁰⁸ directly affect the loss of the network i.e $\frac{dL}{dW}$ where L is our loss function and W are
¹³⁰⁹ the weights of a specific layer. The next step is the backward pass which determines
¹³¹⁰ which weights contribute the most to the loss and finds ways to adjust these weights
¹³¹¹ so that the loss decreases. After the derivative is computed, the last step updates the
¹³¹² weights in the opposite direction of the gradient.

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

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

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

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

¹³¹³ The learning rate is a parameter given to the CNN and it describes the steps the
¹³¹⁴ network takes to update the weights. Higher learning rate equals large steps and a
¹³¹⁵ lower training time, but a learning rate that is too large can mean the CNN never
¹³¹⁶ converges.

¹³¹⁷ Going through backpropagation consists of one training iteration. Once the net-
¹³¹⁸ work completes a specific number of iterations, another parameter given, and runs
¹³¹⁹ over all training images that are split up into batches, the process is considered com-
¹³²⁰ plete. User input parameters, called hyperparameters, help the network converge to

1321 optimal weights for each layer. Batch size, learning rate, and training iteration are just
1322 some of the user input hyperparameters that help. Lastly, to check if the network has
1323 learned, a different set of labeled images are fed to the CNN iteratively through the
1324 training process to see how well it's learning. This process is especially important to
1325 make sure the network architecture isn't being affected by overfitting (memorizing
1326 training input rather than learning).

1327 **6.3 Choosing Hyperparameters**

1328 Convolutional neural networks are a relatively new tools in computer vision. Choosing
1329 hyperparameters for your specific dataset is a non-trivial task. Hyperparameters can
1330 range from the amount of layers and filters per layer in an CNN architecture to the
1331 stride the receptive field of a filter takes, not to mention training hyperparameters
1332 such as learning rate and batch size described above. They're ways to optimize these
1333 hyperparameters via hyperparameter optimization using Bayesian Optimization [?]
1334 but as you can imagine, optimizing an CNN architecture from scratch can be very
1335 computationally intensive. For the purpose of this thesis, two well known CNN
1336 architectures were used, AlexNet [?], which won the ImageNet Large-Scale Visual
1337 Recognition Challenge (ILSVRC) in 2012 and therefore bringing awareness of CNNs,
1338 and GoogleNet [?], which won the ILSVRC in 2014, giving rise to deep networks. Both
1339 AlexNet and GoogleNet architectures were used to train on LArTPC images and their
1340 low level filter weights. Higher level filter weights were randomly initialized before
1341 training so the network can learn high level features of LArTPC image classes. The
1342 AlexNet architecture is shown in figure 6.5 and the GoogleNet architecture is shown
1343 in figure 6.6

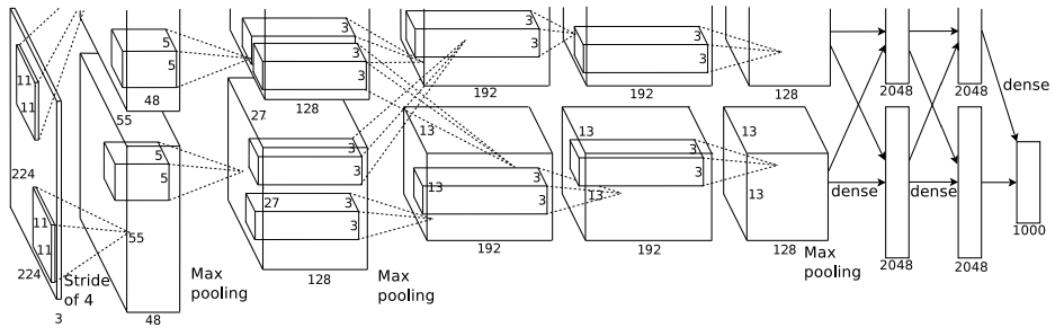


Figure 6.5: Pictoral representation of the AlexNet model. The AlexNet model consists of 5 convolution layers and 3 fully connected layers.

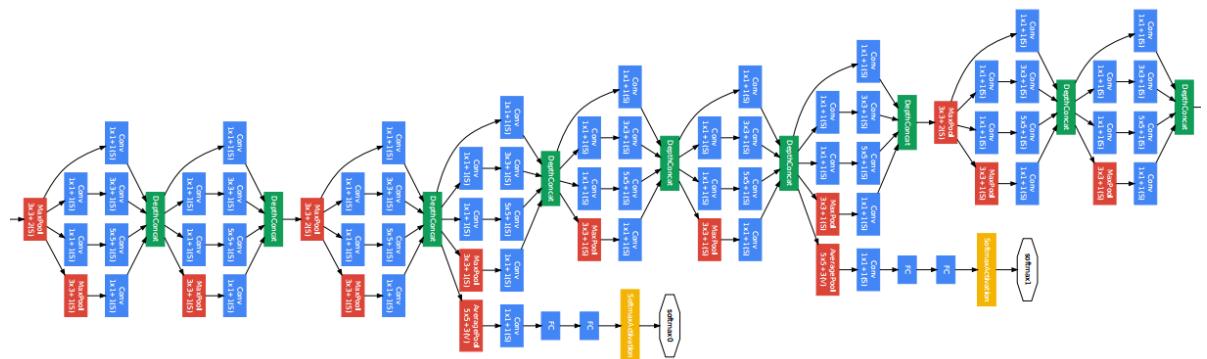


Figure 6.6: Pictoral representation of the GoogleNet model. The GoogleNet model consists of 22 layers. The model implements 9 Inception modules which performs covolution and pooling in parallel and strays away from the basis that CNN layers need to be stacked up sequentially. The GoogleNet model also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

1344 Chapter 7

1345 Training Convolutional Neural 1346 Networks on particles **WORKING** 1347 **TITLE**

1348 Three Convolutional Neural Networks (CNNs) were trained throughout this analysis.
1349 There are differences to each CNN and will be described fully in the next sections but
1350 the main difference are the amount of particle images used for training and validation.
1351 CNN1075 used 1,075 muons and 1,075 pions for training and the same amount of each
1352 particle for validation. CNN10000 used 10,000 muons and 10,000 pions split in half
1353 for testing and training. Lastly CNN100000 had muons, pions, protons, electrons,
1354 and gammas in its training and validation set. Each particle had 20,000 images and
1355 training and validation was split 90% training, 10% validation. This chapter will also
1356 describe the different hardware frameworks used for training beginning on a CPU
1357 and ending on a GPU cluster.

¹³⁵⁸ **7.1 Hardware Frameworks used for Training**

¹³⁵⁹ **7.1.1 Syracuse CPU Machine setup**

¹³⁶⁰ **7.1.2 Syracuse University GPU Cluster Setup**

¹³⁶¹ **7.2 Convolutional Neural Network Training**

¹³⁶² **7.2.1 Image Making Scheme**

¹³⁶³ **Images used for Traing/Validation of Convolutional Neural Networks**

¹³⁶⁴ The μ/π image dataset used to train and validate CNN1075 was created using single
¹³⁶⁵ generated isotropic muons and pions from 0-2 GeV energy range. 2,150 muons and
¹³⁶⁶ 2,150 pions were used for training and testing split 50 %. The images were created
¹³⁶⁷ based on wire number and time tick in the collection plane. Uboonecode reconstruction
¹³⁶⁸ version v05_08_00 was used. The raw ADC value after noise filtering was the wire
¹³⁶⁹ signal. Each collection plane grayscale image was 3456x1600x1 where 6 time ticks
¹³⁷⁰ were pooled into 1 bin. After the image was created, the region of interest (ROI) in
¹³⁷¹ the image was found by using Open CV, a image processing open source software
¹³⁷² package, to scan the image starting from the edges and stopping once a bright spot is
¹³⁷³ encountered. Thi

¹³⁷⁴ The μ/π image dataset used to train and validate the CNN10000 was also created
¹³⁷⁵ using single generated isotropic muons and pions from 0-2 GeV energy range. 10,000
¹³⁷⁶ muons and 10,000 pions were used for training and testing split 50%. Uboonecode
¹³⁷⁷ v06_23_00 was used instead of v05_08_00. Each collection plane grayscale image was
¹³⁷⁸ 3456x1280x1 where 5 time ticks were pooled into 1 bin which is different than the
¹³⁷⁹ previous dataset and was implemented due to the fact that the time ticks of an event
¹³⁸⁰ went from 9400 to 6400 with the change of uboonecode version. The grayscale color
¹³⁸¹ standard is 8bit therefore the ADC value of wire and time tick was also downsampled
¹³⁸² due to the 12bit ADC value MicroBooNE has. To do this, the highest ADC pixel in
¹³⁸³ the image was found and then this was divided by the rest placing all pixel values
¹³⁸⁴ between 0-1. From there, all pixel values are then multiplied by 255. All images
¹³⁸⁵ were made using a LArSoft module. Once the images were created, using and image

1386 manipulation framework called OpenCV images were read into a numpy array and
1387 cropped to the region of interest by only keeping rows and columns where all ADC
1388 values are higher than 0 and then resized it to 224x224 using OpenCV's resize function.
1389 This downsampling of ADC values creates a problem of information loss for example,
1390 a proton which is highly ionizing will have the same brightness as a minimum ionizing
1391 muon by virtue of how the images are created. Issues that arose in CNN1075 that
1392 were fixed in CNN10000 include zero-padding images in X and Y that are smaller
1393 than 224X224 to eliminate over-zooming effect and fixing a bug that shifted pixels
1394 separated by a dead-wire region.

1395 Images were also made from events that passed the cc-inclusive selection 1 filter
1396 right before the 75 cm track length cut and were classified using the CNN10000. The
1397 dataset used to create these images is the same one used in [?], prodgenie_bnb_nu_cosmic_uboone_mc
1398 These images were created using information from the track candidate that passed
1399 the filter. Only wire number and time ticks associated to the track candidate were
1400 drawn on the image to mimic a single particle generated image. These images were
1401 then classified using CNN10000. Two approaches were taken in making these images.
1402 The first was using the image normalization above where the maximum pixel in each
1403 image is used as a normalization constant to get all pixels between 0-1 then multiply
1404 all pixels by 255. As described above, this is the incorrect way to normalize; it should
1405 be normalized by dataset not by event, which is the second way the images were
1406 created. The results of CNN10000 performance are shown in section 7.2.

1407 7.2.2 Training CNN1075

1408 The work shown in these next sections are based on the previous work done described
1409 in [?]. That CNN (now referred to as CNN1075) was trained using single generated
1410 isotropic muons and pions from 0-2 GeV energy range. 1,075 muons and pions were
1411 used to train the network and 1,075 μ/π were used as a validation set. The accuracy is
1412 how well CNN1075 is doing by epoch and was 74.5%. The loss is gradient descent
1413 or minimization of the error of the weights and biases used in each neuron of each
1414 layer of CNN1075 and was 58% with a trend sloping downwards on the loss curve
1415 as well as a trend sloping upward in the accuracy curve. Due to the depth of the
1416 neural network framework, it was necessary to train with a larger dataset and for
1417 more epochs, however, the downward slope of the loss curve is an indication that once
1418 trained for longer with a higher training sample, neural networks can be used for μ/π

¹⁴¹⁹ separation. Updates in the image making and downsampling algorithm were made to
¹⁴²⁰ fix issues that arose in CNN1075.

¹⁴²¹ **7.2.3 Training CNN10000**

¹⁴²² The hyperparameters used for CNN10000 are shown. The batch size for the training
¹⁴²³ and testing as well as the test iter were chosen to encompass the whole training/testing
¹⁴²⁴ image set when doing accuracy/loss calculations. To do this, multiplying the test
¹⁴²⁵ iter by the test batch size give you the amount of images used when calculating
¹⁴²⁶ accuracy/loss curves. For reference, the accuracy and loss are defined as well.

- ¹⁴²⁷ • train_batch_size: 100
- ¹⁴²⁸ • test_batch_size: 100
- ¹⁴²⁹ • test_iter: 100
- ¹⁴³⁰ • test_interval: 100
- ¹⁴³¹ • base_lr: 0.001
- ¹⁴³² • lr_policy: "step"
- ¹⁴³³ • gamma: 0.1
- ¹⁴³⁴ • stepsize: 1000
- ¹⁴³⁵ • display: 100
- ¹⁴³⁶ • max_iter: 10000
- ¹⁴³⁷ • momentum: 0.99
- ¹⁴³⁸ • weight_decay: 0.0005
- ¹⁴³⁹ • snapshot: 100
- ¹⁴⁴⁰ • Accuracy: How often the CNN predicts the truth over total number of images
- ¹⁴⁴¹ • Loss: Error between truth and prediction. Minimize loss by gradient descent to
¹⁴⁴² update weights and biases of CNN

¹⁴⁴³ The same architecure that was used to train CNN1075 was employed on CNN10000,
¹⁴⁴⁴ Imagenet. Caffe [?] was the software package used for both CNNs. The differences

1445 include batch size and test_iter and momentum to account for the larger dataset. Both
 1446 CNNs were trained on a CPU machine, Syracuse01. Further training will be done
 1447 on a GPU cluster stationed at Syracuse University. Figure 7.1 shows the loss and
 1448 accuracy of CNN10000. There is around a 10% increase in accuracy from CNN1075 to
 1449 CNN10000, 85%, and around a 20% decrease in loss, 36%.

