

NEUTRAL CURRENT ELASTIC SCATTERING  
AND THE STRANGE SPIN STRUCTURE OF THE PROTON

BY

KATHERINE WOODRUFF

A dissertation submitted to the Graduate School

in partial fulfillment of the requirements

for the degree

Doctor of Philosophy

Major Subject: Physics

New Mexico State University

Las Cruces New Mexico

November 2018

“Neutral Current Elastic Scattering and the Strange Spin Structure of the Proton,” a dissertation prepared by Katherine Woodruff in partial fulfillment of the requirements for the degree, Doctor of Philosophy, has been approved and accepted by the following:

---

Dr. Luis Cifuentes  
Dean of the Graduate School

---

Dr. Vassili Papavassiliou  
Chair of the Examining Committee

---

Date

Committee in charge:

Dr. Vassili Papavassiliou, Chair

Dr. Stephen Pate

Dr. Igor Vasiliev

Dr. Steven Stochaj

## DEDICATION

I dedicate this work to my parents, Mark and Amy Woodruff, to my sister, Liz Woodruff, and to Bill Engels. Thanks for being such a supportive and loving family.

## ACKNOWLEDGMENTS

I would like to thank my advisor, Vassili Papavassiliou, for all of his help and guidance over the years. I am very grateful to have had such a knowledgeable and insightful advisor to work with and discuss physics and life over cookies.

I would also like to thank the other members of the NMSU group. First, Steve Pate for the weekly guidance and input. Also, the NMSU postdocs, Tia Miceli, who helped me learn the ropes at Fermilab and on MicroBooNE, and Lu Ren, who helped tremendously in the final months getting all of the data that I needed and answering all of my Genie questions. Additionally, Samantha Sword-Fehlberg for the great conversations with Lu and I over lunch and coffee breaks. Finally, Jesus Saenz and Jeongsu Bok for all of the time spent studying for classes and working in the lab in Las Cruces.

I am also grateful for the wonderful people that I have received advice from at Fermilab. Every conversation that I have had with Sam Zeller has been extremely enlightening and encouraging. Similarly, Sowjanya Gollapinni has always been supportive and full of useful advice. I also have to thank Yun-Tse Tsai for her support and teaching over the years.

My thesis wouldn't look the same without the guidance from the cross sections working group conveners, Anne Schukraft, Andy Furmanski, and Joel Mousseau, and the editorial board members, Tim Bolton, Josh Spitz, Taritree Wongjirad, and

Joseph Zennamo. Thank you all the help shaping and fine-tuning the analysis.

Lastly, I have to thank Jessica Esquivel and Bill Engels for the countless hours spent discussing physics, software, stats, and practically everything else.

## VITA

2006-2009 A.S., Linn-Benton Community College, Albany, Oregon

2009-2012 B.S., University of Oregon, Eugene, Oregon

2012-2015 M.S., New Mexico State University, Las Cruces, New Mexico

## PUBLICATIONS

C. Adams *et al.*, "A Deep Neural Network for Pixel-Level Electromagnetic Particle Identification in the MicroBooNE Liquid Argon Time Projection Chamber," *Submitted to: Phys. Rev. D*, 2018.

C. Adams *et al.*, "Comparison of Muon-Neutrino-Argon Multiplicity Distributions Observed by MicroBooNE to GENIE Model Predictions," *Submitted to: Phys. Rev. D*, 2018.

C. Adams *et al.*, "Ionization Electron Signal Processing in Single Phase LAr TPCs II: Data/Simulation Comparison and Performance in MicroBooNE," *JINST*, vol. 13, no. 07, p. P07007, 2018.

C. Adams *et al.*, "Ionization Electron Signal Processing in Single Phase LAr TPCs I: Algorithm Description and Quantitative Evaluation with MicroBooNE Simulation," *JINST*, vol. 13, no. 07, p. P07006, 2018.

R. Acciarri *et al.*, "The Pandora Multi-Algorithm Approach to Automated Pattern Recognition of Cosmic Ray Muon and Neutrino Events in the MicroBooNE Detector," *Eur. Phys. J.*, vol. C78, p. 82, 2018.

R. Acciarri *et al.*, "Measurement of Cosmic Ray Reconstruction Efficiencies in the MicroBooNE LAr TPC Using a Small External Cosmic Ray Counter," *JINST*, vol. 12, no. 12, p. P12030, 2017.

R. Acciarri *at al.*, "Noise Characterization and Filtering in the MicroBooNE Liquid Argon TPC," *JINST*, vol. 12, no. 08, p. P08003, 2017.

R. Acciarri *at al.*, "Michel Electron Reconstruction Using Cosmic-Ray Data from the MicroBooNE LArTPC," *JINST*, vol. 12, no. 09, p. P09014, 2017.

P. Abratenko *at al.*, "Determination of Muon Momentum in the MicroBooNE LArTPC Using an Improved Model of Multiple Coulomb Scattering," *JINST*, vol. 12, no. 10, p. P10010, 2017.

R. Acciarri *at al.*, "Convolutional Neural Networks Applied to Neutrino Events in a Liquid Argon Time Projection Chamber," *JINST*, vol. 12, no. 03, p. P03011, 2017

R. Acciarri *et al.*, "Design and Construction of the MicroBooNE Detector," *JINST*, vol. 12, no. 02, p. P02017, 2017.

## FIELD OF STUDY

Major Field: Experimental Particle Physics

ABSTRACT

NEUTRAL CURRENT ELASTIC SCATTERING  
AND THE STRANGE SPIN STRUCTURE OF THE PROTON

BY

KATHERINE WOODRUFF

Doctor of Philosophy

New Mexico State University

Las Cruces, New Mexico, 2018

Dr. Vassili Papavassiliou, Chair

Neutrinos can be used as a unique and informative probe of the structure within nucleons. This thesis presents the tools and methodology for studying the strange quark spin in the nucleon through neutral current elastic neutrino-proton scattering in the MicroBooNE experiment located at the Fermi National Accelerator Lab.

An automated boosted decision tree based proton identification algorithm is used to select proton in liquid argon TPC data with a 70% efficiency. After a set of protons are selected, a logistic regression model is used to select neutral



current elastic proton interactions in MicroBooNE. Neutral current elastic proton interactions are selected with an 11% efficiency and 30% purity.

The number of selected events measured in MicroBooNE data as a function of  $Q^2$  is compared directly to the number of selected events in MicroBooNE simulation. An event reweighting scheme is used to vary the expected number of events in MicroBooNE simulation based on the values of the physics parameters of interest. The likelihood for any given value of the physical parameters can then be easily calculated and the probability distributions of the physics parameters can be sampled using Markov Chain Monte Carlo.

Future improvements to the detector physics models in MicroBooNE are required for a precise determination of the strange quark spin structure in the proton. With the current level of uncertainty on the detector physics models, we are only able to constrain the net strange quark spin,  $\Delta s$  to the range  $-1.8 < \Delta s < 3.8$  with 95% confidence.

## CONTENTS

LIST OF TABLES . . . . .	xiv
LIST OF FIGURES . . . . .	xxv
1 Introduction . . . . .	1
1.1 The Standard Model . . . . .	2
1.2 The Spin Structure of Nucleons . . . . .	4
1.2.1 Spin Structure Functions . . . . .	4
1.2.2 The Ellis-Jaffe Sum Rule . . . . .	6
1.2.3 Experimental Tests of the Ellis-Jaffe Sum Rule . . . . .	7
1.3 Neutrino Measurements of the Strange Spin Structure . . . . .	10
1.3.1 The BNL E734 Experiment . . . . .	11
1.3.2 The MiniBooNE Experiment . . . . .	15
2 Neutrino-Nucleon interactions . . . . .	19
2.1 Two-particle interactions . . . . .	19
2.2 Electroweak interactions . . . . .	21
2.2.1 The charged and neutral currents . . . . .	22
2.2.2 V–A structure . . . . .	23
2.3 Electroweak Scattering Matrix Elements . . . . .	24
2.3.1 Nucleon matrix elements . . . . .	24

2.3.2	Electromagnetic Scattering . . . . .	25
2.3.3	Vector current . . . . .	27
2.3.4	Axial current . . . . .	28
2.4	Strangeness in the nucleon . . . . .	30
2.4.1	Strange currents . . . . .	30
2.4.2	Strange nucleon matrix elements . . . . .	31
2.5	Neutrino-proton neutral current elastic cross section . . . . .	32
2.6	Determination of the Nucleon Form Factors . . . . .	34
2.6.1	Model-Independent Form Factor Parameterization . . . . .	35
2.6.2	Fits of $z$ Expansion Form Factors to Data . . . . .	36
2.6.3	$z$ Expansion Fit to the Neutral Current Axial Form Factor . . . . .	38
3	The MicroBooNE experiment . . . . .	40
3.1	MicroBooNE Physics Goals . . . . .	40
3.2	The Booster Neutrino Beam . . . . .	42
3.3	MicroBooNE LArTPC . . . . .	46
3.4	MicroBooNE light collection system and event trigger . . . . .	50
4	Simulation and reconstruction . . . . .	53
4.1	Monte Carlo simulation . . . . .	53
4.1.1	Cross section model . . . . .	54
4.1.2	Event generation . . . . .	56
4.1.3	Detector simulation . . . . .	58

4.2	Event reconstruction . . . . .	60
4.2.1	Flash reconstruction . . . . .	60
4.2.2	TPC event reconstruction . . . . .	61
5	Particle Identification and Event Selection . . . . .	66
5.1	Particle Identification . . . . .	66
5.1.1	Reconstructed track features . . . . .	67
5.1.2	Boosted decision trees . . . . .	70
5.1.3	Training . . . . .	72
5.1.4	Performance on a Test Set . . . . .	74
5.1.5	Performance on a Neutrino Data Subset . . . . .	80
5.2	Event Selection . . . . .	92
5.2.1	Optical Pre-selection . . . . .	92
5.2.2	NC Elastic Pre-selection . . . . .	93
5.2.3	Selection Variables . . . . .	97
5.2.4	Logistic Regression . . . . .	104
5.2.5	Comparison of Event Selection on Neutrino Beam Data to Expectation from Simulation and Off-Beam Data . . . . .	108
5.3	Remaining Backgrounds . . . . .	116
5.3.1	Beam Induced Dirt Background . . . . .	117
5.3.2	Beam Induced TPC and Cryostat Background . . . . .	123
5.3.3	Cosmic Background . . . . .	131

5.4	Estimation of Systematic Uncertainty . . . . .	133
5.4.1	Neutrino Beam Flux Uncertainty . . . . .	134
5.4.2	Detector Physics Uncertainty . . . . .	137
5.4.3	Model Uncertainty . . . . .	146
6	Analysis . . . . .	153
6.1	Comparison of Data to Simulation . . . . .	154
6.1.1	Event Reweighting . . . . .	155
6.1.2	Likelihood calculation . . . . .	158
6.2	Strange axial form factor parameter estimation . . . . .	162
6.2.1	Bayesian inference . . . . .	162
6.2.2	Markov Chain Monte Carlo . . . . .	163
6.3	Results . . . . .	166
6.3.1	MCMC Sampling without Systematic Nuisance Parameters	167
6.3.2	MCMC Sampling with Systematic Nuisance Parameters .	171
6.3.3	MCMC Test Metrics . . . . .	179
6.3.4	Distributions of $\Delta s$ and $M_A^s$ . . . . .	182
6.4	Conclusions and Prospects for a Future MicroBooNE $\Delta s$ Measure- ment . . . . .	184
	REFERENCES . . . . .	188

## LIST OF TABLES

1	Breakdown by simulated particle type reconstructed tracks in the gradient-boosted decision tree training set. . . . .	72
2	Number of remaining events in $5 \times 10^{19}$ POT data set after each of the NC elastic pre-selection cuts. . . . .	94
3	Efficiency and purity of simulated NC elastic proton events after each of the NC elastic pre-selection cuts. . . . .	96
4	Contributions to the systematic uncertainty on the number of simulated events passing the NC elastic proton selection. . . . .	153

