

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

16 - * - DRAFT January 26, 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	xxv
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	37
76	5.1 Data and MC Processing Chain	38
77	5.2 Normalization of data and MC	39
78	5.3 Optical Software Trigger and Reconstruction	40
79	5.3.1 Software Trigger	40
80	5.3.2 Flash Reconstruction	41
81	5.3.3 Beam Window	42
82	5.4 TPC Reconstruction	44
83	5.4.1 Hit Reconstruction	45
84	5.4.2 Clustering	45
85	5.4.3 Pandora	46
86	5.4.4 Trackkalmanhit	46
87	5.4.5 Cosmic Hit Removal	46
88	5.4.6 Projection Matching Algorithm	47
89	5.5 Event Selection	47
90	5.5.1 Expected Bakgrounds	48
91	5.5.2 Truth Distributions	51
92	6 Background on Convolutional Neural Networks	59
93	6.1 Image Classification	59
94	6.2 CNN Structure	60
95	6.2.1 Backpropagation	63
96	6.3 Choosing Hyperparameters	64

97	7 Training process of Convolutional Neural Networks	67
98	7.1 Convolutional Neural Network Training	67
99	7.1.1 Creating images using LArTPC data for training/validation of 100 CNNs	68
101	7.1.2 Training CNN1075	70
102	7.1.3 Training CNN10000	73
103	7.1.4 Training CNN100000	74
104	8 Results of Convolutional Neural Networks on particles WORKING TITLE	89
105	8.1 Classification using CNN10000	89
106	8.1.1 Classification of MC data using Selection I Original CC-Inclusive 107 Filter	89
108	8.1.2 Classification of MC data using Selection I Modified CC-Inclusive 109 Filter	93
110	8.1.3 Conclusions and Future Work	101
111	8.2 Classification using CNN100000	103
112	8.2.1 Classification of MC data using Selection I Modified CC-Inclusive 113 Filter	103
114	8.2.2 Classification of MicroBooNE data using Selection I Modified 115 CC-Inclusive Filter	103
116	8.2.3 Comparing two CC-Inclusive Cross Section Selection Filters .	103
117	9 Conclusion	105
118	Bibliography	107

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	4.5	First Neutrino Interaction Candidate Events from MicroBooNE	35
150	4.6	First Neutrino Interaction Candidate Events from MicroBooNE	36
151	5.1	5.1a Track range distribution for selection I. The track range is defined as the 3D distance between the start and end of the muon candidate track. No data is shown below 75 cm due to the track length cut described previously. 5.1b Efficiency of the selected events by process quasi-elastic (QE), resonant (RES), and deep-inelastic (DIS). Statistical uncertainty is shown in the bands and the distributions are a function of true muon momentum. The rise of the efficiency between 0 GeV and 0.5 GeV is due to the minimum track length cut and the decreasing efficiency for higher momentum tracks is caused by the containment requirement.	40
160	5.1a	Track range distribution of selection I	40
161	5.1b	Selection efficiency as a function of the true muon momentum	40
162	5.2	Time distribution of reconstructed optical flashes with a PE value of 50 or more for a sample of BNB unbiased triggered events.	42
164	5.3	Flash time distribution for all flashes (left plot) and flashes > 20PE (right plot). The different curves are as follows: on-beam data (black), off-beam data (red), CORSIKA inTime MC (light blue), BNB only MC (green), and BNB+Cosmic MC (purple). The dashed vertical lines mark the time window that was chosen for each sample	43

169	5.4	Reconstruction chain run on both data and MC. The red stars mean that the algorithm returns reconstructed 3D vertices.	44
170			
171	5.5	Event selection diagram for selection I and selection II. This analysis focused on optimizing selection I. Boxes with the same color across the two selections symbolize similar cuts.	49
172			
173			
174	5.6	MC momentum distributions of the muon originating from a ν_μ CC interaction. Upper left is the momentum distributions of events with a vertex within the FV (green) and the events with a fully contained track (blue) before the selection. The upper right side is the momentum distribution after the selection (red). The lower plot is the selection efficiencies.	52
175			
176			
177			
178			
179			
180	5.7	MC $\cos(\theta)$ distributions of the muon originating from a ν_μ CC interac- tion. Upper left is the $\cos(\theta)$ distributions of events with a vertex within the FV (green) and the events with a fully contained track (blue) before the selection. The upper right side is the $\cos(\theta)$ distribution after the selection (red). The lower plot is the selection efficiencies.	53
181			
182			
183			
184			
185	5.8	MC ϕ distributions of the muon originating from a ν_μ CC interaction. Upper left is the ϕ distributions of events with a vertex within the FV (green) and the events with a fully contained track (blue) before the selection. The upper right side is the ϕ distribution after the selection (red). The lower plot is the selection efficiencies.	54
186			
187			
188			
189			
190	5.9	MC momentum distributions of the muon originating from a ν_μ CC interaction. Upper left is the momentum distributions of events with a vertex within the FV split up into CCQE (red), CCRES (yellow), and CCDIS (green) before selection. The upper right side is the momentum distribution after the selection with the same color schemes. The lower plot is the selection efficiencies for all three interaction types. The definition of QE, RES, and DIS is based on GENIE.	55
191			
192			
193			
194			
195			
196			

197	5.10 MC $\cos(\theta)$ distributions of the muon originating from a ν_μ CC interaction. Upper left is the $\cos(\theta)$ distributions of events with a vertex within the FV split up into CCQE (red), CCRES (yellow), and CCDIS (green) before selection. The upper right side is the $\cos(\theta)$ distribution after the selection with the same color schemes. The lower plot is the selection efficiencies for all three interaction types. The definition of QE, RES, and DIS is based on GENIE.	56
204	5.11 MC ϕ distributions of the muon originating from a ν_μ CC interaction. Upper left is the ϕ distributions of events with a vertex within the FV split up into CCQE (red), CCRES (yellow), and CCDIS (green) before selection. The upper right side is the ϕ distribution after the selection with the same color schemes. The lower plot is the selection efficiencies for all three interaction types. The definition of QE, RES, and DIS is based on GENIE.	57
211	6.1 Pixel representation and visualization of a curve detector filter. As you can see, in the pixel representation, the weights of this filter are greater along a curve we are trying to find in the input image	61
214	6.2 Visualization of filters found in first layer of a CNN.	61
215	6.3 Applying a feature mask over a set of fashion items to extract necessary information for auto-encoding. Unnecessary information for example color or brand emblems are not saved. This feature map is an edge detection mask that leaves only shape information which helps to distinguish between different types of clothes.	62
220	6.4 Pictorial Representation of Convolutional Neural Networks as well as a visual representation on CNN's complexity of layer feature extraction	62
222	6.5 Pictorial representation of the AlexNet model. The AlexNet model consists of 5 convolution layers and 3 fully connected layers.	65

224	6.6	Pictoral representation of the GoogleNet model. The GoogleNet model 225 consists of 22 layers. The model implements 9 Inception modules which 226 performs covolution and pooling in parallel and strays away from 227 the basis that CNN layers need to be stacked up sequentially. The 228 GoogleNet model also doesn't use fully connected layers, instead it 229 uses average pooling which greatly reduces the amount of parameters. 230 GoogleNet has 12x fewer parameters than AlexNet.	65
231	7.1	Accuracy vs. Loss of ImageNet 2-output μ/π sample consisting of 2,150 232 images each.	71
233	7.2	Confusion Matrinx showing accuracy of CNN1075 on testing and train- 234 ing MC data.	72
235	7.3	Accuracy vs. Loss of ImageNet 2-output μ/π sample consisting of 236 10,000 images each.	74
237	7.4	Description of confusion matrix variables: False pion rate = $false\pi/total\pi$ 238 True pion rate = $true\pi/total\pi$ Accuracy = $(true\pi rate + true\mu rate)/2$ 239 Pion prediction value = $true\pi/(true\pi + false\pi)$ Muon prediction value 240 = $true\mu/(true\mu + false\mu)$ 7.4c The probability plot includes muons and 241 pions that are classified as primary particles.	75
242	7.4a	Confusion Matrix showing Accuracy of CNN using training MC 243 data	75
244	7.4b	Confusion Matrix showing Accuracy of CNN using testing MC 245 data	75
246	7.4c	Probability plot of muons and pions from testing set. Images 247 surrounding histogram are a random event from lowest bin and 248 highest bin for each particle.	75
249	7.5	Confusion Matrix of all five particles	76
250	7.6	Training and testing accuracy of CNN trained on 100,000 images of 251 $\mu/\pi/p/\gamma/e$ with 20,000 images of each particle. Each image was a size 252 of 576x576 and the images per particle were split 90% use for training 253 and 10% used for testing the network	77
254	7.7	Training and testing loss oc CNN trained on 100,000 images of $\mu/\pi/p/\gamma/e$	77

255	7.8 t-SNE of CNN	78
256	7.9 Muon Prob	79
257	7.10 Pion Prob	79
258	7.11 Proton Prob	80
259	7.12 Electron Prob	80
260	7.13 Gamma Prob	81
261	7.14 Prob	81
262	7.15 mupi	82
263	7.16 mupi	82
264	7.17 mupi	83
265	7.18 mup	83
266	7.19 mup	84
267	7.20 mup	84
268	7.21 mue	85
269	7.22 mue	85
270	7.23 mue	86
271	7.24 mug	86
272	7.25 mug	87
273	7.26 mug	87
274	8.1 Snapshot of passing rates of Selection I from CC-Inclusive Filter	90
275	8.2 Results of CNN10000 classification of track candidate images output from cc-inclusive filter.	91
277	8.2a Confusion Matrix showing Accuracy of CNN using data with wrong normilazion	91
278		

279	8.2b	Probability plot showing μ/π separation of CNN using wrong normalization	91
280			
281	8.2c	Confusion Matrix showing Accuracy of CNN using data with correct normalization	91
282			
283	8.2d	Probability plot showing μ/π separation of CNN using correct normalization	91
284			
285	8.3	CNN10000 distributions of track candidate images output from Selection I Original cc-inclusive filter with different image data normalizations	92
286			
287	8.3a	Track range distribution of events from Selection I Original passing CNN with 70% accuracy using image data with wrong normalization	92
288			
289	8.3b	Track range distribution of events from Selection I Original passing CNN with 70% accuracy using image data with correct normalization	92
290			
291	8.3c	Momentum distribution of events from Selection I Original passing CNN with 70% accuracy using image data with wrong normalization	92
292			
293	8.3d	Momentum distribution of events from Selection I Original passing CNN with 70% accuracy using image data with correct normalization	92
294			
295			
296	8.4	Snapshot of passing rates of all cuts from Selection I Modified cc-inclusive filter	93
297			
298			
299	8.5	Confusion matrix and probability plot of events passing selection I modified cc-inclusive cuts right before 75cm track length cut	94
300			
301	8.5a	Confusion Matrix for CNN10000 classified events from selection I modified	94
302			
303	8.5b	Probability plot for CNN10000 classified events from selection I modified	94
304			
305	8.6	CNN10000 distributions of track candidate images output from Selection I Modified cc-inclusive filter	95
306			
307			
308			

309	8.6a	Track range distribution of events from Selection I Modified passing CNN with 70% accuracy	95
310			
311	8.6b	Stacked signal and background track range distributions from Selection I Modified passing CNN with 70% accuracy	95
312			
313	8.6c	Stacked signal and background track range distributions from Selection I Modified passing 75 cm track length cut	95
314			
315	8.6d	Stacked signal muons and background muons/pions of track range distributions from Selection I Modified passing CNN with 70% accuracy	95
316			
317			
318	8.6e	Stacked signal muons and background muons/pions of track range distributions from Selection I Modified passing 75 cm track length cut	95
319			
320			
321	8.7	CNN10000 stacked signal/background track range distributions of track candidate images output from Selection I Modified cc-inclusive filter	96
322			
323	8.7a	Stacked signal muons and background muons/pions of track range distributions from Selection I Modified passing CNN with 75% accuracy	96
324			
325			
326	8.7b	Stacked signal muons and background muons/pions of track range distributions from Selection I Modified passing CNN with 80% accuracy	96
327			
328			
329	8.7c	Stacked signal muons and background muons/pions of track range distributions from Selection I Modified passing CNN with 85% accuracy	96
330			
331			
332	8.7d	Stacked signal muons and background muons/pions of track range distributions from Selection I Modified passing CNN with 90% accuracy	96
333			
334			
335	8.8	CNN10000 momentum distributions of track candidate images output from Selection I Modified cc-inclusive filter	97
336			
337	8.8a	Momentum distribution of events from Selection I Modified passing CNN with 70% accuracy	97
338			

339	8.8b	Stacked signal and background momentum distributions from Selection I Modified passing CNN with 70% accuracy	97
340	8.8c	Stacked signal and background momentum distributions from Selection I Modified passing 75 cm track length cut	97
341	8.9	Track distribution comparisons of true CC muons plotted vs true CC muons and pions plotted	99
342	8.9a	Stacked signal μ /background μ and π track range distribution of CNN @ 70%	99
343	8.9b	Stacked signal $\mu\&\pi$ /background $\mu\&\pi$ track range distribution of CNN @ 70%	99
344	8.10	Images of true CC events where the pion was the tagged track candidate	99
345	8.10a	Pion reconstructed track range is less than 75 cm and longer than muon track due to dead wires	99
346	8.10b	Pion reconstructed track range is less than 75 cm and larger than muon reconstructed track	99
347	8.10c	Pion reconstructed track range is greater than 75 cm and larger than muon reconstructed track	99
348	8.11	Snapshot of signal and background event numbers of Selection I modified from cc-inclusive note [?]	102
349	8.12	CNN performance of classified muons and pions compared to the already implemented 75 cm track length cut	102

³⁶⁰ List of tables

³⁶¹ 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
³⁶⁷ 5.1	Passing rates of Selection I. Numbers are absolute event counts and cosmic background is not scaled. The BNB+Cosmic sample contains all events, not just ν_μ CC inclusive. The numbers in brackets give the passing rate wrt the step before (first percentage) and wrt total generated events (second percentage). The BNB+Cosmic MC-Truth column shows how many true ν_μ CC inclusive events are left in the sample. This number includes mis-identifications where a cosmic track is picked by the selection instead of the neutrino interaction in the same event. The cosmic only sample is used just to illustrate the cut efficiency. The last column Signal:Cosmic only gives an estimate of the ν_μ CC events wrt the cosmic only background at each step. For this number, the cosmic background has been scaled as described in section 5.2.	50
³⁷⁹ 5.2	Passing rates for Selection I selection applied to on-beam and off-beam data. The numbers in brackets give the passing rate wrt the step before (first percentage) and wrt the generated events (second percentage). Off-beam data has been scaled with a factor 1.23 to normalize to the on-beam data stream.	50

384	8.1 Comparing passing rates of CNN at different probabilities versus 75	
385	cm track length cut: Numbers are absolute event counts and Cosmic	
386	background is not scaled appropriately. The BNB+Cosmic sample con-	
387	tains all events. The numbers in brackets give the passing rate wrt	
388	the step before (first percentage) and wrt the generated events (second	
389	percentage). In the BNB+Cosmic MC Truth column shows how many	
390	true ν_μ CC-inclusive events (in FV) are left in the sample. This number	
391	includes possible mis-identifications where a cosmic track is picked by	
392	the selection instead of the neutrino interaction in the same event. The	
393	CNN MC True generated events were scaled wrt the MC True generated	
394	events for the 75 cm cut passing rates due to only running over 188,880	
395	generated events versus the 191362 generated events. The last column	
396	Signal:Cosmic only gives an estimate of the ν_μ CC events wrt the cosmic	
397	only background at each step. For this number, the cosmic background	
398	has been scaled as described in [?]. Note that these numbers are not a	
399	purity, since other backgrounds can't be determined at this step. . .	100
400	8.2 Signal and background event numbers at modified selection level with	
401	CNN cut estimated from a BNB+Cosmic sample and Cosmic only sam-	
402	ple normalized to $5 * 10^{19}$ PoT. The last column gives the fraction of	
403	this signal or background type to the total selected events per CNN	
404	probability.	100

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

— Shirley Chisholm

407 **Chapter 1**

408 **Introduction**

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

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

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

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

⁴³² **Chapter 2**

⁴³³ **Neutrinos**

⁴³⁴ **2.1 What are Neutrinos**

⁴³⁵ Neutrinos are one of the fundamental particles which make up the universe. They are
⁴³⁶ also one of the least understood. Neutrinos are not affected by the electromagnetic
⁴³⁷ forces because they do not have electric charge. Neutrinos are affected by a "weak"
⁴³⁸ sub-atomic force of much shorter range than electromagnetism, and are therefore able
⁴³⁹ to pass through great distances in matter without being affected by it. Until the late
⁴⁴⁰ 90's, neutrinos were thought to have no mass. Due to their mass, neutrinos are also
⁴⁴¹ affected by gravity. Neutrinos are created by radioactive decay or nuclear reactions
⁴⁴² such as the ones that happen in the sun, in nuclear reactors or when cosmic rays hit
⁴⁴³ atoms. There are three types of neutrinos, ν_e , ν_μ and ν_τ which correspond to their
⁴⁴⁴ charged lepton pairs.

⁴⁴⁵ As previously stated, neutrinos are very weakly interacting; in fact, neutrinos can
⁴⁴⁶ pass unscathed through a wall of lead several hundred light-years thick. Because
⁴⁴⁷ neutrinos interact so rarely, studying neutrinos requires a massive detector and a
⁴⁴⁸ powerful neutrino source. With that being said, we can only infer their existence when
⁴⁴⁹ they interact in a detector. In a collision, distinct charged particles are produced with
⁴⁵⁰ each type of neutrino. An electron neutrino will create an electron, a muon neutrino
⁴⁵¹ will create a muon, and a tau neutrino will create a tau. The track the particle leaves
⁴⁵² in the detector is how one figures out what type of neutrino interaction was "seen".
⁴⁵³ Liquid Argon Time Projection Chambers are the newest type of detectors being used to
⁴⁵⁴ study neutrinos due to their excellent imaging and particle identification capabilities.

455 2.2 History of Neutrinos

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

480 2.3 Neutrino Oscillations

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

487 2.3.1 Solar Oscillations and the Solar Neutrino Problem

488 The solar neutrino flux derived from Bahcall's Standard Solar Model is shown in figure
 489 2.1. Nuclear fusion and decay processes produce an abundant amount of neutrinos.
 490 The standard solar model predicts that these reactions produce several groups of
 491 neutrinos, each with differing fluxes and energy spectra. The figure also shows the
 492 ranges of detection of existing solar neutrino experiments in different shades of blue
 493 to illustrate that they sample different portions of the solar neutrino energy spectrum.
 494 Three of these experiments, plus a new one, are discussed below.

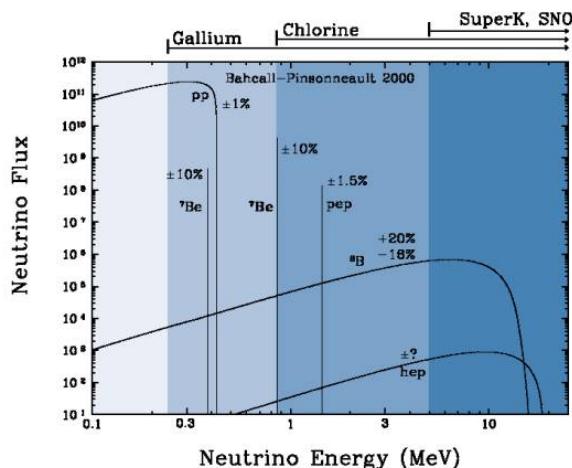


Figure 2.1: The Standard Solar Model

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

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

⁵⁰⁷ The unexplained difference between the measured solar neutrino flux and model
⁵⁰⁸ predictions lead to the Solar Neutrino Problem. [4]



⁵⁰⁹ While it is now known that the Homestake Experiment detected neutrinos, some
⁵¹⁰ physicist were weary of the results. Conclusive evidence of the Solar Neutrino Problem
⁵¹¹ was provided by the Kamiokande-II experiment, a water Cherenkov detector with
⁵¹² a low enough energy threshold to detect neutrinos through neutrino-electron elastic
⁵¹³ scattering. In the elastic scattering interaction the electrons coming out of the point of
⁵¹⁴ reaction strongly point in the direction that the neutrino was traveling, away from the
⁵¹⁵ sun. While the neutrinos observed in Kamiokande-II were clearly from the sun, there
⁵¹⁶ was still a discrepancy between Kamiokande-II and Homestake; The Kamiokande-
⁵¹⁷ II experiment measured about 1/2 the predicted flux, rather than the 1/3 that the
⁵¹⁸ Homestake Experiment saw.

⁵¹⁹ The solution to the solar neutrino problem was finally experimentally determined
⁵²⁰ by the Sudbury Neutrino Observatory(SNO). The Ray Davis's Homestake Experiment
⁵²¹ was only sensitive to electron neutrinos, and the Kamiokande-II Experiment was
⁵²² dominated by the electron neutrino signal. The SNO experiment had the capability to
⁵²³ see all three neutrino flavors. Because of this, it was possible to measure the electron
⁵²⁴ neutrinos and total neutrino flux. The experiment demonstrated that the deficit was
⁵²⁵ due to the MSW effect, the conversion of electron neutrinos from their pure flavor
⁵²⁶ state into the second neutrino mass eigenstate as they passed through a resonance
⁵²⁷ due to the changing density of the sun. The resonance is energy dependent, and is
⁵²⁸ visible near 2MeV. The water cherenkov detectors only detect neutrinos above about
⁵²⁹ 5MeV, while the radiochemical experiments were sensitive to lower energy (0.8MeV
⁵³⁰ for chlorine, 0.2MeV for gallium), and this turned out to be the source of the difference
⁵³¹ in the observed neutrino rates at the two types of experiments. Figure 2.2 shows
⁵³² Homestake, Kamiokande-II and SNO experiments.

⁵³³ MSW Effect

⁵³⁴ The Mikheyev-Smirnov-Wolfenstein effect is a process which acts to modify neu-
⁵³⁵ trino oscillations in matter. The presence of electrons in matter changes the energy

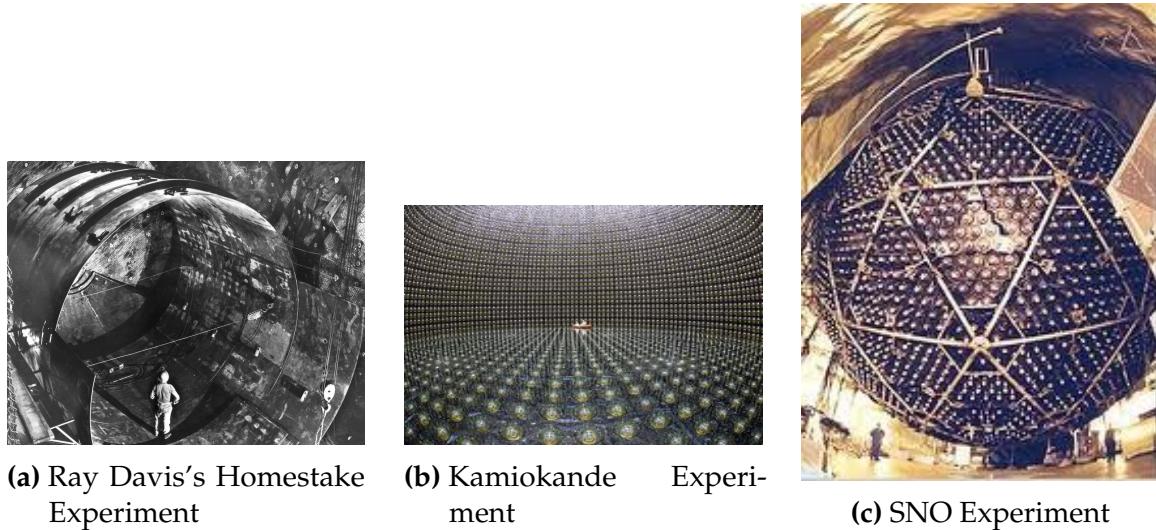


Figure 2.2: Solar Neutrino Experiments

536 levels of the mass eigenstates of neutrinos due to charged current coherent forward
 537 scattering of the electron neutrinos. This coherent forward scattering is similar to
 538 the electromagnetic process with respect to the refractive index of light in a medium.
 539 Because of this MSW Effect, neutrinos in vacuum have a different effective mass than
 540 neutrinos in matter and because neutrino oscillations depend on the squared mass
 541 difference of the neutrinos, the neutrino oscillations are different in matter than in
 542 vacuum. This effect is important at the sun where electron neutrinos are produced.
 543 The neutrinos of high energy leaving the sun are in a vacuum propagation eigenstate
 544 ν_2 that has a very small overlap with the electron neutrino $\nu_e = \nu_1 \cos(\theta) + \nu_2 \sin(\theta)$
 545 seen by the charged current reactions in Kamiokande-II and SNO. The discrepancy of
 546 the deficit between SNO, Kamiokande-II and Homestake is due to the energy of the
 547 solar neutrinos. The MSW effect "turns on" at about 2MeV and at lower energies, this
 548 MSW effect is negligible. [5]

549 **2.3.2 Atmospheric Oscillations and the Atmospheric Neutrino 550 Anomaly**

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

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

554

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

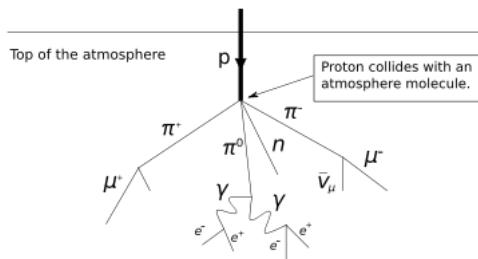


Figure 2.3: Cosmic Ray Shower

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

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

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

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

567 Kamiokande-II has the the capability of measuring the direction of the incoming
 568 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:

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

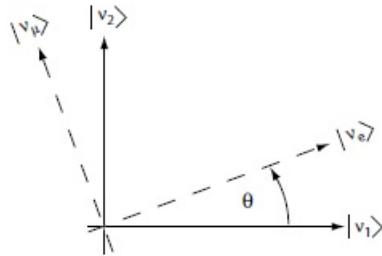


Figure 2.5: The flavor eigenstates are rotated by an angle θ with respect to the mass eigenstates

588 Applying the time evolution operator to ν_μ :

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

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

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

592 because of this,

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

593

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

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

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

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

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

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

⁵⁹⁶ 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:

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

⁵⁹⁷ Finding the Probability for a $\nu_\mu \rightarrow \nu_e$:

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

⁵⁹⁸ Remembering that $\langle \nu_i | \nu_j \rangle = \delta_{ij}$

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

⁵⁹⁹ Taking the absolute value squared gives us:

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

⁶⁰⁰ Since the neutrino is relativistic we can set $p = E_\nu$ and change $x = L$. Also ⁶⁰¹ recognizing the trigonometric relation $(1 - \cos 2\theta)/2 = \sin^2\theta$ the above equation ⁶⁰² becomes:

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

⁶⁰³ All that's left to do now is re-introduce \hbar and c doing this we get:

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

⁶⁰⁴ 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

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

629 **2.3.5 Reactor Oscillation**

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

637 **Chapter 3**

638 **The MicroBooNE Experiment**

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

644 **3.1 Liquid argon time projection chambers**

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

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

685 3.2 The MicroBooNE Time Projection Chamber

686 MicroBooNE (Micro Booster Neutrino Experiment) is a 89 ton active volume (180 ton
687 total mass) LArTPC which is then inserted into a cylindrical cryostat on axis of the
688 Booster Neutrino Beam (BNB) stationed at Fermilab in Batavia, Illinois. Understanding
689 LArTPC technology and detector physics is necessary to build a LArTPC the size of
690 DUNE, and MicroBooNE has made many advances in developing this technology [7]
691 [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.

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

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

754 3.3 MicroBooNE's Physics Goals

755 3.3.1 The low-energy excess

756 The primary goal of the MicroBooNE experiment is to study and investigate the low-
757 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.