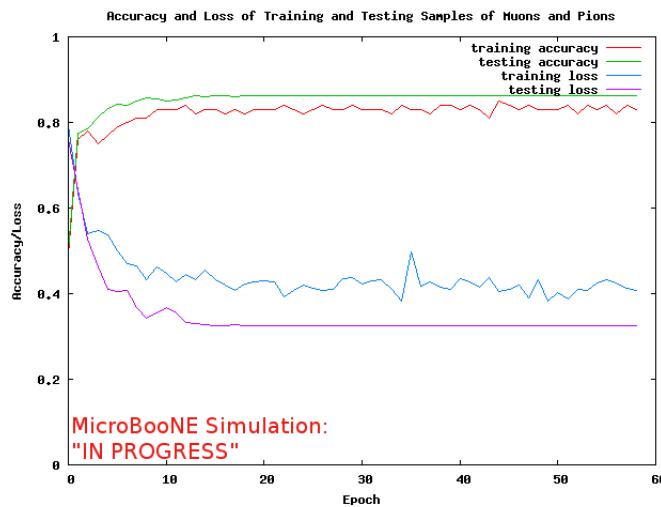
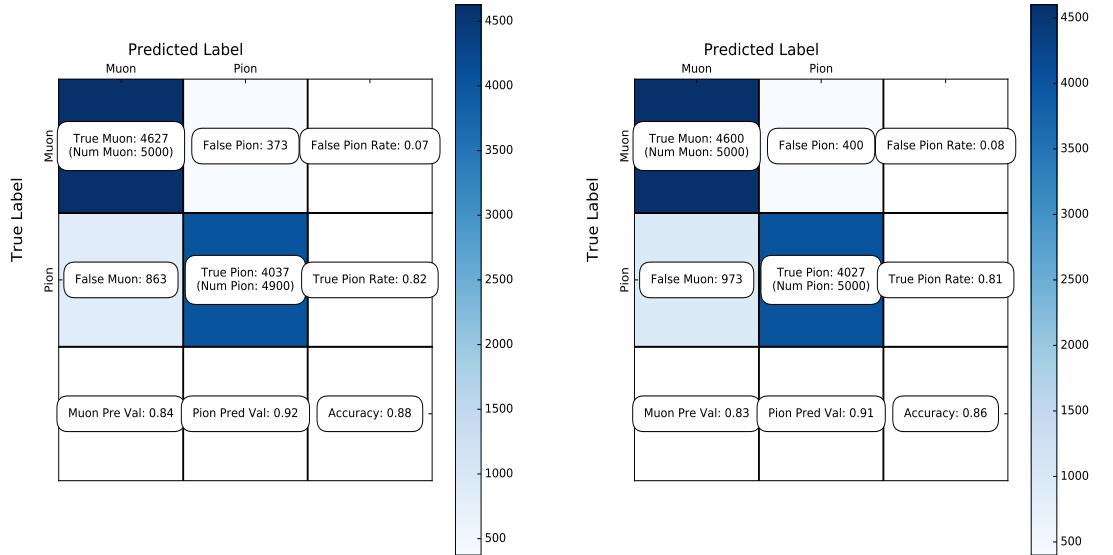


Figure 7.1: Accuracy vs. Loss of ImageNet 2-output μ/π sample consisting of 10000 images each.

1450 Figure 7.2 show a breakdown of μ/π separation for CNN10000. It also shows
 1451 the network is not being overtrained due to the Accuracy of both the training and
 1452 testing datasets being within .01% of eachother. The CNN is doing a very good job of
 1453 classifying true muons as muons, and our loss increase from CNN1075 is due to the
 1454 increase in accurately classifying pions as pions.

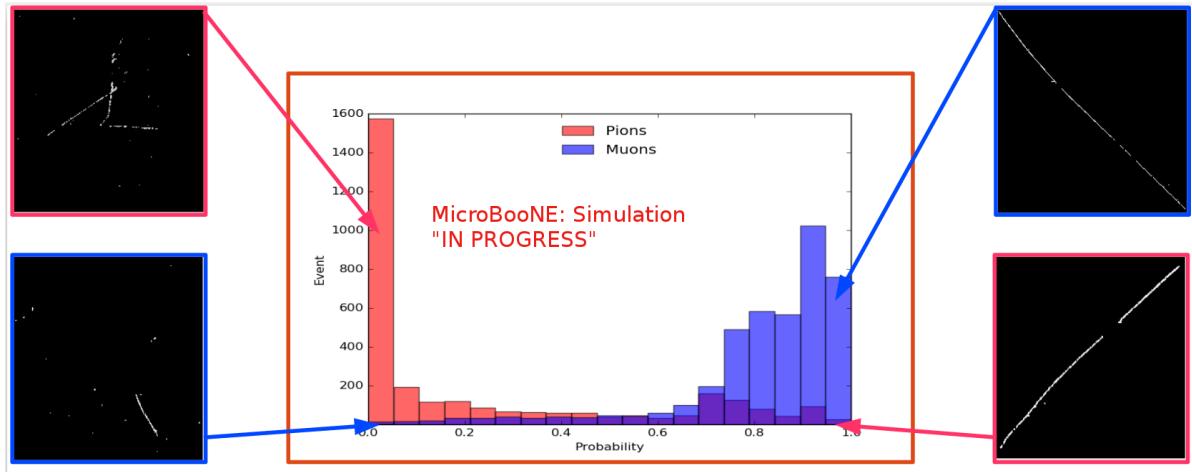
1455 7.2.4 Training CNN10000

1456 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.

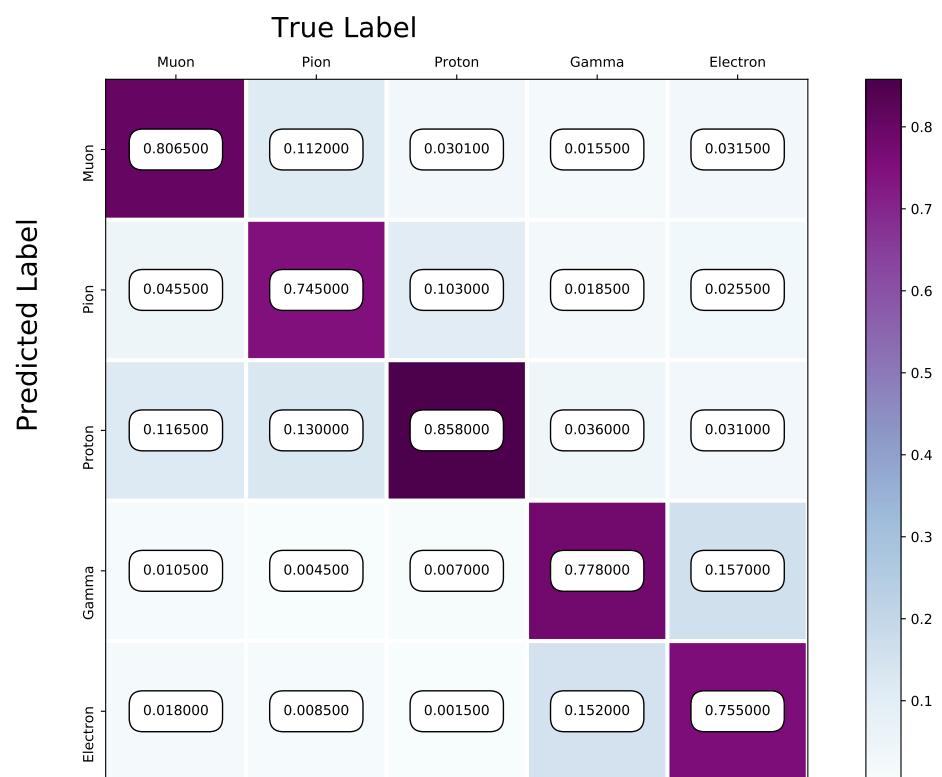


Figure 7.3: Confusion Matrix of all five particles

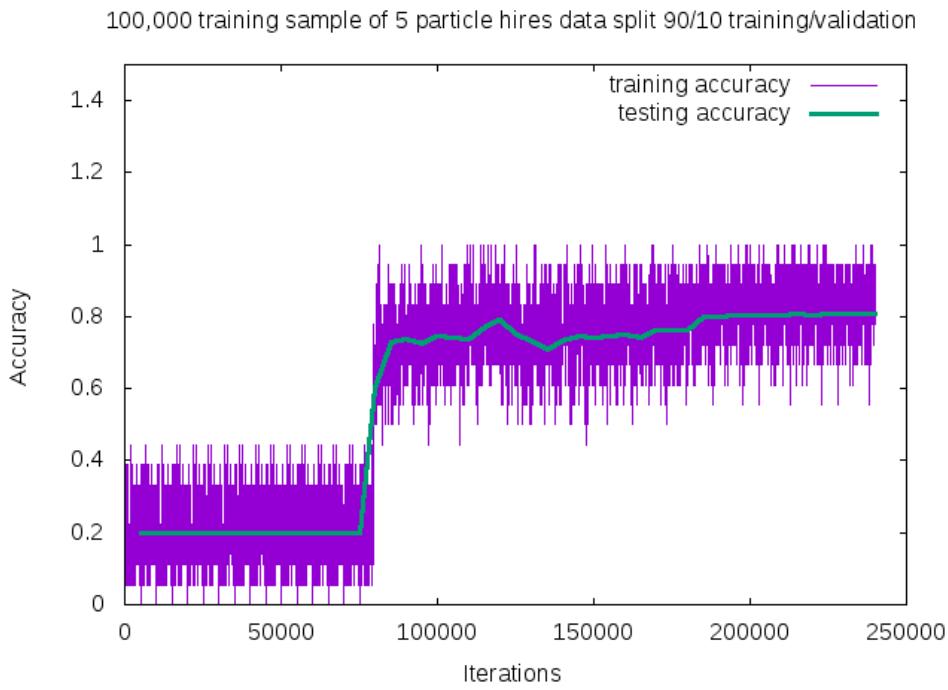


Figure 7.4: Training and testing accuracy of CNN trained on 100,000 images of $\mu/\pi/p/\gamma/e$ with 20,000 images of each particle. Each image was a size of 576x576 and the images per particle were split 90% use for training and 10% used for testing the network

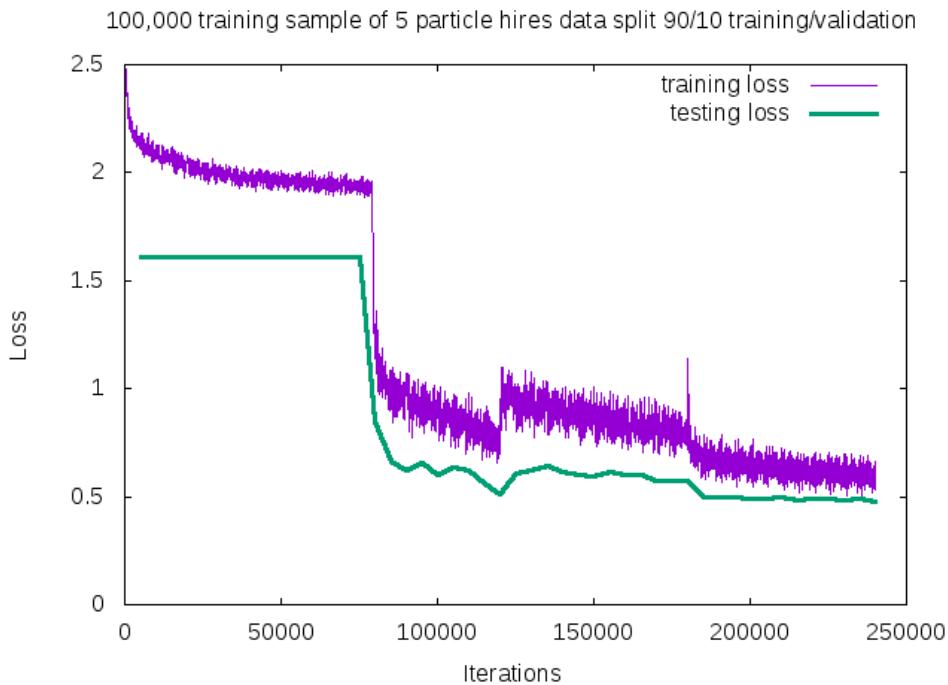


Figure 7.5: Training and testing loss of CNN trained on 100,000 images of $\mu/\pi/p/\gamma/e$