## LIST OF FIGURES

1	The standard model of particle physics taken from [1]. . . . .	3
2	The integral of the $g_1(x)$ spin structure function measured by the EMC experiment [2]. . . . .	8
3	Schematic of the BNL E734 detector [3]. . . . .	12
4	Results from the Brookhaven E734 measurement of the neutral current elastic cross section. . . . .	13
5	Schematic of the MiniBooNE detector [4]. . . . .	16
6	Ratio of the neutrino-proton NC elastic cross section to the neutrino- nucleon NC elastic cross section measured in MiniBooNE [5]. . . . .	17
7	Neutral current elastic cross section results from Ref. [6] . . . . .	18
8	Feynman diagram of two-fermion scattering. . . . .	20
9	Feynman diagrams of elastic neutrino-nucleon scattering. . . . .	25
10	Feynman diagram of neutrino-proton neutral current elastic scat- tering. . . . .	33
11	Observed low energy electron-like excess in MiniBooNE. . . . .	41
12	Predicted neutrino beam flux at MicroBooNE. . . . .	44
13	Top view of all simulated neutrino interaction positions. . . . .	45
14	Side view of all simulated neutrino interaction positions. . . . .	46

15	Representation of the operational principle of the MicroBooNE LArTPC. . . . .	49
16	Expected number of muon neutrino interactions in the MicroBooNE TPC based on Monte Carlo simulation as a function of POT. . . .	50
17	Fractional rate of flashes surrounding BNB spill window. . . . .	51
18	Rendering of the simulated MicroBooNE geometry. . . . .	59
19	Graphical example of a decision tree. . . . .	71
20	Area-normalized histograms of decision tree proton identification scores for simulated protons and other simulated proton tracks. .	75
21	Area-normalized histogram of decision tree proton identification scores for simulated proton tracks from NC elastic proton interactions.	76
22	Heatmap showing the fraction of each class that is made up of a given particle type. . . . .	77
23	The efficiency of simulated neutrino-induced proton tracks correctly classified as protons as a function of true proton kinetic energy. .	78
24	The efficiency of simulated neutrino-induced proton tracks correctly classified as protons as a function of true proton angle. . . . .	79
25	Two-dimensional efficiency for true proton $\cos(\theta_p)$ versus true proton kinetic energy. . . . .	80



26	Comparison of the decision tree proton scores between a subset of MicroBooNE neutrino data and a combination of MicroBooNE neutrino simulation and off-beam data. . . . .	83
27	Breakdown of the different particle track types in neutrino data and simulation as a function of the number of hits on the collection plane. . . . .	85
28	Breakdown of the different particle track types in neutrino data and simulation as a function of the track straightness. . . . .	85
29	Breakdown of the different particle track types in neutrino data and simulation as a function of the track length. . . . .	86
30	Breakdown of the different particle track types in neutrino data and simulation as a function of the track start $dE/dx$ . . . . .	86
31	Breakdown of the different particle track types in neutrino data and simulation as a function of the end to start $dE/dx$ ratio. . . . .	87
32	Breakdown of the different particle track types in neutrino data and simulation as a function of the track truncated total $dE/dx$ . . . . .	87
33	Breakdown of the different particle track types in neutrino data and simulation as a function of the truncated average $dE/dx$ . . . . .	88
34	Breakdown of the different particle track types in neutrino data and simulation as a function of the track starting $y$ position. . . . .	88

35	Breakdown of the different particle track types in neutrino data and simulation as a function of the track ending $y$ position. . . . .	89
36	Breakdown of the different particle track types in neutrino data and simulation as a function of the track starting $z$ position. . . . .	89
37	Breakdown of the different particle track types in neutrino data and simulation as a function of the track ending $z$ position. . . . .	90
38	Breakdown of the different particle track types in neutrino data and simulation as a function of the track $\cos(\theta)$ angle. . . . .	90
39	Breakdown of the different particle track types in neutrino data and simulation as a function of the track $\phi$ angle. . . . .	91
40	Efficiency of optical pre-selection on simulated NC elastic proton events. . . . .	93
41	Efficiency of simulated NC elastic proton events as a function of true $Q^2$ after the pre-selection cuts. . . . .	97
42	Reconstructed track length of the longest proton candidate track in the events after all pre-selection cuts. . . . .	98
43	Decision tree proton score after the pre-selection cuts. . . . .	100
44	Distance to the next closest track after the pre-selection cuts. . . . .	100
45	Distance to the beam flash in $z$ after the pre-selection cuts. . . . .	101
46	Distance to the beam flash in $y$ after the pre-selection cuts. . . . .	101

47	Whether or not the track is forward going after the pre-selection cuts. . . . .	102
48	Distance between muon track and flash after the pre-selection cuts.	102
49	Distance between pion track and flash after the pre-selection cuts.	103
50	Logistic regression NC elastic events selection score. . . . .	106
51	Efficiency and purity of the NC elastic proton event selection given several different cut values on the logistic regression score. . . . .	107
52	Efficiency of the NC elastic proton event selection on simulated NC elastic proton events as a function of true $Q^2$ . . . . .	108
53	Selected NC elastic proton events as a function of reconstructed $Q_p^2$ .	111
54	Selected NC elastic proton events as a function of reconstructed $\cos(\theta_p)$ . . . . .	112
55	Selected NC elastic proton events as a function of reconstructed $\phi_p$ .	113
56	Selected NC elastic proton events as a function of reconstructed $x$ position. . . . .	114
57	Selected NC elastic proton events as a function of reconstructed $y$ position. . . . .	115
58	Selected NC elastic proton events as a function of reconstructed $z$ position. . . . .	116
59	Remaining background dirt events after NC elastic proton event selection as a function of $Q^2$ . . . . .	118

60	Comparison of the cosmic proton flux predictions from three different generator models in MicroBooNE [7]. . . . .	119
61	Position of simulated neutrino interactions which produce a neutron that enters and interacts in the MicroBooNE TPC. . . . .	121
62	Comparison between neutrino beam data and the expectation from simulation and off-beam data of the $z$ distribution of reconstructed flashes <i>without</i> the contribution from neutrino interactions in the dirt. . . . .	122
63	Distribution of the $z$ position of reconstructed flashes in MicroBooNE.	123
64	Comparison between neutrino beam data and the expectation from simulation and off-beam data of the $z$ distribution of reconstructed flashes <i>with</i> the contribution from neutrino interactions in the dirt.	124
65	Remaining background cryostat events after NC elastic proton event selection as a function of $Q^2$ . . . . .	126
66	Remaining CC TPC events background after NC elastic proton event selection as a function of $Q^2$ . . . . .	127
67	Remaining NC TPC background events after NC elastic proton event selection as a function of $Q^2$ . . . . .	128
68	Remaining background NC TPC background events broken down by interaction type after NC elastic proton event selection as a function of $Q^2$ . . . . .	129

69	Comparison of the number of CC MEC events in MicroBooNE simulation between the two models as a function of true $Q^2$ . . . . .	131
70	Corrected number of CC MEC events as a function of true $Q^2$ with one standard deviation systematic uncertainty bars. . . . .	132
71	Remaining in-time cosmic background events after NC elastic proton event selection as a function of $Q^2$ . . . . .	133
72	Neutrino flux prediction with systematic uncertainty at MicroBooNE [8].	134
73	The uncertainty on the number of simulated events due to the neutrino beam flux uncertainty as a function of $E_\nu$ . . . . .	135
74	Two-dimensional histogram showing the relationship between true $Q^2$ and true $E_\nu$ in simulated NC elastic neutrino-proton interactions in MicroBooNE. . . . .	136
75	Fractional uncertainty on the number of simulated NC elastic neutrino-proton interactions due to the beam flux uncertainty as a function of $Q^2$ . . . . .	137
76	The uncertainty on the number of simulated events due to the neutrino beam flux uncertainty as a function of $Q^2$ . . . . .	138
77	The NC elastic proton selection on simulation and off-beam data with statistical and systematic uncertainty due to the uncertainty on the neutrino beam flux. . . . .	139

78	Comparison of an event display of the second induction plane using the default detector physics parameters and using the estimated dynamic induced charge on the wires. . . . .	140
79	Change in the number of simulated events passing the NC elastic proton selection as a function of reconstructed $Q^2$ between the default detector physics and with the estimated dynamic induced charge on the wires. . . . .	141
80	The NC elastic proton selection on simulation and off-beam data with statistical and systematic uncertainty due to the dynamic induced charge on the wires. . . . .	142
81	Change in the number of simulated events passing the NC elastic proton selection as a function of reconstructed $Q^2$ between the default detector physics and with the space charge correction. . .	143
82	Change in the number of simulated events passing the NC elastic proton selection as a function of reconstructed $Q^2$ when the single PE background rate is varied. . . . .	144
83	The NC elastic proton selection on simulation and off-beam data with statistical and systematic uncertainty due to the single PE background rate. . . . .	145

84	The NC elastic proton selection on simulation and off-beam data with statistical and systematic uncertainty due to the uncertainty on the MEC model. . . . .	147
85	The systematic uncertainty due to Pauli blocking on simulated CCQE events as a function of true $Q^2$ using the GENIE Reweighting tool. . . . .	148
86	The NC elastic proton selection on simulation and off-beam data with statistical and systematic uncertainty due to Pauli blocking.	149
87	The NC elastic proton selection on simulation and off-beam data with statistical and systematic uncertainty due to the nucleon rescattering probability in the argon nucleus. . . . .	150
88	The change in the number of NC elastic events with different nuclear momentum and FSI models as a function of reconstructed $Q^2$ . . . . .	151
89	The NC elastic proton selection on simulation and off-beam data with statistical and systematic uncertainty due to the number of dirt interactions. . . . .	152
90	The NC elastic proton selection on simulation and off-beam data with statistical and systematic uncertainty. . . . .	154
91	Calculated likelihood values at discrete $a_0^s$ , $a_1^s$ , and $a_2^s$ with zero systematic uncertainty. . . . .	161

92	Two-dimensional views of the multivariate normal proposal distribution of the strange axial form factor parameters. . . . .	169
93	MCMC chains of the strange axial form factor parameters after 50,000 steps with the systematic parameters held to zero. . . . .	170
94	Two-dimensional views of the 50,000 MCMC samples of the posterior distributions of the strange axial form factor parameters with the systematic parameters held to zero. . . . .	171
95	MCMC chains of the strange axial form factor parameters after 200,000 steps with the systematic parameters included. . . . .	175
96	Two-dimensional views of the 200,000 MCMC samples of the posterior distributions of the strange axial form factor parameters with the systematic parameters included. . . . .	176
97	Two-dimension histogram of the MCMC samples of the posterior distribution for the $a_0^s$ and $a_1^s$ parameters. . . . .	177
98	One-dimensional histograms of the strange axial form factor parameter samples. . . . .	178
99	The posterior predictive distribution of the number of selected events compared to Run I neutrino-beam data as a function of reconstructed $Q^2$ . . . . .	179
100	Two-dimensional histogram of the MCMC samples transformed to $\Delta s$ and $M_A^s$ . . . . .	183



101	One-dimensional histogram of the MCMC samples transformed to $\Delta s$ with 68% and 95% credible intervals. . . . .	184
102	One-dimensional histogram of the MCMC samples transformed to $M_A^s$ with 68% and 95% credible intervals. . . . .	185

## 1 Introduction

Ten years after the discovery of the proton by Rutherford in 1920 [9] Pauli postulated “a particle that cannot be detected”, which later became known as the neutrino. The neutrino was eventually detected by Reines and Cowan in 1956 [10]. Our understanding of both the proton and the neutrino has radically changed since they were discovered. In the 1960s Gell-Mann and Zweig independently proposed that the proton had an internal structure composed of a new particle called quarks [11, 12, 13]. This quark substructure was first detected in 1968 at the Stanford Linear Accelerator Lab [14]. Also in 1968, Davis built a detector to measure neutrinos coming from the sun and detected fewer than predicted by theory [15, 16]. This was later confirmed in 1998 by the Super-Kamiokande collaboration [17] to be due to the fact that neutrinos are able to oscillate into other flavors, violating lepton number conservation.