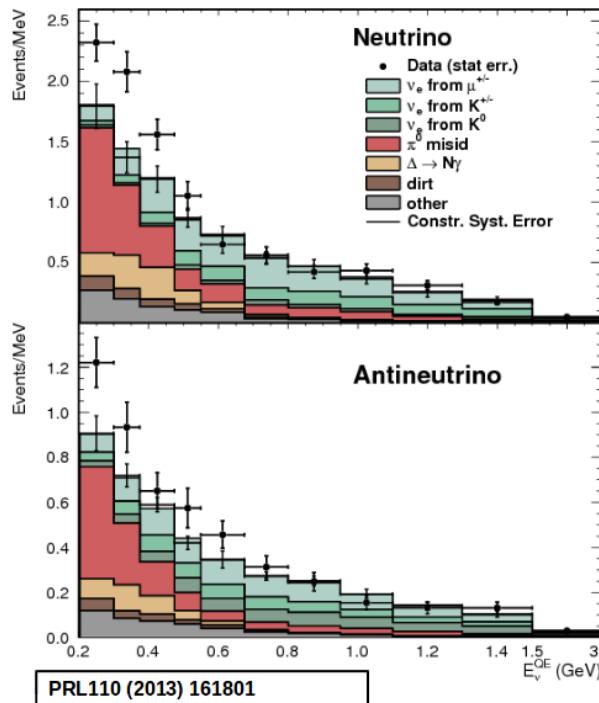


Figure 3.1: Low Energy excess seen in MiniBooNE

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.

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

791 3.3.3 Liquid argon detector development

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

796 3.4 The Booster Neutrino Beam

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

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

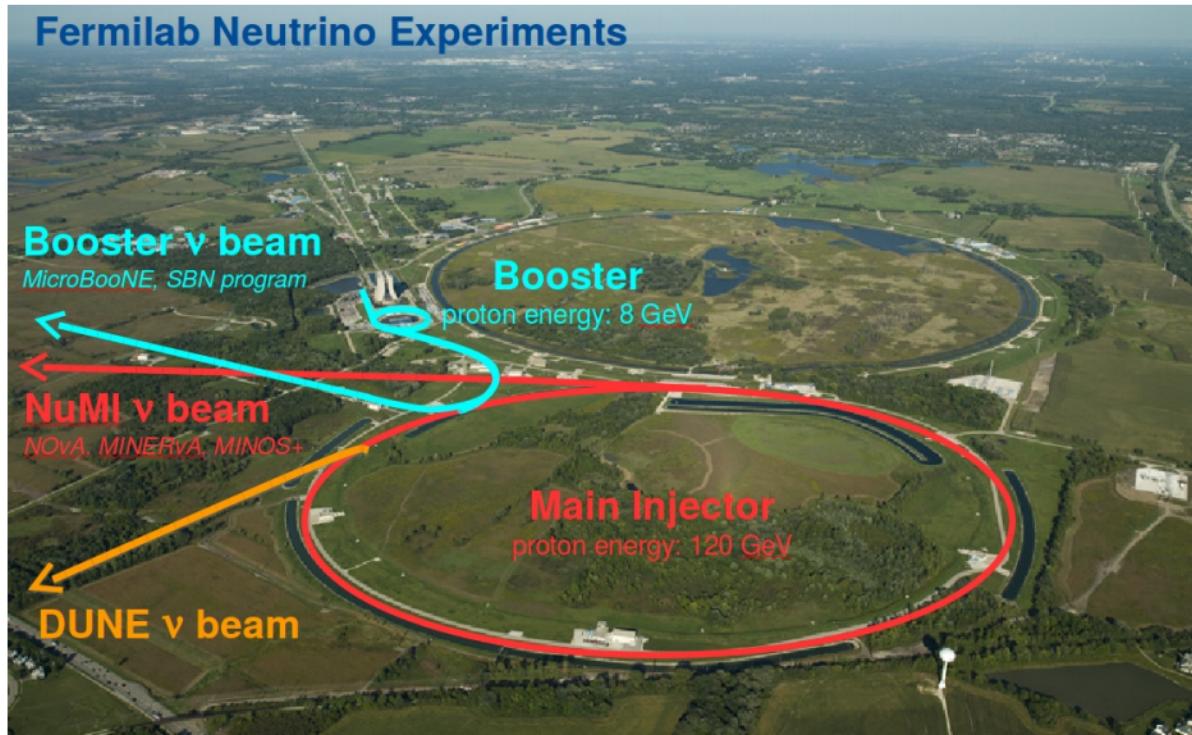


Figure 3.2: Arial view of the Main Injector and the Booster Neutrino Beam at Fermilab

806 3.4.1 Creating the Booster Neutrino Beam

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

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

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

841 **3.5 Event Reconstruction**

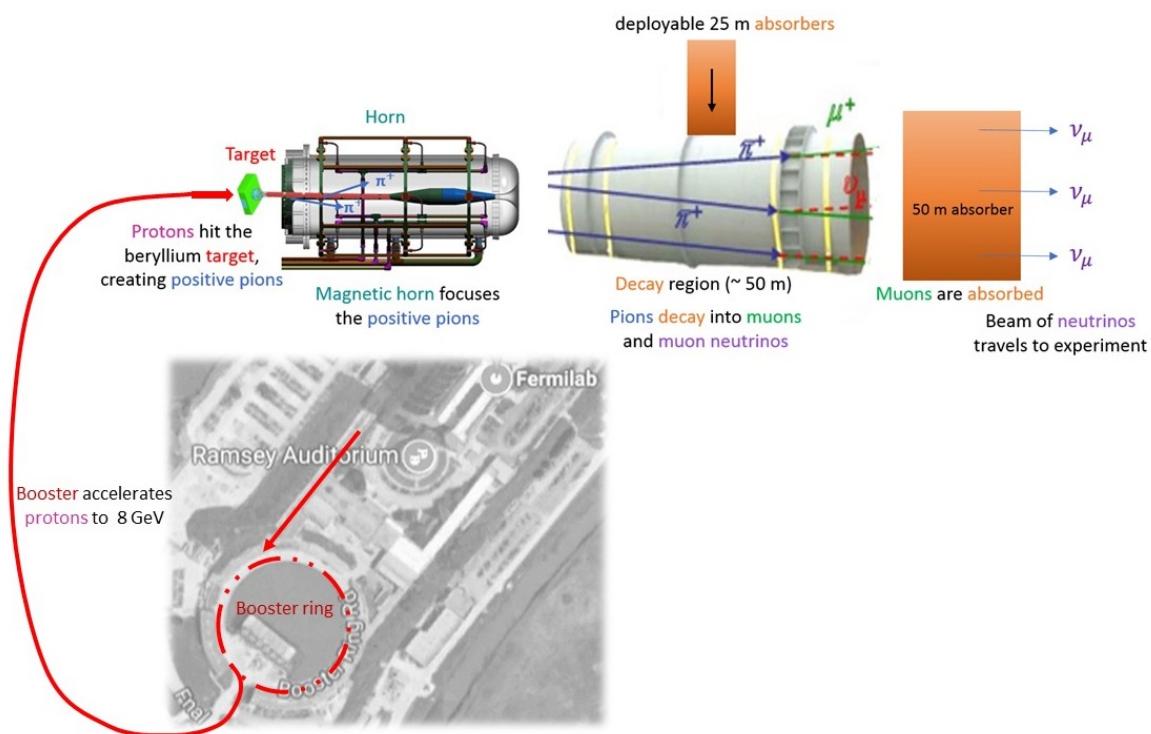


Figure 3.3: Aerial view of the Main Injector and the Booster Neutrino Beam at Fermilab

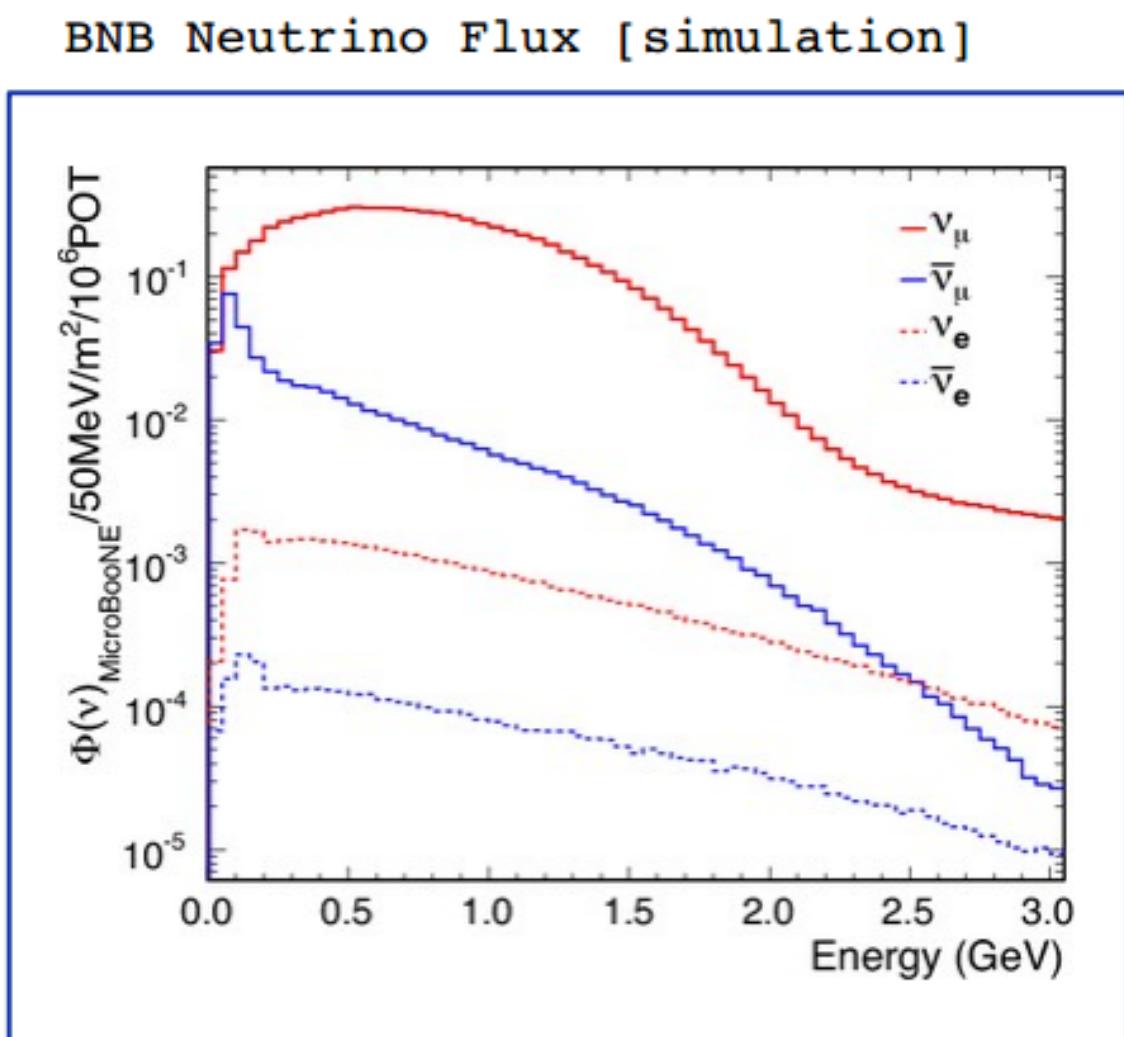


Figure 3.4: Energy spectrum of the Booster Neutrino Beam at Fermi National Laboratories

842 **Chapter 4**

843 **Neutrino Identification: Finding**
844 **MicroBooNE's first Neutrinos**

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

859 **4.1 Flash Finding**

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

863 **4.1.1 Flash Reconstruction**

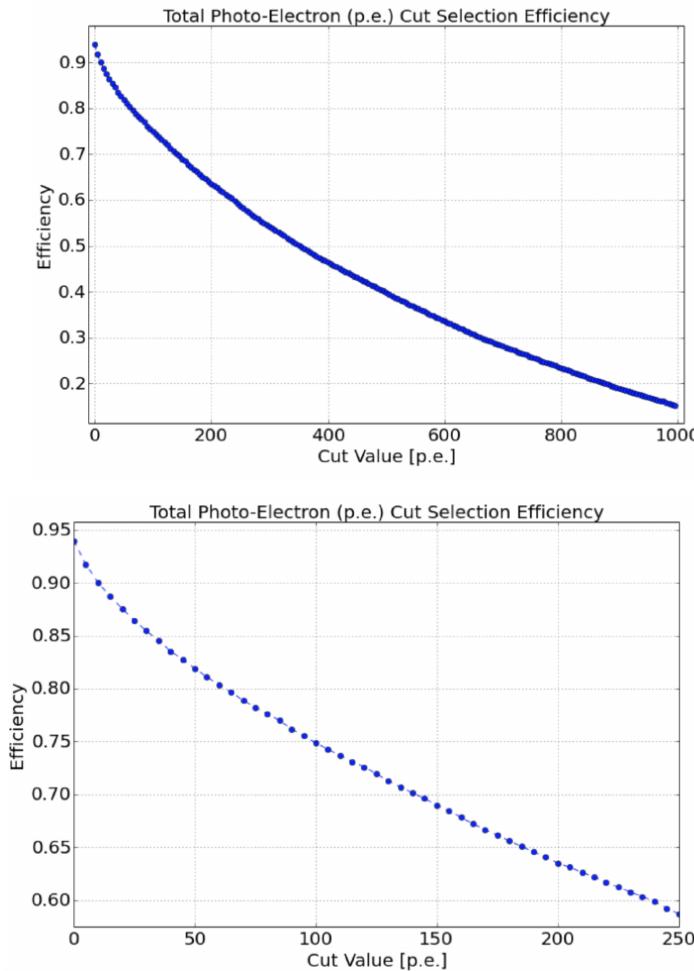
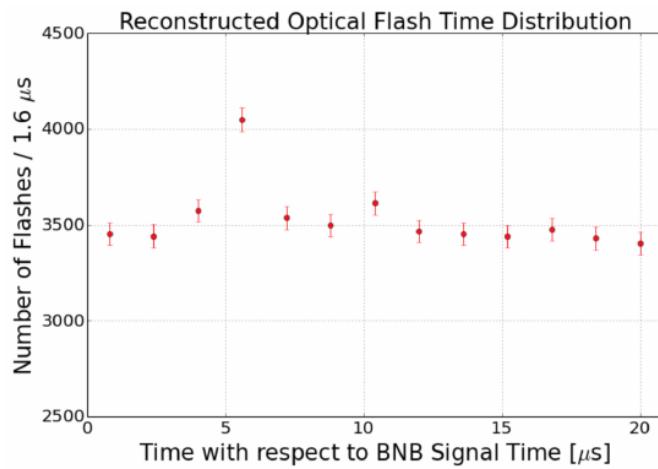


Figure 4.1: Efficiency for selecting beam events as a function of minimum total PE cut for all PE cuts as well as zoomed into interesting region.

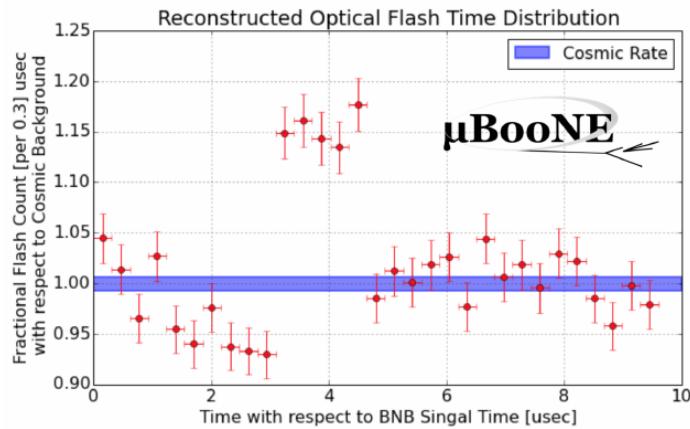
864 A flash is described as a collection of light seen at the same time within the detector.
865 They are then reconstructed by identifying signal from the PMTs above a specific
866 photoelectron (PE) threshold. These signals are called optical hits. Optical hits from
867 all the PMTs are then accumulated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

⁸⁷⁴ cut of 50 PE was deemed sufficient for this analysis. Figure 4.1 show the total PE
⁸⁷⁵ versus the selection efficency of selecting neutrino beam events.

⁸⁷⁶ 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.

⁸⁷⁷ It is necessary to get the specific time from flashes if one uses flashes to filter out
⁸⁷⁸ neutrino interactions coincident with the neutrino beam spill period and background.
⁸⁷⁹ Before a filter can be applied, an understanding of the timing of the trigger and PMT

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

4.1.3 Event Rates

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

4.2 TPC Topology Selection

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

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

906 4.2.1 Cosmic Tagging

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

- 911 • 1: The track is tagged as entering or exiting the TPC
- 912 • 0.95: The track is a delta ray associated with a tagged track
- 913 • 0.5: The track is either entering or exiting, but not both
- 914 • 0.4: The track is entering or exiting through the Z boundary
- 915 • 0: The track isn't tagged

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

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

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

931 4.2.2 2D Cluster Selection

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

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

941 **Primary Cuts**

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

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

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

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

961 **Secondary Cuts**

962 The secondary cuts look to match long, low-angle clusters with short, high-charge
 963 clusters. Only clusters that have passed previous cuts are used. First clusters with
 964 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.

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

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

977 4.2.3 3D Tracks and vertices Selection

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

- 981 • d : distance between the start points of the two tracks.
- 982 • d_1 : distance between vertex and start of track 1.
- 983 • d_2 : distance between vertex and start of track 2.

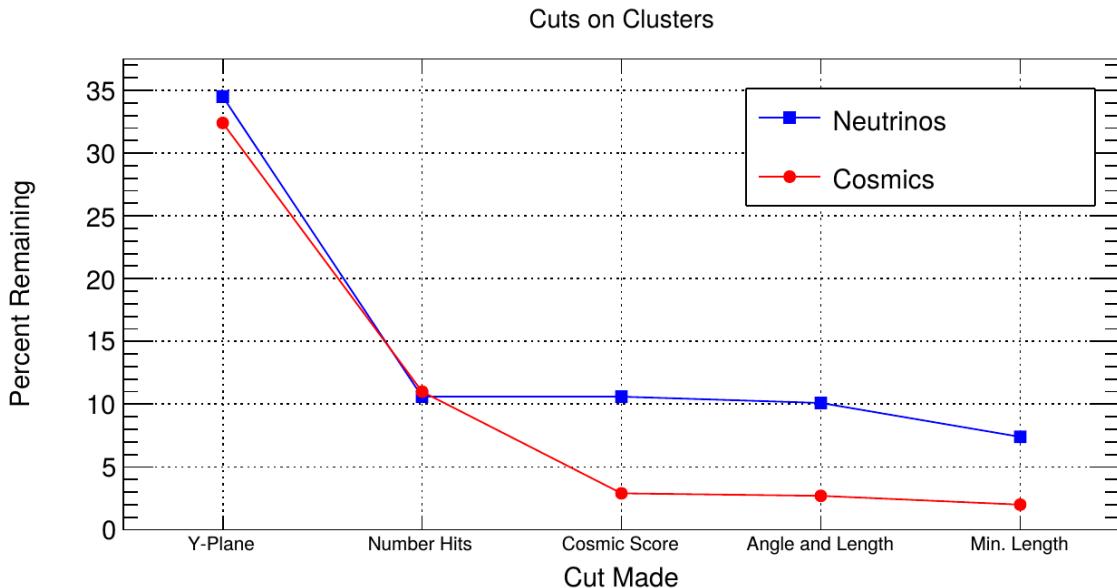


Figure 4.3: Percent of good clusters remaining for neutrinos and cosmics after the primary cuts were applied. This is relative to total number of initial clusters.

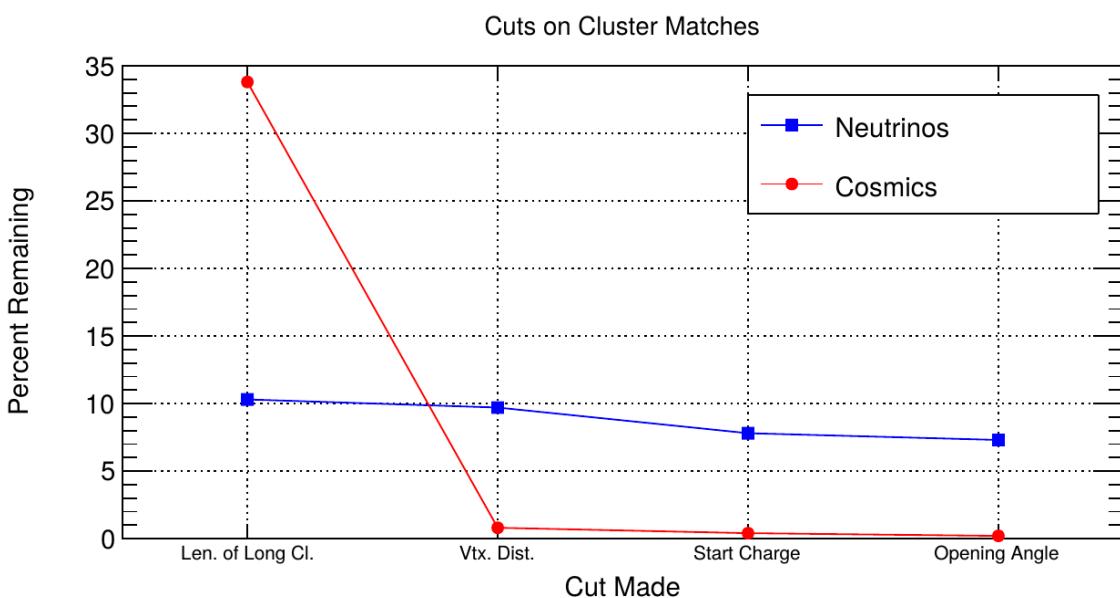


Figure 4.4: Percent matched cluster pairs remaining for neutrinos and cosmics after secondary cuts applied. This is relative to the number of events that contain clusters which pass the primary cuts.

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

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

994 **4.2.4 TPC Updates**

995 After doing a visual hand-scanning of the first beam data processed with the filters
996 detailed above, the events passing had a larger contamination of background than
997 expected. This was mainly in part due to the reconstruction performing better on
998 simulation than on data. Due to this, additional cuts on both streams needed to be
999 implemented in order to increase signal/background ratio. These cuts were added on
1000 top of the filters described above and further reduce the event count.

1001 **2D Filter Updates**

1002 The main background observed in the 2D filter were Michel events, where the muon
1003 and electron formed two connected clusters. These events were rejected by comparing
1004 the start and end charge deposition of the long cluster (i.e muon particle). The start
1005 charge deposition must be less than the end charge deposition. This cut is implemented
1006 because muons have a higher ionizaiton loss at the end.

1007 **3D Filter Updates**

1008 It was seen that cosmic tracks can often originate or end at the same point, therefore
1009 faking a signal. Cosmic tracks, however, are mostly vertical. By requiring the angle
1010 of the longer track have a cosine greater than 0.85 with respect to the z-axis as well

₁₀₁₁ as requiring the longer track to have a length greater than 10 cm, we can reduce this
₁₀₁₂ background.

₁₀₁₃ **4.3 Conclusion**

₁₀₁₄ After proccesing these filters in parallel, it was shown that the 3D filter had a higher
₁₀₁₅ purity than the 2D filter because of the higher cosmic rejection being used due to 3D
₁₀₁₆ reconstruction. The 2D filter is blind to track entering/exiting from the top or bottom
₁₀₁₇ of the TPC. Although the 3D filter had a higher purity, the 2D filter was still able to
₁₀₁₈ find identifiable events in data that were used as public event displays. A sample of
₁₀₁₉ event displays are shown in figures 4.5 and 4.6.

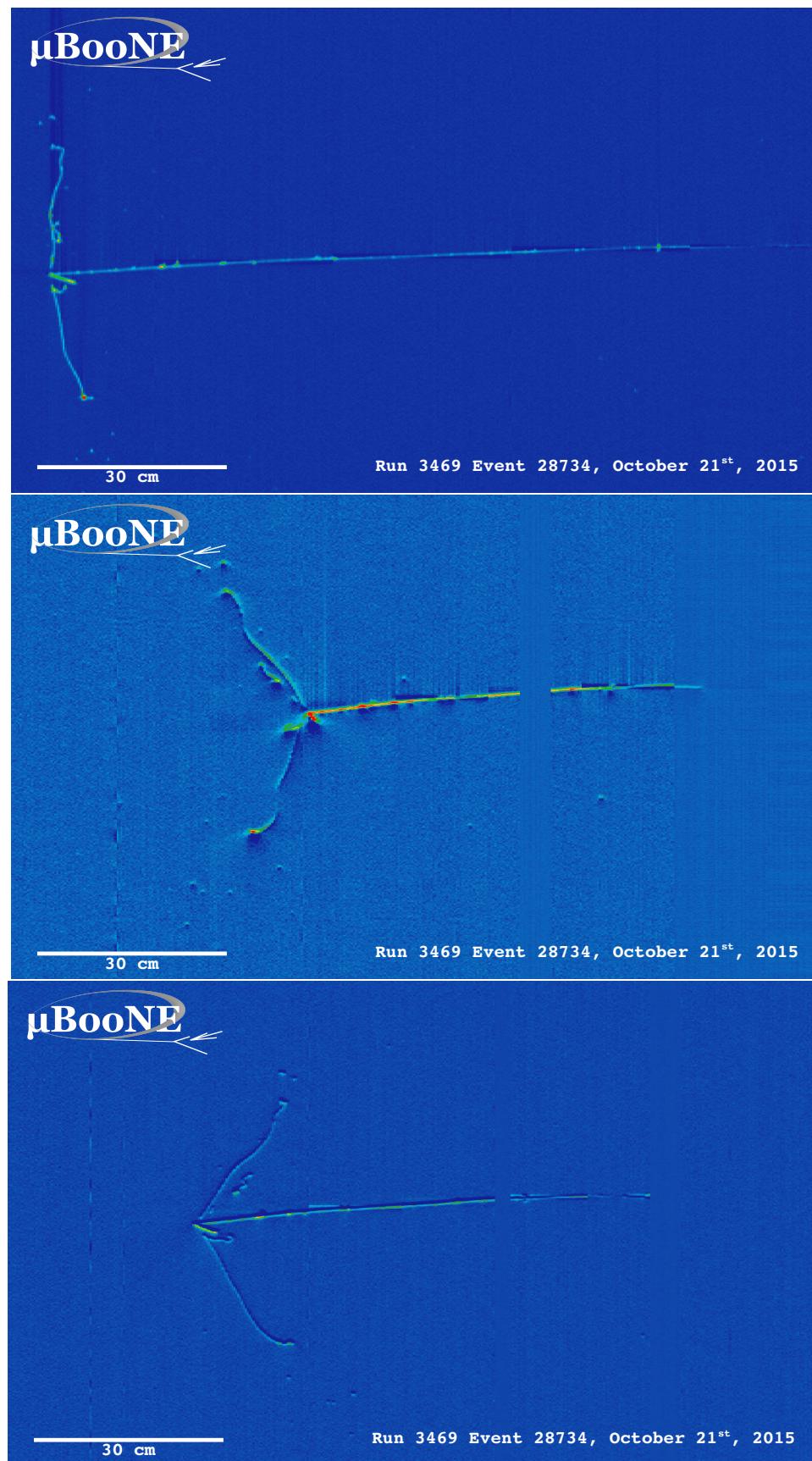


Figure 4.5: First Neutrino Interaction Candidate Events from MicroBooNE

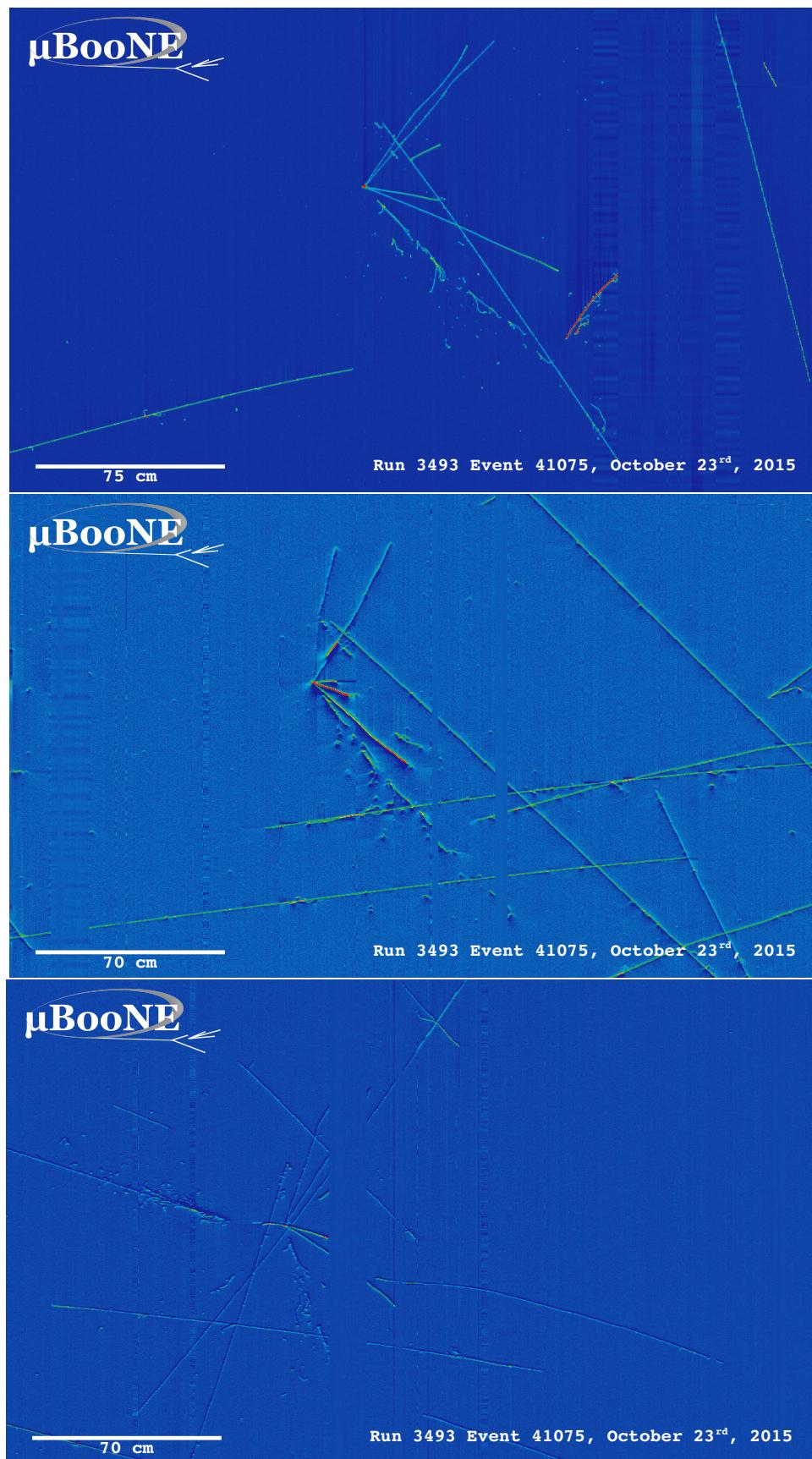


Figure 4.6: First Neutrino Interaction Candidate Events from MicroBooNE

1020 **Chapter 5**

1021 **CC-Inclusive Cross Section Selection**

1022 **Filter**

1023 The CC-Inclusive cross-section selection I and selection I modified filters used in this
1024 analysis will be described in the following sections below. These filters are an expansion
1025 of the Neutrino ID filter. The work done in this thesis was to further improve these
1026 selections by increasing both efficiency and purity as well as increasing acceptance
1027 without further affecting the kinematic distributions of the selected neutrino events.