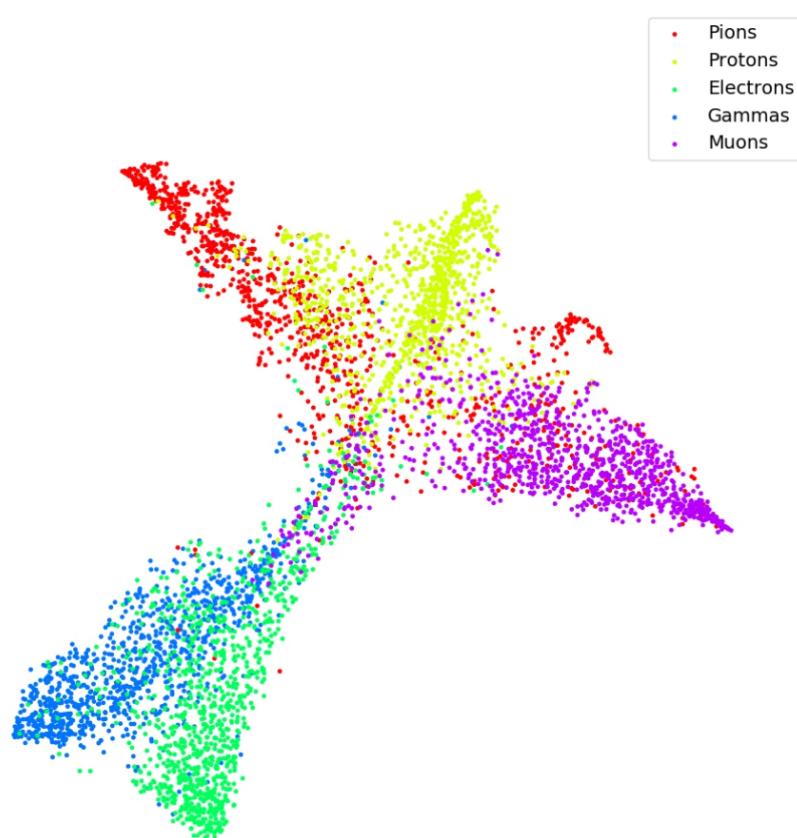


Figure 7.6: t-SNE of CNN

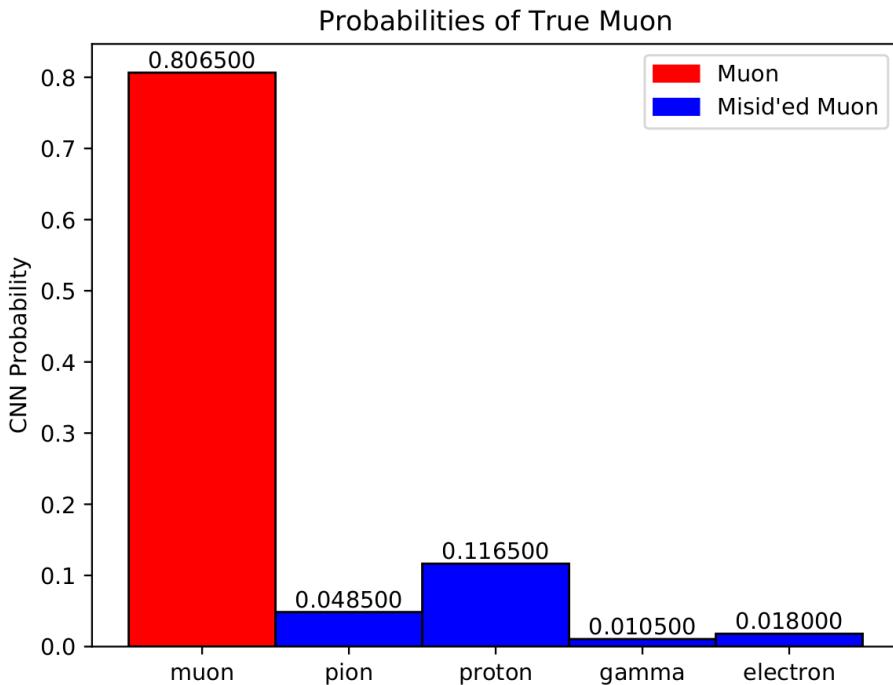


Figure 7.7: Muon Prob

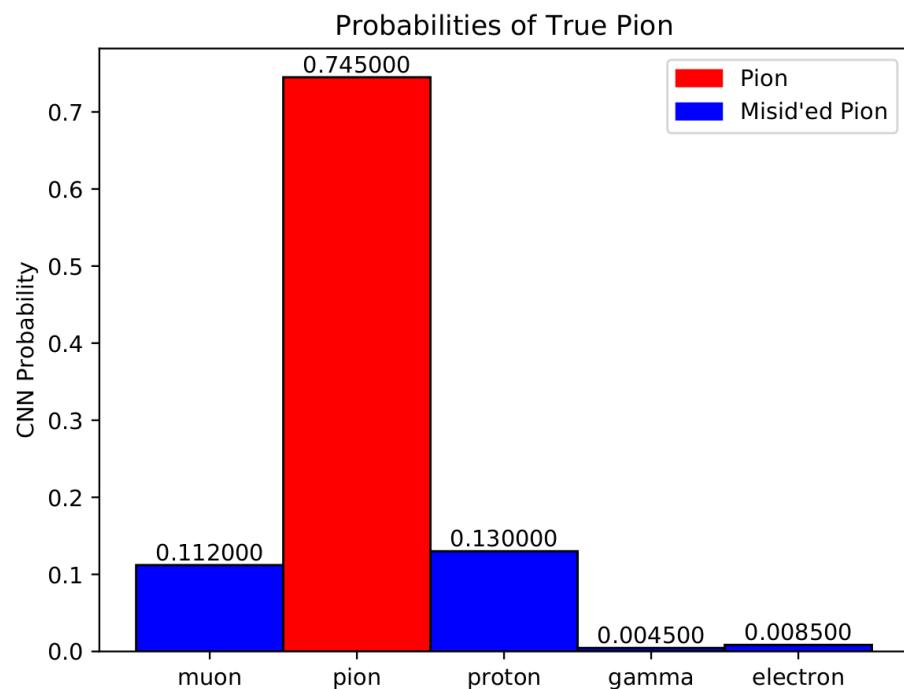


Figure 7.8: Pion Prob

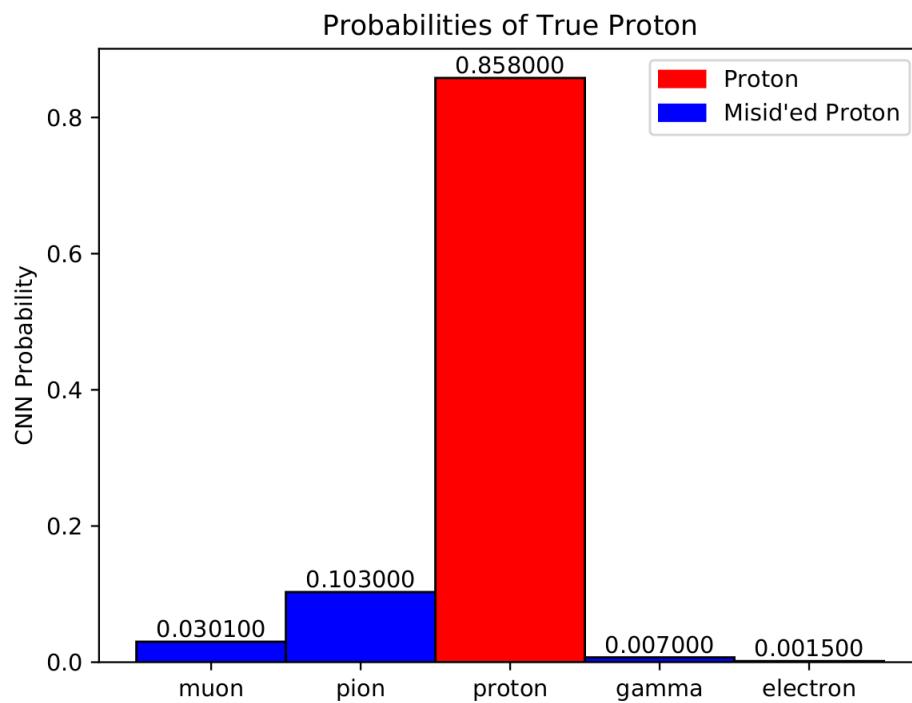


Figure 7.9: Proton Prob

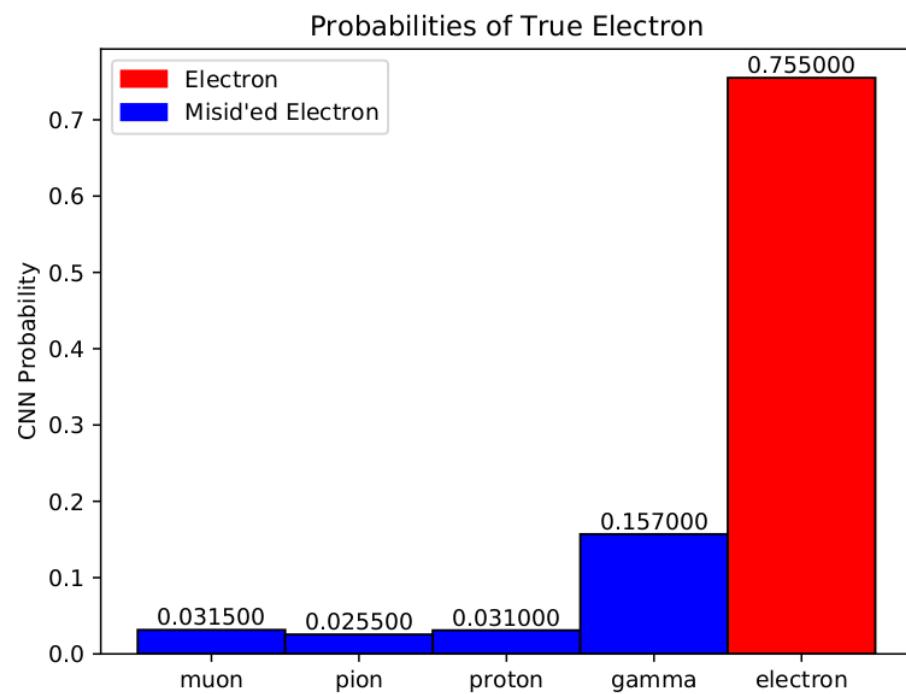
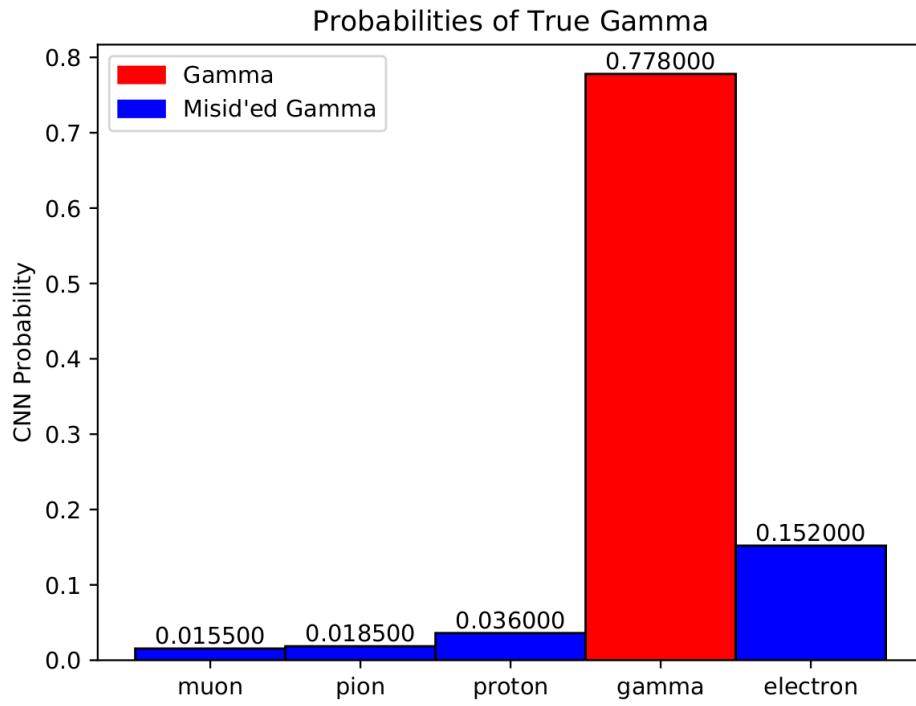
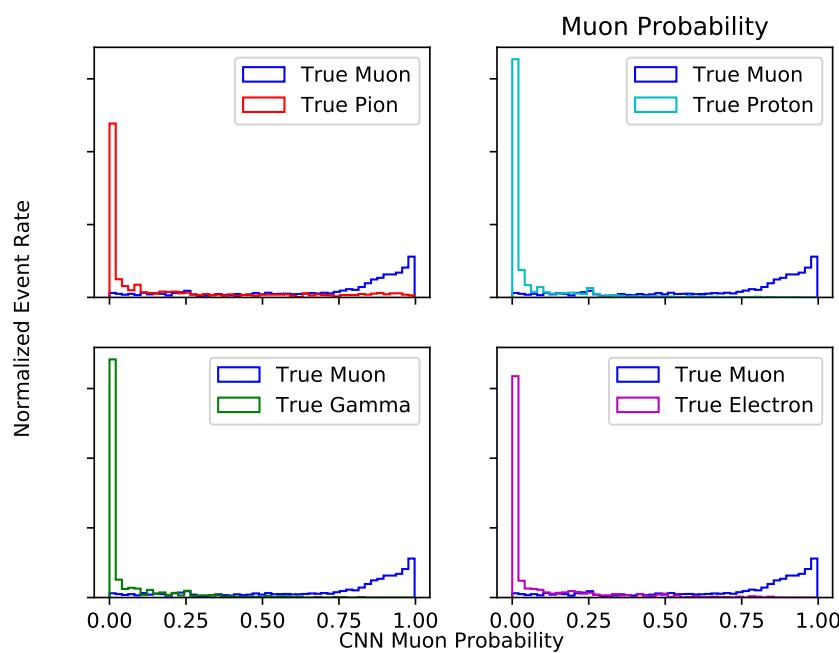