Two decades later in 1987, the European Muon Collaboration (EMC) discovered that the spin of the quarks in the proton only makes up a small percentage of the total spin of the proton [18], referred to as the proton spin puzzle. In addition to the total quark spin contribution being smaller than expected, the EMC also saw evidence that the net spin contribution from the strange quark-antiquark pairs in the proton’s quark-gluon sea was not zero as expected but actually negative[2].

In 1997 the Liquid Scintillator Neutrino Detector (LSND) collaboration announced that it had observed an anomaly in the neutrino oscillation spectrum [19] which might suggest a fourth generation of neutrinos that do not interact via the standard model, called sterile neutrinos. This announcement led to the MiniBooNE experiment which was unable to verify the LSND anomaly, but observed a new anomaly at low neutrino energy [20, 21], known as the low energy excess, which may also suggest sterile neutrinos.

The MicroBooNE experiment [22] was designed to be able to investigate the MiniBooNE low energy excess. Interestingly, the design parameters required for MicroBooNE to shed light on the MiniBooNE excess happen to make it an ideal experiment for studying the mystery of strange quark spin structure in the proton that was found by the EMC.

## 1.1 The Standard Model

The standard model of particle physics, illustrated in Fig. 1, classifies all of the known elementary particles and describes the fundamental forces between them. The elementary particles in the standard model are divided into the particles with half-integer spin, called fermions, that make up matter and the particles with whole-integer spin, called bosons, that carry the forces between particles.

The fermions are further divided based on what type of forces they interact with. The six quarks (and six antiquarks) can interact via all three standard

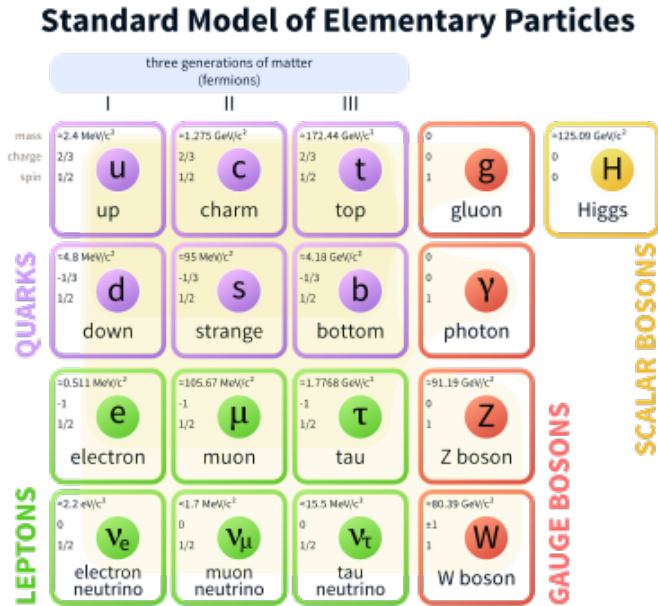


Figure 1: The standard model of particle physics taken from [1].

model forces. The quarks compose protons and neutrons and are the only fermions that can interact via the strong force. The six leptons (and six anti-leptons) can all interact via the weak force, and the charged leptons (electrons, muons, and tau leptons) can interact via the electromagnetic force, but the neutral neutrinos cannot. The quarks and leptons are also divided into three “generations”, also illustrated in Fig. 1. Almost all matter is composed of the first generation fermions which include electrons and up and down quarks.

The bosons consist of the gluon, which carries the strong force, the photon, which carries the electromagnetic force, the  $W^\pm$  and  $Z^0$  bosons, which carry the weak force, and the Higgs boson, which gives the particles their mass. As their

names suggest, the strong force is much stronger than the other forces, the weak force is the weakest force in the standard model, and the electromagnetic force is in between.

Neutrinos are so difficult to detect because they can only interact via the weakest force. The probability of a given neutrino interacting with any other particle is very low. This is also what allows us to use neutrinos to probe the internal quark and gluon structure of protons and neutrons.

## 1.2 The Spin Structure of Nucleons

Protons and neutrons, called nucleons, are composed of three up and down quarks and the gluons that bind the quarks together. These gluons carry enough energy to split into short-lived quark-antiquark pairs inside the nucleon making up a quark-gluon sea. The sea is made up of gluons and up, down, and strange quarks and antiquarks. The momentum, electromagnetic, and spin structure of each component of the quark-gluon sea is an active area of research.

### 1.2.1 Spin Structure Functions

In inclusive lepton-nucleon deep inelastic scattering (DIS), it is useful to parameterize the scattering cross section in terms of nucleon structure functions  $F_1(x)$ ,  $F_2(x)$ ,  $g_1(x)$ , and  $g_2(x)$ . In the QCD parton model [23],  $x$  is the fraction of the nucleon's momentum carried by the quarks, and  $g_1$  and  $g_2$ , the spin structure

functions, parameterize the polarized DIS cross section [24].

The  $g_1$  spin structure function can be written as a combination of the spin contribution from each of the quark flavors [25],

$$g_1(x) = \frac{1}{2} \sum_q e_q^2 \Delta q(x), \quad (1)$$

where  $q$  is the quark flavor ( $q = u, d, s$ ),  $e_q$  is the electric charge of the quark, and  $\Delta q$  is the contribution of the quark spin to the nucleon spin,

$$\Delta q(x) = (q^\uparrow + \bar{q}^\uparrow)(x) - (q^\downarrow + \bar{q}^\downarrow)(x). \quad (2)$$

Here  $q^\uparrow(x)$  ( $q^\downarrow(x)$ ) is the probability of finding a quark with momentum fraction  $x$  with its spin in the same (opposite) direction of the nucleon's spin, and  $\bar{q}^\uparrow(x)$  ( $\bar{q}^\downarrow(x)$ ) is the probability of finding an antiquark with momentum fraction  $x$  with its spin in the same (opposite) direction of the nucleon's spin. Integrating over the quark spin structure gives the net quark spin contribution to the nucleon spin

$$\Delta q = \int_0^1 [(q^\uparrow + \bar{q}^\uparrow)(x) - (q^\downarrow + \bar{q}^\downarrow)(x)] dx. \quad (3)$$

### 1.2.2 The Ellis-Jaffe Sum Rule

The Ellis-Jaffe sum rule [26] relates the integral of the  $g_1$  spin structure function to the axial charges [24],

$$g_A = \Delta u - \Delta d \quad (4)$$

$$g_A^{(8)} = \Delta u + \Delta d - 2\Delta s \quad (5)$$

$$g_A^{(0)} = \Delta u + \Delta d + \Delta s. \quad (6)$$

where  $g_A$  is the isovector axial charge,  $g_A^{(8)}$  is the  $SU(3)$  octet axial charge, and  $g_A^{(0)}$  is the flavor-singlet axial charge. For the proton, the integral of the  $g_1$  spin structure function is

$$\begin{aligned} S_p &= \int_0^1 dx g_{1p}(x) \\ &= \int_0^1 dx \left[ \frac{4}{18} \Delta u(x) + \frac{1}{18} \Delta d(x) + \frac{1}{18} \Delta s(x) \right], \\ &= \frac{4}{18} \Delta u + \frac{1}{18} \Delta d + \frac{1}{18} \Delta s, \\ &= \frac{g_A}{12} + \frac{g_A^{(8)}}{38} + \frac{g_A^{(0)}}{9}. \end{aligned} \quad (7)$$

The axial charges can be determined through experimental measurements. The isovector axial charge,  $g_A$ , can be obtained in neutron  $\beta$ -decay [27]. Assuming  $SU(3)$  symmetry,  $g_A^{(8)}$  can be obtained through hyperon  $\beta$ -decay. If the net strange contribution to the nucleon spin is assumed to be negligible, the flavor-singlet charge is equal to the  $SU(3)$  octet charge,

$$\Delta s \sim 0 \Rightarrow g_A^{(0)} = g_A^{(8)}. \quad (8)$$

### 1.2.3 Experimental Tests of the Ellis-Jaffe Sum Rule

One of the first experiments to test the Ellis-Jaffe sum rule through inclusive DIS was the European Muon Collaboration (EMC) at CERN in 1989 [18, 2]. The EMC scattered polarized muons off a polarized proton target and detected the scattered muon in a forward muon spectrometer. Figure 2 shows the EMC measurement of the integral of the  $g_1$  spin structure function as a function of the lower bound on the integral,  $x_m$ , and the expected value of the integral at  $x_m = 0$  from the Ellis-Jaffe sum rule. There is a significant discrepancy between the measured value and the value expected from theory. Assuming that the discrepancy comes from the assumption that  $\Delta s = 0$ , and not the  $SU(3)$  symmetry assumption, the extracted non-zero value of  $\Delta s$  to resolve the difference is

$$\Delta s_{\text{EMC}} = -0.095 \pm 0.016 \pm 0.023. \quad (9)$$

This implies not only that the overall spin polarization of the strange quarks and antiquarks in the nucleon sea is nonzero, but that they are polarized in the opposite direction of the proton.

After the EMC result, many subsequent polarized target inclusive DIS experiments tested the Ellis-Jaffe sum rule over the next few decades. See Refs. [28] and [25] for detailed reviews. Several inclusive DIS polarized target experiments were performed at SLAC [29, 30, 31] with a polarized electron beam, at CERN with the polarized muon beam by the Spin Muon Collaboration (SMC) [32, 33] and by



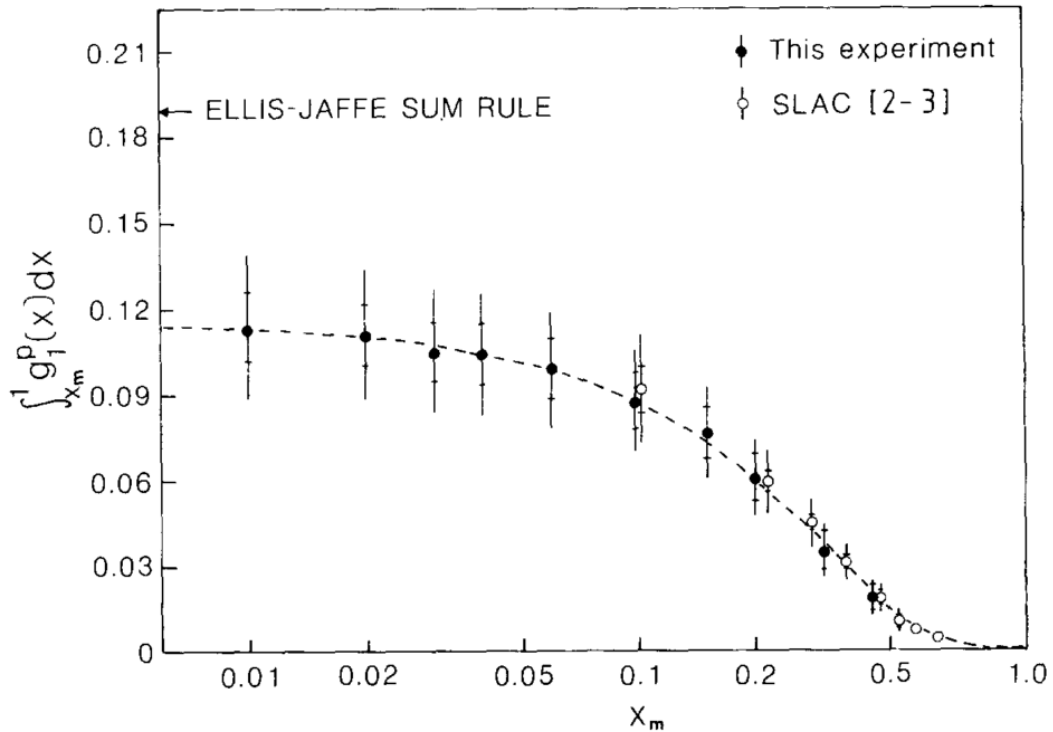


Figure 2: The integral of the  $g_1(x)$  spin structure function measured by the EMC experiment [2].

COMPASS [34], the polarized electron or positron HERA beam at DESY by the HERMES [35, 36, 37] collaboration. A more recent measurements of the violation of the Ellis-Jaffe sum rule from polarized muon inclusive DIS off a polarized target in the COMPASS experiment in 2007 gives [34]  $\Delta_{s\text{COMPASS}} = -0.08 \pm 0.01 \pm 0.02$ , and from the HERMES experiment in 2007 gives [37]  $\Delta_{s\text{HERMES}} = -0.085 \pm 0.013 \pm 0.008 \pm 0.009$ .