1028 MicroBooNE requires fully automated event reconstruction and selection algorithms for use in the many physics measurements being worked on to date due to
1029 the large data rate MicroBooNE receives. Being able to automatically pluck out the
1030 neutrino interaction among a sea of cosmics proved to be challenging but was accomplished.
1031 MicroBooNE has developed two complementary and preliminary selection
1032 algorithms to select charged-current $\nu_\mu - Ar$ interactions. Both are fully automated
1033 and cut based. The results of this thesis will focus on selection I and selection I modified
1034 and will focus on further improving these algorithms using Convolutional Neural
1035 Network (CNN) implementations. These selections identify the muon from a neutrino
1036 interaction without biasing towards track multiplicity. To combat cosmic and neutral
1037 current background, the analysis is strongly biased towards forward-going long tracks
1038 which are contained. This limits phase space and reduces acceptance.
1039

1040 5.1 Data and MC Processing Chain

1041 The data used for this analysis were based on hardware and software triggers. Events
1042 used came from the *BNB_INCLUSIVE* and *EXT_BNB_INCLUSIVE* streams and were
1043 used for signal and background. The *BNB_INCLUSIVE* stream is chosen by requiring
1044 that the hardware trigger bit is fired and that the event passed an optical software
1045 trigger within a BNB spill timing window. The *EXT_BNB_INCLUSIVE* stream requires
1046 the EXT hardware trigger to fire as well as pass the same optical software trigger
1047 within a BNB spill size timing window similar to the *BNB_INCLUSIVE*.

1048 The two MC samples used in this analysis and for determining selection efficiencies
1049 and purities were GENIE BNB neutrino interactions with CORSIKA cosmic ray overlay
1050 within the readout window and inTime CORSIKA cosmic rays. The MC samples
1051 generated used *uboonecode v04_36_00* and are based on the following packages:

- 1052 • larsoft v04_36_00
- 1053 • GEANT v04_09_06_p04d
- 1054 • GENIE v02_08_06d
- 1055 • GENIE xsec v02_08_06a
- 1056 • pandora v02_03_0a
- 1057 • CORSIKA v07_4003

1058 Both data and MC samples were processed using the same reconstruction release,
1059 *uboonecode v05_08_00* and the fcl files used for reconstruction are listed below:

- 1060 • MC fcl files
 - 1061 – reco_uboone_mcc7_driver_stage1.fcl
 - 1062 – reco_uboone_mcc7_driver_stage2.fcl
- 1063 • Data fcl files
 - 1064 • reco_uboone_data_Feb2016_driver_stage1.fcl
 - 1065 • reco_uboone_data_Feb2016_driver_stage2.fcl

1066 On top of the hardware and software triggers, the data also had to pass more
1067 criteria to be identified as part of the good run list. The criteria is detailed below.

- 1068 • **Detector conditions:** the detector has to be in a good operating condition. The
1069 detector conditions are read from the slow monitoring database and are required
1070 to be within the alarm thresholds. The variables of interest for events passing
1071 the good run list criteria include DAQ, PMT, HV, Drift HV, wire bias, electron
1072 lifetime and detector power. These conditions need to be met on a run-by-run
1073 basis in order to pass the selection.
- 1074 • **Data quality:** normal and stable behavior for basic reconstruction quantities.
1075 These reconstruction variables include average number of tracks, hits, and flashes
1076 in each event, the average length of tracks, the average amplitude and area of
1077 hits, the average PE and the average spread of each one of these quantities.
- 1078 • **Beam Conditions:** the BNB must be on and stable and the POT per spill needs
1079 to above the intensity threshold. Beam quality conditions include checking the
1080 fraction of proton beam interacting within the target, the horn current, and the
1081 intensity of protons per spill. The final sample is $5 * 10^{19}$ and a per-spill intensity
1082 of $4 * 10^{12}$
- 1083 • **Run processed:** the full run must be processed completely without missing
1084 subruns or crashes in the data processing.

1085 5.2 Normalization of data and MC

1086 The off-beam sample is used to measure beam unrelated backgrounds. For normalization
1087 one needs the total number of BNB spills (N_{BNB}) and the total number of external
1088 triggers. The BNB spills used need to pass the beam quality cuts. The normalization
1089 factor is then N_{BNB}/N_{EXT} which is 1.23.

1090 To normalize generated BNB MC events to POT, we used the following:

- 1091 • $5 * 10^{19} POT = 41524.3$ generated events

1092 where this scaling factor only applies to mcc7 generated events. The inTime cosmic
1093 sample is normalized with respect to the open cosmic sample so an understanding
1094 of both is necessary. The POT per beam spill for mcc7 BNB samples is $5 * 10^{12}$. To
1095 calculate how many spills are necessary to produce a specific POT one would multiply
1096 the total POT by the average 1/POT per spill. For a total POT of $5 * 10^{19}$ the amount
1097 of spills necessary is $\frac{5 * 10^{19}}{5 * 10^{12}} = 1 * 10^7$. This is only one in ~ 241 events therefore each

1098 cosmic event needs to be scaled up by a factor of 240.8 when comparing to BNB
 1099 MC. For inTime cosmics however, two filters are applied to reduce computing and
 1100 processing time and only leave cosmics that will interact within the detector. The
 1101 passing rate after these two filters is 0.02125, therefore the total inTime cosmic scaling
 1102 factor to compare inTime cosmics to BNB is $0.02125 * 240.8 = 5.12$.

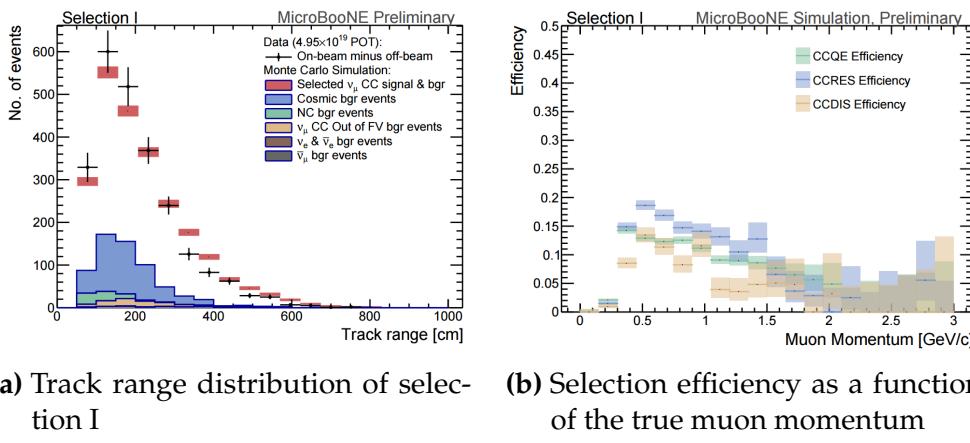


Figure 5.1: 5.1a Track range distribution for selection I. The track range is defined as the 3D distance between the start and end of the muon candidate track. No data is shown below 75 cm due to the track length cut described previously. 5.1b Efficiency of the selected events by process quasi-elastic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

1103 5.3 Optical Software Trigger and Reconstruction

1104 5.3.1 Software Trigger

1105 Most of the BNB spills from the accelerator do not have a neutrino interaction in
 1106 MicroBooNE. To save computation resources and reduce data-rates, we require a
 1107 burst of light in the light collection system in coincidence with the $1.6 \mu\text{s}$ beam spill.
 1108 Requiring light activity in coincidence with the beam spill eliminates the vast majority
 1109 of triggers with no neutrino interaction in the detector, however, it doesn't guarantee
 1110 the activity in the detector is a neutrino interaction since a cosmic ray can interact in
 1111 coincidence with the beam spill as well.

1112 To implement this, a software trigger was used on the PMT waveforms to decide
1113 whether or not to keep that event. The software trigger is implemented after the event
1114 builder combines data from the PMTs and triggers into a single event. The software
1115 trigger uses the digitized output of the 32 PMT channels in the light collection system.
1116 Only the waveform region in coincidence with the beam spill is used to search for
1117 possible triggers. For each PMT, a waveform is found by taking the difference of ADC
1118 values is calculated between t and $t + s$. This waveform is then scanned for ADC
1119 values above a threshold X_0 . Once an ADC is above this threshold, a discriminator
1120 window is opened for a fixed number of time ticks (W_0). If the ADC count within this
1121 window W_0 is greater than a second larger threshold X_3 , a final window of width W_3
1122 is opened. The max ADC value within this final window is set as the peak amplitude
1123 for the PMT and then summed across all 32 PMTs and set to the variable PHMAX. The
1124 software trigger places a final cut on the PHMAX variable to decide whether or not
1125 to keep the event. The thresholds were found by the Trigger task force using Monte
1126 Carlo Studies and are as follows:

- 1127 • $X_0 = 5$ ADC
- 1128 • $X_3 = 10$ ADC
- 1129 • $W_0 = 6$ Ticks
- 1130 • $W_3 = 6$ Ticks
- 1131 • PHMAX cut = 130 ADC

1132 5.3.2 Flash Reconstruction

1133 MicroBooNE collects light from each of the 32 PMTs either in a continuous readout
1134 window of $23.4 \mu\text{s}$ activated by a beam gate signal on the trigger board, or in discrimi-
1135 nated pulses of $\sim 1 \mu\text{s}$ duration activated if the ADC count for any PMT goes above 80
1136 ADC count. These two formats are saved as output waveforms and put onto an event.
1137 Additionally, each PMT can provide two output streams, high-gain (~ 20 ADC/PE)
1138 and low-gain (~ 2 ADC/PE) channels. The first step in the reconstruction is to merge
1139 both these channels into a “saturation corrected waveform” which uses information
1140 from the low-gain waveform to correct for saturating high-gain pulses.

1141 The saturation corrected waveform in the continuous readout window is used to
1142 reconstruct optical hits. Each PMT’s waveform is scanned for hits then a threshold

1143 based hit reconstruction algorithm is applied which requires pulses of a minimum
1144 area in order to be reconstructed. Each reconstructed hit is associated to a PMT, a time
1145 in μs , and a PE count.

1146 Once hits are reconstructed for all 32 PMTs, all PMT information is then combined
1147 into optical flashes which represent optical information seen by the PMTs from interac-
1148 tions in the detector. Each flash has information on total light seen per interaction, the
1149 distribution of the light across all 32 PMTs, the flash time with respect to the trigger
1150 time of the flash, and lastly, the spacial information of the flash in Y-Z plane of the
1151 detector. These flashes are reconstructed by requiring that there is a $\sim 1 \mu\text{s}$ coincidence
1152 between the reconstructed hits in all 32 PMTs. The total PE is summed up among
1153 all coincident hits across the PMTs and if the total PE is greater than 2 PE, a flash is
1154 reconstructed. There are also safe guards in place to take care of late scintillation light.

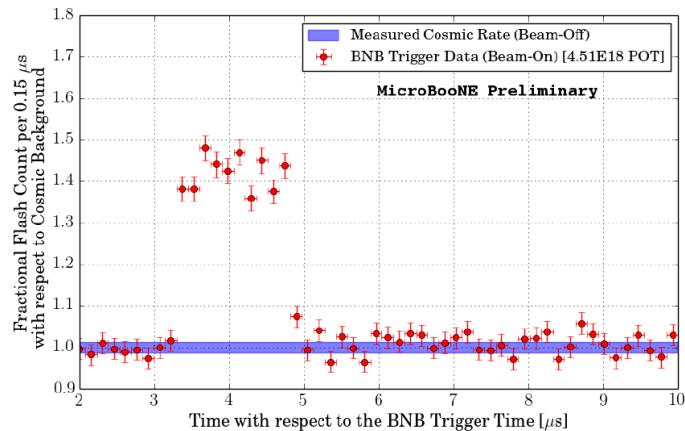


Figure 5.2: Time distribution of reconstructed optical flashes with a PE value of 50 or more for a sample of BNB unbiased triggered events.

1155 Figure 5.2 shows the time distribution of reconstructed optical flashes using the
1156 BNB continuous stream. You can see a clear excess in coincidence with the expected
1157 arrival time of neutrinos. The same flash reconstruction that was used in the cc-
1158 inclusive filter detailed here was used to create this plot in data.

1159 5.3.3 Beam Window

1160 Figure 5.3 shows the distribution of flashes for on-beam, off-beam and various MC
1161 samples. The software trigger has been applied to these samples. The pile-up seen just
1162 after 0 μs is a feature of the flash finding algorithm and consists of low PE flashes and

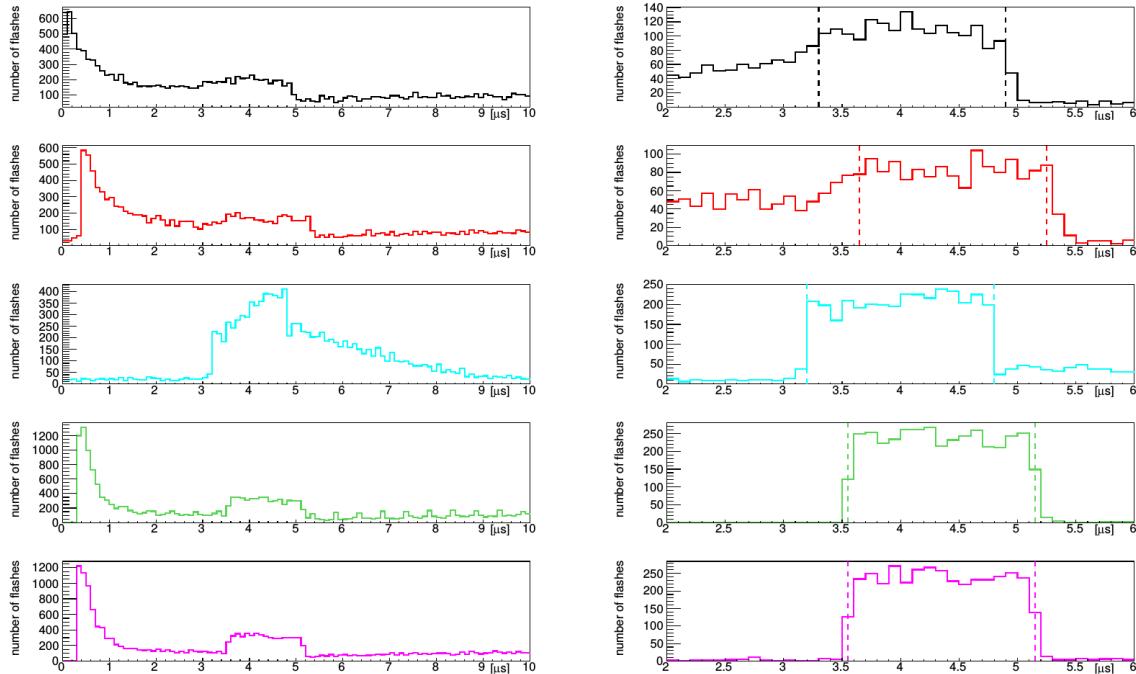


Figure 5.3: Flash time distribution for all flashes (left plot) and flashes $> 20\text{PE}$ (right plot). The different curves are as follows: on-beam data (black), off-beam data (red), CORSIKA inTime MC (light blue), BNB only MC (green), and BNB+Cosmic MC (purple). The dashed vertical lines mark the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

1163 is removed in the second column of distributions with a low 20 PE threshold cut. The
1164 plots show that the time window for the distributions are shifted a small amount from
1165 each-other. This is caused by different hardware configurations per sample. Using
1166 these distributions, the windows chosen per sample are as follows:

- 1167 • On-Beam: 3.3 to 4.9 μ s
- 1168 • Off-Beam: 3.65 to 5.25 μ s
- 1169 • CORSIKA inTime: 3.2 to 4.8 μ s
- 1170 • BNB only: 3.55 to 5.15 μ s
- 1171 • BNB+Cosmic: 3.55 to 5.15 μ s

1172 Each window has a width of 1.6 μ s.

1173 5.4 TPC Reconstruction

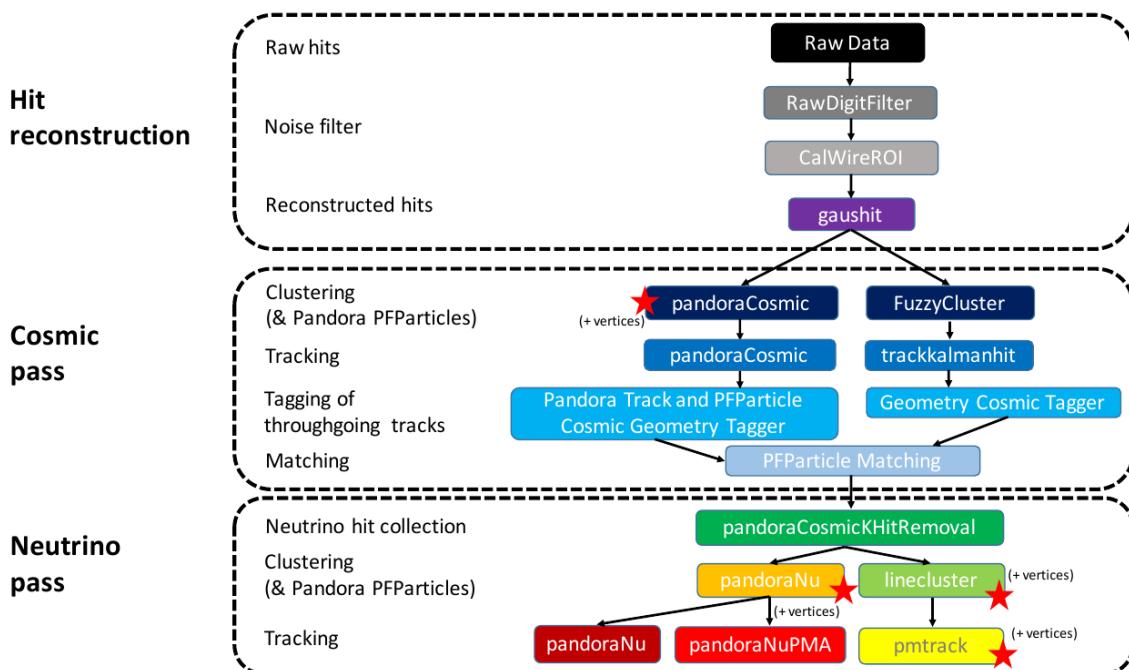


Figure 5.4: Reconstruction chain run on both data and MC. The red stars mean that the algorithm returns reconstructed 3D vertices.

1174 Figure 5.4 summarizes the reconstruction chain applied to both MC and data for
1175 this analysis. After the hit reconstruction, a cosmic pass is applied which removes all
1176 hits associated to through-going tracks. A description of these TPC reconstruction
1177 algorithms will be detailed below.

1178 **5.4.1 Hit Reconstruction**

1179 The waveforms used for hit reconstruction consist of charge deposited on the sense
1180 wire in drift time. The first step in hit reconstruction is to pass the waveforms through
1181 a filtering algorithm to filter out the noise introduced from the electronics. The input
1182 waveforms are also truncated from 9600 time ticks to 6400 time ticks in this first step
1183 to reduce the data footprint of these waveforms.

1184 Once noise filtering is complete, a deconvolution algorithm is applied to the wave-
1185 forms to remove the drift field and electronics response, therefore leaving only the
1186 ionized electrons kicked off the argon atoms by an incident track. During this process,
1187 Region of Interests (ROI) are identified and cut out of the waveforms to further reduce
1188 the data volume.

1189 The hit finding algorithm then finds candidate peaks in these ROI's and fits the
1190 peaks to Gaussian curves. These Gaussian shaped peaks are now called hits and
1191 represent the charge deposition on a wire by the incoming track. These hit objects
1192 have a peak time and width and are the basic object input to further algorithms down
1193 the reconstruction chain.

1194 **5.4.2 Clustering**

1195 There are multiple clustering algorithms used in this analysis. The main purpose of all
1196 the clustering algorithms is to associate hits together in 2D space to create objects like
1197 tracks, vertices and showers. For the fuzzy cluster algorithm, three steps are used to
1198 achieve this. The first step is to associate hits to each-other using a fuzzy clustering
1199 algorithm which gives each hit a degree of belonging to the cluster. Second, a Hough
1200 transform is used to find hits associated to candidate tracks and showers within each
1201 of the clusters found in the first step. The last step merges smaller candidate tracks
1202 and showers into large clusters. The last step also associates unclustered hits into

1203 nearby objects which helps shower reconstruction. The result is a set of clusters made
1204 up of associate hits that represent tracks or showers per plane.

1205 The pandora algorithm utilizes it's own clustering algorithm and will be detailed
1206 in the next section. The last clustering algorithm is called linecluster. The linecluster
1207 algorithm reconstructs 2D linear clusters per plane by fitting a line onto nearby hits
1208 which is then extrapolated to neighboring wires. 2D vertices are found per plane by
1209 using the intersection points of the ends of nearby clusters. These 2D vertices are then
1210 matched in time across all three planes to get a 3D vertex in space.

1211 5.4.3 Pandora

1212 5.4.4 Trackkalmanhit

1213 The trackkalmanhit algorithm takes 2D clusters returned from the fuzzy cluster algo-
1214 rithm and outputs track objects. There are no hierarchy structure as it is in pandora,
1215 each track is independent. There also is no vertex reconstruction with this algorithm
1216 as well.

1217 5.4.5 Cosmic Hit Removal

1218 The Pandora algorithm is applied to the events twice, the first to remove downward
1219 going tracks primarily from cosmic ray muon like particles. The second pass only runs
1220 on a subset of hits that aren't associated with cosmic ray muon tracks.

1221 After the first pass, the output of PFParticle hierarchy is then passed to a cosmic
1222 ray tagger to look through all hits to determine start and end points. If the start or
1223 end point trajectories are consistent with entering or exiting the TPC, then these hits
1224 are removed from the second pass. Hits are considered entering or exiting the TPC
1225 if the drift time are outside of the neutrino drift window or outside of the fiducial
1226 volume of the TPC. The fiducial volume was based on a montecarlo study and is 20
1227 cm from the top or bottom of the TPC and 10 cm from the TPC ends. Hits associated
1228 with candidate cosmic ray tracks are removed from the input hit collection and the
1229 remaining hits are passed to the neutrino optimized pass of Pandora.

1230 5.4.6 Projection Matching Algorithm

1231 The projection matching algorithm (PMA) was inherited from ICARUS and has been
1232 implemented in LArSoft. PMA differs from traditional LArSoft 3D reconstruction
1233 algorithms. Most 3D reconstruction attempts to match 2D objects from all three planes
1234 by drift time, while the PMA algorithm projects a track hypothesis on each plane
1235 then the distance between this projection and the hits on each plane is minimized
1236 simultaneously. More information can be found in [?].

1237 5.5 Event Selection

1238 The first requirement for selecting ν_μ CC events is that the event has at least one
1239 scintillation light flash in the beam trigger window with more than 50 PE on all PMTs
1240 combined. From the flashes that pass, the most intense is chosen and considered to be
1241 originating from a neutrino interaction and will be the only flash used in further cuts.

1242 Vertices are then required to have at least one reconstructed track start or endpoint
1243 within a 5 cm radius. Showers associated with a vertex do not pass this cut. All
1244 tracks associated with a vertex are then used to calculate a track length weighted
1245 average of the θ -angle. Of all the vertices that do pass, only the vertex with the most
1246 forward going θ -angle average of all associated tracks is considered the neutrino vertex
1247 candidate. The most forward going θ -angle average is chosen by picking the largest
1248 track range weighted average of $|\cos(\theta)|$, seeing as $\cos(\theta) = 1$ is the beam direction.
1249 Next, it is required that the reconstructed neutrino vertex candidate be within the
1250 fiducial volume as well as within the drift time starting at t_0 . The fiducial volume
1251 boundaries chosen are 10 cm from the edges of the TPC in x and z which is the drift
1252 direction and beam direction respectively, and 20 cm from the edges of the TPC in y
1253 which is the vertical direction. For all further cuts, only the longest track associated
1254 with the neutrino vertex candidate and this track is assumed to be the muon candidate
1255 of the neutrino event.

1256 The next cut requires the position of the flash in the z-direction and the track z-
1257 projection to be compared. This basic flash matching algorithm is rudimentary and a
1258 placeholder for a more sophisticated algorithm. The z-position of the flash needs to be
1259 within 80 cm to the z-positions of track start or endpoints. If the flash is between the
1260 track start and endpoint, the distance of the flash to the track is considered to be 0 cm.

1261 Lastly, the track needs to be fully contained within the fiducial volume and have a
 1262 track range greater than 75 cm. The range is the 3D distance between the track's start
 1263 and endpoint. The length cut was optimized to remove NC background that contain
 1264 a pion due to the pion interaction rate to be ~ 70 cm. A track that makes all the cuts
 1265 is considered to be the muon of a ν_μ CC event. The list of cuts for this selection is
 1266 described below:

- 1267 1. At least one flash > 50 PE within the beam gate.
- 1268 2. At least one track within 5 cm around a vertex.
- 1269 3. Vertex with flattest tracks is chosen to be vertex candidate.
- 1270 4. Vertex candidate in fiducial volume.
- 1271 5. Longest track associated with vertex candidate is chosen to be track candidate.
- 1272 6. Longest track is within 80 cm (z-axis only) of the flash.
- 1273 7. Longest track is fully contained.
- 1274 8. Longest track is greater than 75 cm.

1275 The event selection scheme can also be seen in figure 5.5. Table 5.1 lists the passing
 1276 rates for MC events for the selection scheme described above. Table 5.2 lists the passing
 1277 rates for on-beam and off-beam data for the selection scheme. The normalization
 1278 factors applied between on-beam and off-beam data are described in section 5.2.

1279 5.5.1 Expected Backgrounds

1280 Most of the selected background events will be of cosmic origin. There are two types
 1281 of cosmic background, one triggered by a cosmic-ray event occurring in the beam
 1282 gate time window, the other triggered by a beam induced interaction in the cryostat
 1283 followed by a misidentification of a cosmic event as a neutrino event. The first
 1284 cosmic background can be subtracted from the selected events using the off-beam
 1285 BNBEXT sample normalized to the on-beam. The second cosmic background events
 1286 are modeled by MC by using BNB+Cosmic MC sample.