Figure 7.10: Electron Prob

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

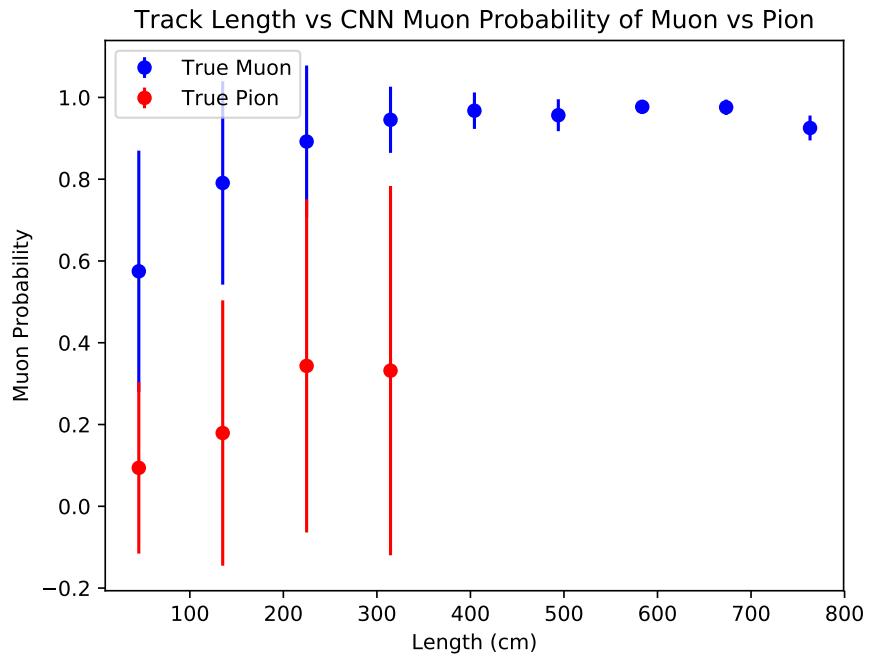


Figure 7.13: mupi

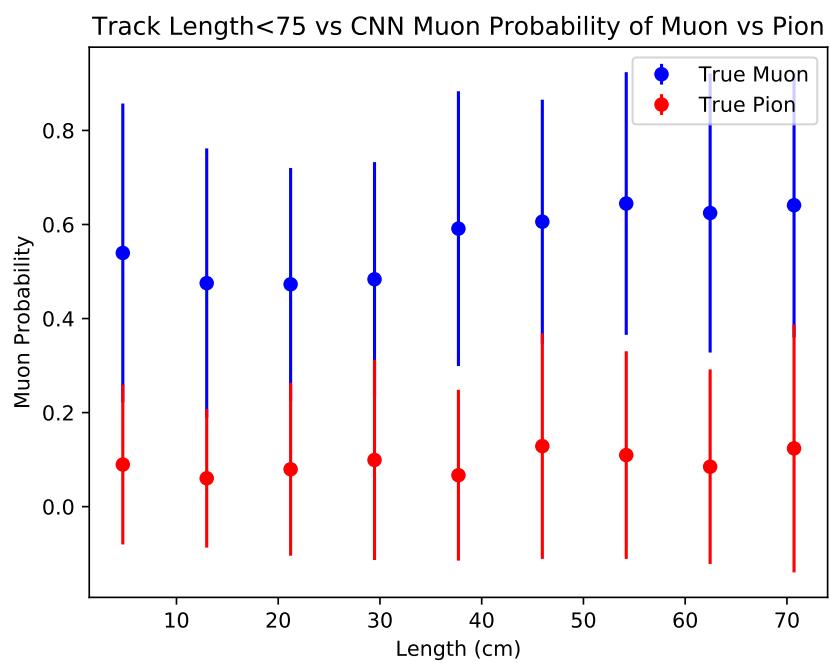
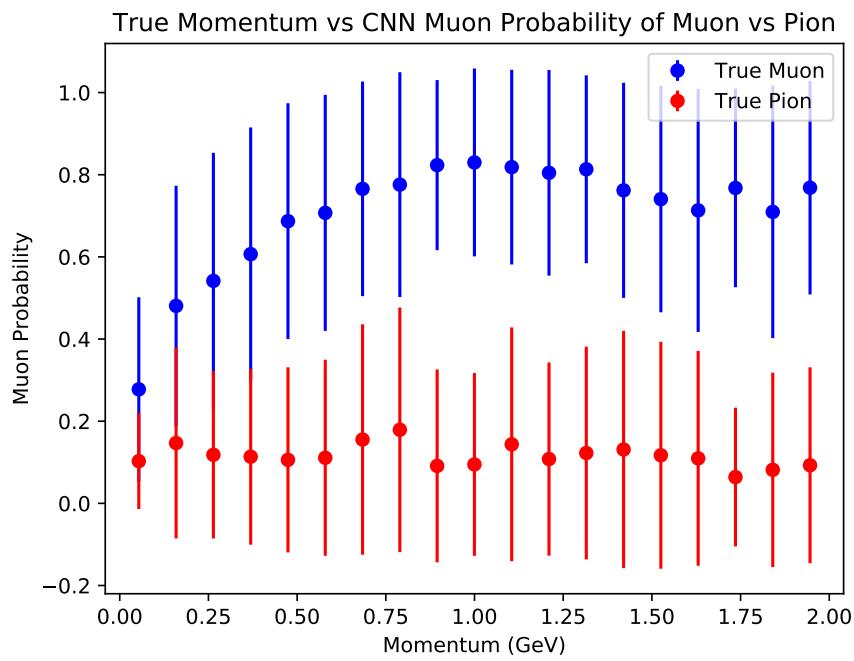
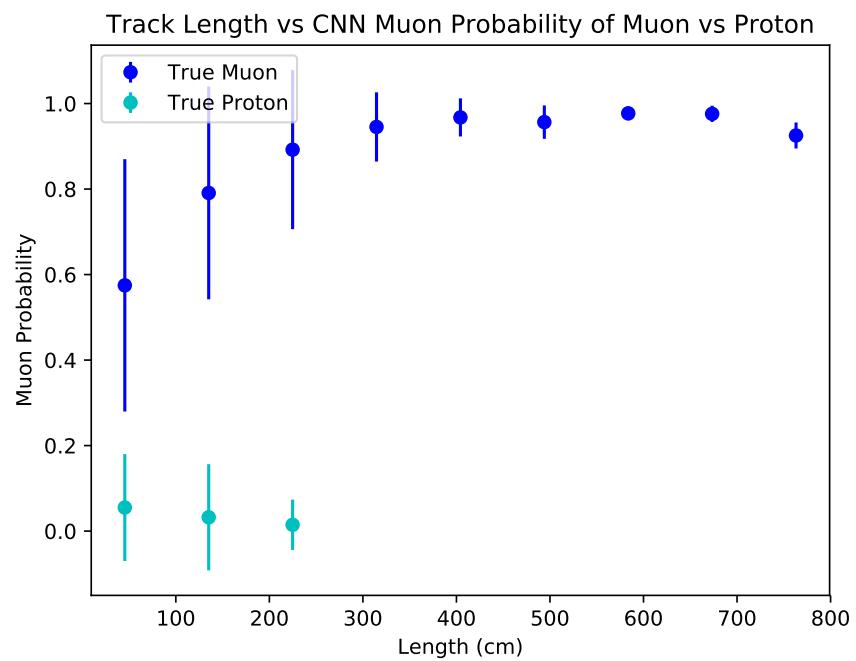