The nucleon spin structure can also be studied through semi-inclusive deep inelastic scattering (SIDIS). In SIDIS, in addition to detecting the scattered lep-

ton, at least one of the final state pions or kaons is detected. If the detected hadron has a large enough fraction of the energy transfer, it can be assumed that it contains the quark that was struck by the lepton [25]. A factor is included in the measured spin structure functions that describes the probability of a struck quark producing a pion or kaon with the measured energy fraction. This factor is called a fragmentation function, and it can be used to reconstruct individual quark flavor contributions to the nucleon spin. Several experiments have made measurements of the strange quark spin through SIDIS including COMPASS [38, 39] at CERN and HERMES [40, 41, 42] at DESY. Measurements of the strange quark polarization in the nucleon through SIDIS tend to favor much smaller values of  $\Delta s$  that are consistent with zero. While these results depend less on  $SU(3)$  flavor symmetry than inclusive DIS results, they do depend strongly on the choice of fragmentation functions. Results from global analyses of both DIS and SIDIS data [43, 44, 45, 46, 47] tend to give negative values of  $\Delta s$ .

In addition to the experimental effort, there has been a parallel effort to calculate the standard model prediction of the nucleon structure using lattice QCD. Early lattice QCD calculations of  $\Delta s$  [48, 49] suggested negative values at a similar scale to what is measured in inclusive DIS. The value found in Ref. [49] was  $\Delta s = -0.12 \pm 0.01$ . More recent calculations [50, 51, 52, 53] give results of  $\Delta s$  much closer to zero, but still negative. The newer results are more similar to the values measured in SIDIS. For example,  $\Delta s = -0.031 \pm 0.017$  was found in

Ref. [51] and  $\Delta s = -0.018 \pm 0.006$  was found in Ref. [53].

### 1.3 Neutrino Measurements of the Strange Spin Structure

Since neutrinos only interact via the weak force, neutrino-nucleon elastic scattering is sensitive to the weak currents and are great tools for measuring the axial form factor,  $G_A(Q^2)$ . See Refs. [54] and [55] for detailed reviews of the many measurements of  $G_A(Q^2)$  through charged current quasi-elastic (CCQE) scattering. Neutral current (NC) elastic neutrino-nucleon scattering ( $\nu N \rightarrow \nu N$ ) specifically is sensitive to the NC form factor  $G_A^{NC}(Q^2)$  which contains contributions from the up, down, and strange quarks to the spin structure of the nucleon ( $G_A(Q^2)$  only contains contribution from the up and down quarks).

At the limit where the negative four-momentum transfer squared,  $Q^2$ , goes to zero, the NC axial form factor becomes a combination of the net spin contribution from each of the quarks to the nucleon spin [25],

$$G_A^{NC}(Q^2 = 0) = \frac{1}{2}(\Delta u - \Delta d - \Delta s) \quad (10)$$

The reconstructed  $Q^2$  is determined entirely from the recoil nucleon kinetic energy using

$$\begin{aligned} Q_N^2 &= -q^2 = -(p'_N - p_N)^2 \\ &= -(E'_N - E_N)^2 + (\mathbf{p}'_N - \mathbf{p}_N)^2 \\ &= 2T_N M_N, \end{aligned} \quad (11)$$

where  $p$  is four-momentum,  $E$  is energy,  $\mathbf{p}$  is three-momentum,  $M$  is mass,  $T$  is

kinetic energy determined by the length of the track, the  $N$  subscript represents the nucleon in the neutrino-nucleon interaction, the prime represents the final state, and the nucleon momentum in the nucleus is assumed to be small compared to the final nucleon momentum. This means that the ability to measure the axial form factor at low  $Q^2$  in NC elastic neutrino-nucleon scattering depends on the experimental nucleon energy threshold.

Two previous neutrino experiments have performed a measurement of  $\Delta s$  through neutral current elastic neutrino-nucleon scattering. The first was the E734 experiment [3] at Brookhaven National Lab (BNL) in 1987, and the second was the MiniBooNE experiment [5] at Fermilab in 2010.

### 1.3.1 The BNL E734 Experiment

The main target and detector of the E734 experiment was 170 tons and was made of a combination of liquid scintillator cells and proportional drift tubes (PDTs). The liquid scintillator composed 80% of the target and was used for calorimetry and timing, while the PDTs were used for position information. Additionally, there was a electromagnetic shower counter and a muon spectrometer just downstream of the main detector. The full detector schematic is shown in Fig. 3. The E734 detector sat in a neutrino beam at BNL that could run in either neutrino or antineutrino mode with a mean energy of 1.3 GeV for neutrino and 1.2 GeV for antineutrinos. Note that we are using natural units with the speed of light,  $c$ , set

equal to one.

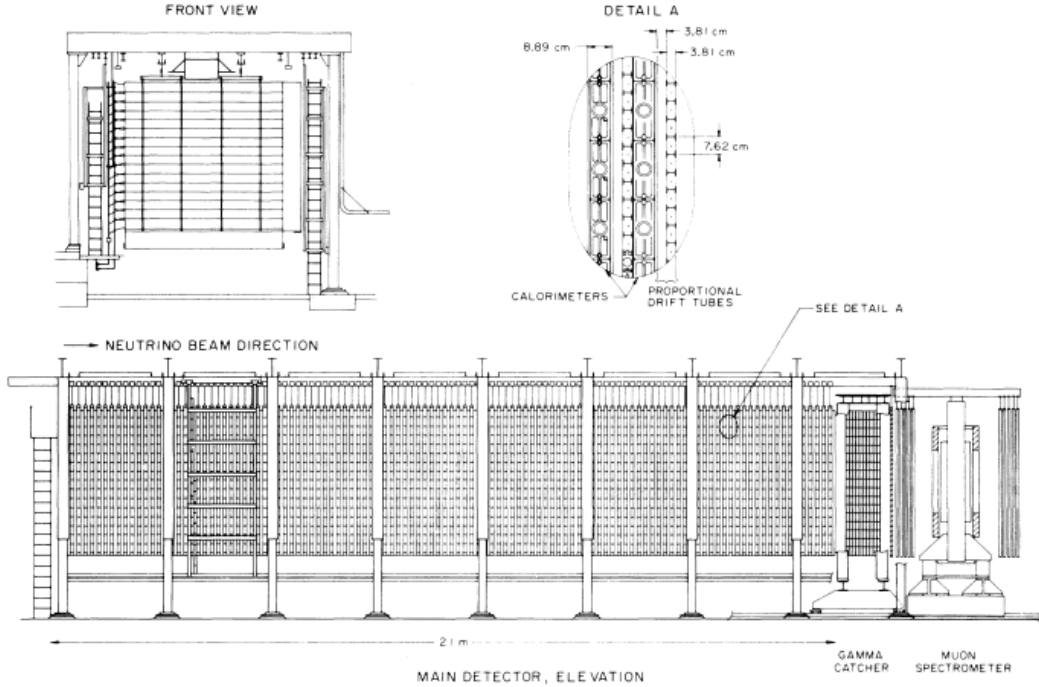


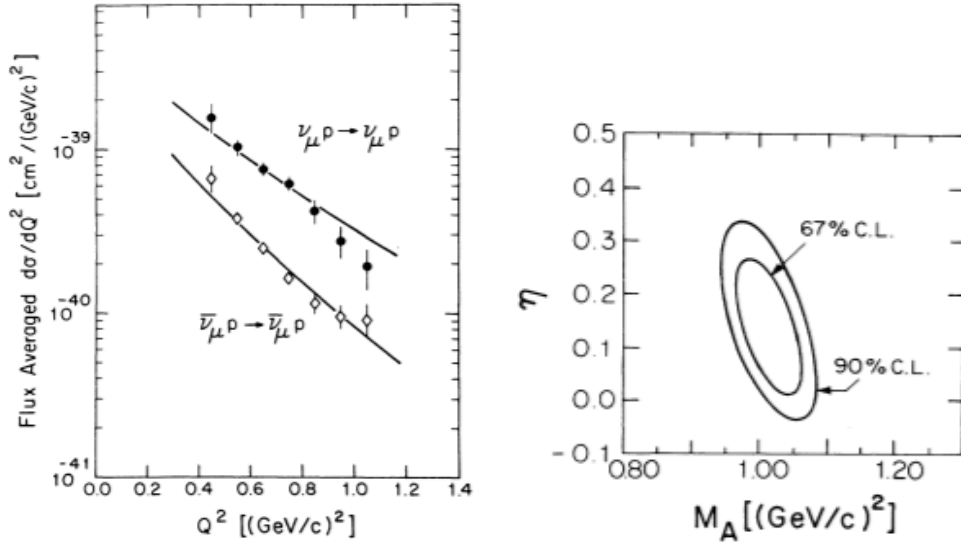
Figure 3: Schematic of the BNL E734 detector [3].

A simultaneous fit to the neutrino-proton and antineutrino-proton neutral current elastic cross sections in the range between  $Q^2 = 0.45 \text{ GeV}^2$  and  $Q^2 = 1.05 \text{ GeV}^2$  was performed to extract the neutral current axial form factor. In the parameter estimation, the NC axial form factor was assumed to have the form

$$G_A^{NC}(Q^2) = \frac{1}{2} \frac{g_A}{(1 + Q^2/M_A^2)^2} (1 + \eta), \quad (12)$$

where  $g_A$  is the weak coupling constant,  $M_A$  is the axial mass, and  $\eta$  is a factor that encodes the difference between the charged current axial form factor and the strange axial form factor. This form assumes that both parts of the form

factor have the exact same shape. If the difference is only due to the net spin contribution of the strange quark,  $\Delta s$ , then  $\Delta s = -\eta g_A$  which was found to be  $-0.15 \pm 0.09$  in this analysis. Figure 4a shows the measured data and the cross section fits. Figure 4b shows the bounds on the axial mass  $M_A$  and  $\eta$  from the fit.



(a) Measured neutrino-proton and antineutrino-proton cross sections. (b) Extracted neutral current axial form factor parameters.

Figure 4: Results from the Brookhaven E734 measurement of the neutral current elastic cross section.

A later analysis of the E734 NC elastic data was performed [56] in which the strange part of the electric and magnetic form factors was not assumed to be zero. Four fits to the neutrino-proton and antineutrino-proton cross section data were

performed. In the first fit, only the axial mass was allowed to vary and the strange quark contribution to the electric, magnetic, and axial form factors were all held fixed at zero. In the second, the strange contribution to the electric and magnetic form factors were fixed at zero, but the axial mass and the strange axial form factor were allowed to vary. In the third fit, all three strange form factors and the axial mass were allowed to vary, and in the last fit, the strange form factors were all allowed to vary, but the axial mass was held to the world average at the time,  $M_A = 1.032 \pm 0.036$  GeV. The same form of the axial form factor in Eq. 12 was assumed and the strange electric and magnetic form factors were assumed to have the same shape as the charged current electric and magnetic form factors. The extracted value of  $\Delta s$  ranged from  $\Delta s = -0.13 \pm 0.09$  in the second fit to  $\Delta s = -0.21 \pm 0.10$  in the fourth fit. Each of the extracted  $\Delta s$  values is consistent with the original measurement and with a  $\Delta s$  being negative. Additionally, a strong correlation between  $\Delta s$  and  $M_A$  was again observed. In the first fit with  $\Delta s$  fixed at zero, a best fit to the data was found when  $M_A = 1.086 \pm 0.015$  GeV which is very consistent with the original results in Ref. [3] shown in Fig. 4b.