1287 Other backgrounds originate from neutrino beam contaminants. A major contribu-
 1288 tion in this sector is by neutral current neutrino events for example a charged pion track
 1289 misidentified as a muon. Another contribution are ν_e -like and anti-muon-neutrino

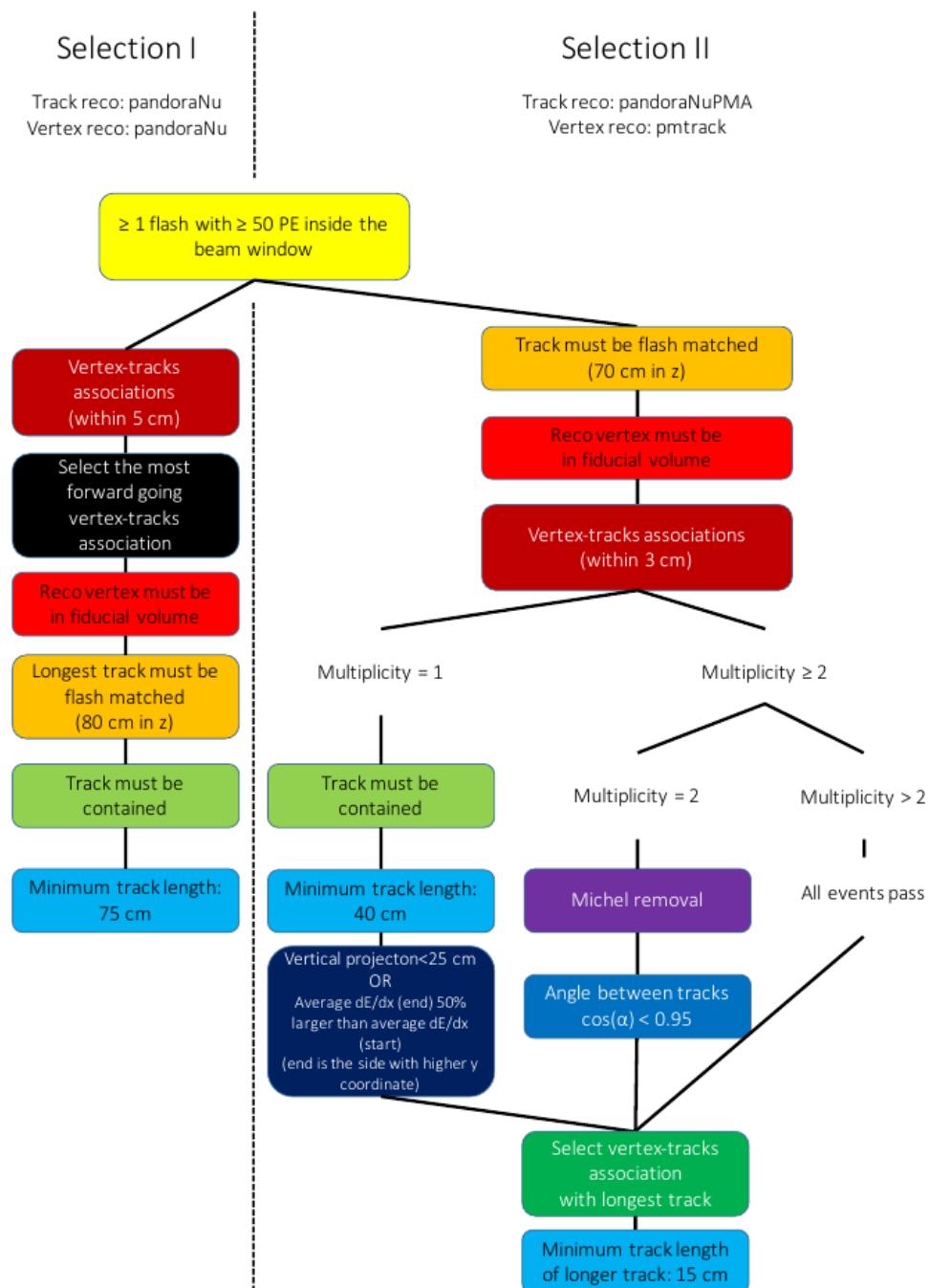


Figure 5.5: Event selection diagram for selection I and selection II. This analysis focused on optimizing selection I. Boxes with the same color across the two selections symbolize similar cuts.

	BNB+Cosmic Selection	BNB+ Cosmic MC-Truth	Cosmic Only	Signal:Cosmic Only
Generated Events	191362	45273	4804	1:22
≥ 1 flash with ≥ 50 PE	136219 (71%/71%)	44002 (97%/97%)	2970 (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

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

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

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

1290 events. These beam related backgrounds are an order of magnitude smaller than the
1291 cosmic misidentification backgrounds. These backgrounds can not be subtracted and
1292 are estimated using MC truth.

1293 The efficiency and purity of Selection I are calculated below:

- 1294 • Efficiency: Number of selected true ν_μ CC events divided by the number of
1295 expected true ν_μ CC events with interaction in the FV.

1296 – $(12.3 \pm 3.4) \%$

- 1297 • Purity: Number of selected true ν_μ CC events divided by the sum of itself and
1298 the number of all backgrounds.

1299 – $(53.8 \pm 4.4) \%$

1300 **5.5.2 Truth Distributions**

1301 The truth distributions of MC truth variables before and after the selection are detailed
1302 in this section. The overall efficiencies are calculated for all ν_μ CC signal events
1303 with a true interaction within the fiducial volume and a fully contained muon track
1304 originating from said vertex. Figures 5.6 through 5.8 detail the truth distributions for
1305 muon momentum, $\cos(\theta)$ and ϕ and figures 5.9 through 5.11 detail the total efficiency
1306 of the selection for charged current quasi elastic (CCQE) events, charged current
1307 resonant (CCRES) events, and charged current deep inelastic (CCDIS) events.

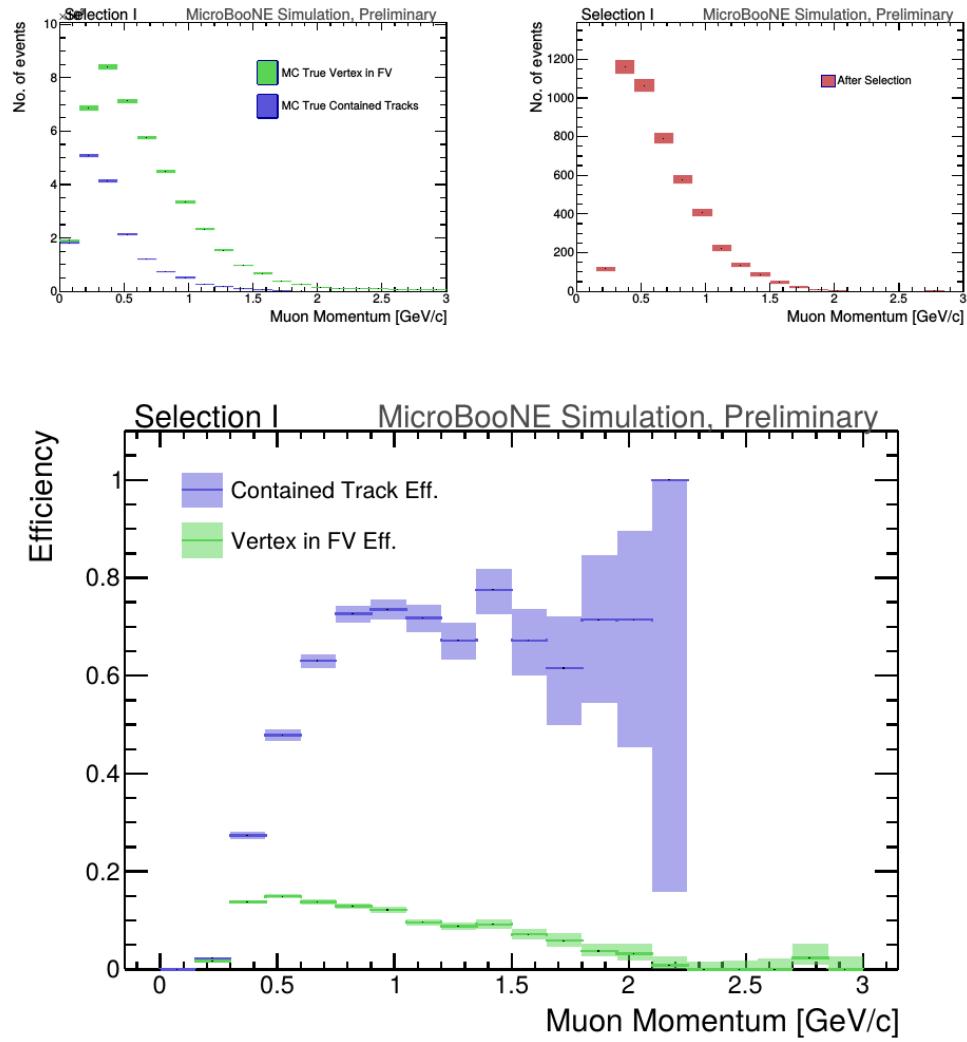


Figure 5.6: MC momentum distributions of the muon originating from a ν_μ CC interaction. Upper left is the momentum distributions of events with a vertex within the FV (green) and the events with a fully contained track (blue) before the selection. The upper right side is the momentum distribution after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

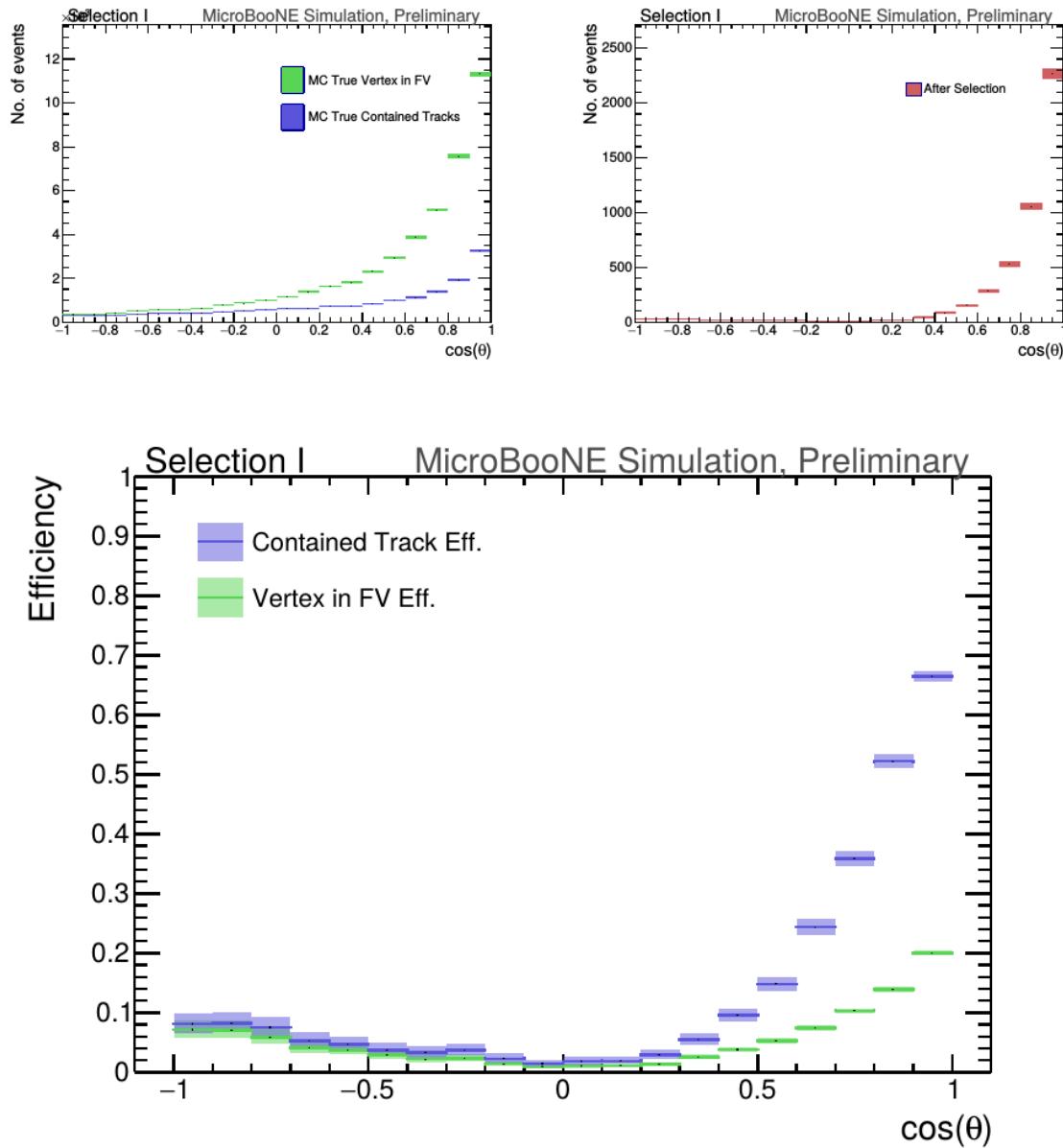


Figure 5.7: MC $\cos(\theta)$ distributions of the muon originating from a ν_μ CC interaction. Upper left is the $\cos(\theta)$ distributions of events with a vertex within the FV (green) and the events with a fully contained track (blue) before the selection. The upper right side is the $\cos(\theta)$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

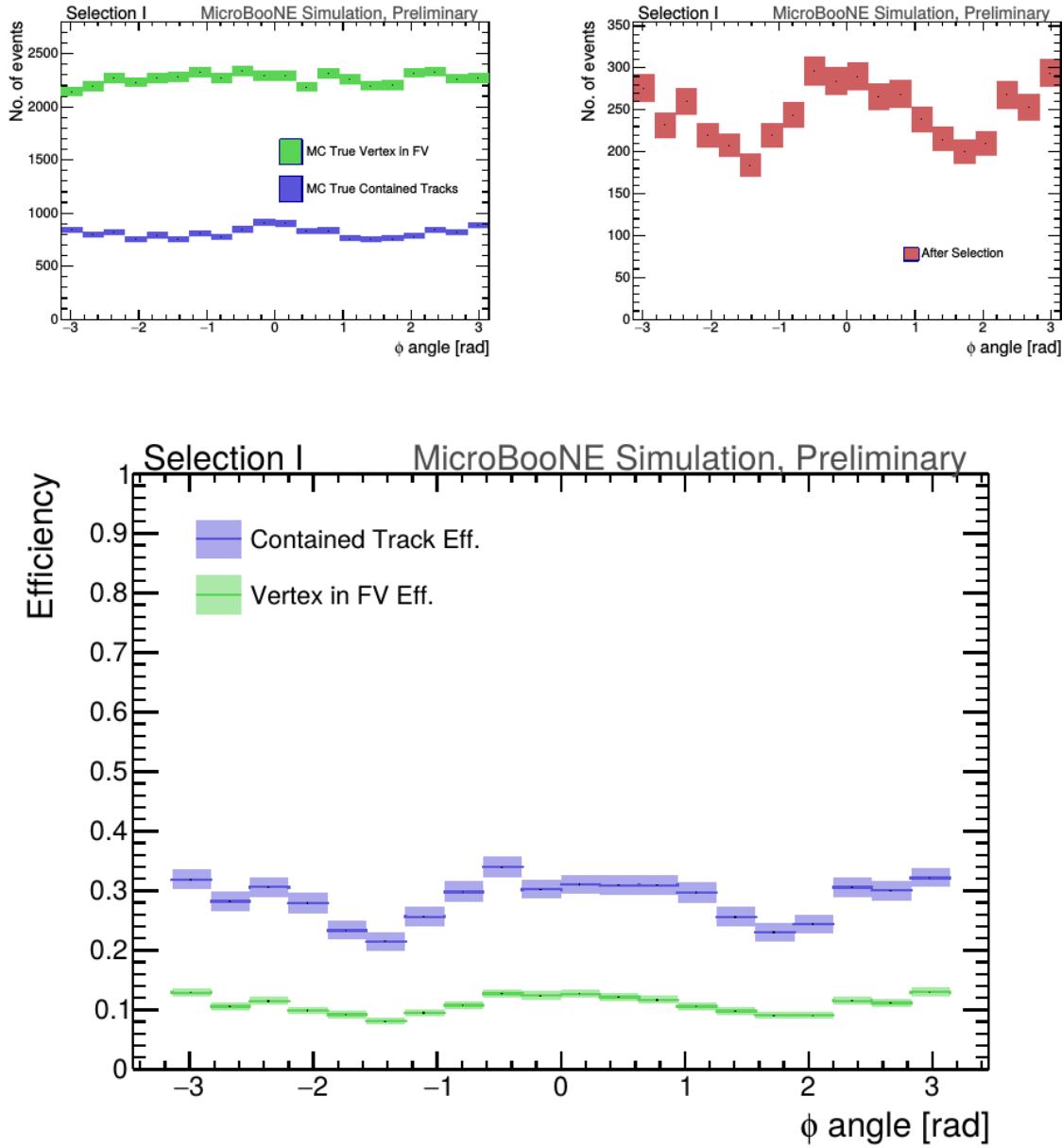


Figure 5.8: MC ϕ distributions of the muon originating from a ν_μ CC interaction. Upper left is the ϕ distributions of events with a vertex within the FV (green) and the events with a fully contained track (blue) before the selection. The upper right side is the ϕ distribution after the selection (red).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

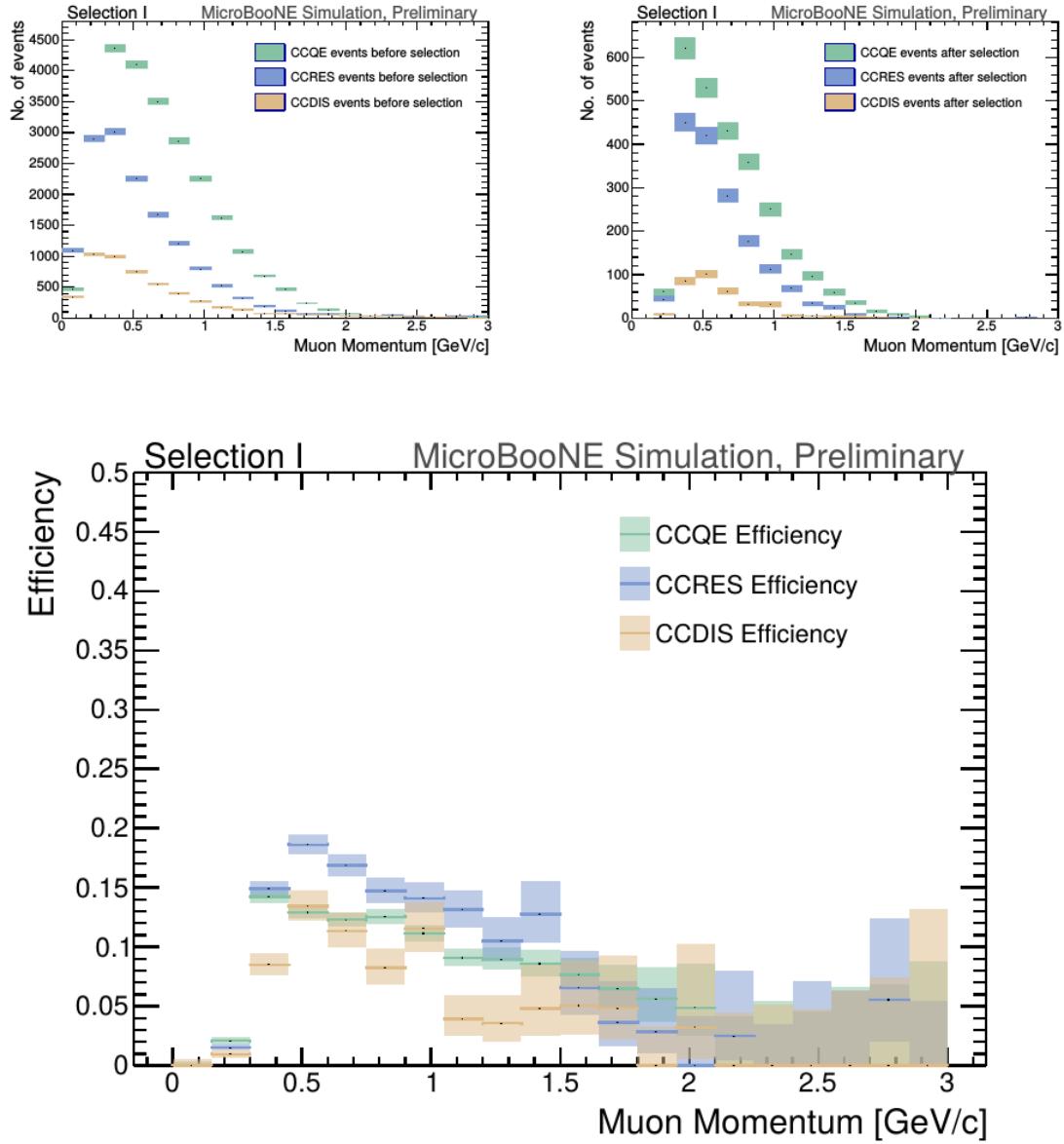


Figure 5.9: MC momentum distributions of the muon originating from a ν_μ CC interaction. Upper left is the momentum distributions of events with a vertex within the FV split up into CCQE (red), CCRES (yellow), and CCDIS (green) before selection. The upper right side is the momentum distribution after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

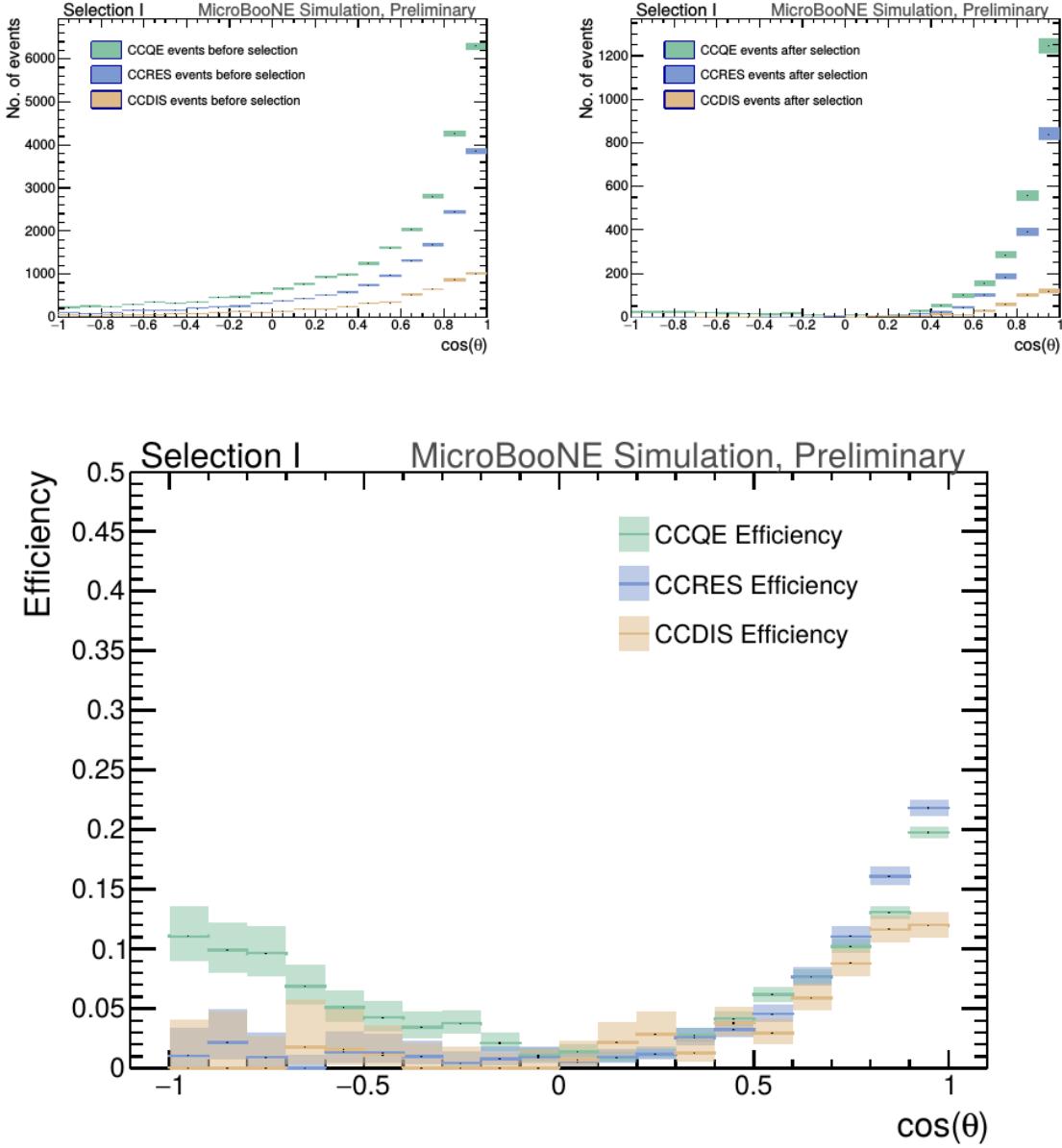


Figure 5.10: MC $\cos(\theta)$ distributions of the muon originating from a ν_μ CC interaction. Upper left is the $\cos(\theta)$ distributions of events with a vertex within the FV split up into CCQE (red), CCRES (yellow), and CCDIS (green) before selection. The upper right side is the $\cos(\theta)$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

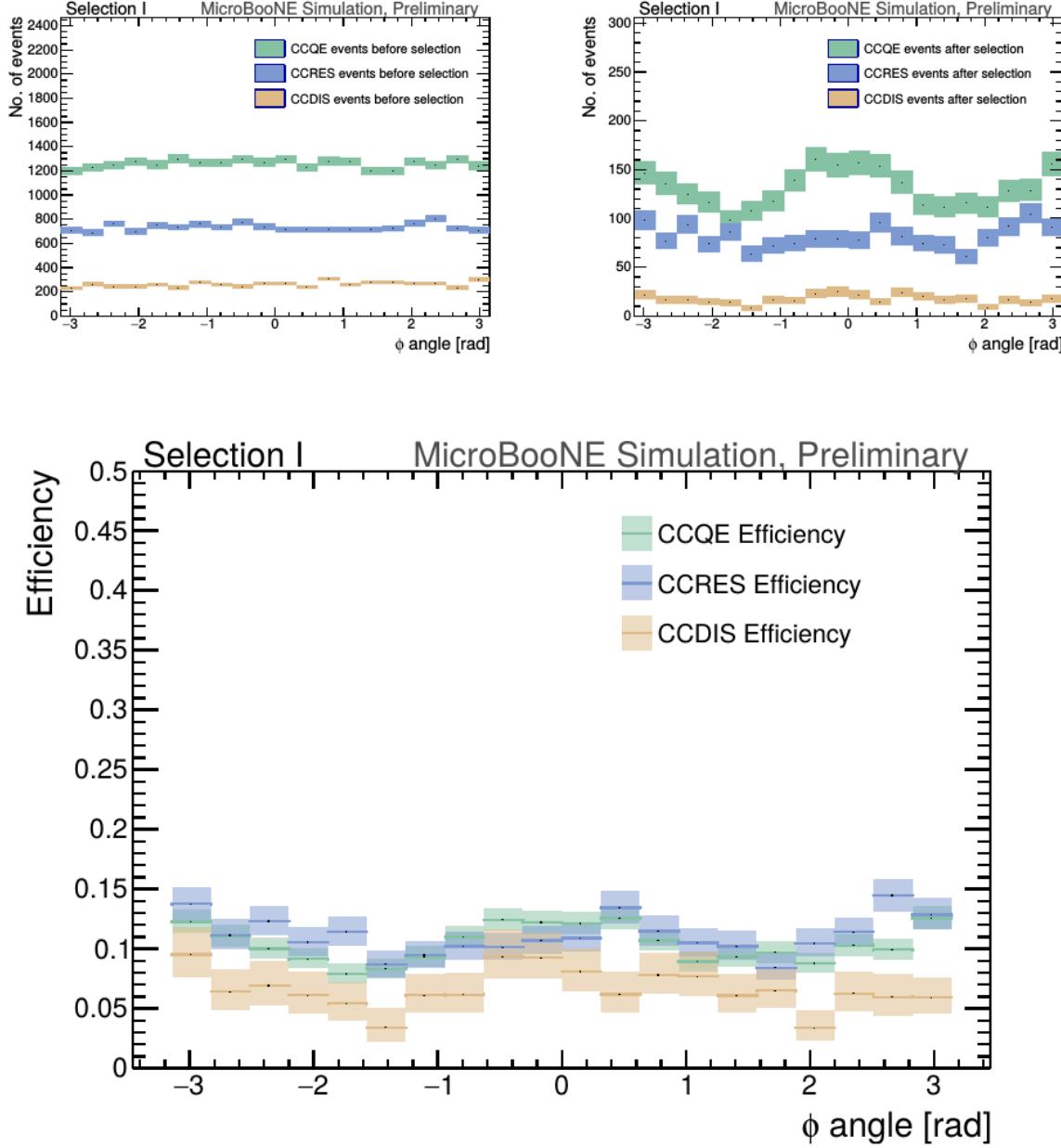


Figure 5.11: MC ϕ distributions of the muon originating from a ν_μ CC interaction. Upper left is the ϕ distributions of events with a vertex within the FV split up into CCQE (red), CCRES (yellow), and CCDIS (green) before selection. The upper right side is the ϕ distribution after the selection with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

1308 **Chapter 6**

1309 **Background on Convolutional Neural
1310 Networks**

1311 Convolutional neural networks (CNNs) have been one of the most influential inno-
1312 vations in the field of computer vision. Neural networks became popular in 2012
1313 when Alex Krizhevsky used them to win that year's ImageNet competition [?] by
1314 dropping the error from 26% to 15%. Since then, many companies are using deep
1315 learning including Facebook's tagging algorithms, Google for their photo search and
1316 Amazon for product recommendations. For the purpose of this thesis CNNs were
1317 used for image classification, specifically, images of varying particles created using
1318 LArTPC data.