Figure 7.14: mupi

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

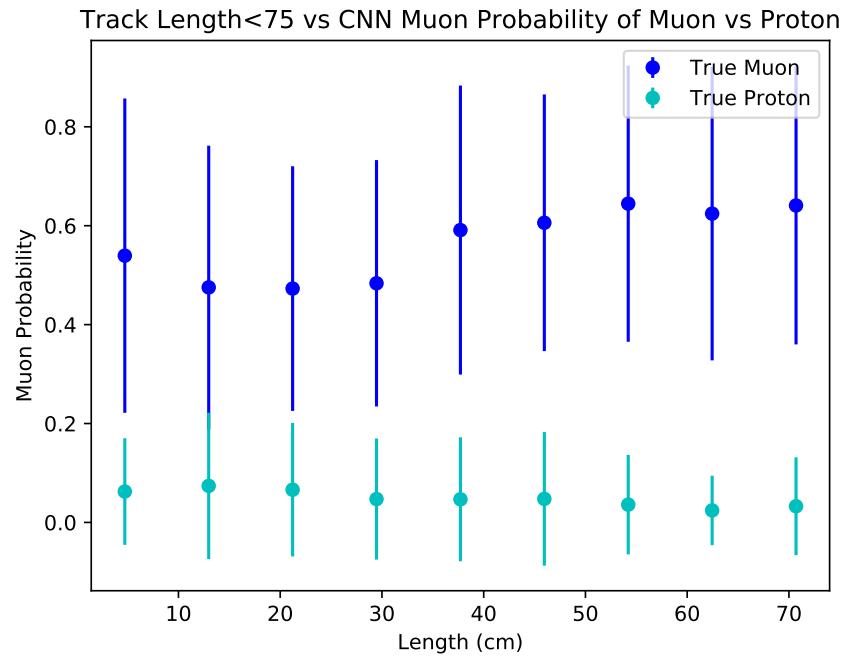


Figure 7.17: mup

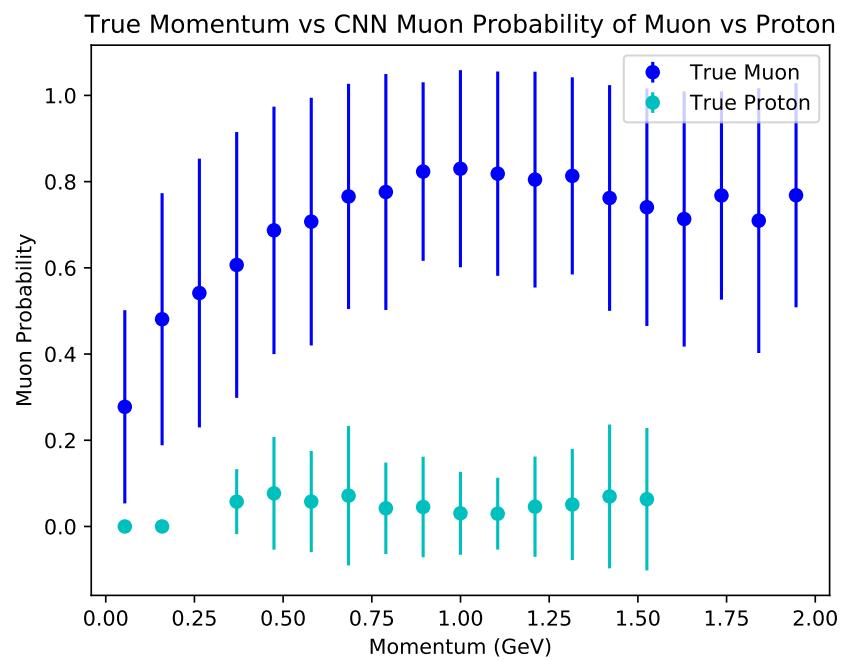
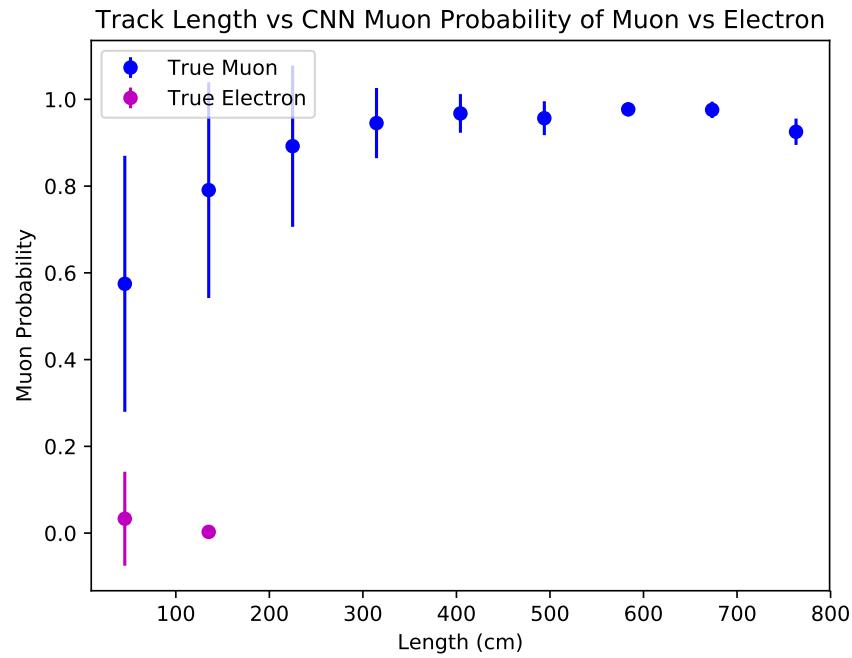
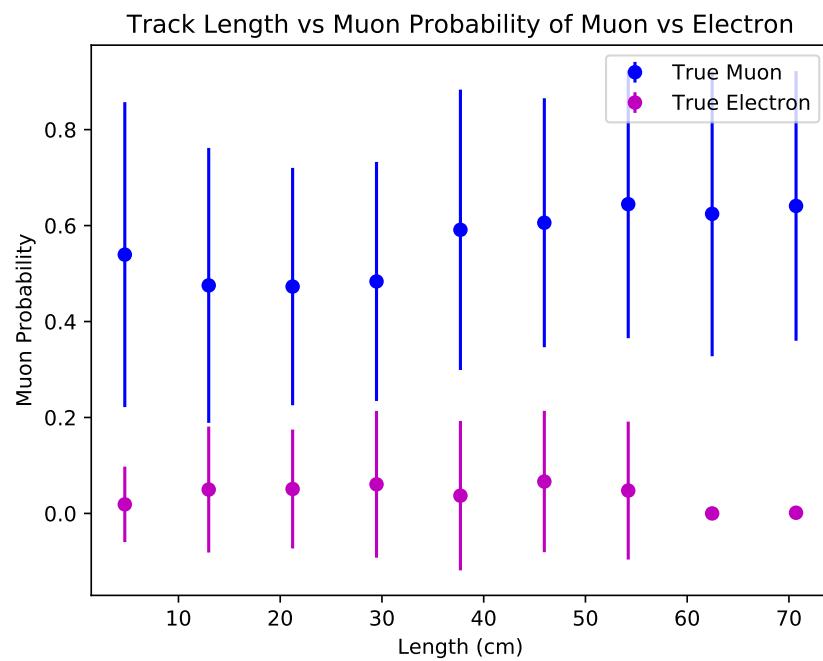


Figure 7.18: mup

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

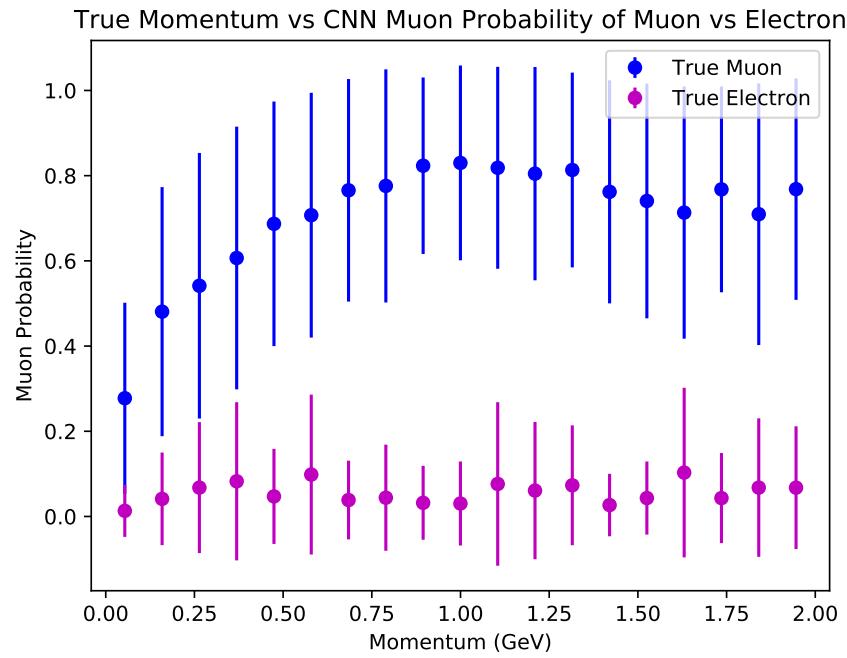


Figure 7.21: mue

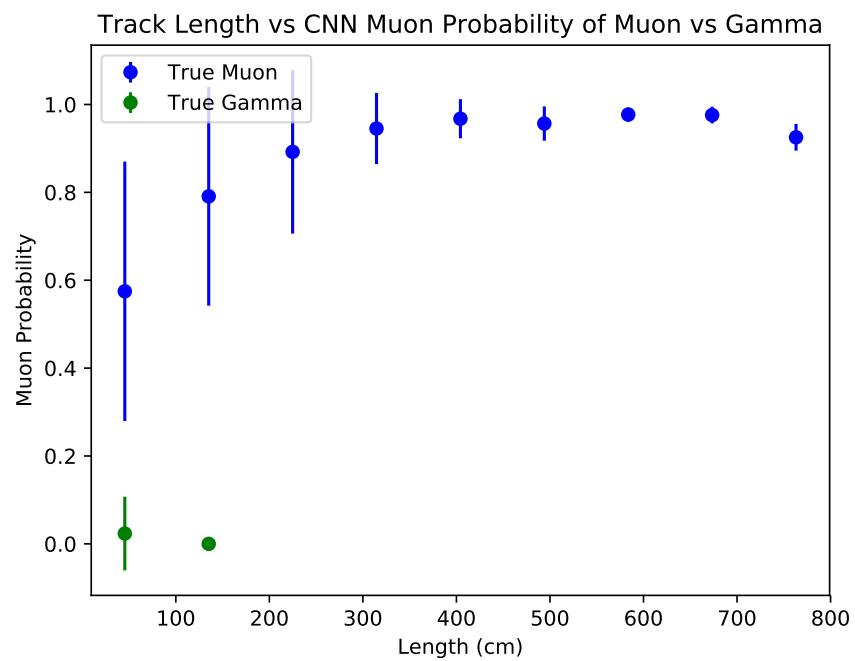


Figure 7.22: mug

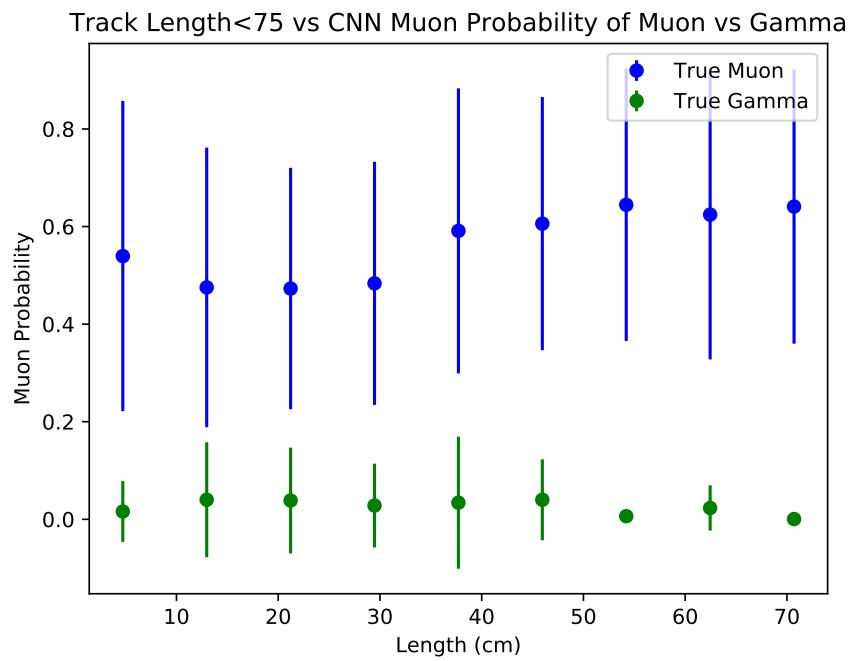


Figure 7.23: mug

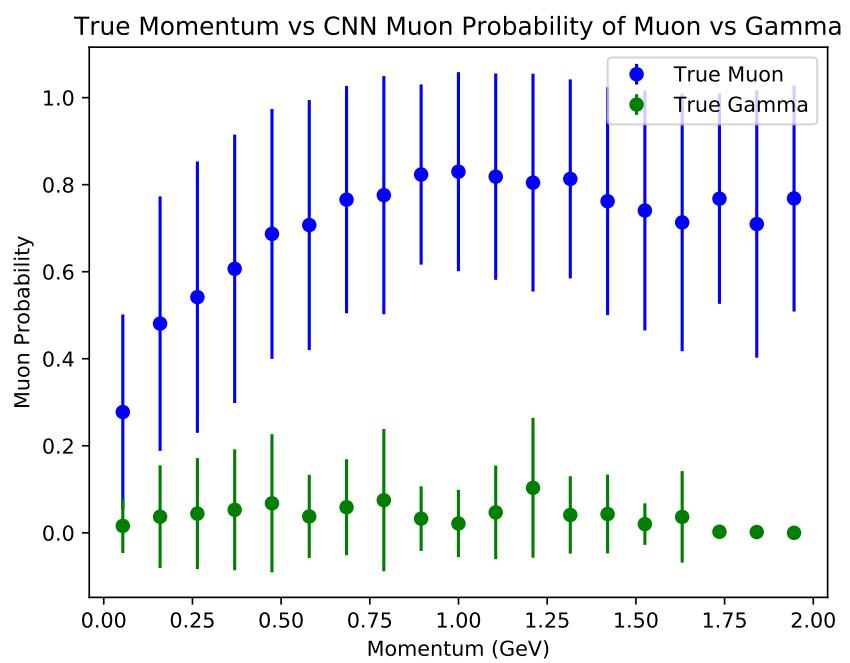


Figure 7.24: mug

¹⁴⁵⁷ Chapter 8

¹⁴⁵⁸ Results of Convolutional Neural ¹⁴⁵⁹ Networks on particles **WORKING** ¹⁴⁶⁰ **TITLE**

¹⁴⁶¹ 8.1 Classification using CNN10000

¹⁴⁶² 8.1.1 Classification of MC data using Selection I Original ¹⁴⁶³ CC-Inclusive Filter

¹⁴⁶⁴ The next step that was taken was to use CNN10000 to classify track candidate images
¹⁴⁶⁵ that were identified by the selection I original cc-inclusive filter described in [?].
¹⁴⁶⁶ Passing rates for each cut in cc-inclusive filter are show in figure 8.1. For the incorrect
¹⁴⁶⁷ image making normalization dataset, out of 188,880 events, 7438 passed the cut right
¹⁴⁶⁸ before 75 cm track length cut which is 3.9% of total data. Discrepancies in passing rates
¹⁴⁶⁹ are due to grid submission issues, however, this dataset is used to check if changes
¹⁴⁷⁰ in image making normalization affects μ/π separation probability due to CNN10000
¹⁴⁷¹ being trained with incorrectly image making normalized data. For the second dataset
¹⁴⁷² with correct image making normalization, out of 188,880 events, 9552 events passed the
¹⁴⁷³ cut right before the 75 cm track length cut which is 5.1% passing rate and is comparable
¹⁴⁷⁴ to figure 8.1. In time cosmics were also run over for efficiency and purity calculations.
¹⁴⁷⁵ Out of 14395 in time cosmic events, 175 passed the cut right before the 75 cm track
¹⁴⁷⁶ 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