An additional analysis of the E734 data considering ratios of neutral current elastic interactions to charged current elastic interactions was performed [57]. Specifically, they looked at the neutrino-antineutrino asymmetry

$$\mathcal{A}_p(Q^2) = \frac{\left(\frac{d\sigma}{dQ^2}\right)_{\nu p \rightarrow \nu p} - \left(\frac{d\sigma}{dQ^2}\right)_{\bar{\nu} p \rightarrow \bar{\nu} p}}{\left(\frac{d\sigma}{dQ^2}\right)_{\nu n \rightarrow \mu^- p} - \left(\frac{d\sigma}{dQ^2}\right)_{\bar{\nu} p \rightarrow \mu^+ n}}. \quad (13)$$

This asymmetry has an enhanced sensitivity to the strange axial and magnetic form factors. It was found that the experimental uncertainty was too large to determine  $\Delta s$  and that a large factor of the uncertainty was due to the uncertainty on the axial mass. This analysis also assumed the dipole form of the axial form factor in Eq. 12.

A global analysis of the strange quark contribution to the electromagnetic and axial form factors performed in Ref. [58] included E734 NC elastic neutrino-proton scattering data. The analysis combined the E734 data with parity-violating elastic polarized-electron-proton scattering data from the G0 [59] and HAPPEX [60] experiments. The electron-proton data are sensitive to the strange contribution to the electric and magnetic form factors but not very sensitive to the strange axial form factor. The analysis found the extracted  $\Delta s$  to be consistent with negative values.

### **1.3.2 The MiniBooNE Experiment**

The main target and detector of the MiniBooNE experiment was 800 tons of scintillator oil in a 12.2 m diameter spherical tank. Charged particles from the neutrino interactions in the mineral oil produced Cerenkov light which was collected by 1520 8-inch PMTs surrounding the oil. A schematic of the detector is shown in Fig. 5. MiniBooNE sat in the Booster Neutrino Beam (BNB) at Fermilab that can run in either neutrino or antineutrino mode with an average neutrino



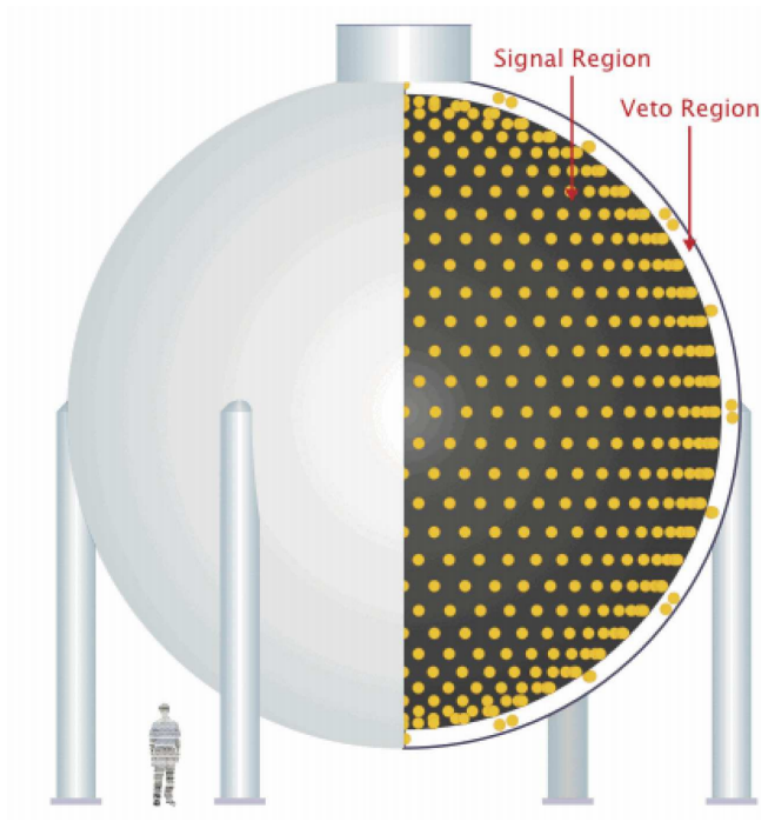


Figure 5: Schematic of the MiniBooNE detector [4].

energy of  $\sim 800$  MeV [61].

The MiniBooNE collaboration performed a  $\Delta s$  fit to the ratio of the neutrino-proton to the neutrino-nucleon NC elastic cross section in the proton kinetic energy between  $T = 350$  MeV and  $T = 800$  MeV. This corresponds to a range of  $Q^2 = 0.66$  GeV<sup>2</sup> to  $Q^2 = 1.5$  GeV<sup>2</sup>. Figure 6 shows the measured ratio and the fits to the data. In the analysis, the dipole shape in Eq. 12 was used for the axial form factor. When the value of the axial mass was held to  $M_A = 1.35$  GeV a value of  $\Delta s = 0.08 \pm 0.26$  was found, and when the axial mass was held to  $M_A = 1.23$  GeV

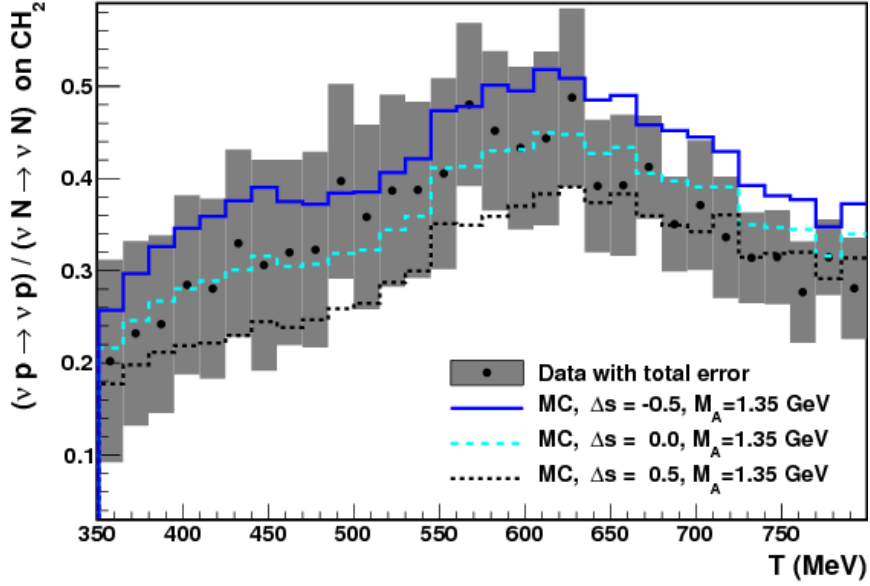


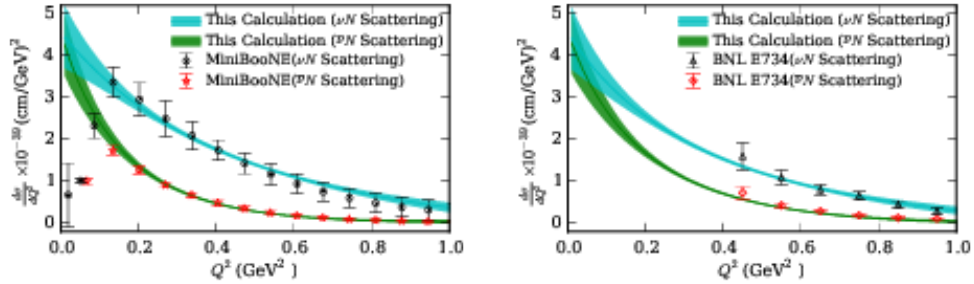
Figure 6: Ratio of the neutrino-proton NC elastic cross section to the neutrino-nucleon NC elastic cross section measured in MiniBooNE [5].

a value of  $\Delta s = 0.00 \pm 0.30$ . Both of these values are consistent with the E734 measurement and with zero.

A later analysis of the MiniBooNE data was performed which included a two-body current contribution to the cross section [62]. The inclusion of the two-body current was done using the NuWro Monte Carlo neutrino event generator [63]. The original MiniBooNE analysis used the NUANCE Monte Carlo neutrino event generator [64]. A simultaneous extraction of  $\Delta s$  and the axial mass from the asymmetry  $\mathcal{A}_p(Q^2)$  assuming the dipole form of the axial form factor in Eq. 12 and including two-body currents resulted in an axial mass value of  $M_A = 1.1^{+0.13}_{-0.15}$  GeV and  $\Delta s = -0.4^{+0.5}_{-0.3}$ . This result of  $\Delta s$  is consistent both with the original Mini-

BooNE result and with zero.

Very recently, another re-analysis of the MiniBooNE result was performed using updated lattice QCD calculations of the strange electric and magnetic form factors and the measured MiniBooNE NC elastic nucleon cross section [6]. Notably, this analysis used a  $z$ -expansion fit to the NC axial form factor (described in Sec. 2.6.1). A value of  $\Delta s = -0.196 \pm 0.127 \pm 0.041$  was found. The result was used to predict the BNL E734 NC elastic  $\nu p$  and  $\bar{\nu} p$  measurements and was found to be consistent. The NC elastic cross sections calculated using the results are shown compared to MiniBooNE data in Fig. 7a and compared to E734 data in Fig. 7.



(a) Extracted NC elastic cross section compared to MiniBooNE data. (b) Extracted NC elastic cross section compared to E734 data.

Figure 7: Neutral current elastic cross section results from Ref. [6]

## 2 Neutrino-Nucleon interactions

This section describes the mathematical foundation for the analysis. It gives a derivation of the neutral current elastic neutrino-proton cross section, how the cross section depends on the vector and axial form factors, and the relationship of the neutral current form factors to the ones measured in charged current scattering. The section ends with a discussion of the shape of the form factors that are used in the analysis.

### 2.1 Two-particle interactions

The differential cross section for two-fermion scattering, shown in Fig. 8, is given by Fermi's Golden Rule [65]

$$\sigma = \int \frac{(2\pi)^4 |\mathcal{M}|^2}{4\sqrt{(p_1 \cdot p_2)^2 - m_1^2 m_2^2}} \times d\Phi_2(p_1 + p_2; p'_1, p'_2), \quad (14)$$

where  $d\Phi_2(p_1 + p_2; p'_1, p'_2)$  is an element of two-body phase space given by

$$d\Phi_2(p_1 + p_2; p'_1, p'_2) = \delta^4(p_1 + p_2 - p'_1 - p'_2) \frac{d^3 \mathbf{p}'_1}{(2\pi)^3 2E'_1} \frac{d^3 \mathbf{p}'_2}{(2\pi)^3 2E'_2}, \quad (15)$$

$p_1$  and  $p_2$  are the incoming four-momenta of the fermions  $f_1$  and  $f_2$ , respectively,  $p'_1$  and  $p'_2$  are the outgoing fermion momenta,  $m_1$  and  $m_2$  are the fermion masses,  $E'_1$  and  $E'_2$  are the outgoing fermion energies, and  $\mathcal{M}$  is the scattering amplitude.

Combining Eqns. 14 and 15 gives

$$\sigma = \int \frac{|\mathcal{M}|^2}{64\pi^2} \frac{\delta^4(p_1 + p_2 - p'_1 - p'_2)}{E'_1 E'_2 \sqrt{(p_1 \cdot p_2)^2 - m_1^2 m_2^2}} d\mathbf{p}'_1 d\mathbf{p}'_2. \quad (16)$$

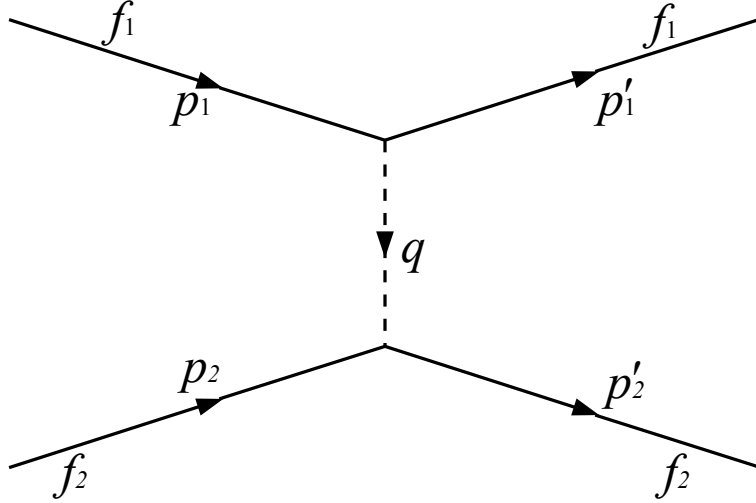


Figure 8: Feynman diagram of two-fermion scattering.