1319 **6.1 Image Classification**

1320 Image classification is the process of inputting an image into the CNN and receiving a
1321 probability of classes that best describes what is happening in the image. As humans,
1322 image classification is something that is learned at a very young age and is easy to
1323 do without much effort. This is also apparent when hand-scanning LArTPC images.
1324 After learning what a neutrino event looks like in MicroBooNE, it is relatively easy
1325 to recognize simple neutrino events from cosmic ray background as well as highly
1326 ionizing particles like protons from minimum ionizing particles like muons. The very
1327 detailed images LArTPC detectors output are prime candidates for input images into
1328 a CNN. CNNs mimic a human's ability to classify objects by creating an architecture
1329 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.

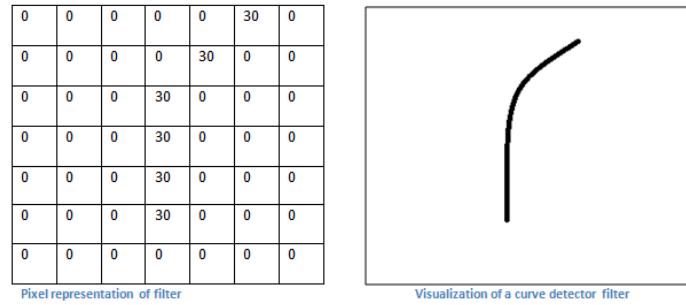


Figure 6.1: Pixel representation and visualization of a curve detector filter. As you can see, in the pixel representation, the weights of this filter are greater along a curve we are trying to find in the input image

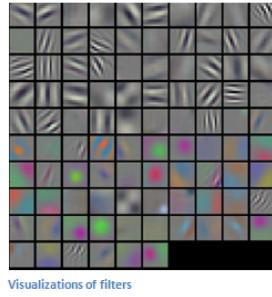


Figure 6.2: Visualization of filters found in first layer of a CNN.

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.

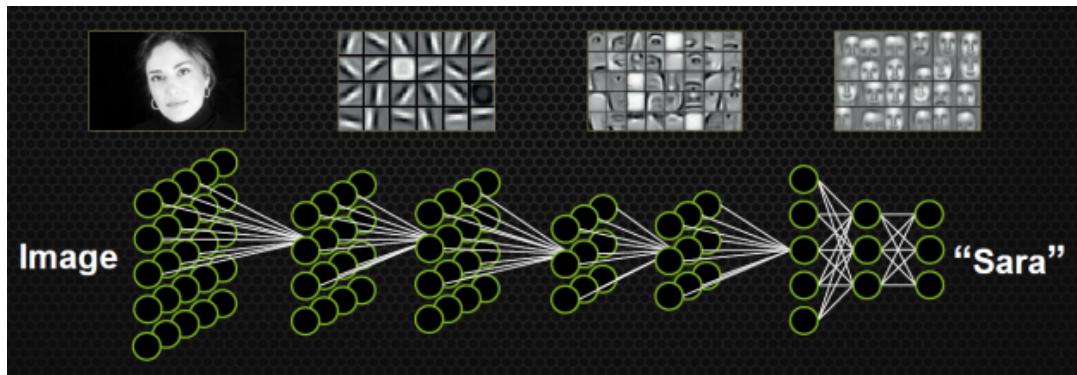


Figure 6.4: Pictorial Representation of Convolutional Neural Networks as well as a visual representation on CNN's complexity of layer feature extraction

1379 There are other layers in a CNN architecture that will not be covered in the scope
 1380 of this thesis but in a general sense, these layers are interspersed between convolution
 1381 layers to preserve dimensionality and control overfitting of the network. The last layer
 1382 is called a fully connected layer and it's job is to output an N dimensional vector where
 1383 N is the number of classes the network has been trained on. Each number in this vector
 1384 represents the probability that the input image is a certain class. Fully connected layers
 1385 use the feature maps of the high level features to compute the products between the
 1386 weights of the previous layer to get the probabilities of each class. These weights are
 1387 then adjusted through the training process using backpropagation.

1388 6.2.1 Backpropagation

1389 A CNN at it's onset has weights that are randomized. The filters themselves don't
 1390 know how to pull out identifying information per class. For a neural network to learn,
 1391 it must be trained on a training set that is labeled. Backpropagation has four seperate
 1392 steps: foward pass, loss function, backward pass and updating weights. In the forward
 1393 pass, a training image is passed through the whole network. All of our weights at this
 1394 time are randomly initialized so the output for the first image will have no preference
 1395 to a specific class. A common loss function is mean squared error (MSE):

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

1396 If we assume that the MSE is the loss of our CNN, the goal would be that our
 1397 predicted label (output of CNN) is the same as our training label. To do this, we need
 1398 to minimize the loss function. To do this, it is necessary to find out which weights most
 1399 directly affect the loss of the network i.e $\frac{dL}{dW}$ where L is our loss function and W are
 1400 the weights of a specific layer. The next step is the backward pass which determines
 1401 which weights contribute the most to the loss and finds ways to adjust these weights
 1402 so that the loss decreases. After the derivative is computed, the last step updates the
 1403 weights in the opposite direction of the gradient.

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

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

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

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

1404 The learning rate is a parameter given to the CNN and it describes the steps the
 1405 network takes to update the weights. Higher learning rate equals large steps and a
 1406 lower training time, but a learning rate that is too large can mean the CNN never
 1407 converges.

1408 Going through backpropagation consists of one training iteration. Once the net-
 1409 work completes a specific number of iterations, another parameter given, and runs
 1410 over all training images that are split up into batches, the process is considered com-
 1411 plete. User input parameters, called hyperparameters, help the network converge to

1412 optimal weights for each layer. Batch size, learning rate, and training iteration are just
1413 some of the user input hyperparameters that help. Lastly, to check if the network has
1414 learned, a different set of labeled images are fed to the CNN iteratively through the
1415 training process to see how well it's learning. This process is especially important to
1416 make sure the network architecture isn't being affected by overfitting (memorizing
1417 training input rather than learning).

1418 6.3 Choosing Hyperparameters

1419 Convolutional neural networks are a relatively new tools in computer vision. Choosing
1420 hyperparameters for your specific dataset is a non-trivial task. Hyperparameters can
1421 range from the amount of layers and filters per layer in an CNN architecture to the
1422 stride the receptive field of a filter takes, not to mention training hyperparameters
1423 such as learning rate and batch size described above. They're ways to optimize these
1424 hyperparameters via hyperparameter optimization using Bayesian Optimization [?]
1425 but as you can imagine, optimizing an CNN architecture from scratch can be very
1426 computationally intensive. For the purpose of this thesis, two well known CNN
1427 architectures were used, AlexNet [?], which won the ImageNet Large-Scale Visual
1428 Recognition Challenge (ILSVRC) in 2012 and therefore bringing awareness of CNNs,
1429 and GoogleNet [?], which won the ILSVRC in 2014, giving rise to deep networks. Both
1430 AlexNet and GoogleNet architectures were used to train on LArTPC images and their
1431 low level filter weights. Higher level filter weights were randomly initialized before
1432 training so the network can learn high level features of LArTPC image classes. The
1433 AlexNet architecture is shown in figure 6.5 and the GoogleNet architecture is shown
1434 in figure 6.6

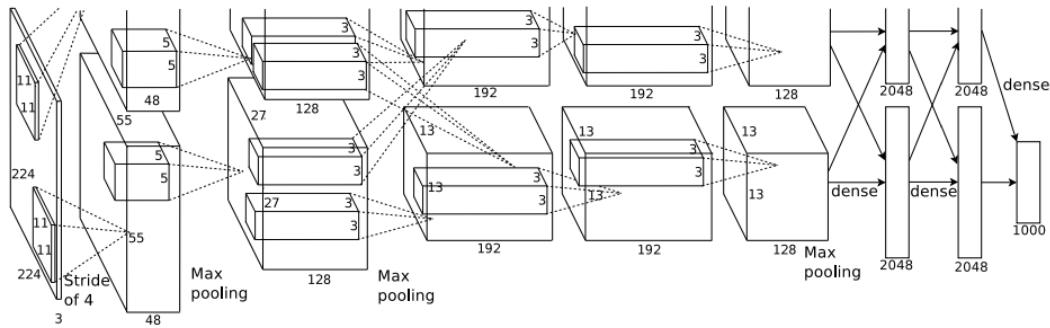


Figure 6.5: Pictoral representation of the AlexNet model. The AlexNet model consists of 5 convolution layers and 3 fully connected layers.

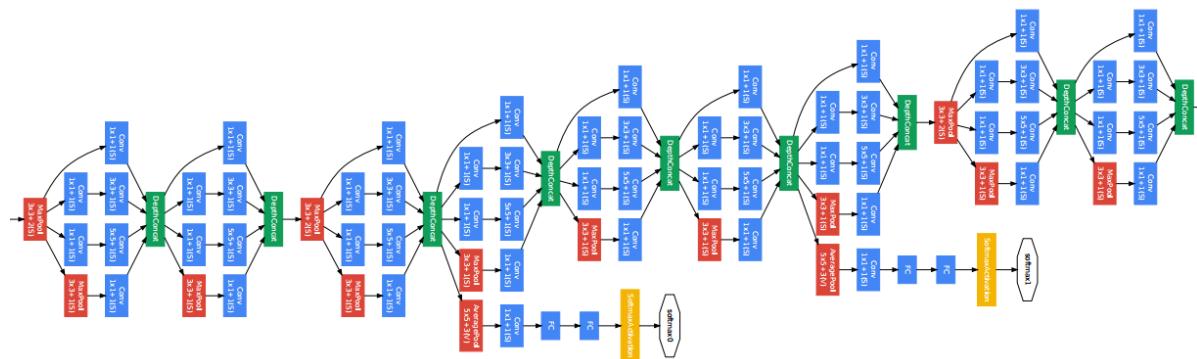


Figure 6.6: Pictoral representation of the GoogleNet model. The GoogleNet model consists of 22 layers. The model implements 9 Inception modules which performs covolution and pooling in parallel and strays away from the basis that CNN layers need to be stacked up sequentially. The GoogleNet model also 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.

¹⁴³⁵ Chapter 7

¹⁴³⁶ Training process of Convolutional ¹⁴³⁷ Neural Networks

¹⁴³⁸ Three Convolutional Neural Networks (CNNs) were trained throughout this analysis.
¹⁴³⁹ There are differences to each CNN and will be described fully in the next sections but
¹⁴⁴⁰ the main difference are the amount of particle images used for training and validation.
¹⁴⁴¹ CNN1075 used 1,075 muons and 1,075 pions for training and the same amount of each
¹⁴⁴² particle for validation. CNN10000 used 10,000 muons and 10,000 pions split in half
¹⁴⁴³ for testing and training. Lastly CNN100000 had muons, pions, protons, electrons,
¹⁴⁴⁴ and gammas in its training and validation set. Each particle had 20,000 images and
¹⁴⁴⁵ training and validation was split 90% training, 10% validation. This chapter will also
¹⁴⁴⁶ describe the different hardware frameworks used for training beginning on a CPU
¹⁴⁴⁷ and ending on a GPU cluster.

¹⁴⁴⁸ 7.1 Convolutional Neural Network Training

¹⁴⁴⁹ The first training iteration, CNN1075, was a proof of concept. This CNN was trained
¹⁴⁵⁰ on my local machine for ~4-5 weeks. The batch size had to be very small as well as the
¹⁴⁵¹ image size due to the lack of computation resources. The second iteration of training,
¹⁴⁵² CNN10000, was trained on a Fermilab stationed Syracuse University machine. This
¹⁴⁵³ machine had 6 TB of disk space, 6 cores at 2.1 GHz and 32 GB of RAM. The use of
¹⁴⁵⁴ this machine allowed me to increase the training sample as well as the batch size and
¹⁴⁵⁵ hence further increase the accuracy of the neural network. Lastly, the CNN100000 was
¹⁴⁵⁶ trained using two GTX 1080 Ti GPUs with 11GB of memory on a node on the Syracuse

¹⁴⁵⁷ University GPU cluster, SUrge, that has 8 cores and 16GB of memory. This increase in
¹⁴⁵⁸ memory as well as the capability to use 2 GPUs drastically cut down on training time
¹⁴⁵⁹ from \sim 4-5 weeks to \sim 8 hours. SUrge also allowed for hyperparameter optimization
¹⁴⁶⁰ by being able to run multiple training iterations over the two GPUs. Lastly, SUrge
¹⁴⁶¹ allowed for training over higher resolution images and a larger particle class of 5
¹⁴⁶² particles vs 2 particles.

¹⁴⁶³ **7.1.1 Creating images using LArTPC data for training/validation of**
¹⁴⁶⁴ **CNNs**

¹⁴⁶⁵ The μ/π image dataset used to train and validate CNN1075 was created using single
¹⁴⁶⁶ generated isotropic muons and pions from 0-2 GeV energy range. 2,150 muons and
¹⁴⁶⁷ 2,150 pions were used for training and testing split 50%. The images were created
¹⁴⁶⁸ using LArSoft, a liquid argon software, and were based on wire number and time
¹⁴⁶⁹ tick in the collection plane. Uboonecode reconstruction version v05_08_00 was used.
¹⁴⁷⁰ The raw ADC value after noise filtering was the wire signal. Each collection plane
¹⁴⁷¹ greyscale image was 3456x1600x1 where 6 time ticks were pooled into 1 bin. After
¹⁴⁷² the image was created, the region of interest (ROI) in the image was found by using
¹⁴⁷³ Open CV, a image processing open source software package, to scan the image starting
¹⁴⁷⁴ from the edges and stopping once a bright pixel is encountered. At this step, the XY
¹⁴⁷⁵ ROI can be larger or smaller than the necessary size of a training image and the XY
¹⁴⁷⁶ ratio of the image is not kept. This ROI is then resized to an image of 224x224x1. The
¹⁴⁷⁷ greyscale color standard is 8bit therefore the ADC value of wire and time tick was also
¹⁴⁷⁸ downsampled due to the 12bit ADC value MicroBooNE has. To do this, the highest
¹⁴⁷⁹ ADC pixel in the image was found and then this was divided by the rest placing all
¹⁴⁸⁰ pixel values between 0-1. From there, all pixel values are then multiplied by 255.

¹⁴⁸¹ The μ/π image dataset used to train and validate the CNN10000 was also created
¹⁴⁸² using single generated isotropic muons and pions from 0-2 GeV energy range. 10,000
¹⁴⁸³ muons and 10,000 pions were used for training and testing split 50%. Uboonecode
¹⁴⁸⁴ v06_23_00 was used instead of v05_08_00. Each collection plane greyscale image was
¹⁴⁸⁵ 3456x1280x1 where 5 time ticks were pooled into 1 bin which is different than the
¹⁴⁸⁶ previous dataset and was implemented due to the fact that the time ticks of an event
¹⁴⁸⁷ went from 9400 to 6400 with the change of uboonecode version. Issues that arose in
¹⁴⁸⁸ CNN1075 that were fixed in CNN10000 include zero-padding images in X and Y that

1489 are smaller than 224X224 to eliminate over-zooming effect and fixing a bug that shifted
1490 pixels separated by a dead-wire region.

1491 The $\mu/\pi/p/e/\gamma$ image dataset used to train and validate the CNN100000 were
1492 created using single generated isotropic particles with energy range from 0-2 GeV.
1493 20,000 of each particle were used for training and were split 90/10 between training
1494 and testing sets. Uboonecode v06_23_00 was used for these images. The collection
1495 plane greyscale iamge had the same dimensions as CNN10000, 3456x1280x1 and
1496 the ROI algorithm was the same except for resizing these images to 576x576. A
1497 major change other than the higher resolution images was the treatment of the ADC
1498 values. In the first two image making schemes, the highest pixel value was found
1499 per image and the image was then normalized by that. The issue arising from this
1500 ADC normalization wasn't inherent in μ/π training due to the fact that both particles
1501 are minimum ionizing particles in liquid argon, however, when dealing with a larger
1502 particle class, it was necessary to try and make sure energy deposition by each particle
1503 was preserved. The energy deposition in a particle image corresponds to the ADC
1504 value or pixel brightness. To preserve energy deposition, the ADC float value was
1505 passed straight to the image rather than doing any image normalization. This then
1506 makes sure that minimum ionizing particles like muons and pions appear dimmer
1507 than highly ionizing particles like protons.

1508 Images were also made from BNB+Cosmic events that passed the cc-inclusive
1509 selection 1 filter right before the 75 cm track length cut and were classified using
1510 the CNN10000. The dataset used to create these images is the same one used in [?],
1511 *prodgenie_bnb_nu_cosmic_uboone_mcc7_reco2*. These images were created using infor-
1512 mation from the track candidate that passed the filter. Only wire number and time
1513 ticks associated to the track candidate were drawn on the image to mimic a single
1514 particle generated image. These images were then classified using CNN10000. Two
1515 approaches were taken in making these images. The first was using the image nor-
1516 malization above where the maximum pixel in each image is used as a normalization
1517 constant to get all pixels between 0-1 then multiply all pixels by 255. As described
1518 above, this is the incorrect way to normalize. The second way the images were created
1519 was by passing the ADC float to the image. The results of CNN10000 performance are
1520 shown in section 7.1.

1521 Lastly, multiple BNB+Cosmic images per event were made for CNN100000 by
1522 reducing many of selection I cuts to try and let the CNN do particle as well as event

1523 identification. This image making scheme used for CNN100000 will be described in
1524 more detail in later sections.

1525 7.1.2 Training CNN1075

1526 The results of CNN1075 are described in this section. The accuracy is how well
1527 CNN1075 is doing by epoch and was 74.5%. The loss is gradient descent or mini-
1528 mization of the error of the weights and biases used in each neuron of each layer of
1529 CNN1075 and was 58% with a trend sloping downwards on the loss curve as well as a
1530 trend sloping upward in the accuracy curve. The accuracy and loss of CNN1075 are
1531 shown in figure 7.1. Due to the depth of the neural network framework, it was neces-
1532 sary to train with a larger dataset and for more epochs, however, the downward slope
1533 of the loss curve is an indication that once trained for longer with a higher training
1534 sample, neural networks can be used for μ/π separation. The hyperparameters used
1535 to train CNN1075 are detailed below:

- 1536 • train_batch_size: 50
- 1537 • test_batch_size: 50
- 1538 • test_iter: 50
- 1539 • test_interval: 50
- 1540 • base_lr: 0.01
- 1541 • lr_policy: "step"
- 1542 • gamma: 0.1
- 1543 • stepsize: 200
- 1544 • display: 50
- 1545 • max_iter: 5000
- 1546 • momentum: 0.9
- 1547 • weight_decay: 0.0005
- 1548 • snapshot: 100

Loss/Accuracy of CNN trained on 2150 images

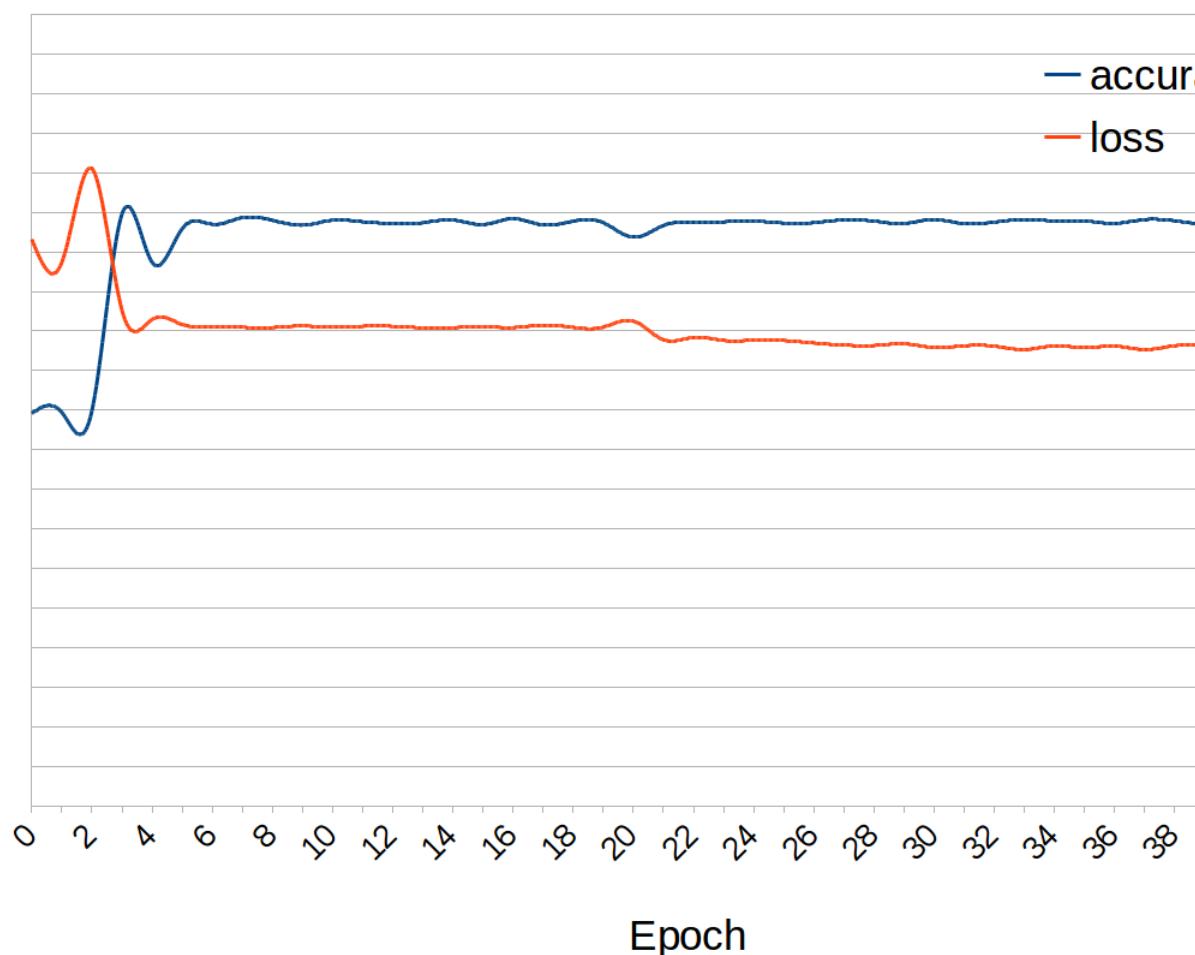


Figure 7.1: Accuracy vs. Loss of ImageNet 2-output μ/π sample consisting of 2,150 images each.

1549 The confusion matrices shown in figure 7.2 show the accuracy for both the training
 1550 and testing datasets. The fact that these two have similar accuracies is important
 1551 because if the training dataset had a much higher accuracy, that indicates an over-
 1552 training of the training sample which means the neural network didn't learn features
 1553 to separate muons from pions, it just memorized what was in the training dataset.
 1554 Also note that the neural network does a better job of identifying muons than pions.
 1555 This can be attributed to the more complex event scenes pions tend to leave in the
 1556 detector due to pion interacting more in LAr than muons do. The CNN may do better
 at identifying pions with a larger training sample.

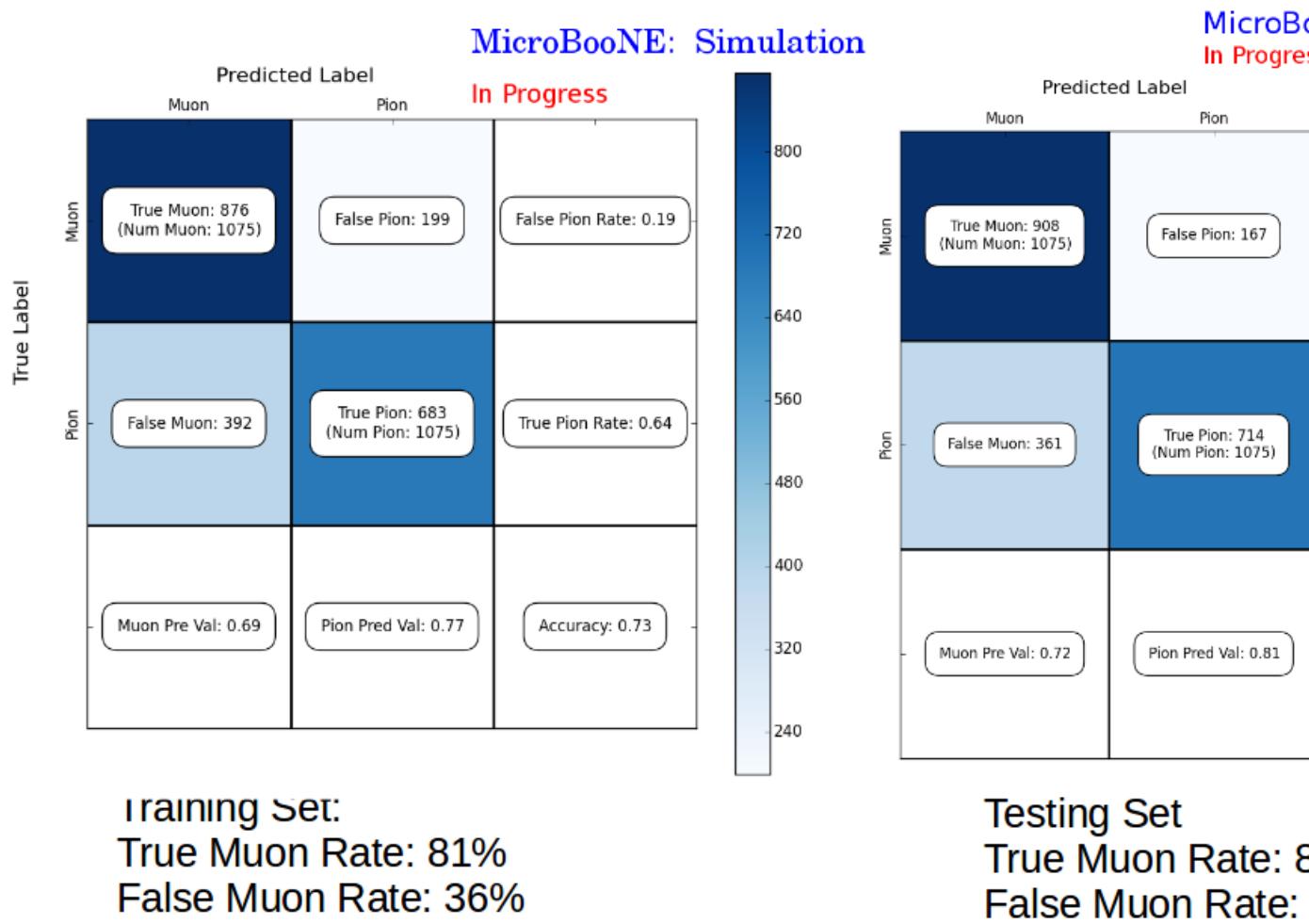


Figure 7.2: Confusion Matrices showing accuracy of CNN1075 on testing and training MC data.