1477 Figures 8.2a, 8.2b, 8.2c and 8.2d show the accuracy and μ/π separation of both the
 1478 correct and incorrect normalized images. The confusion matrices are only composed
 1479 of μ/π data. Other particles passed the cc-inclusive filter before the 75 cm track length
 1480 cut and were all mis-id'ed as muons. Since CNN10000 has not seen any particles
 1481 other than muons and pions, it makes sense that those get mis-id'ed. Figures 8.2b
 1482 and 8.2d don't have μ/π separation comparable to 7.2c, but 8.2b does skew to higher
 1483 probabilities compared to 8.2d. This is to be expected and further work on quantifying
 1484 the performance of CNN10000 should use the incorrect image making normalization. It
 1485 is also expected that the separation isn't as defined as the testing dataset for CNN10000.
 1486 CNN10000 was trained and tested using single particle muons and pions and the track
 1487 candidate dataset come from BNB+Cosmic events, not to mentions all track candidates
 1488 have passed the cc-inclusive filter that tags "muon-like" tracks therefore the pions
 1489 in this sample look much closer in muon topology than the network has seen. Also,
 1490 these images were made from wire and time ticks associated to hits from the track
 1491 candidate that passed the cc-inclusive filter. This is different from the training images
 1492 where a bounding box was drawn over the total μ or π interaction. Spurious energy
 1493 deposition from a $\pi - Ar$ interaction is most likely not included in the BNB+Cosmic
 1494 images due to the tracking algorithm. To remedy this, the neural network needs to
 1495 see more "muon-like" pions and muons and pions from a neutrino interaction passing
 1496 the cc-inclusive filter as well as a larger particle variety including protons, photons

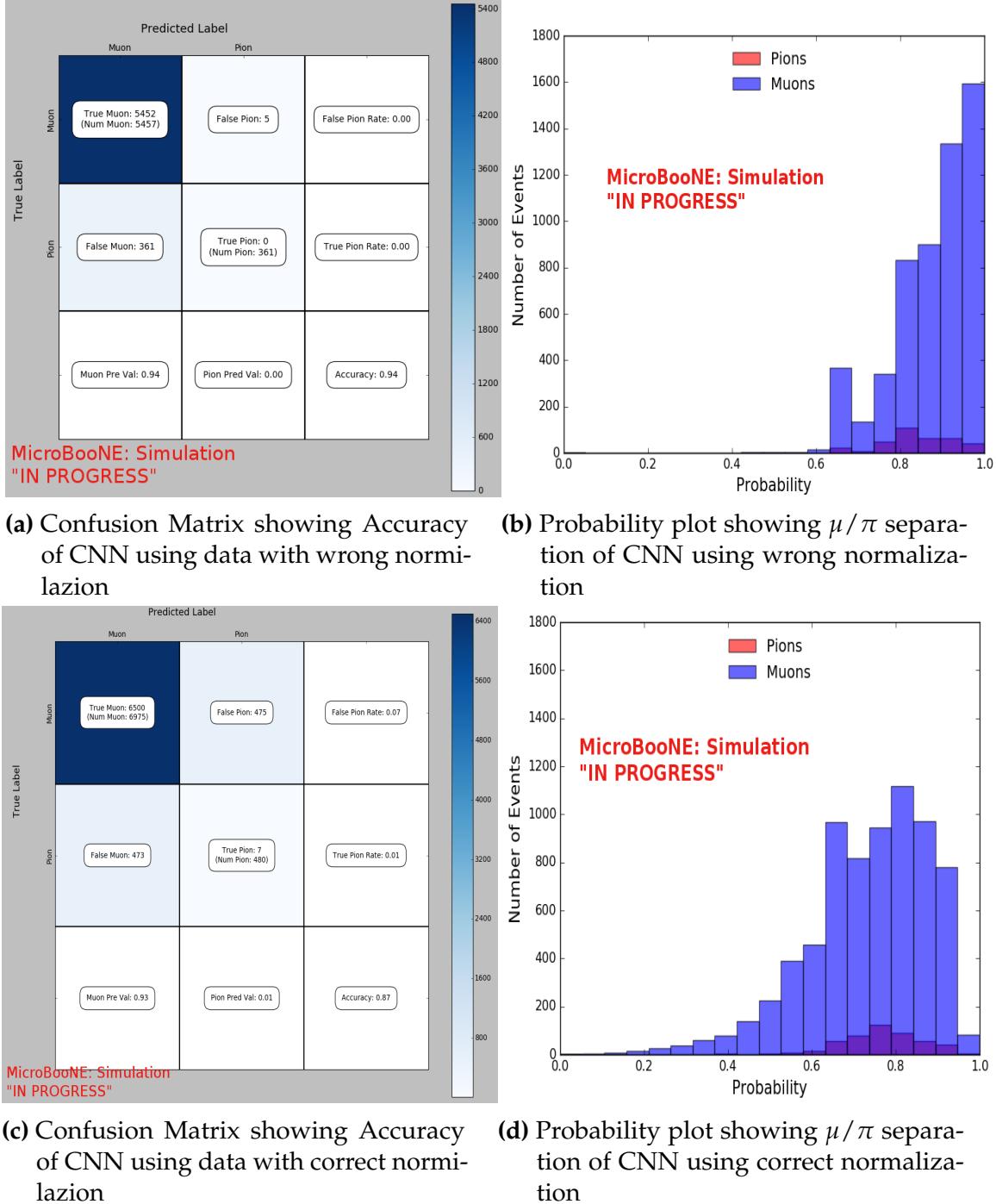


Figure 8.2: Results of CNN10000 classification of track candidate images output from cc-inclusive filter.

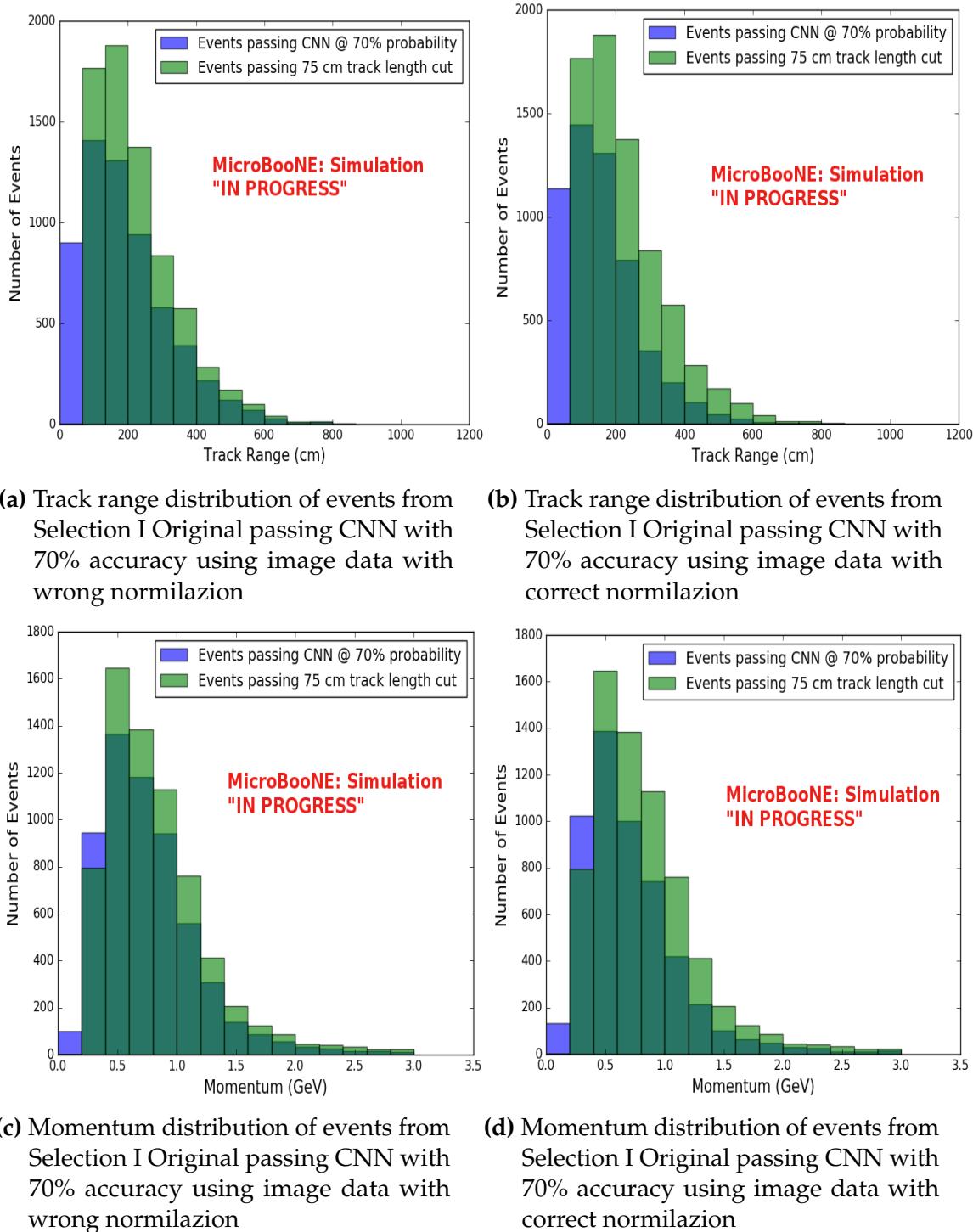


Figure 8.3: CNN10000 distributions of track candidate images output from Selection I Original cc-inclusive filter with different image data normalizations

1497 and electrons. Although μ/π separation is lacking, CNN10000 does an excellent job of
1498 classifying muons and using higher CNN probability can increase purity. Figures 8.3a,
1499 8.3b, 8.3c and 8.3d show the track and momentum distributions for these two datasets.
1500 In both sets you have an increase in data in the bin below 75 cm and at bins below 0.5
1501 GeV. These distributions were made with events classified with 70% probability of
1502 being a muon regardless of true particle type.

1503 **8.1.2 Classification of MC data using Selection I Modified**
1504 **CC-Inclusive Filter**

1505 **Table 5: Selection I: Modified** The table shows passing rates for the above
1506 described event selection. Numbers are absolute event counts and Cosmic back-
1507 ground is not scaled appropriately. The BNB+Cosmic sample contains all
1508 events, not just ν_μ CC inclusive. The selected events are further broken up
1509 in the following subsection. The numbers in brackets give the passing rate wrt
1510 the step before (first percentage) and wrt the generated events (second per-
1511 centage). In the BNB+Cosmic Mc Truth column shows how many true ν_μ CC
1512 inclusive events (in FV) are left in the sample. This number includes possible
1513 mis-identifications where a cosmic track is picked by the selection instead of the
1514 neutrino interaction in the same event. The cosmic only sample has low statis-
1515 tics, but please note that it is not used in any plots, it is just for illustrating the
1516 cut efficiency. The last column Signal:Cosmic only gives an estimate of the ν_μ
1517 CC events wrt the cosmic only background at each step. For this number, the
1518 cosmic background has been scaled as described in chapter 2. Note that this
1519 numbers is not a purity, since other backgrounds can't be determined at this
1520 step.