The scattering amplitude is given by the matrix element of the scattering matrix,  $S$ , between the final and initial states ( $\mathcal{M} = \langle f|S|i\rangle$ ). The general form of  $S$  is [65]

$$S = \sum_{n=0}^{\infty} \frac{(-i)^n}{n!} \int \cdots \int d^4x_1 d^4x_2 \cdots d^4x_n T\{\hat{\mathcal{H}}'_I(x_1)\hat{\mathcal{H}}'_I(x_2)\cdots\hat{\mathcal{H}}'_I(x_n)\}, \quad (17)$$

where  $\hat{\mathcal{H}}'_I(x_i)$  is the interaction Hamiltonian density.

The matrix element can more easily be determined using Feynman calculus. For a massive, vector-boson propagator the matrix element for this interaction is given by [65]

$$\mathcal{M} = \langle f|S|i\rangle = \langle p_1'|J^\mu(0)|p_1\rangle \frac{i}{q^2 - M_V^2} (-g_{\mu\nu} + q_\mu q_\nu/M_V^2) \langle p_2'|J^\mu(0)|p_2\rangle, \quad (18)$$

where  $q$  is the four-momentum carried by the vector-boson propagator,  $M_V$  is the mass of the propagator,  $J^\mu(0)$  is the probability current operator, and  $g_{\mu\nu}$  is the metric tensor.

## 2.2 Electroweak interactions

The charged current,  $j_{CC}^\mu$ , which corresponds to the exchange of a  $W^\pm$  boson, and the neutral current,  $j_{NC}^\mu$ , which corresponds to the exchange of the  $Z^0$  boson are given by [66]

$$j_{CC}^\mu = \sum_f \bar{\psi}_f \gamma^\mu (1 - \gamma_5) \frac{1}{2} (\tau_1 + i\tau_2) \psi_f \quad (19)$$

$$j_{NC}^\mu = \sum_f \bar{\psi}_f \gamma^\mu (1 - \gamma_5) \frac{1}{2} (\tau_3) \psi_f - 2 \sin^2(\theta_W) j_{em}^\mu \quad (20)$$

where  $j_{em}^\mu$  is the electromagnetic current,  $\psi_f$  are the weak isospin doublets, and  $\tau_i$  are the Pauli matrices

$$\tau_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \tau_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \tau_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (21)$$

The lepton weak isospin doublets are [66]

$$\psi_e = \begin{pmatrix} \hat{\nu}_e \\ \hat{e}^- \end{pmatrix}, \quad \psi_\mu = \begin{pmatrix} \hat{\nu}_\mu \\ \hat{\mu}^- \end{pmatrix}, \quad \psi_\tau = \begin{pmatrix} \hat{\nu}_\tau \\ \hat{\tau}^- \end{pmatrix}, \quad (22)$$

and the quark weak isospin doublets are

$$\psi_1 = \begin{pmatrix} \hat{u} \\ \hat{d}' \end{pmatrix}, \quad \psi_2 = \begin{pmatrix} \hat{c} \\ \hat{s}' \end{pmatrix}, \quad \psi_3 = \begin{pmatrix} \hat{t} \\ \hat{b}' \end{pmatrix}, \quad (23)$$

where  $d'$ ,  $s'$ , and  $b'$  represent the “mixed” states

$$\begin{pmatrix} \hat{d}' \\ \hat{s}' \\ \hat{b}' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} \hat{d} \\ \hat{s} \\ \hat{b} \end{pmatrix}. \quad (24)$$

The matrix,  $V$ , is the Cabibbo-Kobayashi-Maskawa matrix. These doublets contain the allowed weak transitions where  $V_{ab}$  is the probability of a transition from a quark with flavor  $a$  to a quark with flavor  $b$ .

### 2.2.1 The charged and neutral currents

The linear combination of Pauli matrices in the charged current

$$\frac{1}{2}\tau^+ = \frac{1}{2}(\tau_1 + \tau_2), \quad (25)$$

acts as an “isospin raising matrix” and corresponds to the exchange of a  $W^+$  boson. For the leptons, this gives [66]

$$\begin{aligned} j_{CC}^\mu(\text{leptons}) &= \sum_{l=e,\mu,\tau} \begin{pmatrix} \bar{\nu}_l \\ \bar{l} \end{pmatrix} \gamma^\mu (1 - \gamma_5) \frac{1}{2} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \hat{\nu}_l \\ \hat{l} \end{pmatrix} \\ &= \sum_{l=e,\mu,\tau} \bar{\nu}_l \gamma^\mu (1 - \gamma_5) \frac{1}{2} \hat{l}, \end{aligned} \quad (26)$$

and, similarly, for the quarks we get

$$j_{CC}^\mu(\text{quarks}) = \bar{u} \gamma^\mu (1 - \gamma_5) \frac{1}{2} \hat{d}' + \bar{c} \gamma^\mu (1 - \gamma_5) \frac{1}{2} \hat{s}' + \bar{t} \gamma^\mu (1 - \gamma_5) \frac{1}{2} \hat{b}', \quad (27)$$

with the total charged current being  $j_{CC}^\mu = j_{CC}^\mu(\text{leptons}) + j_{CC}^\mu(\text{quarks})$ .

In neutral current scattering,  $\frac{1}{2}\tau_3$  gives the weak isospin which acts as a “weak charge”. The electromagnetic current is given by

$$j_{em}^\mu = \sum_f Q_f \bar{f} \gamma^\mu \hat{f} \quad (28)$$

where the sum is over the fermion flavors, and  $Q_f$  is the electric charge of  $f$ . So, the total neutral current is

$$\begin{aligned}
j_{NC}^\mu &= \sum_{l=e,\mu,\tau} \left( \bar{\nu}_l \gamma^\mu (1 - \gamma_5) \frac{1}{2} \nu_l - \bar{l} \gamma^\mu (1 - \gamma_5) \frac{1}{2} l + \sin^2 \theta_W \bar{l} \gamma^\mu l \right) \\
&+ \sum_{q=u,c,t} \left( \bar{q} \gamma^\mu (1 - \gamma_5) \frac{1}{2} q - \sin^2 \theta_W \frac{2}{3} \bar{q} \gamma^\mu q \right) \\
&+ \sum_{q=d,s,b} \left( -\bar{q} \gamma^\mu (1 - \gamma_5) \frac{1}{2} q + \sin^2 \theta_W \frac{1}{3} \bar{q} \gamma^\mu q \right).
\end{aligned} \tag{29}$$

### 2.2.2 V–A structure

We can separate the currents into their vector and pseudovector, or axial vector, components. The terms that contain just  $\gamma^\mu$  behave like vectors under a parity transformation [67]

$$\hat{\mathbf{P}} \hat{\psi}(\mathbf{x}, t) \hat{\mathbf{P}}^{-1} \gamma^\mu \hat{\mathbf{P}} \psi(\hat{\mathbf{x}}, t) \hat{\mathbf{P}}^{-1} = -\hat{\mathbf{P}} \hat{\psi}(-\mathbf{x}, t) \hat{\mathbf{P}}^{-1} \gamma^\mu \hat{\mathbf{P}} \psi(-\hat{\mathbf{x}}, t) \hat{\mathbf{P}}^{-1} \tag{30}$$

where  $\hat{\mathbf{P}}$  is the parity operator

$$\mathbf{P} : \mathbf{x} \rightarrow -\mathbf{x}, t \rightarrow t. \tag{31}$$

The terms that contain  $\gamma^\mu \gamma_5$  behave like axial vectors under a parity transformation

$$\hat{\mathbf{P}} \hat{\psi}(\mathbf{x}, t) \hat{\mathbf{P}}^{-1} \gamma^\mu \gamma_5 \hat{\mathbf{P}} \psi(\hat{\mathbf{x}}, t) \hat{\mathbf{P}}^{-1} = +\hat{\mathbf{P}} \hat{\psi}(-\mathbf{x}, t) \hat{\mathbf{P}}^{-1} \gamma^\mu \gamma_5 \hat{\mathbf{P}} \psi(-\hat{\mathbf{x}}, t) \hat{\mathbf{P}}^{-1}. \tag{32}$$

The charged current is

$$\begin{aligned}
j_{CC}^\mu &= \frac{g}{\sqrt{2}} \left[ \sum_{l=e,\mu,\tau} \bar{\nu}_l (\gamma^\mu - \gamma^\mu \gamma_5) \frac{1}{2} l \right. \\
&\quad \left. + \bar{u} (\gamma^\mu - \gamma^\mu \gamma_5) \frac{1}{2} d' + \bar{c} (\gamma^\mu - \gamma^\mu \gamma_5) \frac{1}{2} s' + \bar{t} (\gamma^\mu - \gamma^\mu \gamma_5) \frac{1}{2} b' \right],
\end{aligned} \tag{33}$$



and the neutral current becomes

$$j_{NC}^\mu = \frac{g}{2 \cos \theta_W} \left[ \sum_{l=e,\mu,\tau} \left( \tilde{\nu}_l (g_V^l \gamma^\mu - g_A^l \gamma^\mu \gamma_5) \hat{\nu}_l + \tilde{l} (g_V^l \gamma^\mu - g_A^l \gamma^\mu \gamma_5) \hat{l} \right) + \sum_{q=u,d,c,s,t,b} \left( -\tilde{q} (g_V^q \gamma^\mu - g_A^q \gamma^\mu \gamma_5) \hat{q} \right) \right]. \quad (34)$$

where

$$g_V^f = \frac{1}{2} \tau_3^f - 2 \sin^2 \theta_W Q_f, \quad g_A^f = \frac{1}{2} \tau_3^f. \quad (35)$$

### 2.3 Electroweak Scattering Matrix Elements

To calculate the charged-current neutrino-nucleon scattering matrix element, as shown in Fig. 9a, we need to determine the lepton CC matrix element,  ${}_l \langle k' | J_{CC}^\mu(0) | k \rangle_{\nu_l}$ , and the nucleon CC matrix element,  ${}_p \langle p' | J_{CC}^\mu(0) | p \rangle_n$ , and for neutral-current, as shown in Fig. 9b, we need the lepton NC matrix element,  ${}_{\nu_l} \langle k' | J_{NC}^\mu(0) | k \rangle_{\nu_l}$ , and the nucleon NC matrix element,  ${}_p \langle p' | J_{NC}^\mu(0) | p \rangle_p$ . Leptons are point-like particles, so their matrix elements are straight-forward

$$\begin{aligned} {}_l \langle k' | J_{CC}^\mu(0) | k \rangle_{\nu_l} &= -i \frac{G_F}{\sqrt{2}} \bar{u}(k') (\gamma^\mu - \gamma^\mu \gamma_5) u(k), \\ {}_{\nu_l} \langle k' | J_{NC}^\mu(0) | k \rangle_{\nu_l} &= -\frac{G_F}{\sqrt{2}} \bar{u}(k') (\gamma^\mu - \gamma^\mu \gamma_5) u(k). \end{aligned} \quad (36)$$

#### 2.3.1 Nucleon matrix elements

Since nucleons have a finite structure, the nucleon matrix elements have a more complicated form. The current is corrected by form factors which encode the longitudinal nucleon structure.

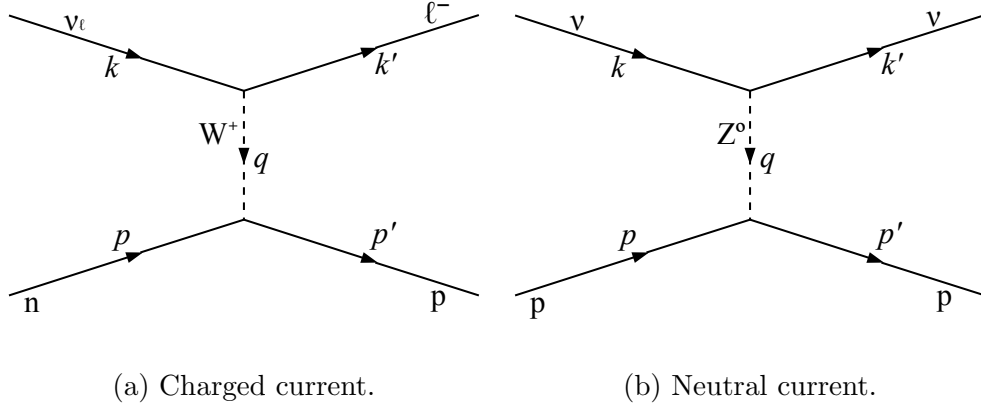


Figure 9: Feynman diagrams of elastic neutrino-nucleon scattering.