1558 7.1.3 Training CNN10000

1559 The hyperparameters used for CNN10000 are shown below. The batch size for the
1560 training and testing as well as the test iter were chosen to encompass the whole
1561 training/testing image set when doing accuracy/loss calculations. To do this, multi-
1562 plying the test iter by the test batch size gives you the amount of images used when
1563 calculating accuracy/loss curves.

```
1564     • train_batch_size: 100  
1565     • test_batch_size: 100  
1566     • test_iter: 100  
1567     • test_interval: 100  
1568     • base_lr: 0.001  
1569     • lr_policy: "step"  
1570     • gamma: 0.1  
1571     • stepsize: 1000  
1572     • display: 100  
1573     • max_iter: 10000  
1574     • momentum: 0.99  
1575     • weight_decay: 0.0005  
1576     • snapshot: 100
```

1577 The same architecture that was used to train CNN1075 was employed on CNN10000,
1578 Imagenet. Caffe [?] was the software package used for both CNNs. The differences
1579 include batch size and test_iter and momentum to account for the larger dataset.
1580 Figure 7.3 shows the loss and accuracy of CNN10000. There is around a 10% increase
1581 in accuracy from CNN1075 to CNN10000, 85%, and around a 20% decrease in loss,
1582 36%.

1583 Figure 7.4 show a breakdown of μ/π separation for CNN10000. It also shows
1584 the network is not being overtrained due to the Accuracy of both the training and
1585 testing datasets being within .01% of each other. The CNN is doing a very good job of

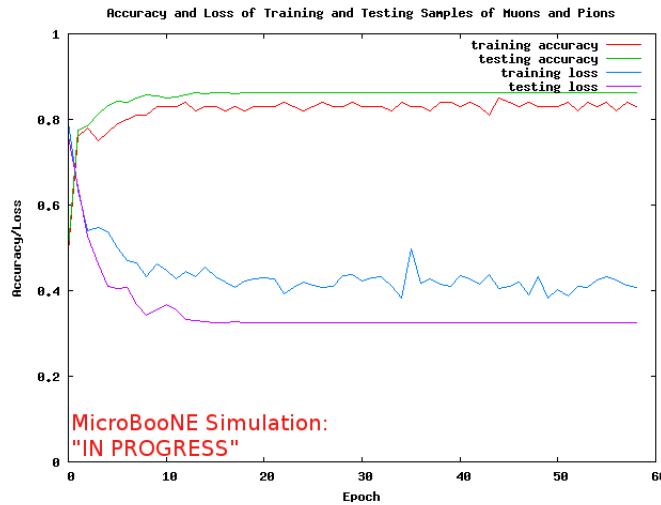
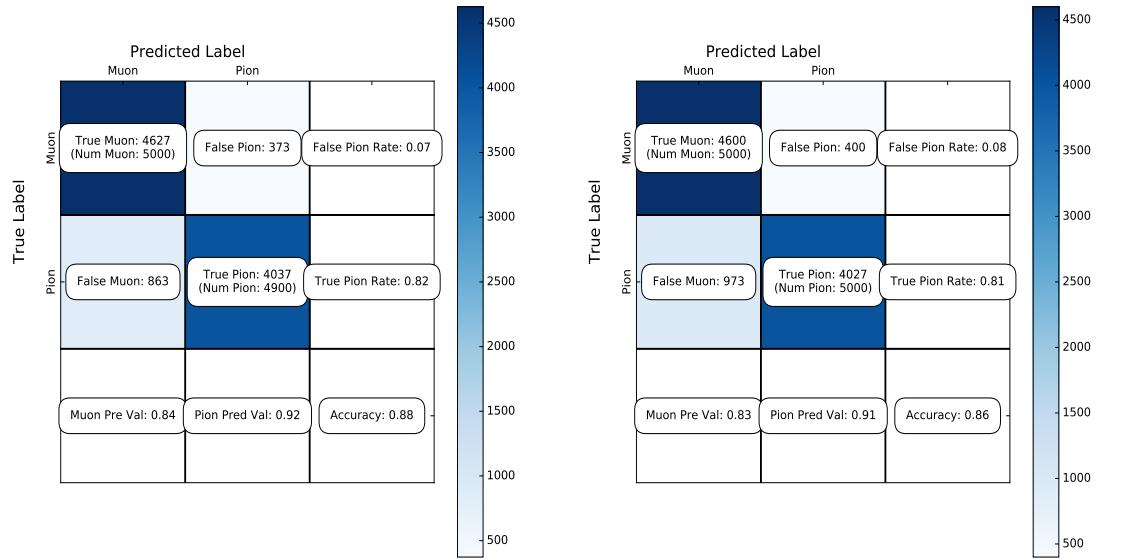


Figure 7.3: Accuracy vs. Loss of ImageNet 2-output μ/π sample consisting of 10,000 images each.

1586 classifying true muons as muons, and our loss increase from CNN1075 is due to the
 1587 increase in accurately classifying pions as pions.

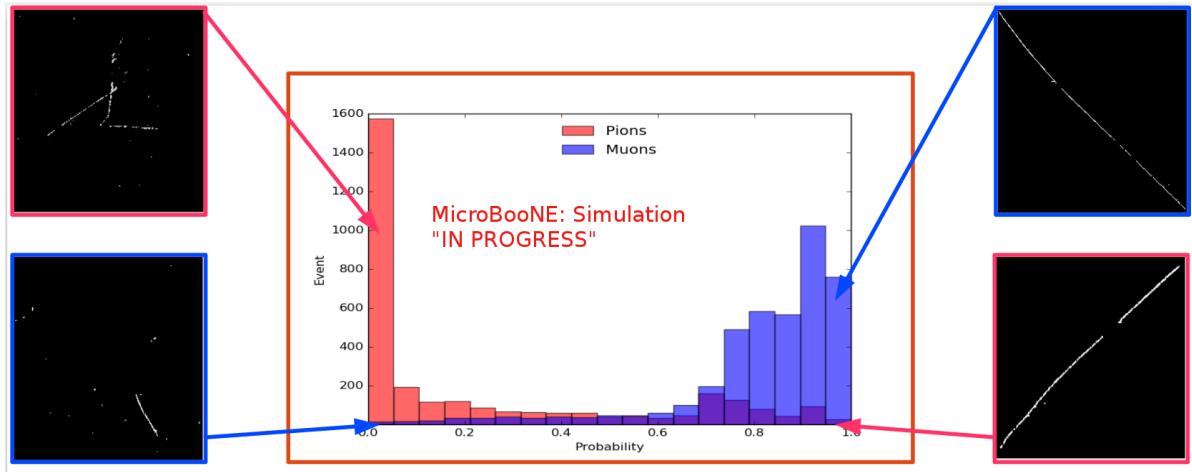
1588 7.1.4 Training CNN100000

1589 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 MC data

(b) Confusion Matrix showing Accuracy of CNN using testing MC 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 7.4: Description of confusion matrix variables: False pion rate = $false\pi / total\pi$ True pion rate = $true\pi / total\pi$ Accuracy = $(true\pi rate + true\mu rate) / 2$ Pion prediction value = $true\pi / (true\pi + false\pi)$ Muon prediction value = $true\mu / (true\mu + false\mu)$

7.4c The probability plot includes muons and pions that are classified as primary particles.

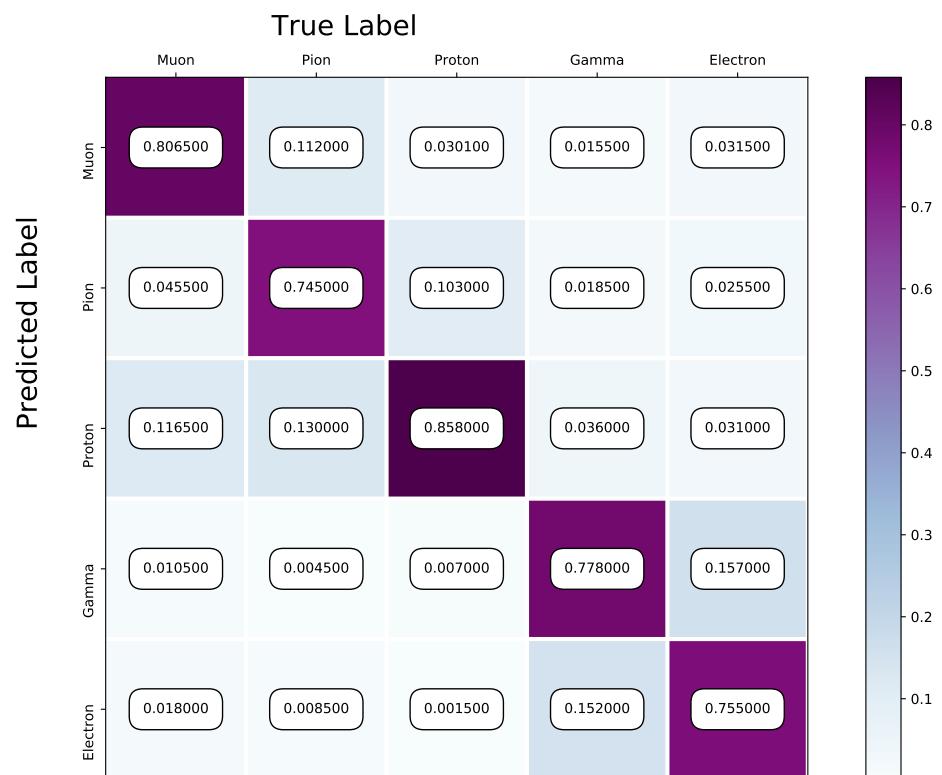


Figure 7.5: Confusion Matrix of all five particles

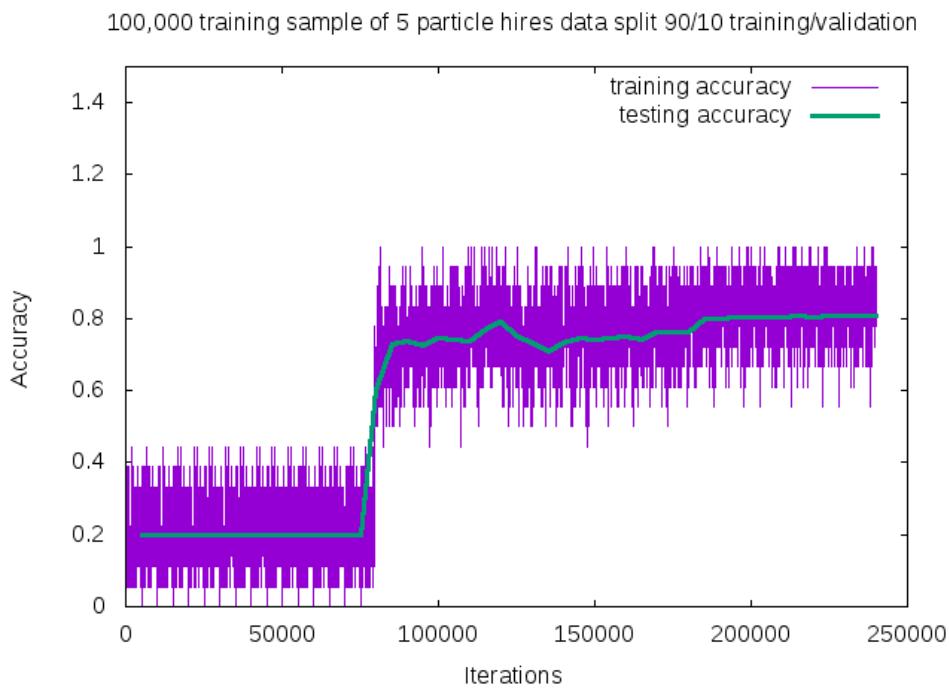


Figure 7.6: Training and testing accuracy of CNN trained on 100,000 images of $\mu/\pi/p/\gamma/e$ with 20,000 images of each particle. Each image was a size of 576x576 and the images per particle were split 90% use for training and 10% used for testing the network

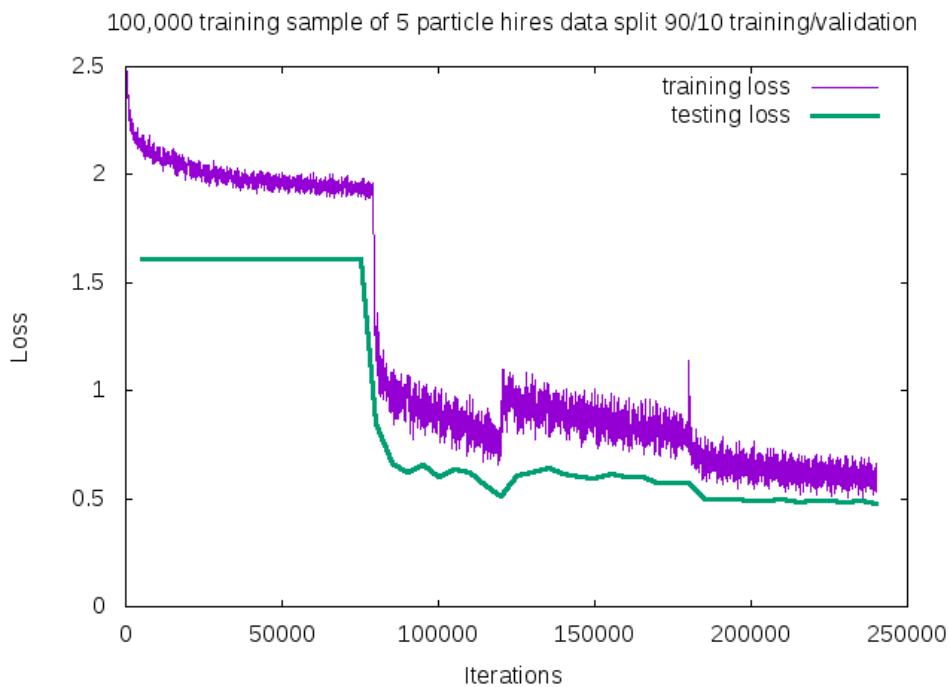


Figure 7.7: Training and testing loss of CNN trained on 100,000 images of $\mu/\pi/p/\gamma/e$

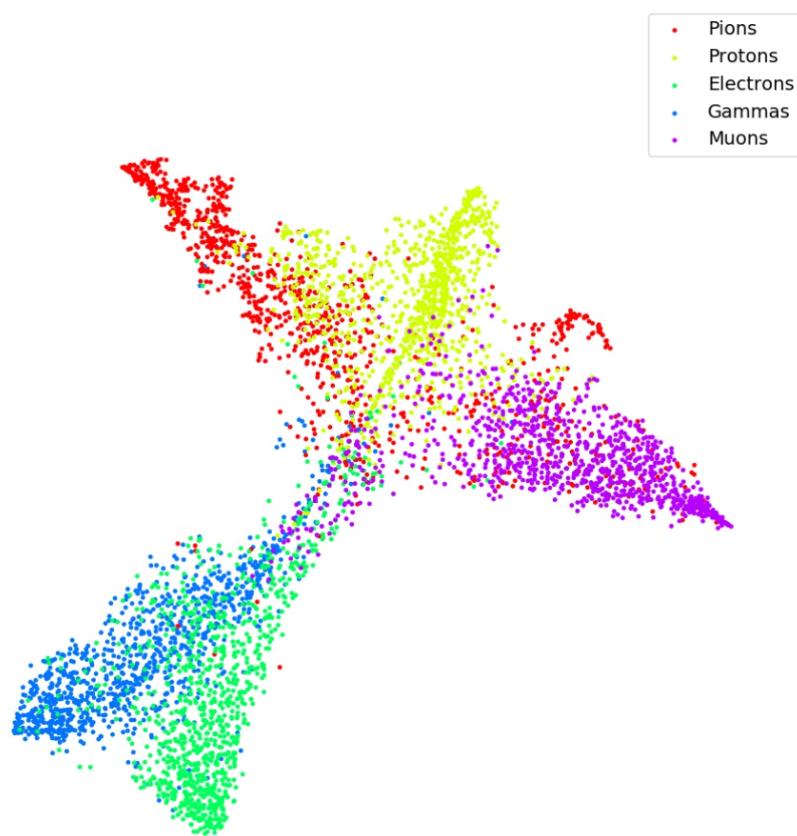


Figure 7.8: t-SNE of CNN

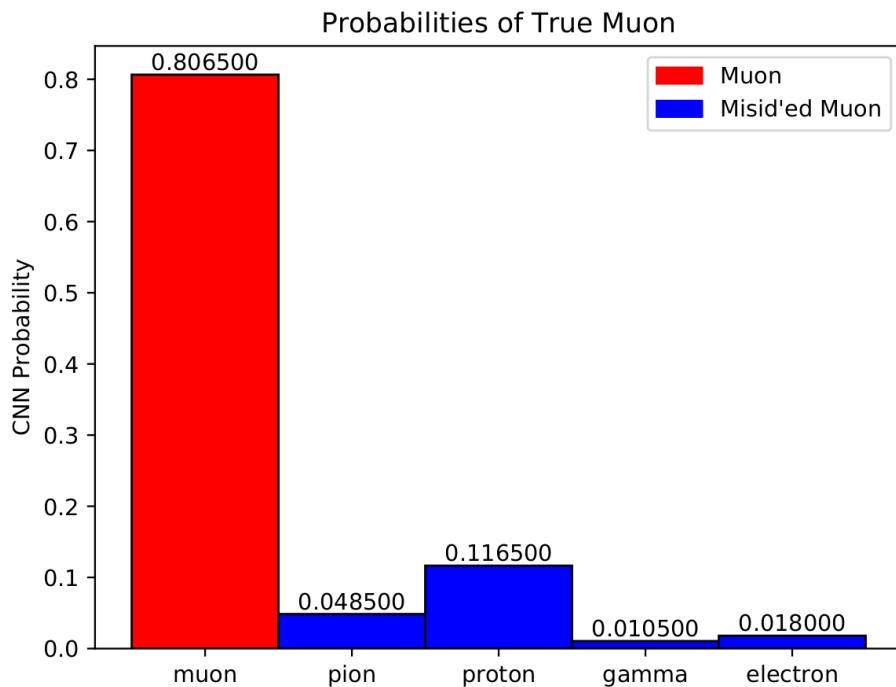


Figure 7.9: Muon Prob

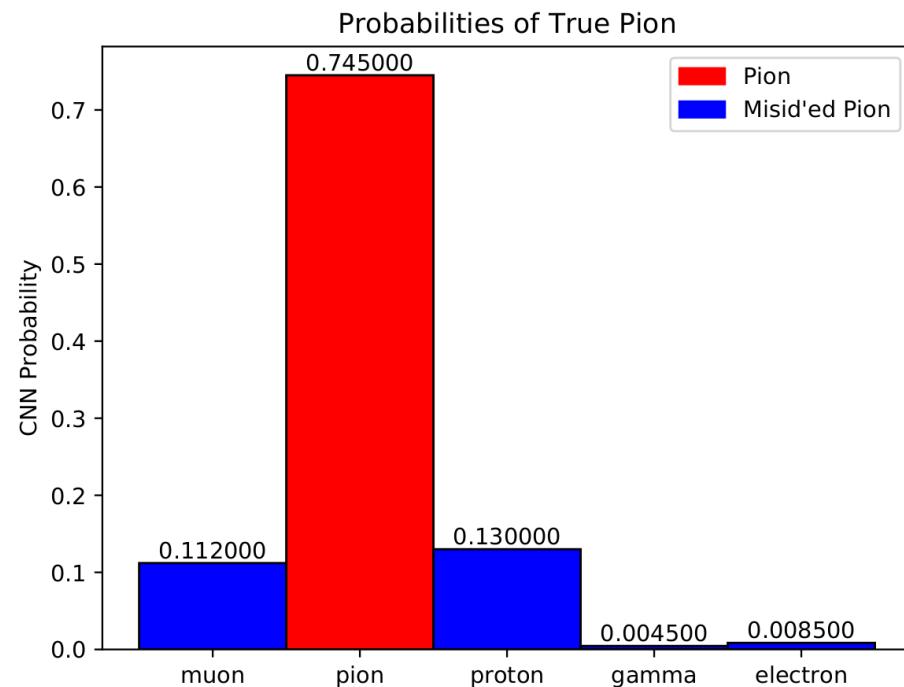


Figure 7.10: Pion Prob

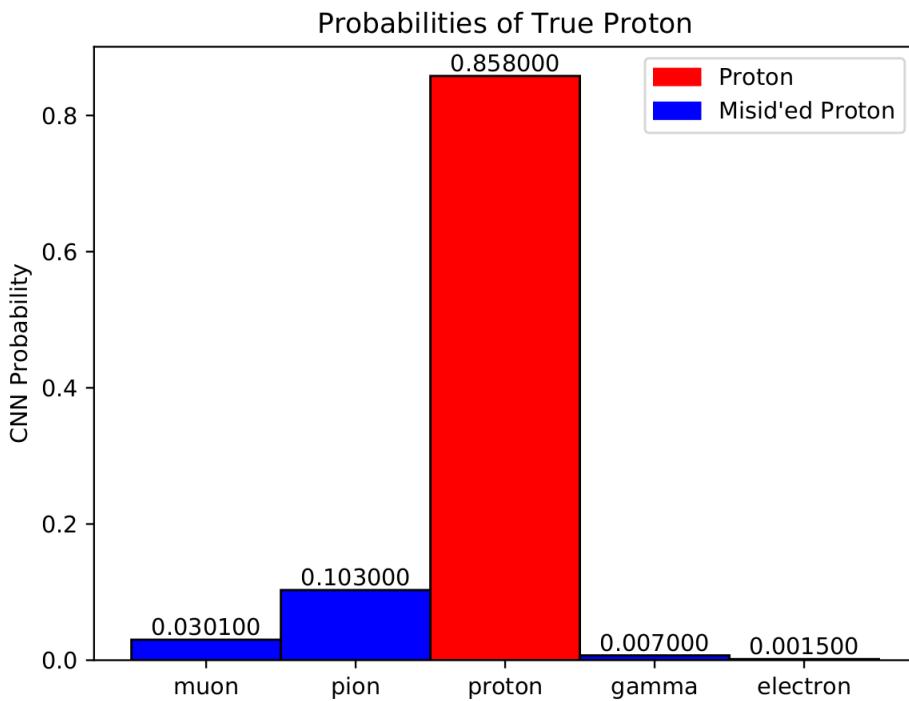


Figure 7.11: Proton Prob

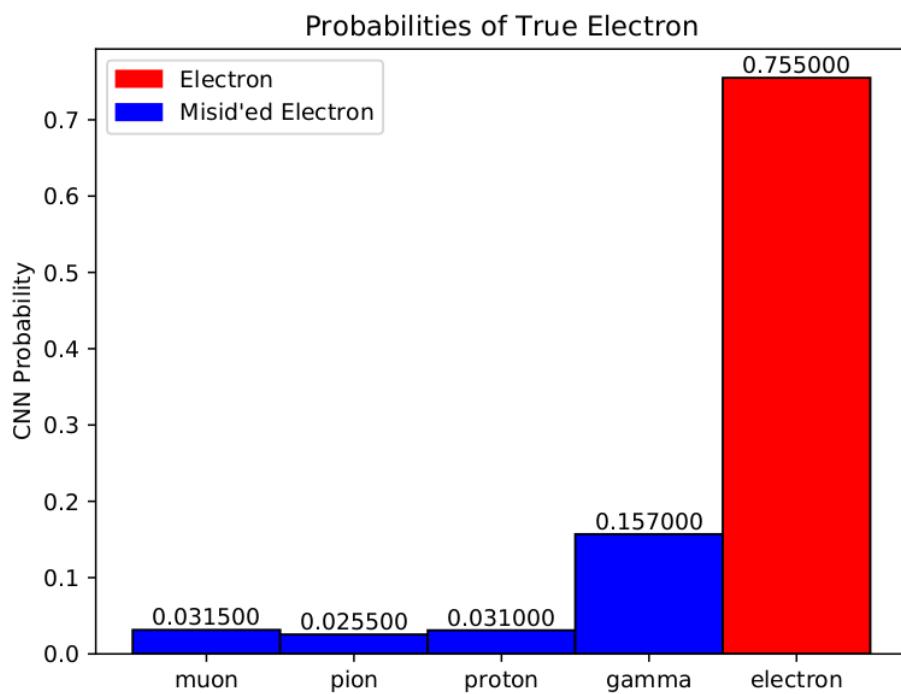
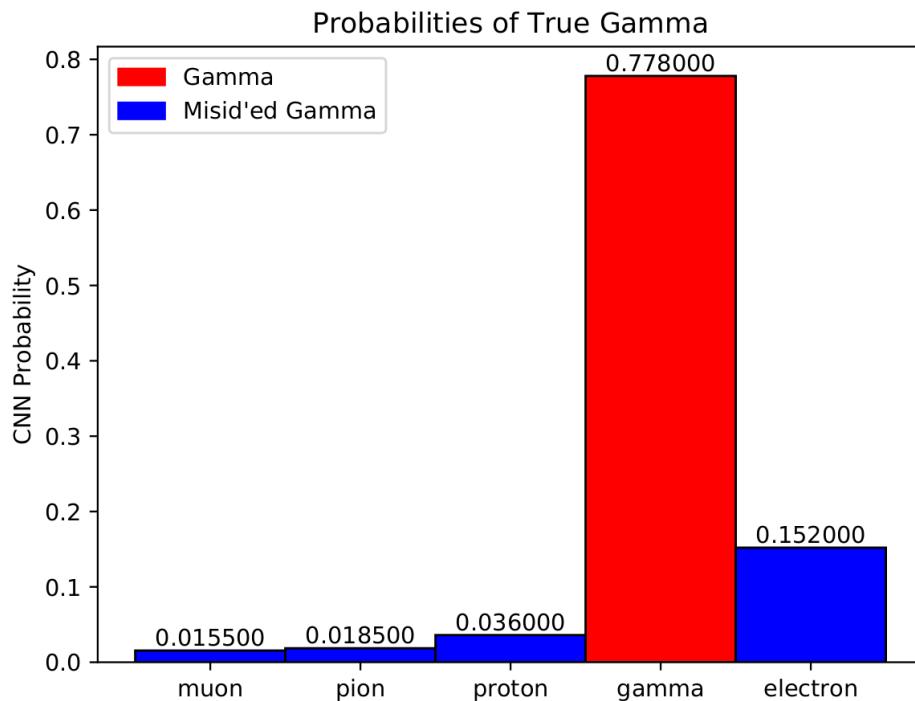
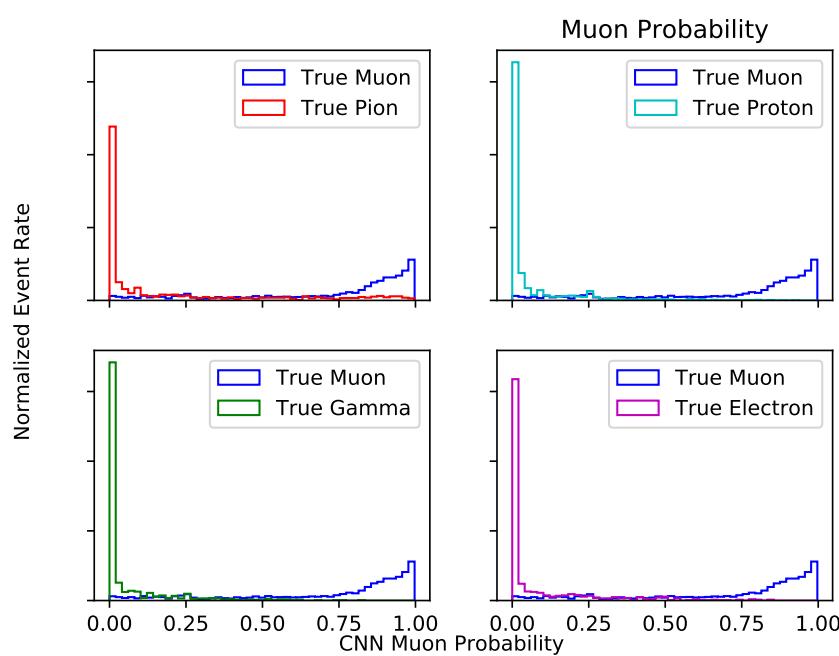
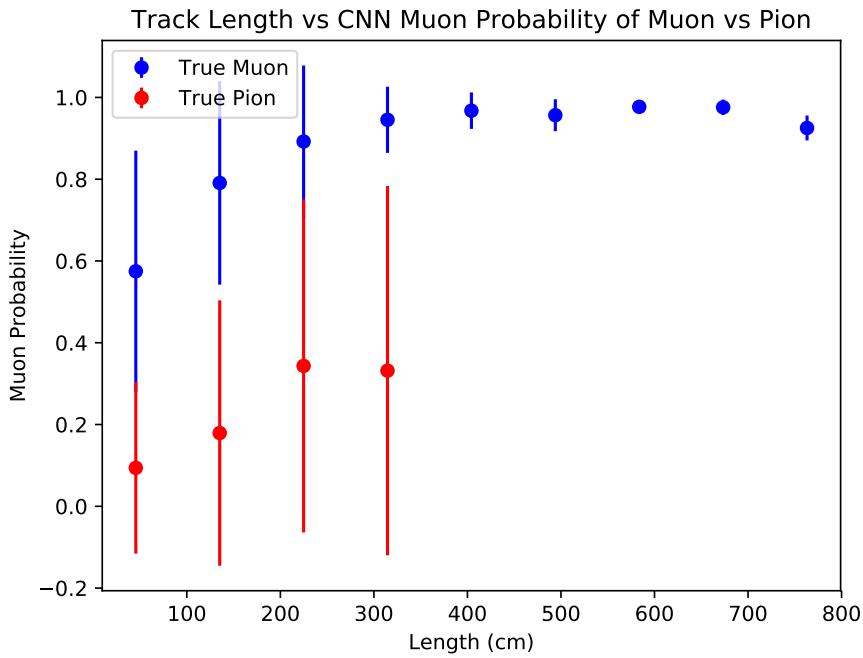
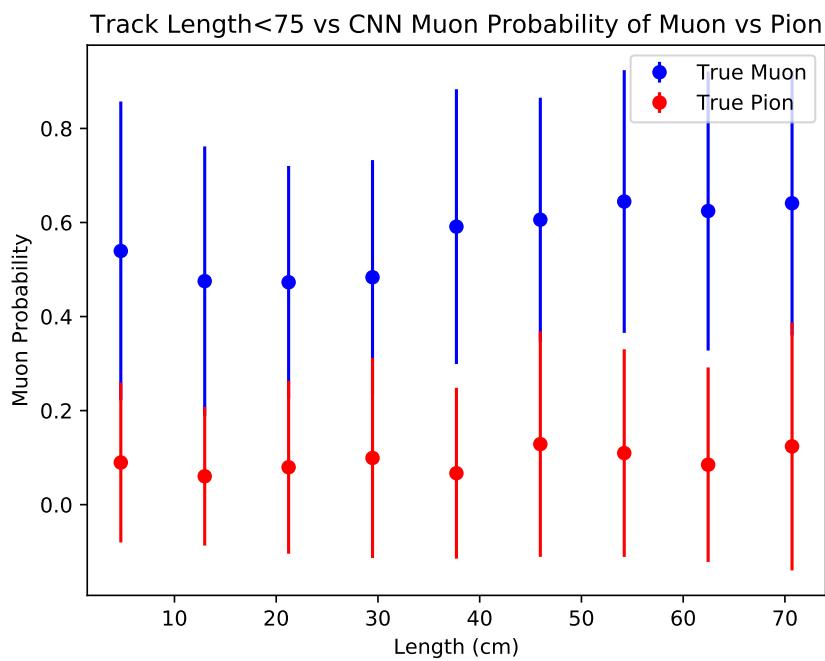
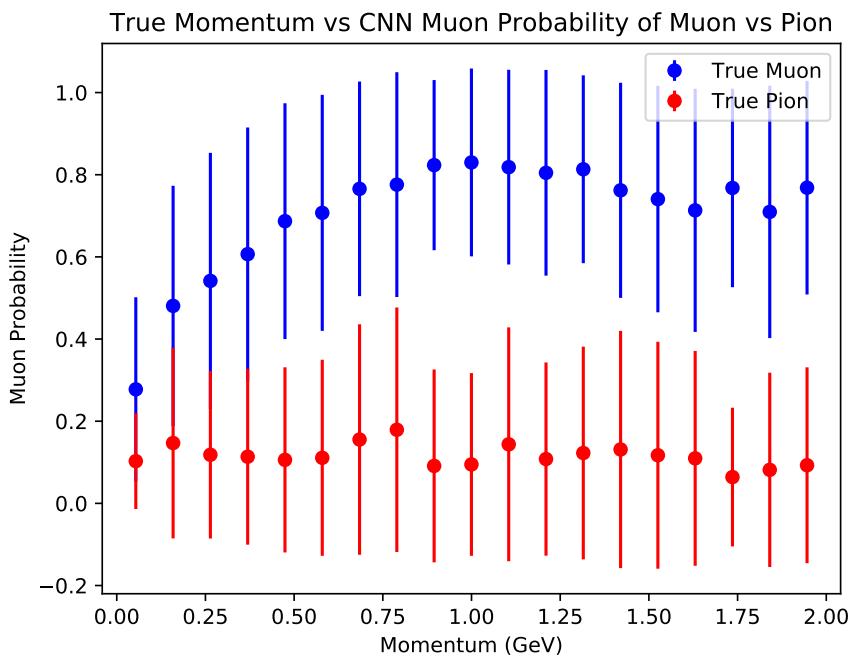
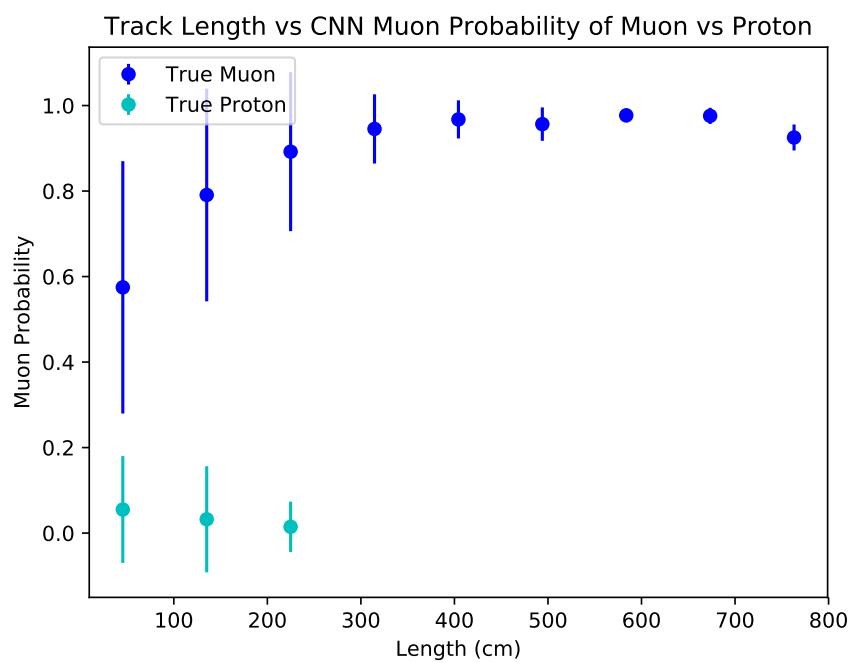