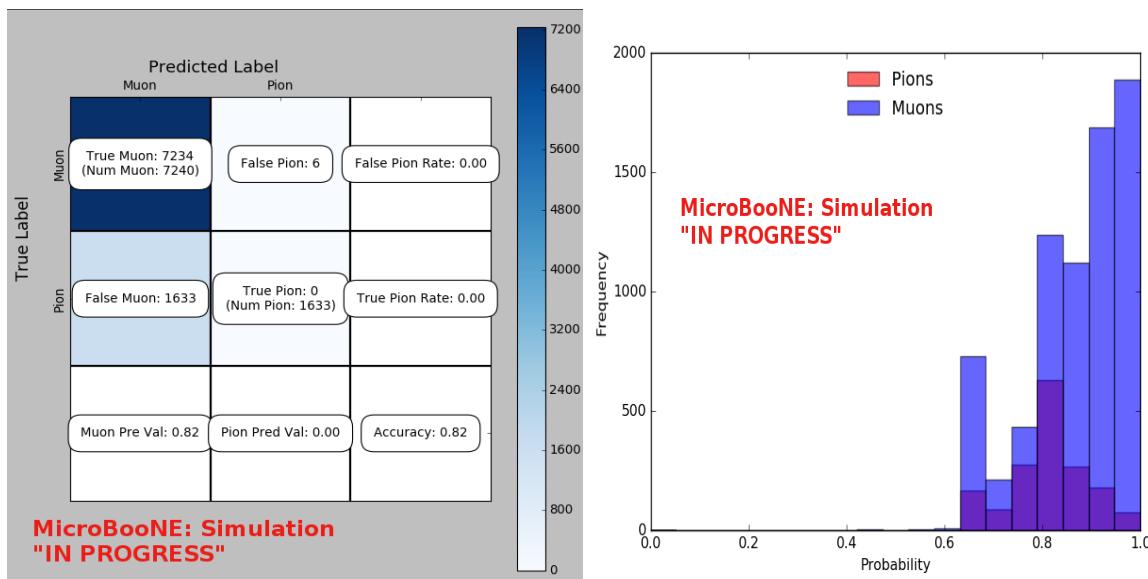
	BNB + Cosmic		Cosmic only	Signal: Cosmic only
	Selection	MC-Truth		
Generated events	191362	45273	4804	1:22
≥ 1 flash with ≥ 50 PE	136219 (71%/71%)	44002 (97%/97%)	2979 (62%/62%)	1:14
≥ 1 track within 5 cm of vertex	135830 (99%/71%)	43974 (99%/97%)	2975 (99%/62%)	1:14
vertex candidate in FV	79112 (58%/41%)	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

1505 CNN10000 was also used to classify track candidate images that were identified by
1506 the selection I modified cc-inclusive filter described in [?]. Passing rates for each cut in
1507 this filter are shown in figure 8.4. As seen in section 8.1.1, wrong image normalization
1508 had a higher muon classification probability so all work done using selection I modified
1509 cc-inclusive filter was done using this normalization. Out of 188,880 events, 19,112
1510 passed the cut right before the 75 cm track length cut which is a 10.1% passing rate and
1511 comparable to the 10% passing rate shown in figure 8.4. In time cosmics were also run
1512 over, out of 14,606 in time cosmics events, 302 passed the cut right before the 75 cm
1513 track length cut which is a 2.1% passing rate comparable to the 2.7% passing rate in the
1514 cc-inclusive tech-note. Figures 8.5a and 8.5b show the accuracy and μ/π separation.
1515 Both plots are only composed of muons and pions and like selection I original data,

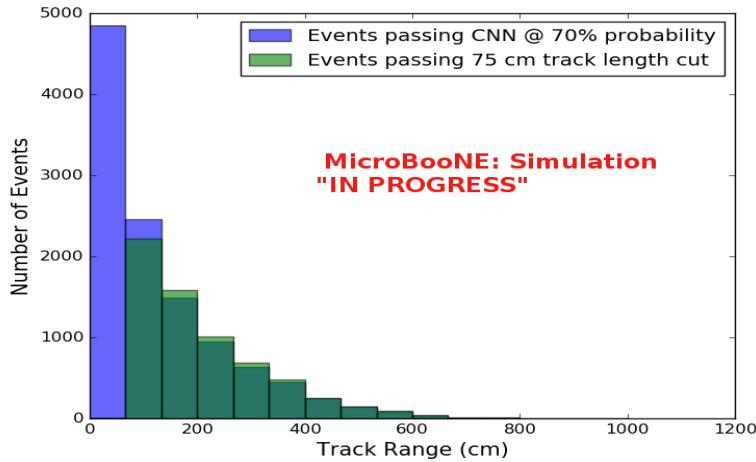
1516 all other particles were id'ed as muons. Also like selection I original data, muons are
 1517 being identified at a very high rate. Figure 8.6a shows the track range distributions
 1518 of all events from selection I modified being classified by the CNN as a muon with a
 1519 probability of 70% regardless of true particle type. We get entries for the CNN curve
 1520 in the lowest bin and none for the 75 cm curve. To see how many true CC events
 1521 were identified by CNN10000 breaking down figure 8.6a by event type was necessary.
 1522 Figures 8.6b and 8.6c show track range distributions separated by signal and various
 1523 backgrounds. Particle type was not taken into consideration in these plots so true CC
 1524 event images can be any track candidate particle passing selection I modified cut right
 1525 before track length cut including pions and protons.

1526 To gain an even deeper understanding on how CNN10000 is performing, plotting
 1527 these distributions with only muons and pions was done due to the fact that CNN10000
 1528 was trained with only those particles for μ/π separation. Figures 8.6d-8.7d show the
 1529 stacked histograms of signal and background of the track range distributions with
 1530 varying CNN probabilities starting from 70% and ending at 90% probability. With
 1531 higher probabilities we get a purer sample in the lower bin but we end up losing
 1532 events as well. Momentum distributions for all signal/background events are shown
 1533 in figure 8.8.

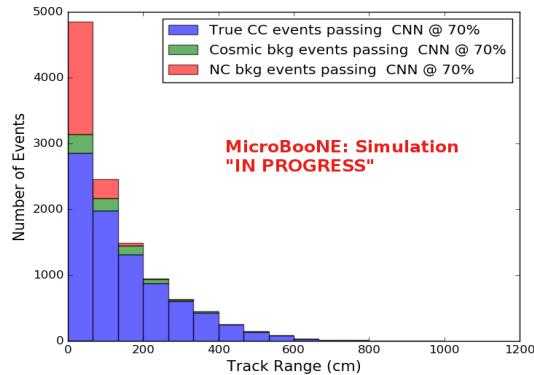


(a) Confusion Matrix for CNN10000 classified events from selection I modified (b) Probability plot for CNN10000 classified events from selection I modified

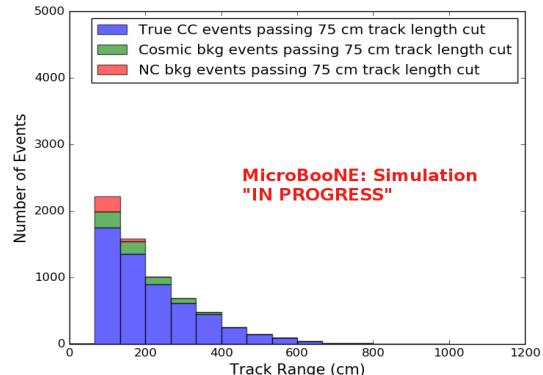
Figure 8.5: Confusion matrix and probability plot of events passing selection I modified cc-inclusive cuts right before 75cm track length cut



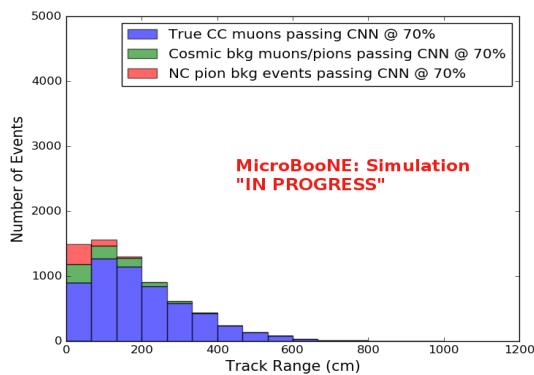
(a) Track range distribution of events from Selection I Modified passing CNN with 70% accuracy



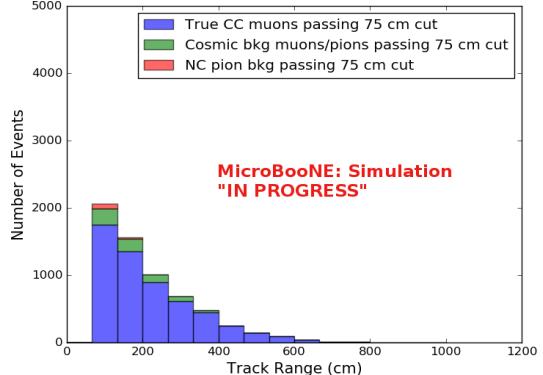
(b) Stacked signal and background track range distributions from Selection I Modified passing CNN with 70% accuracy



(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

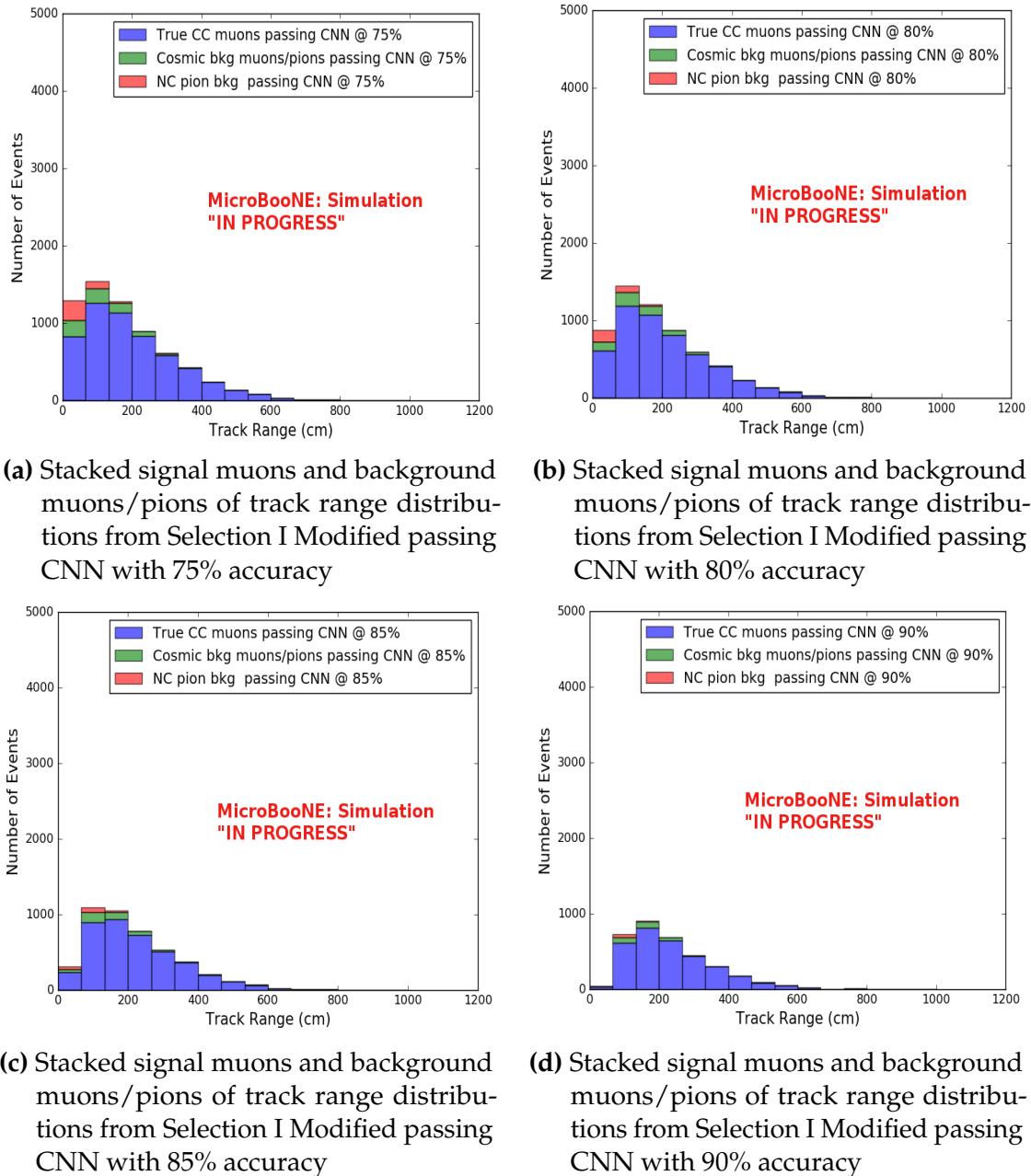
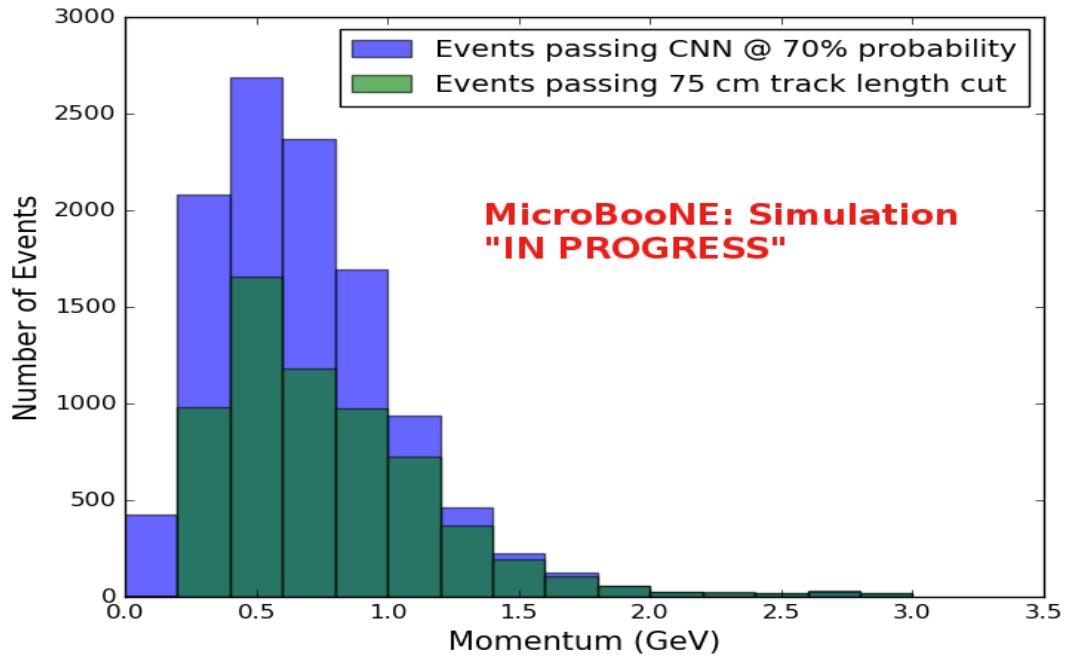
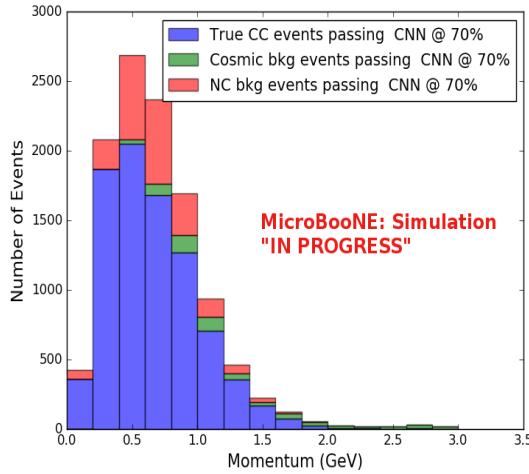