If we define the vector and axial parts of the nucleon currents by

$$\begin{aligned}
 v_i^\mu &= \bar{\psi}_f \gamma^\mu \frac{1}{2} \tau_i \hat{\psi}_f, \\
 a_i^\mu &= \bar{\psi}_f \gamma^\mu \gamma_5 \frac{1}{2} \tau_i \hat{\psi}_f,
 \end{aligned}
 \tag{37}$$

where  $\hat{\psi}$  are the quark doublets and  $\tau_i$  are the Pauli matrices still, then the nucleon charged and neutral current become

$$j_{CC}^\mu = v_+^\mu - a_+^\mu,
 \tag{38}$$

$$j_{NC}^\mu = v_3^\mu - a_3^\mu - 2 \sin^2 \theta_W j_{em;q}^\mu,$$

where  $v_+^\mu = v_1^\mu + i v_2^\mu$  and  $a_+^\mu = a_1^\mu + i a_2^\mu$ . The nucleon matrix elements become

$${}_p \langle p' | J_{CC}^\mu | p \rangle_n = {}_p \langle p' | V_+^\mu | p \rangle_n - {}_p \langle p' | A_+^\mu | p \rangle_n,
 \tag{39}$$

$${}_p \langle p' | J_{NC}^\mu | p \rangle_p = {}_p \langle p' | V_3^\mu | p \rangle_p - {}_p \langle p' | A_3^\mu | p \rangle_p - {}_p \langle p' | J_{em}^\mu | p \rangle_p,$$

where  $J^\mu$ ,  $V^\mu$ , and  $A^\mu$  are current operators.

### 2.3.2 Electromagnetic Scattering

It is easiest to first find an equation for  $j_{em;q}^\mu$  in terms of the electric and magnetic form factors. The nucleon electromagnetic current should have the same vector

form as the electromagnetic current for point-like particles. The available physical variables to construct the nucleon EM current are  $p$ ,  $p'$ ,  $\gamma^\mu$ . The most general form is [67]

$$\Gamma^\mu = \gamma^\mu \cdot A + (p'^\mu + p^\mu) \cdot B + (p'^\mu - p^\mu) \cdot C, \quad (40)$$

where  $\Gamma^\mu$  is given by the equation  ${}_p\langle p'|J_{em}^\mu(0)|p\rangle_p = \bar{u}(p')\Gamma^\mu u(p)$ , and  $A$ ,  $B$ , and  $C$  are arbitrary form factors.  $\Gamma^\mu$  is constrained further by the Ward identity [68],  $q_\mu\Gamma^\mu = 0$ , which guarantees current conservation. The  $\gamma^\mu$  and  $(p'^\mu + p^\mu)$  terms in  $q_\mu\Gamma^\mu$  go to zero, but the  $(p'^\mu - p^\mu)$  term does not, so  $C = 0$ . Using the Gordon identity [69], the general form for the nucleon EM current matrix element is

$${}_p\langle p'|J_{em}^\mu(0)|p\rangle_p = \bar{u}(p') \left[ \gamma^\mu F_1(Q^2) + \frac{i\sigma^{\mu\nu}q_\nu}{2M} F_2(Q^2) \right] u(p). \quad (41)$$

The form factors  $F_1$  and  $F_2$  are the Dirac and Pauli form factors, respectively, and they are functions of  $Q^2 = -q^2$ , the negative four-momentum transfer. They can be transformed into the Sachs form factors,  $G_E$  and  $G_M$  by the relationships

$$G_E(Q^2) = F_1(Q^2) - \frac{Q^2}{4M^2} F_2(Q^2), \quad G_M(Q^2) = F_1(Q^2) + F_2(Q^2). \quad (42)$$

The electric form factor,  $G_E$  encodes the longitudinal electric charge structure of the nucleon and the magnetic form factor,  $G_M$ , encodes the longitudinal magnetic structure. At the limit when  $Q^2$  goes to zero, the Sachs form factors become the net charge and magnetic moment of the nucleon.

$$\begin{aligned} G_{E;p}(Q^2 = 0) &= 1, & G_{M;p}(Q^2 = 0) &= \mu_p, \\ G_{E;n}(Q^2 = 0) &= 0, & G_{M;n}(Q^2 = 0) &= \mu_n, \end{aligned} \quad (43)$$

where  $\mu_p$  and  $\mu_n$  are the proton and neutron magnetic moments.

### 2.3.3 Vector current

The quark part of the electromagnetic current,  $j_{em;q}^\mu$ , can be written as [67]

$$\begin{aligned} j_{em;q}^\mu &= \bar{\psi}_f Q_f \gamma^\mu \hat{\psi}_f \\ &= \bar{\psi}_f \left( \tau_3 + \frac{1}{6} \right) \gamma^\mu \hat{\psi}_f \\ &= v_3^\mu + v_0^\mu, \end{aligned} \quad (44)$$

where  $v_0^\mu = \frac{1}{6} \bar{\psi}_f \gamma^\mu \hat{\psi}_f$ . Now we can write the nucleon electromagnetic current matrix element in terms of the vector and isoscalar parts

$$\begin{aligned} {}_{\text{p(n)}} \langle p' | J_{em}^\mu(0) | p \rangle_{\text{p(n)}} &= {}_{\text{p(n)}} \langle p' | V_3^\mu(0) + V_0^\mu(0) | p \rangle_{\text{p(n)}} \\ &= {}_{\text{p(n)}} \langle p' | V_3^\mu(0) | p \rangle_{\text{p(n)}} + {}_{\text{p(n)}} \langle p' | V_0^\mu(0) | p \rangle_{\text{p(n)}}, \end{aligned} \quad (45)$$

Since  $V_3^\mu$  behaves as a vector under the charge symmetry operator and  $V_0^\mu$  behaves as a scalar, the following equations are true

$$\begin{aligned} {}_{\text{p}} \langle p' | V_3^\mu(0) | p \rangle_{\text{p}} &= - {}_{\text{n}} \langle p' | V_3^\mu(0) | p \rangle_{\text{n}}, \\ {}_{\text{p}} \langle p' | V_0^\mu(0) | p \rangle_{\text{p}} &= + {}_{\text{n}} \langle p' | V_0^\mu(0) | p \rangle_{\text{n}}. \end{aligned} \quad (46)$$

Then

$${}_{\text{p}} \langle p' | V_3^\mu(0) | p \rangle_{\text{p}} = \frac{1}{2} \left[ {}_{\text{p}} \langle p' | J_{em}^\mu(0) | p \rangle_{\text{p}} - {}_{\text{n}} \langle p' | J_{em}^\mu(0) | p \rangle_{\text{n}} \right], \quad (47)$$

$${}_{\text{p}} \langle p' | V_0^\mu(0) | p \rangle_{\text{p}} = \frac{1}{2} \left[ {}_{\text{p}} \langle p' | J_{em}^\mu(0) | p \rangle_{\text{p}} + {}_{\text{n}} \langle p' | J_{em}^\mu(0) | p \rangle_{\text{n}} \right]. \quad (48)$$

Combining equations 41 and 47 gives

$${}_{\text{p}} \langle p' | V_3^\mu(0) | p \rangle_{\text{p}} = \bar{u}(p') \left[ \gamma^\mu F_1^V(Q^2) + \frac{i\sigma^{\mu\nu} q_\nu}{2M} F_2^V(Q^2) \right] u(p), \quad (49)$$

where the vector form factors,  $F_1^V$  and  $F_2^V$  are defined as

$$\begin{aligned} F_1^V(Q^2) &= \frac{1}{2} (F_{1,p}(Q^2) - F_{1,n}(Q^2)) , \\ F_2^V(Q^2) &= \frac{1}{2} (F_{2,p}(Q^2) - F_{2,n}(Q^2)) . \end{aligned} \tag{50}$$

The isoscalar nucleon matrix element is then

$$\langle p' | V_0^\mu(0) | p \rangle = \bar{u}(p') \left[ \gamma^\mu \frac{1}{2} (F_{1,p}(Q^2) + F_{1,n}(Q^2)) + \frac{i\sigma^{\mu\nu} q_\nu}{2M} \frac{1}{2} (F_{2,p}(Q^2) + F_{2,n}(Q^2)) \right] u(p) \tag{51}$$

Under the conserved vector current (CVC) hypothesis [70, 71] the “raising” and “lowering” vector currents,  $v_\pm^\mu$ , in the charged current interactions are the same as the vector part of the electromagnetic current,  $v_3^\mu$ .

### 2.3.4 Axial current

The axial current nucleon matrix elements can also be parameterized in terms of form factors. Starting with the charged axial current, the most general form contains an axial, a pseudoscalar, and a tensor term, [67]

$${}_p \langle p' | A_+^\mu(0) | p \rangle_n = \bar{u}(p') \left[ \gamma^\mu \gamma^5 G_A(Q^2) + \frac{q^\mu \gamma^5}{2M} G_P^{CC}(Q^2) + \frac{i\sigma^{\mu\nu} q_\nu \gamma^5}{2M} G_T^{CC}(Q^2) \right] u(p) , \tag{52}$$

where  $G_A$ ,  $G_P^{CC}$ , and  $G_T^{CC}$  are the axial form factors. From isospin symmetry, we have

$${}_p \langle p' | A_+^\mu(0) | p \rangle_n = {}_n \langle p' | A_-^\mu(0) | p \rangle_p \equiv {}_p \langle p | A_+^\mu(0) | p' \rangle_n^* , \tag{53}$$

which implies that  $G_T^{CC}(Q^2) = 0$ . In quasi-elastic scattering the pseudoscalar term containing  $G_P^{CC}(Q^2)$  is proportional to the lepton mass and can be ignored

in neutral current scattering since  $m_\nu \approx 0$ . The charged current axial matrix element is then just

$${}_p\langle p' | A_+^\mu(0) | p \rangle_n = \bar{u}(p') \gamma^\mu \gamma^5 G_A(Q^2) u(p). \quad (54)$$

The form factor  $G_A$  encodes the longitudinal spin structure of the nucleon due to the spin of the up and down quarks and is referred to as the charged current axial form factor, or just the axial form factor.

The most general form for the neutral axial current nucleon matrix element is similarly

$${}_p\langle p' | A_3^\mu(0) | p \rangle_n = \bar{u}(p') \left[ \gamma^\mu \gamma^5 G_A^{NC}(Q^2) + \frac{q^\mu \gamma^5}{2M} G_P^{NC}(Q^2) + \frac{i\sigma^{\mu\nu} q_\nu \gamma^5}{2M} G_T^{NC}(Q^2) \right] u(p). \quad (55)$$

Just as in the charged-current case, the tensor part is zero,  $G_T^{NC}(Q^2) = 0$ , and we can again neglect  $G_P^{NC}$  which is proportional to the lepton mass. The neutral axial current matrix element is

$${}_p\langle p' | A_3^\mu(0) | p \rangle_n = \bar{u}(p') \gamma^\mu \gamma^5 G_A^{NC}(Q^2) u(p). \quad (56)$$

which can be related to the charged current axial form factor through isospin symmetry. The relationship between the neutral and charged axial current operators is

$$[I^k, A_j^\mu] = i\epsilon^{kjl} A_l^\mu, \quad (57)$$

where  $I^k$  is the total isospin operator, and  $\epsilon^{kjl}$  is the antisymmetric tensor. From

this relationship, we get

$${}_p\langle p'|A_3^\mu(0)|p\rangle_p = \frac{1}{2} {}_p\langle p'|A_+^\mu(0)|p\rangle_n, \quad (58)$$

which implies

$$G_A^{NC}(Q^2) = \frac{1}{2} G_A(Q^2), \quad (59)$$

assuming only contributions from up and down quarks.