Figure 7.12: Electron Prob

**Figure 7.13:** Gamma Prob**Figure 7.14:** Prob

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

**Figure 7.17:** mupi**Figure 7.18:** mup

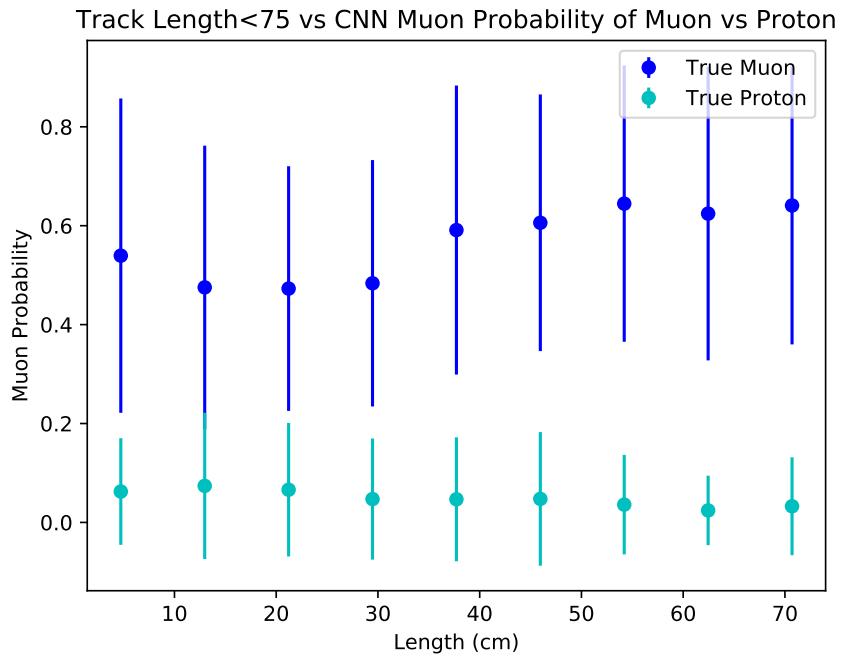


Figure 7.19: mup

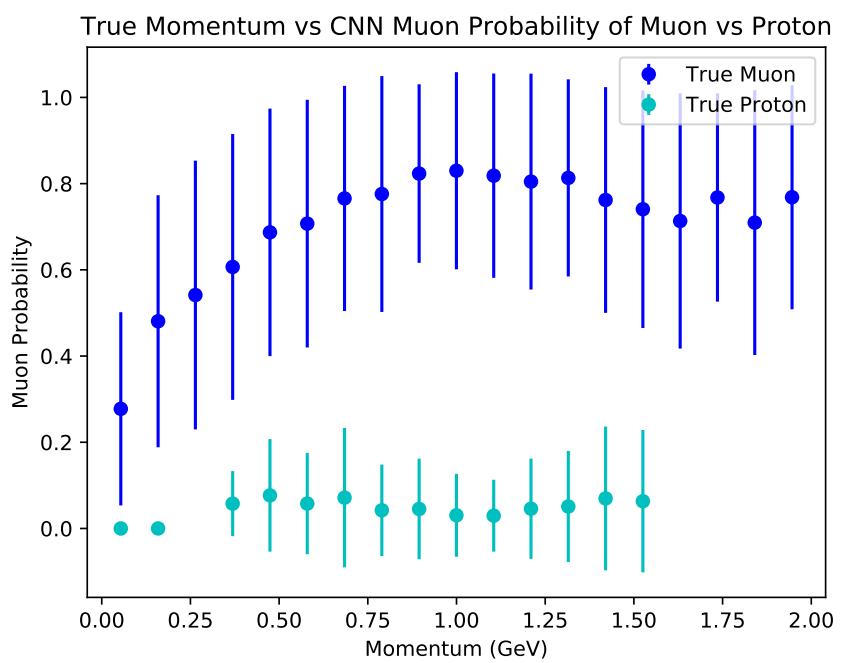


Figure 7.20: mup

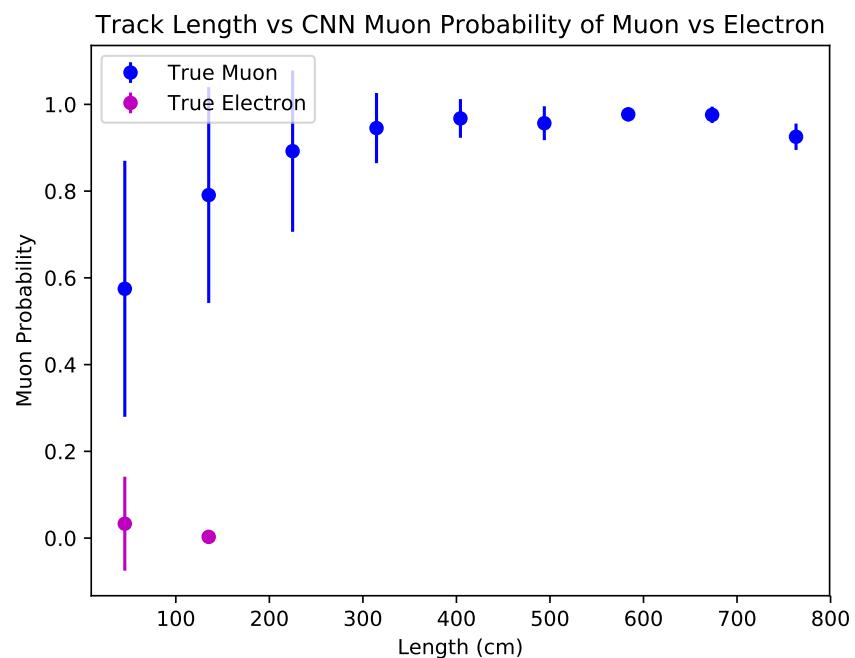


Figure 7.21: mue

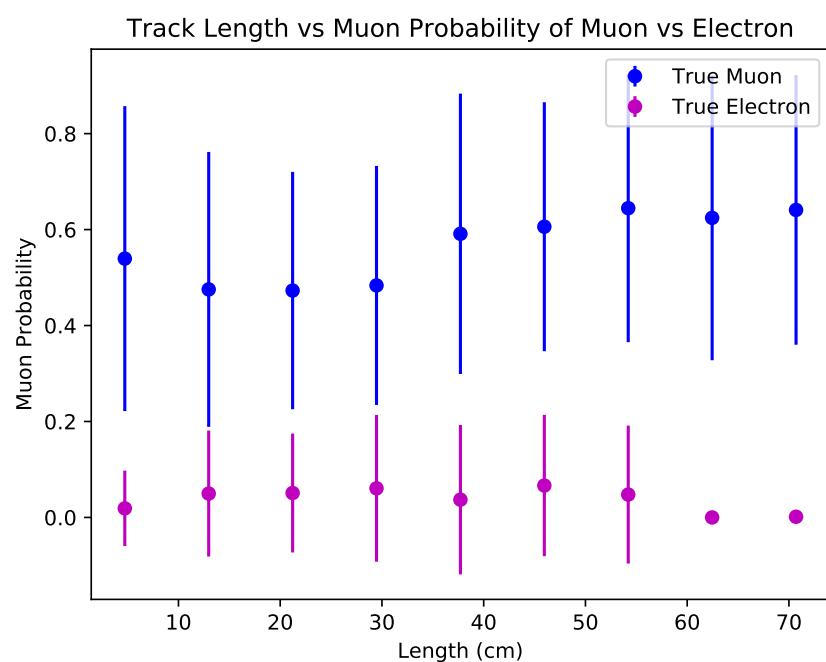


Figure 7.22: mue

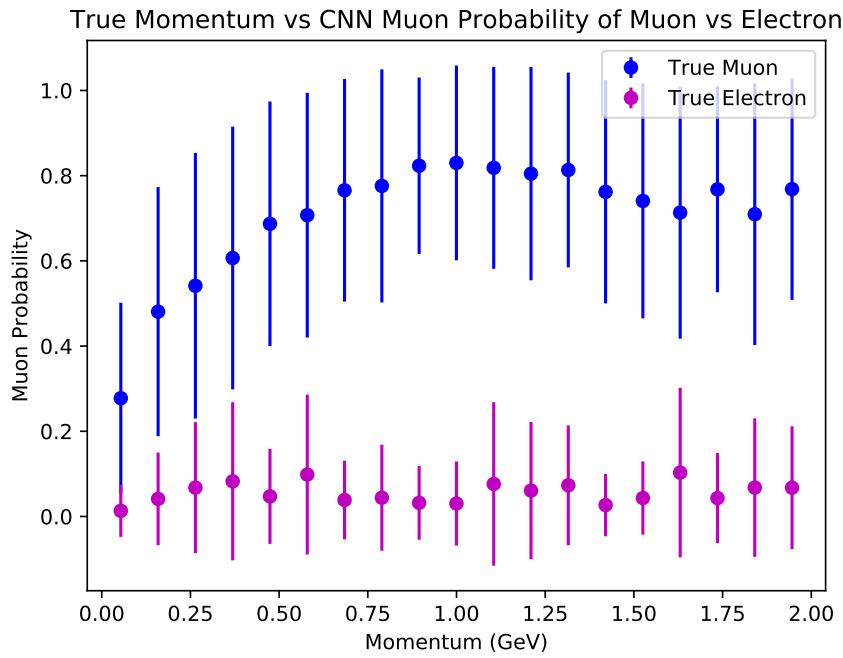


Figure 7.23: mue

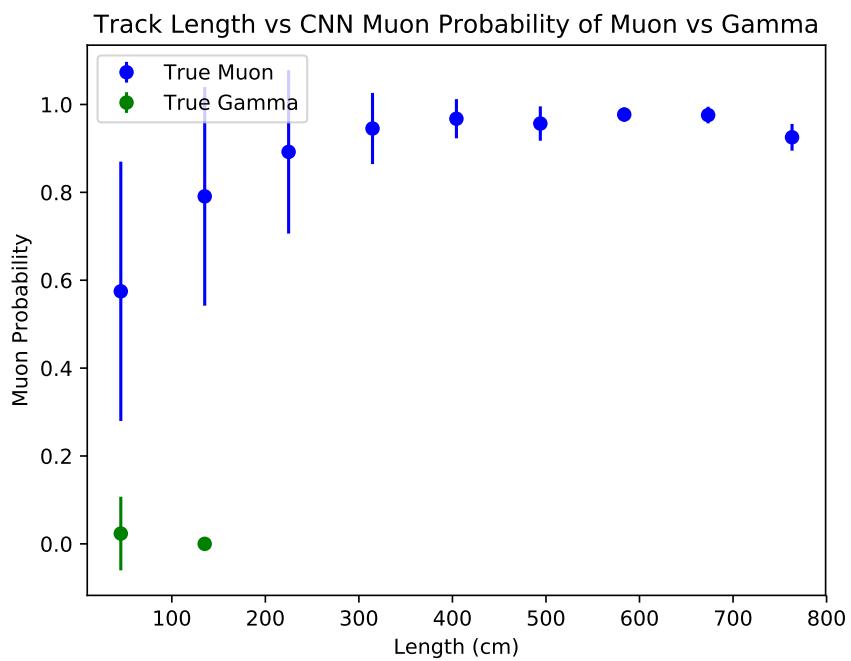


Figure 7.24: mug

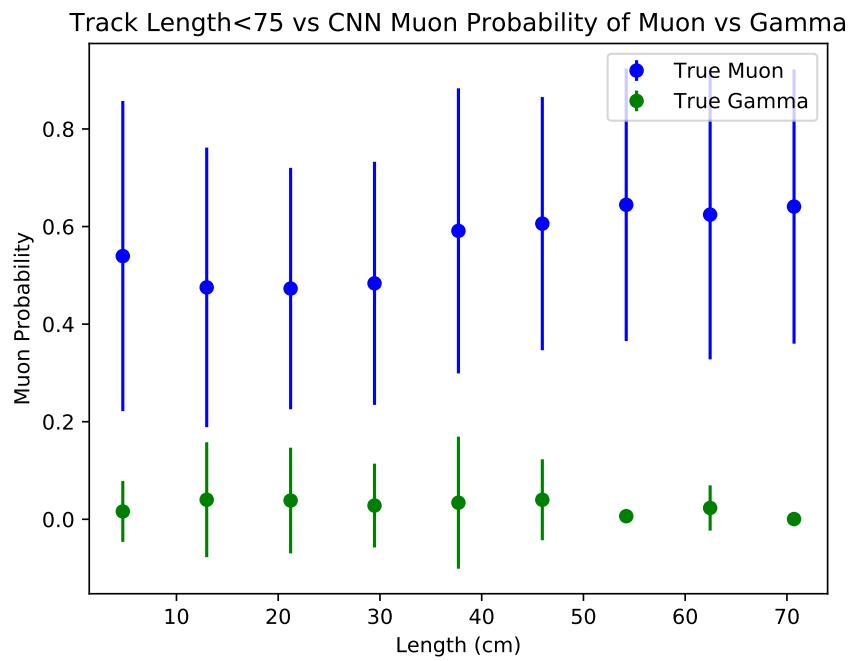


Figure 7.25: mug

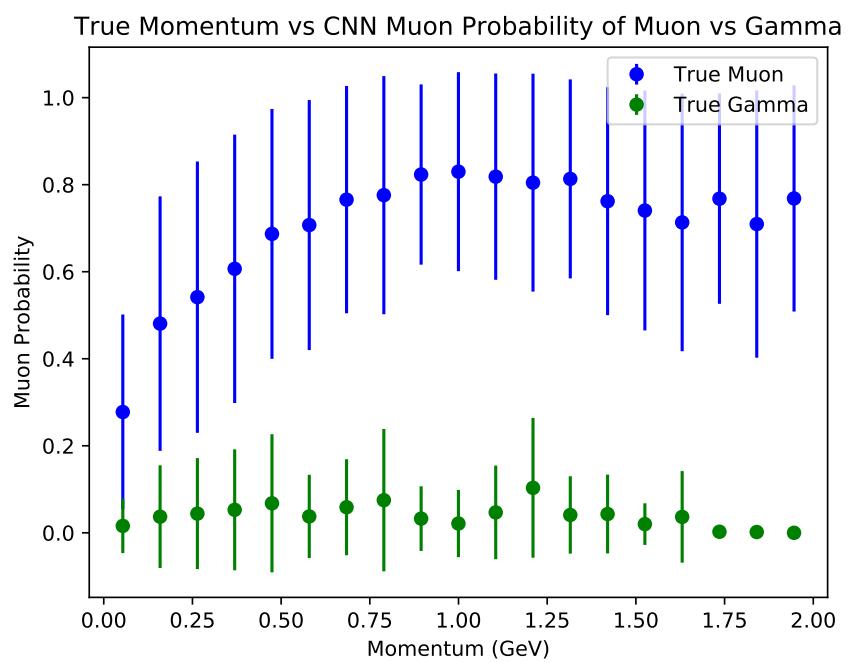


Figure 7.26: mug

1590 **Chapter 8**

1591 **Results of Convolutional Neural
1592 Networks on particles WORKING
1593 TITLE**

1594 **8.1 Classification using CNN10000**

1595 **8.1.1 Classification of MC data using Selection I Original
1596 CC-Inclusive Filter**

1597 The next step that was taken was to use CNN10000 to classify track candidate images
1598 that were identified by the selection I original cc-inclusive filter described in [?].
1599 Passing rates for each cut in cc-inclusive filter are show in figure 8.1. For the incorrect
1600 image making normalization dataset, out of 188,880 events, 7438 passed the cut right
1601 before 75 cm track length cut which is 3.9% of total data. Discrepancies in passing rates
1602 are due to grid submission issues, however, this dataset is used to check if changes
1603 in image making normalization affects μ/π separation probability due to CNN10000
1604 being trained with incorrectly image making normalized data. For the second dataset
1605 with correct image making normalization, out of 188,880 events, 9552 events passed the
1606 cut right before the 75 cm track length cut which is 5.1% passing rate and is comparable
1607 to figure 8.1. In time cosmics were also run over for efficiency and purity calculations.
1608 Out of 14395 in time cosmic events, 175 passed the cut right before the 75 cm track
1609 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

1610 Figures 8.2a, 8.2b, 8.2c and 8.2d show the accuracy and μ/π separation of both the
 1611 correct and incorrect normalized images. The confusion matrices are only composed
 1612 of μ/π data. Other particles passed the cc-inclusive filter before the 75 cm track length
 1613 cut and were all mis-id'ed as muons. Since CNN10000 has not seen any particles
 1614 other than muons and pions, it makes sense that those get mis-id'ed. Figures 8.2b
 1615 and 8.2d don't have μ/π separation comparable to 7.4c, but 8.2b does skew to higher
 1616 probabilities compared to 8.2d. This is to be expected and further work on quantifying
 1617 the performance of CNN10000 should use the incorrect image making normalization. It
 1618 is also expected that the separation isn't as defined as the testing dataset for CNN10000.
 1619 CNN10000 was trained and tested using single particle muons and pions and the track
 1620 candidate dataset come from BNB+Cosmic events, not to mentions all track candidates
 1621 have passed the cc-inclusive filter that tags "muon-like" tracks therefore the pions
 1622 in this sample look much closer in muon topology than the network has seen. Also,
 1623 these images were made from wire and time ticks associated to hits from the track
 1624 candidate that passed the cc-inclusive filter. This is different from the training images
 1625 where a bounding box was drawn over the total μ or π interaction. Spurious energy
 1626 deposition from a $\pi - Ar$ interaction is most likely not included in the BNB+Cosmic
 1627 images due to the tracking algorithm. To remedy this, the neural network needs to
 1628 see more "muon-like" pions and muons and pions from a neutrino interaction passing
 1629 the cc-inclusive filter as well as a larger particle variety including protons, photons

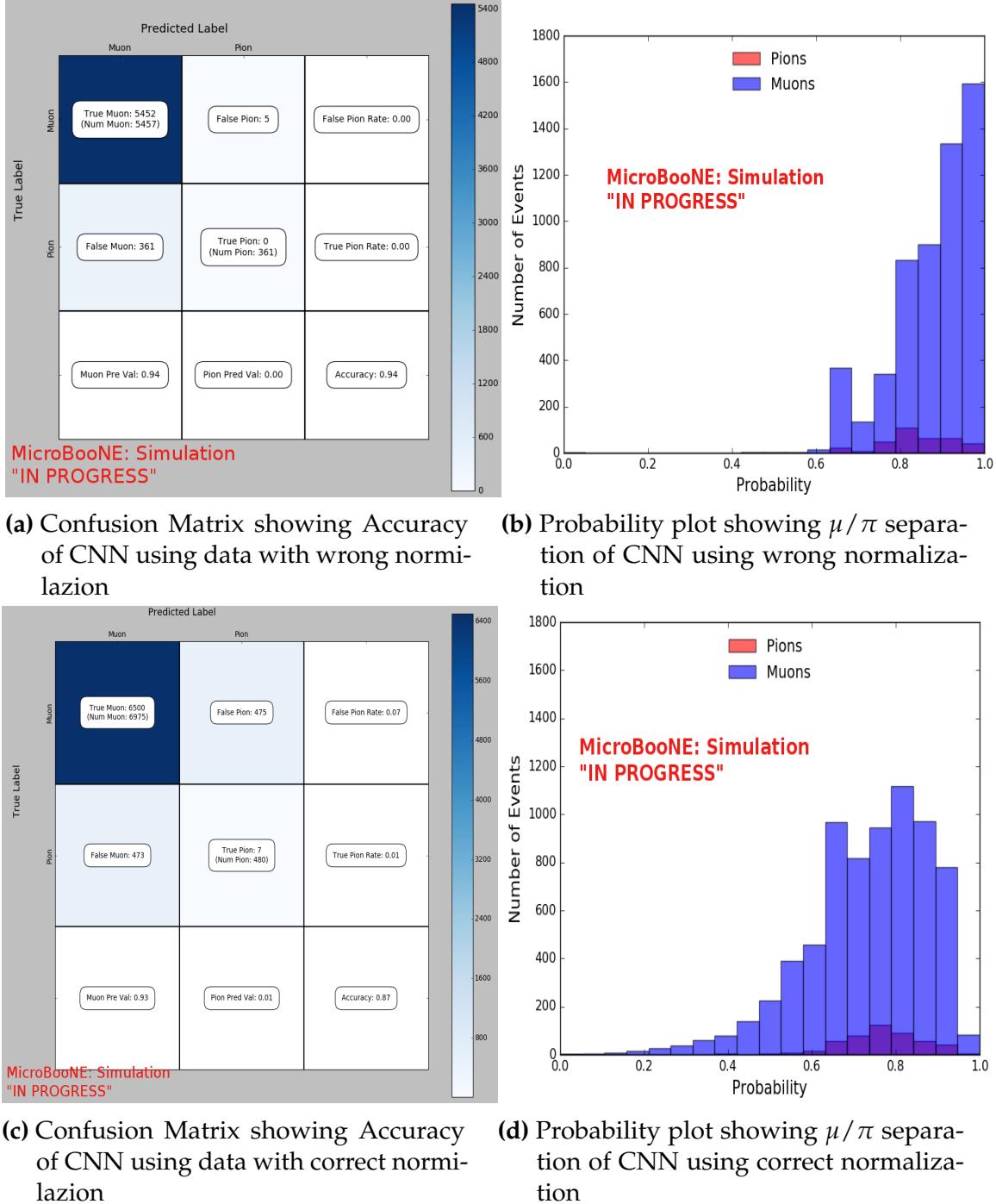


Figure 8.2: Results of CNN10000 classification of track candidate images output from cc-inclusive filter.