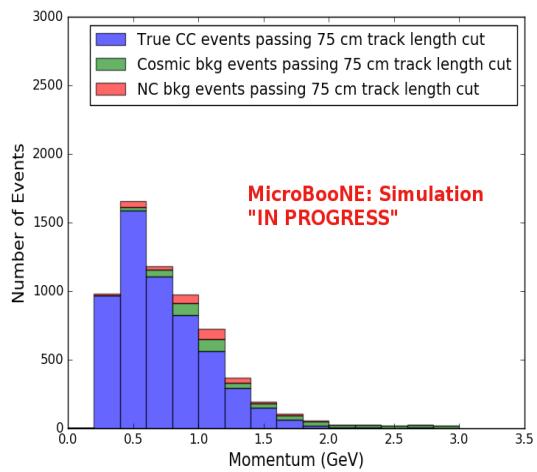
Figure 8.7: CNN10000 stacked signal/background track range distributions of track candidate images output from Selection I Modified cc-inclusive filter



(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.

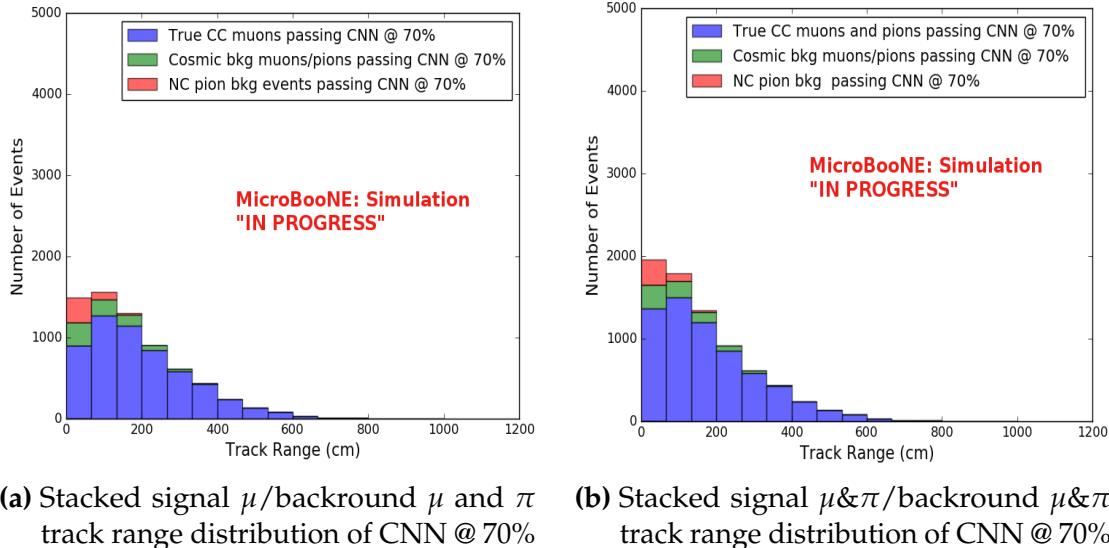


Figure 8.9: Track distribution comparisons of true CC muons plotted vs true CC muons and pions plotted

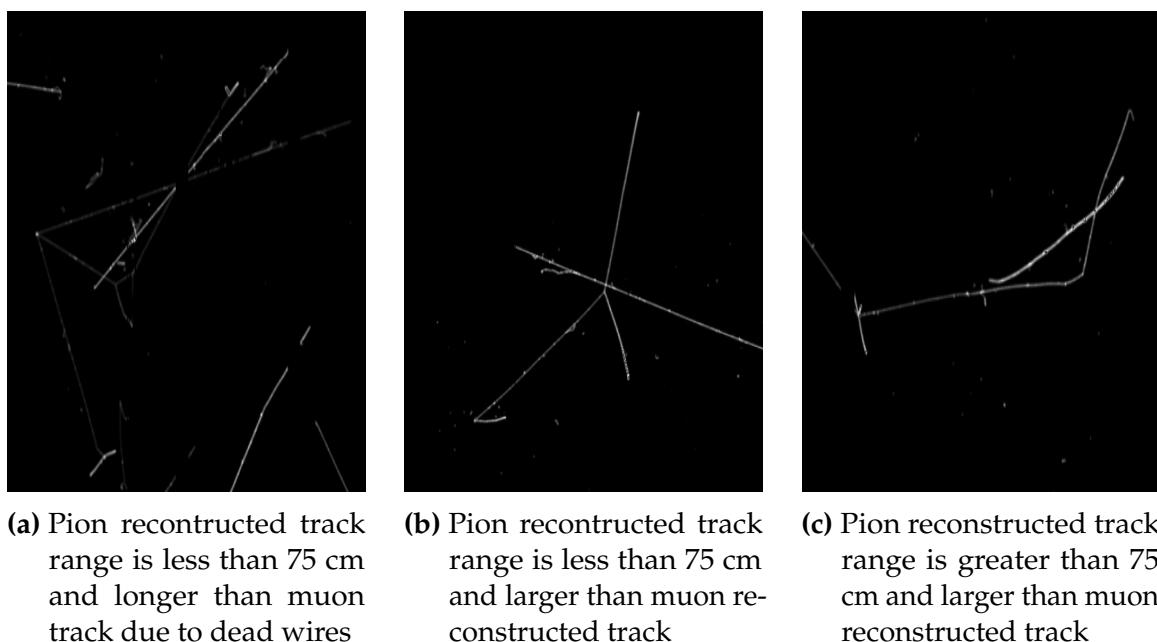


Figure 8.10: Images of true CC events where the pion was the tagged track candidate

		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

1567 a MC true CC event passing rate of 14% compared to the 13% passing rate of the 75
 1568 cm track length cut. The Signal:Cosmic Only background is also reduced from 1:0.6
 1569 to 1:0.4 The total passing rate is also higher than the 75 cm cut, 3.6% vs 4.0%. Table
 1570 8.2 shows the breakdown of signal and backgrounds for the CNN at the different
 1571 probabilities. We have a 61% signal passing rate with the CNN cut @ 83% versus the
 1572 53.8% signal passing rate of the 75 cm cut.

1573 Based on these numbers, the following performance values of the modified selec-
 1574 tion with 75 cm cut versus modified selection with CNN @ 83% probability cut were
 1575 calculated:

- 1576 • Efficiency: Number of selected true ν_μ CC events divided by the number of
 1577 expected true ν_μ CC events with interaction in the FV.
 - 1578 – Selection I modified: 13%
 - 1579 – Selection I modified with CNN cut @ 83% probability: 14%
- 1580 • Purity: Number of selected true ν_μ CC events divided by sum of itself and the
 1581 number of all backgrounds.
 - 1582 – Selection I modified: 53.8%
 - 1583 – Selection I modified with CNN cut @ 83% probability: 61%

1584 Lastly, figure 8.12 shows a more representative performance of the CNN. Due to
 1585 the fact that the CNN was trained on muons and pions, showing the performance
 1586 of CC muon events versus NC pion events with respect to CNN probability gives a
 1587 better picture of how the network is performing. Figure 8.12 shows that at 83% we
 1588 are below the 75 cm cut NC pion threshold and still above the CC muon threshold.
 1589 Using 83% probability not only reduced the NC pion background, it also dramatically
 1590 reduced the in time cosmics and cosmics in the BNB.

1591 8.1.3 Conclusions and Future Work

1592 It was shown that even though CNN10000 was trained with single particle generated
 1593 muons and pions, it performs fairly well at classifying track candidate images from
 1594 BNB+Cosmic events. Events have been regained below the 75 cm track length cut and
 1595 the momentum and track range distributions have similar shapes to the distributions of
 1596 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 [?]

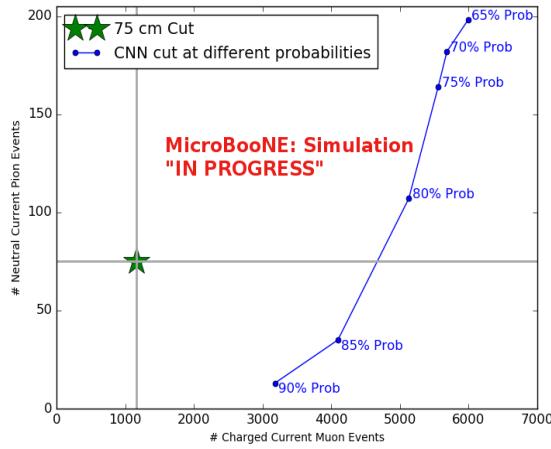


Figure 8.12: CNN performance of classified muons and pions compared to the already implemented 75 cm track length cut

1597 I modified events before 75 cm track length cut with the CNN at 83% probability and
1598 are 14% and 62% respectively. Although the CNN doesn't have separation between
1599 muons and pions and although all particles passing CNN are classified as muon,
1600 increasing CNN probability allows us to increase the purity as well as maintain an
1601 efficiency comparable to the 75 cm track length cut all while recovering events below
1602 that 75 cm cut. Out of the 6142 events that passed the CNN @ 83% 1470 events were
1603 below the 75 cm cut, a recovery of 3.3% of data with a purity of 15%. Although
1604 these numbers are low, it is an improvement from the selection I modified in both total
1605 efficiency and purity and an increase in phase space by recovering these events.

1606 **8.2 Classification using CNN100000**

1607 All future classifications will be done using Selection I Modified CC-Inclusive Filter
1608 because it has a higher efficiency and purity than Selection I Original CC-Inclusive
1609 Filter. To reiterate, CNN100000 was trained using 20,000 images of each $\mu/\pi/p/\gamma/e$.
1610 The results of using CNN100000 to classify BNB+Cosmics will be outlined below.

1611 **8.2.1 Classification of MC data using Selection I Modified
1612 CC-Inclusive Filter**

1613 **8.2.2 Classification of MicroBooNE data using Selection I Modified
1614 CC-Inclusive Filter**

1615 **8.2.3 Comparing two CC-Inclusive Cross Section Selection Filters**

¹⁶¹⁶ **Chapter 9**

¹⁶¹⁷ **Conclusion**

¹⁶¹⁸ Your Conclusions here.

¹⁶¹⁹

¹⁶²⁰ Bibliography

- ¹⁶²¹ [1] Wikipedia, Neutrino, <http://wikipedia.org/wiki/Neutrino>, 2013.
- ¹⁶²² [2] Wikipedia, Neutrino oscillation, http://en.wikipedia.org/wiki/Neutrino_oscillation, 2013.
- ¹⁶²⁴ [3] B. N. Laboratory, Neutrinos and nuclear chemistry, <http://www.chemistry.bnl.gov/sciandtech/sn/default.htm>, 2010.
- ¹⁶²⁶ [4] K. Heeger, Big world of small neutrinos, <http://conferences.fnal.gov/lp2003/forthepublic/neutrinos/>.
- ¹⁶²⁸ [5] A. Y. Smirnov, p. 23 (2003), hep-ph/0305106.
- ¹⁶²⁹ [6] The MicroBooNE Collaboration, R. e. a. Acciari, (2015), arXiv:1512.06148.
- ¹⁶³⁰ [7] The MicroBooNE Collaboration, R. e. a. Acciari, (2016), arXiv:1705.07341.
- ¹⁶³¹ [8] The MicroBooNE Collaboration, R. e. a. Acciari, (2017), arXiv:1704.02927.
- ¹⁶³² [9] The MicroBooNE Collaboration, R. e. a. Acciari, JINST **12** (2017).