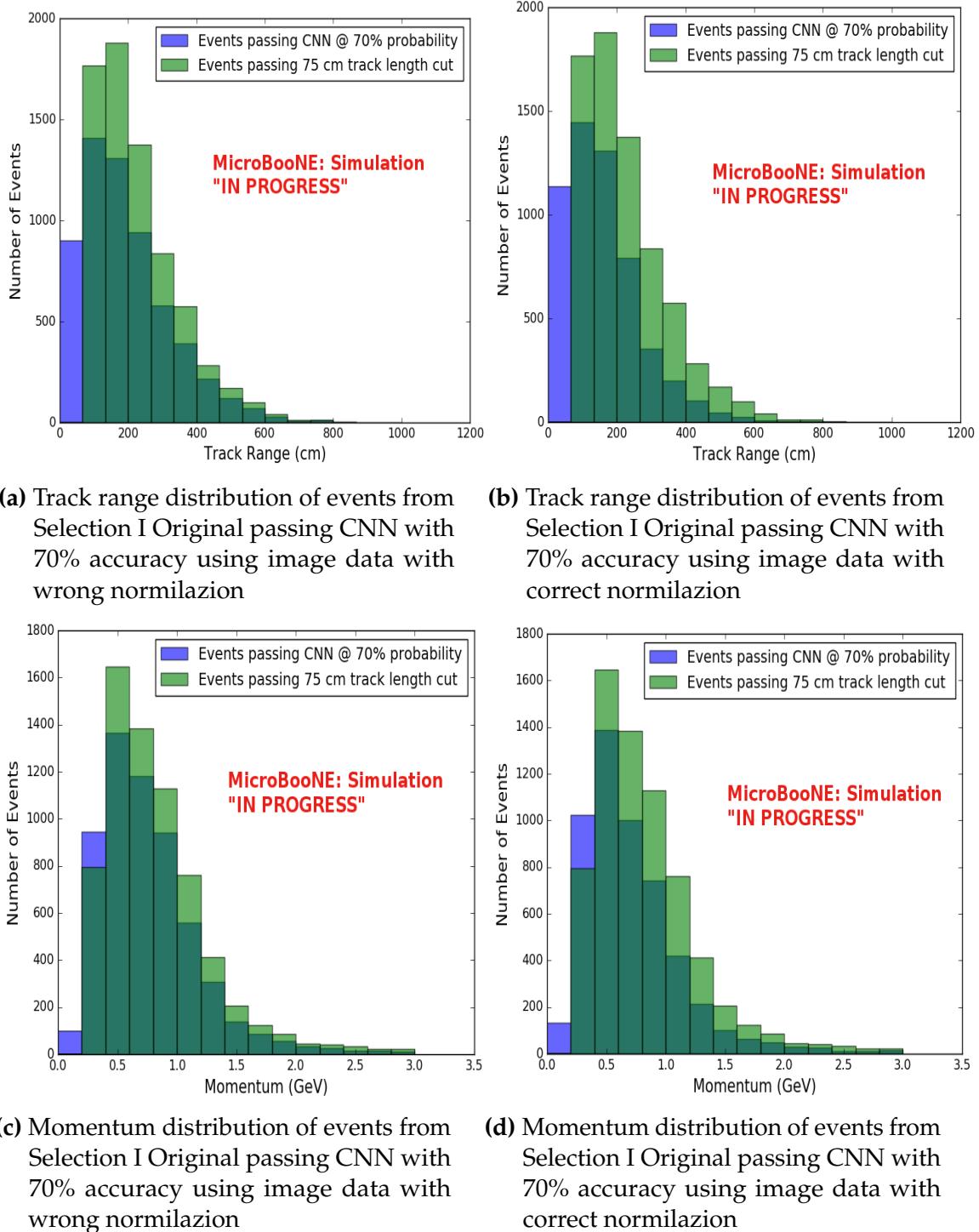


Figure 8.3: CNN10000 distributions of track candidate images output from Selection I Original cc-inclusive filter with different image data normalizations

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

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

Table 5: **Selection I: Modified** The table shows passing rates for the above described event selection. Numbers are absolute event counts and Cosmic background is not scaled appropriately. The BNB+Cosmic sample contains all events, not just 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.

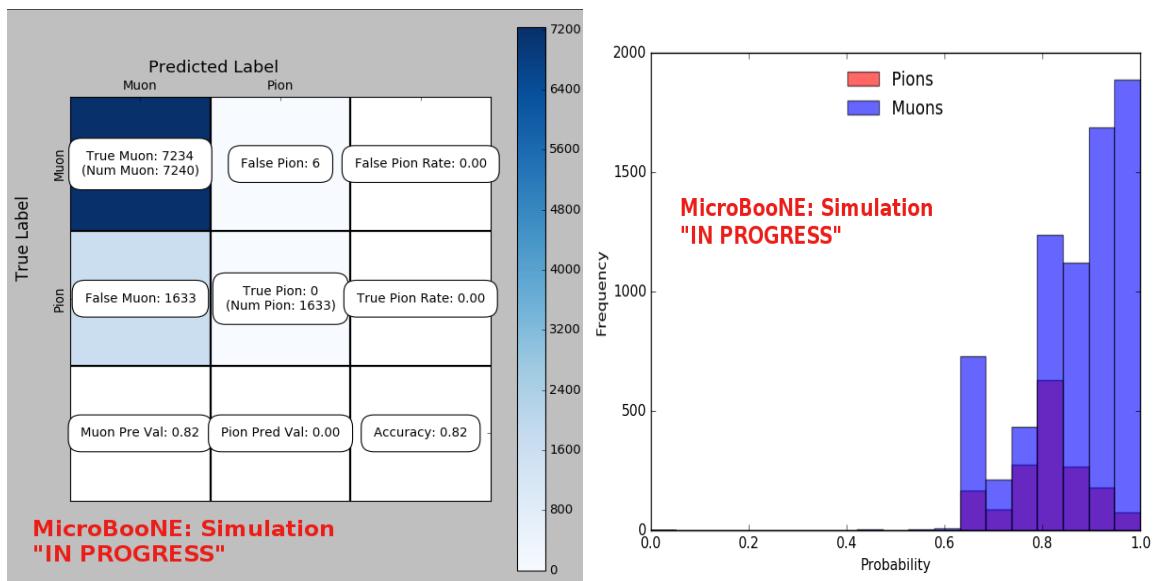
	BNB + Cosmic Selection		Cosmic only	Signal: Cosmic only
	MC-Truth			
Generated events	191362	45273	4804	1:22
≥ 1 flash with ≥ 50 PE	136219 (71%/71%)	44002 (97%/97%)	2979 (62%/62%)	1:14
≥ 1 track within 5 cm of vertex	135830 (99%/71%)	43974 (99%/97%)	2975 (99%/62%)	1:14
vertex candidate in FV	79112 (58%/41%)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

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

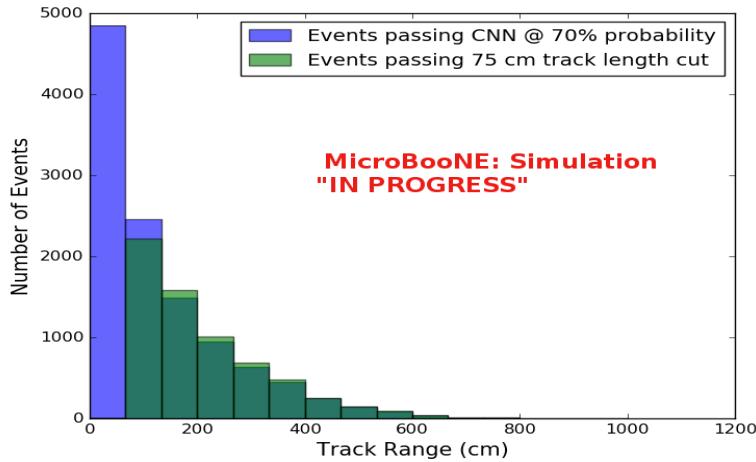
1649 all other particles were id'ed as muons. Also like selection I original data, muons are
 1650 being identified at a very high rate. Figure 8.6a shows the track range distributions
 1651 of all events from selection I modified being classified by the CNN as a muon with a
 1652 probability of 70% regardless of true particle type. We get entries for the CNN curve
 1653 in the lowest bin and none for the 75 cm curve. To see how many true CC events
 1654 were identified by CNN10000 breaking down figure 8.6a by event type was necessary.
 1655 Figures 8.6b and 8.6c show track range distributions separated by signal and various
 1656 backgrounds. Particle type was not taken into consideration in these plots so true CC
 1657 event images can be any track candidate particle passing selection I modified cut right
 1658 before track length cut including pions and protons.

1659 To gain an even deeper understanding on how CNN10000 is performing, plotting
 1660 these distributions with only muons and pions was done due to the fact that CNN10000
 1661 was trained with only those particles for μ/π separation. Figures 8.6d-8.7d show the
 1662 stacked histograms of signal and background of the track range distributions with
 1663 varying CNN probabilities starting from 70% and ending at 90% probability. With
 1664 higher probabilities we get a purer sample in the lower bin but we end up losing
 1665 events as well. Momentum distributions for all signal/background events are shown
 1666 in figure 8.8.

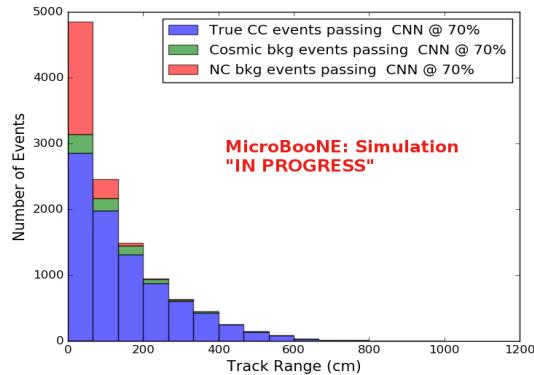


(a) Confusion Matrix for CNN10000 classified events from selection I modified (b) Probability plot for CNN10000 classified events from selection I modified

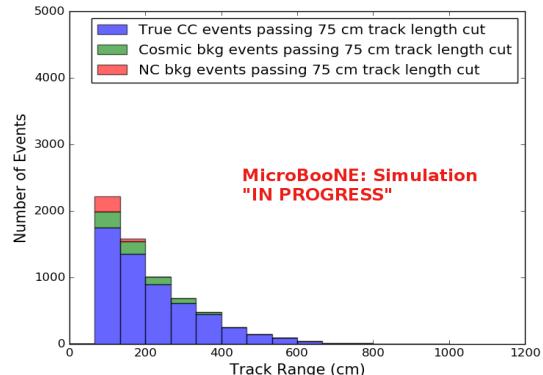
Figure 8.5: Confusion matrix and probability plot of events passing selection I modified cc-inclusive cuts right before 75cm track length cut



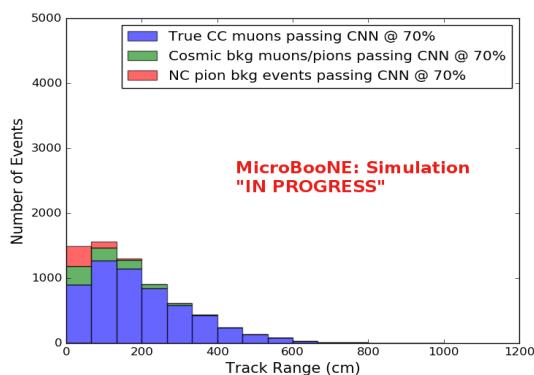
(a) Track range distribution of events from Selection I Modified passing CNN with 70% accuracy



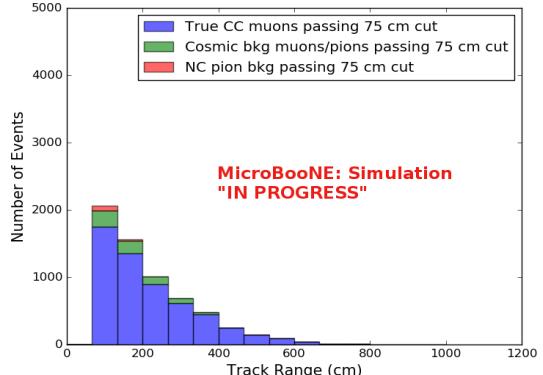
(b) Stacked signal and background track range distributions from Selection I Modified passing CNN with 70% accuracy



(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

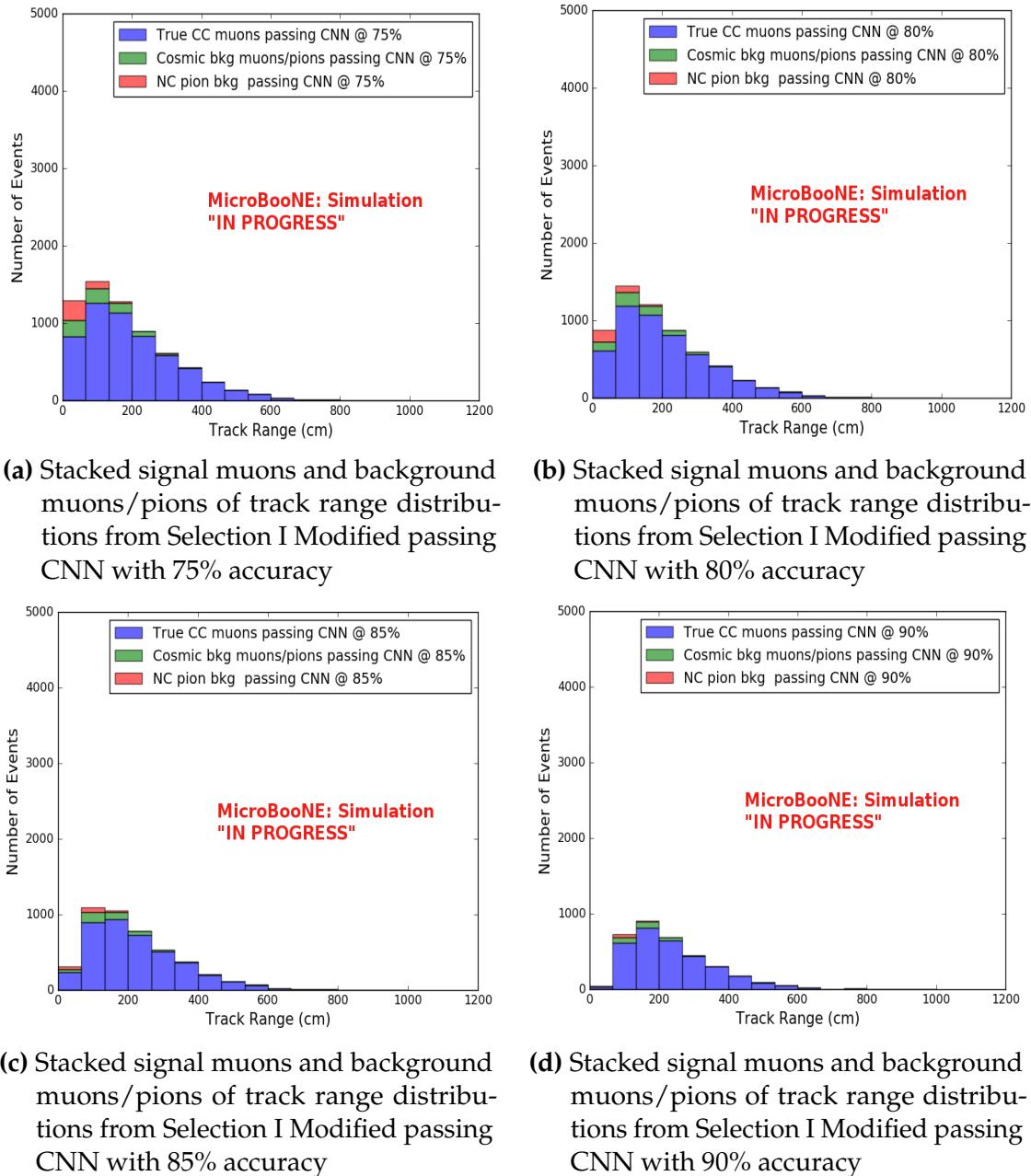
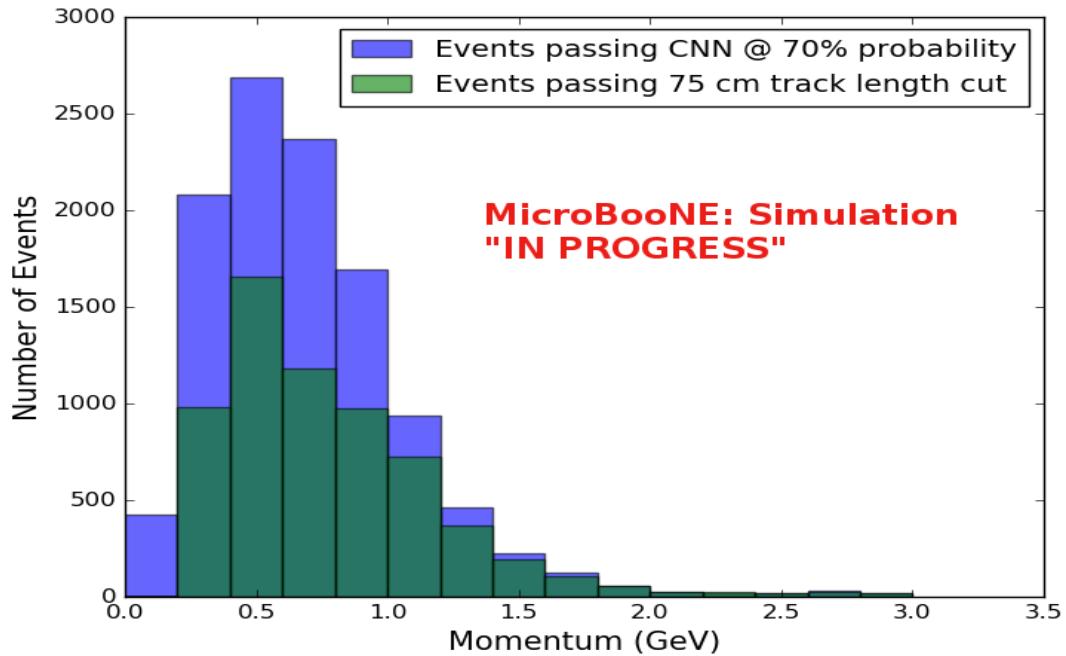
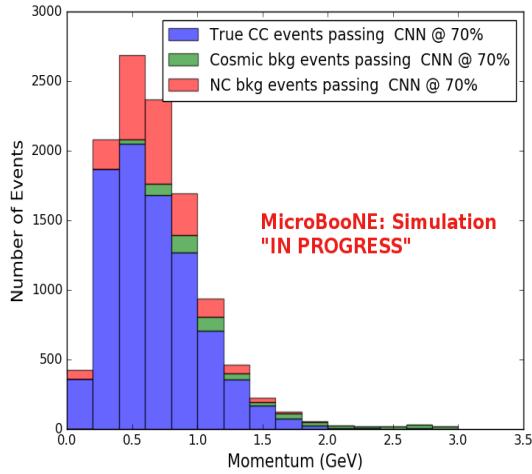


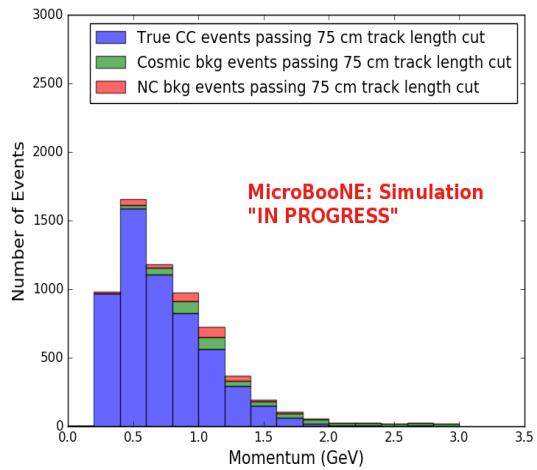
Figure 8.7: CNN10000 stacked signal/background track range distributions of track candidate images output from Selection I Modified cc-inclusive filter



(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.

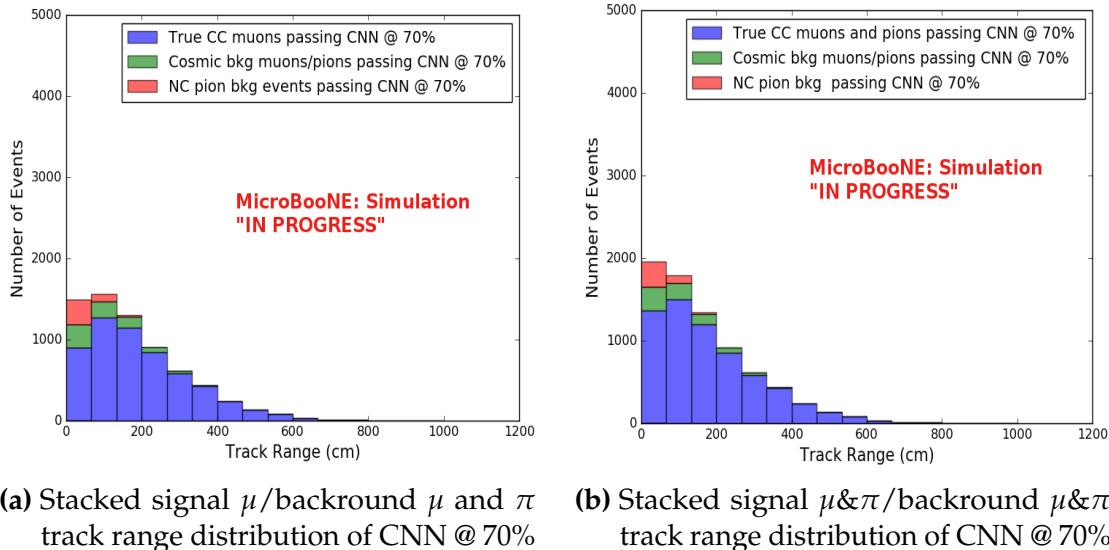


Figure 8.9: Track distribution comparisons of true CC muons plotted vs true CC muons and pions plotted

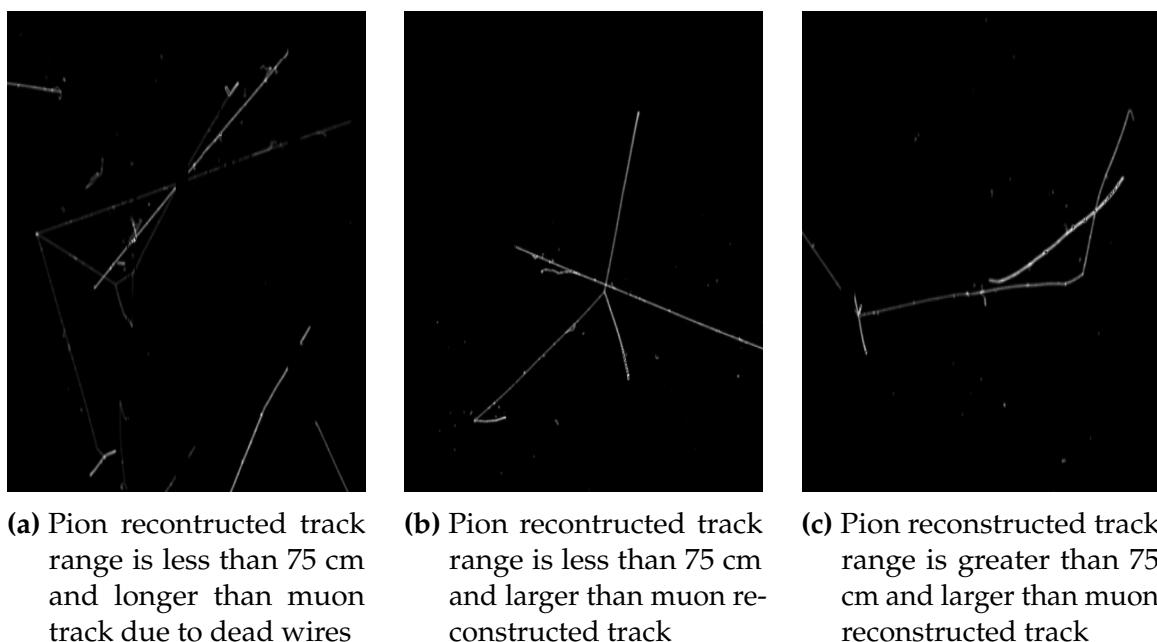


Figure 8.10: Images of true CC events where the pion was the tagged track candidate

		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	
Backgrounds		10605(35%)	#Events(Fraction) passing CNN @ 83% Probability
	Cosmics Only Events	13573(45%)	
	Cosmics in BNB Events	2249(7.4%)	
	NC Events	3412(11%)	
	ν_e and $\bar{\nu}_e$ Events	139(0.5%)	
	$\bar{\nu}_\mu$ Events	97(0.3%)	

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

1700 a MC true CC event passing rate of 14% compared to the 13% passing rate of the 75
1701 cm track length cut. The Signal:Cosmic Only background is also reduced from 1:0.6
1702 to 1:0.4 The total passing rate is also higher than the 75 cm cut, 3.6% vs 4.0%. Table
1703 8.2 shows the breakdown of signal and backgrounds for the CNN at the different
1704 probabilities. We have a 61% signal passing rate with the CNN cut @ 83% versus the
1705 53.8% signal passing rate of the 75 cm cut.

1706 Based on these numbers, the following performance values of the modified selec-
1707 tion with 75 cm cut versus modified selection with CNN @ 83% probability cut were
1708 calculated:

- 1709 • Efficiency: Number of selected true ν_μ CC events divided by the number of
1710 expected true ν_μ CC events with interaction in the FV.
 - 1711 – Selection I modified: 13%
 - 1712 – Selection I modified with CNN cut @ 83% probability: 14%
- 1713 • Purity: Number of selected true ν_μ CC events divided by sum of itself and the
1714 number of all backgrounds.
 - 1715 – Selection I modified: 53.8%
 - 1716 – Selection I modified with CNN cut @ 83% probability: 61%

1717 Lastly, figure 8.12 shows a more representative performance of the CNN. Due to
1718 the fact that the CNN was trained on muons and pions, showing the performance
1719 of CC muon events versus NC pion events with respect to CNN probability gives a
1720 better picture of how the network is performing. Figure 8.12 shows that at 83% we
1721 are below the 75 cm cut NC pion threshold and still above the CC muon threshold.
1722 Using 83% probability not only reduced the NC pion background, it also dramatically
1723 reduced the in time cosmics and cosmics in the BNB.

1724 8.1.3 Conclusions and Future Work

1725 It was shown that even though CNN10000 was trained with single particle generated
1726 muons and pions, it performs fairly well at classifying track candidate images from
1727 BNB+Cosmic events. Events have been regained below the 75 cm track length cut and
1728 the momentum and track range distributions have similar shapes to the distributions of
1729 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 [?]

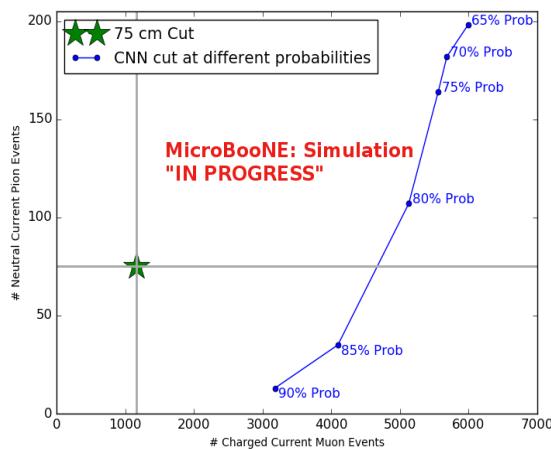


Figure 8.12: CNN performance of classified muons and pions compared to the already implemented 75 cm track length cut

1730 I modified events before 75 cm track length cut with the CNN at 83% probability and
1731 are 14% and 62% respectively. Although the CNN doesn't have separation between
1732 muons and pions and although all particles passing CNN are classified as muon,
1733 increasing CNN probability allows us to increase the purity as well as maintain an
1734 efficiency comparable to the 75 cm track length cut all while recovering events below
1735 that 75 cm cut. Out of the 6142 events that passed the CNN @ 83% 1470 events were
1736 below the 75 cm cut, a recovery of 3.3% of data with a purity of 15%. Although
1737 these numbers are low, it is an improvement from the selection I modified in both total
1738 efficiency and purity and an increase in phase space by recovering these events.

1739 **8.2 Classification using CNN100000**

1740 All future classifications will be done using Selection I Modified CC-Inclusive Filter
1741 because it has a higher efficiency and purity than Selection I Original CC-Inclusive
1742 Filter. To reiterate, CNN100000 was trained using 20,000 images of each $\mu/\pi/p/\gamma/e$.
1743 The results of using CNN100000 to classify BNB+Cosmics will be outlined below.

1744 **8.2.1 Classification of MC data using Selection I Modified 1745 CC-Inclusive Filter**

1746 **8.2.2 Classification of MicroBooNE data using Selection I Modified 1747 CC-Inclusive Filter**

1748 **8.2.3 Comparing two CC-Inclusive Cross Section Selection Filters**

¹⁷⁴⁹ **Chapter 9**

¹⁷⁵⁰ **Conclusion**

¹⁷⁵¹ Your Conclusions here.

¹⁷⁵²

¹⁷⁵³ Bibliography

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