## 2.4 Strangeness in the nucleon

If contributions to the nucleon from quarks heavier than the strange are neglected, the quark part of the charged and neutral currents can be separated between the light quarks and the strange quark, [67]

$$\begin{aligned} j_{CC}^\mu(\text{quarks}) &= \bar{u}(\gamma^\mu - \gamma^\mu\gamma_5)\frac{1}{2}\tau_\pm\hat{u} - \bar{d}(\gamma^\mu - \gamma^\mu\gamma_5)\frac{1}{2}\tau_\pm\hat{d} \\ &\equiv \bar{N}(\gamma^\mu - \gamma^\mu\gamma_5)\frac{1}{2}\tau_\pm\hat{N}, \end{aligned} \quad (60)$$

$$\begin{aligned} j_{NC}^\mu(\text{quarks}) &= \bar{N}(\gamma^\mu - \gamma^\mu\gamma_5)\frac{1}{2}\tau_3\hat{N} - \bar{s}(\gamma^\mu - \gamma^\mu\gamma_5)\frac{1}{2}\tau_3\hat{s} \\ &\quad - 2\sin^2\theta_W j_{em;q}^\mu. \end{aligned} \quad (61)$$

### 2.4.1 Strange currents

If we redefine the neutral vector currents as

$$\begin{aligned} v_3^\mu &\equiv \bar{N}\gamma^\mu\frac{1}{2}\tau_3\hat{N} \\ a_3^\mu &\equiv \bar{N}\gamma^\mu\gamma_5\frac{1}{2}\tau_3\hat{N}, \end{aligned} \quad (62)$$

and define the strange part of the currents as

$$\begin{aligned} v_s^\mu &\equiv \bar{\hat{s}}\gamma^\mu\hat{s} \\ a_s^\mu &\equiv \bar{\hat{s}}\gamma^\mu\gamma_5\hat{s}, \end{aligned} \tag{63}$$

we can combine Eqs. 60, 61, 62, and 63 to get

$$j_{CC}^\mu(\text{quarks}) = (v_3^\mu - a_3^\mu), \tag{64}$$

$$j_{NC}^\mu(\text{quarks}) = (v_3^\mu - a_3^\mu) - \frac{1}{2}(v_s^\mu - a_s^\mu) - 2\sin^2\theta_W j_{em;q}^\mu. \tag{65}$$

The quark part of the electromagnetic current can also be separated into light and heavy quark components. Separating the vector and isoscalar terms from Eq. 44, into quark components gives

$$j_{em;q}^\mu = (v_3^\mu + v_0^\mu) - \frac{1}{2}(v_s^\mu + v_{0s}^\mu), \tag{66}$$

where  $v_3^\mu$  and  $v_s^\mu$  are as defined in Eqs. 62 and 63, and  $v_0^\mu$  and  $v_{0s}^\mu$  are defined as

$$\begin{aligned} v_0^\mu &\equiv \frac{1}{6}\bar{\hat{N}}\gamma^\mu\hat{N}, \\ v_{0s}^\mu &\equiv -\frac{1}{3}\bar{\hat{s}}\gamma^\mu\hat{s}. \end{aligned} \tag{67}$$

The quark part of the neutral current becomes

$$j_{NC}^\mu(\text{quarks}) = (1 - 2\sin^2\theta_W)(v_3^\mu - \frac{1}{2}v_s^\mu) - (a_3^\mu - \frac{1}{2}a_s^\mu) - 2\sin^2\theta_W(v_0^\mu - \frac{1}{2}v_{0s}^\mu). \tag{68}$$

#### 2.4.2 Strange nucleon matrix elements

The single-nucleon matrix elements for the charged current becomes

$${}_p\langle p' | J_{CC}^\mu(0) | p \rangle_n = {}_p\langle p' | V_3^\mu(0) | p \rangle_n - {}_p\langle p' | A_3^\mu(0) | p \rangle_n. \tag{69}$$



In terms of the vector and axial form factors, the matrix element is

$${}_p\langle p'|J_{CC}^\mu(0)|p\rangle_n = \bar{u}(p') \left[ \gamma^\mu F_1^V(Q^2) + \frac{i\sigma^{\mu\nu}q_\nu}{2M} F_2^V(Q^2) - \gamma^\mu \gamma_5 G_A(Q^2) \right] u(p). \quad (70)$$

The single-nucleon matrix elements for the neutral current becomes

$$\begin{aligned} {}_p\langle p'|J_{NC}^\mu(0)|p\rangle_p &= (1 - 2\sin^2\theta_w)({}_p\langle p'|V_3^\mu(0)|p\rangle_p - \frac{1}{2}{}_p\langle p|V_s^\mu(0)|p\rangle_p) \\ &\quad - ({}_p\langle p'|A_3^\mu(0)|p\rangle_p - \frac{1}{2}{}_p\langle p|A_s^\mu(0)|p\rangle_p) \\ &\quad - 2\sin^2\theta_w({}_p\langle p'|V_0^\mu(0)|p\rangle_p - \frac{1}{2}{}_p\langle p|V_{0s}^\mu(0)|p\rangle_p). \end{aligned} \quad (71)$$

If we ignore the strange components of the Dirac and Pauli form factors ( $F_1$  and  $F_2$ ), the matrix element is

$$\begin{aligned} {}_p\langle p'|J_{NC}^\mu(0)|p\rangle_p &= \bar{u}(p') \left[ (1 - \sin^2\theta_w) \left\{ \gamma^\mu F_1^{NC}(Q^2) + \frac{i\sigma^{\mu\nu}q_\nu}{2M} F_2^{NC}(Q^2) \right\} \right. \\ &\quad \left. - \gamma^\mu \gamma_5 G_A^{NC}(Q^2) - 2\sin^2\theta_w \left\{ \gamma^\mu F_1^p(Q^2) + \frac{i\sigma^{\mu\nu}q_\nu}{2M} F_2^p(Q^2) \right\} \right] u(p), \end{aligned} \quad (72)$$

where

$$F_{1,2}^{NC}(Q^2) = \frac{1}{2}F_{1,2}^V(Q^2), \quad (73)$$

and

$$G_A^{NC} = \frac{1}{2}G_A(Q^2) - \frac{1}{2}G_A^s(Q^2). \quad (74)$$

The neutral current axial form factor,  $G_A^{NC}$ , represents the longitudinal spin structure of the nucleon due to all three of the quark flavors (up, down, and strange).

## 2.5 Neutrino-proton neutral current elastic cross section

We now have everything to calculate the neutrino-proton neutral current elastic cross section. Figure 10 shows the Feynman diagram of the interaction. The neu-

trino is represented by the letter  $\nu$  with incoming four-momentum  $k$  and outgoing four-momentum  $k'$ . The proton is represented by the letter  $p$  with incoming four-momentum  $p$  and outgoing four-momentum  $p'$ . The four-momentum transferred by the  $Z^0$  boson is  $q$  with  $q = k - k' = p' - p$ .

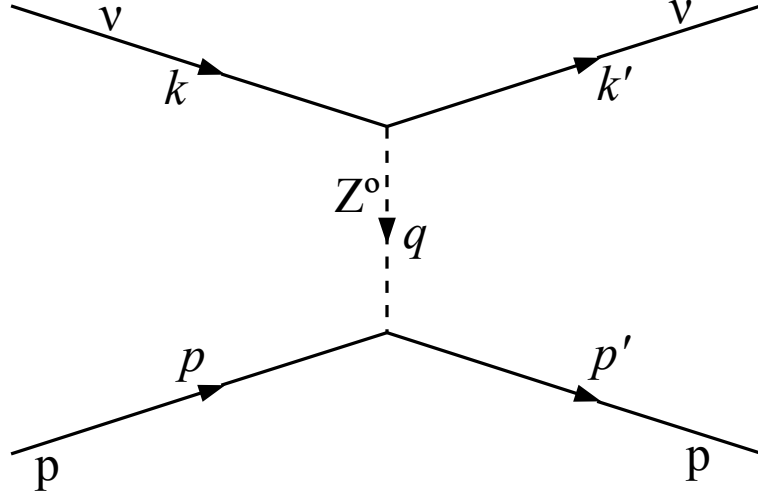


Figure 10: Feynman diagram of neutrino-proton neutral current elastic scattering.

Combining Eqs. 16, 18, 36, and 72, averaging over the spin states, and setting the outgoing neutrino mass to zero gives (in the Llewellyn-Smith formalism) [72]

$$\left(\frac{d\sigma}{dQ^2}\right)^{NC} = \frac{G_F^2 M_p^2}{8\pi E_\nu^2} \left[ A - \frac{(4M_p E_\nu - Q^2)}{M_p^2} B + \frac{(4M_p E_\nu - Q^2)^2}{M_p^4} C \right] \quad (75)$$

with

$$A = 4\tau \left[ (1 + \tau)(G_A^{NC})^2 - (1 - \tau)(F_2^{NC})^2 + \tau(1 - \tau)(F_2^{NC})^2 + 4\tau F_1^{NC} F_2^{NC} \right], \quad (76)$$

$$B = 4\tau [G_A^{NC}(F_1^{NC} + F_2^{NC})], \quad (77)$$

$$C = \frac{1}{4} [(G_A^{NC})^2 + (F_1^{NC})^2 + (F_2^{NC})^2], \quad (78)$$

where  $E_\nu$  is the incoming neutrino energy,  $M_P$  is the proton mass,  $\tau = \frac{Q^2}{4M_P}$ , and the form factors are functions of  $Q^2$ .

## 2.6 Determination of the Nucleon Form Factors

Up to this point, we have only parameterized the electromagnetic and weak currents in terms of the vector and axial form factors, but have said nothing about their form. The actual form factors are determined empirically from experiment. To determine the form factors from experimental data some parameterization must be chosen.

There are many existing parameterizations of the electromagnetic form factors which have been fit to electromagnetic scattering data. Since there are much more data available from electromagnetic scattering than from weak scattering, the parameterizations of the electromagnetic form factors can be more precise with more parameters. See [73] for an extensive review of the nucleon electromagnetic form factors.

The most common parameterization of the axial form factor, the dipole form, has only one free parameter known as the axial mass,  $M_A$ ,

$$G_A^{(dipole)}(Q^2) = \frac{G_A(0)}{(1 + Q^2/M_A^2)^2}. \quad (79)$$

The dipole form of the axial form factor was motivated by early models of the electromagnetic form factors which are no longer used. There hasn't been strong

physical motivation for new axial form factor parameterizations since not as much data exist as for the electromagnetic form factors.

Constraining the shape of the axial form factor to one parameter can lead to both overconfidence in the uncertainty of the measurement of the form factor and disagreement of the parameters between different experimental measurements at different ranges of negative four-momentum transfer squared [74].

### 2.6.1 Model-Independent Form Factor Parameterization

Both the electromagnetic and axial form factors can be determined using a model-independent parameterization referred to as  $z$  expansion [75].

The  $z$  expansion parameterization is made by mapping  $Q^2$  onto a domain where the form factor,  $F$  is analytic

$$z(Q^2, t_{cut}, t_0) = \frac{\sqrt{t_{cut} + Q^2} - \sqrt{t_{cut} - t_0}}{\sqrt{t_{cut} + Q^2} + \sqrt{t_{cut} - t_0}}, \quad (80)$$

where  $t_{cut}$  is the leading threshold for states that can be produced by the vector or axial current. The form factor is analytic everywhere where  $Q^2 \geq -t_{cut}$ . In the case of the electromagnetic (vector) form factors,  $t_{cut} = (2m_\pi)^2$  (two-pion threshold). In the axial form factor case,  $t_{cut} = (3m_\pi)^2$  (three-pion threshold) [76]. The parameter  $t_0$  is an arbitrary number that can be chosen to minimize the absolute value of  $z$  [77].

Since  $F(z(Q^2))$  is analytic by definition of  $z$ , and  $z$  can be constrained to be

less than unity by choice of  $t_0$  and the finite range of  $Q^2$  for a given experiment, the Taylor series of  $F(z(Q^2))$  around zero will converge to the true  $F(Q^2)$

$$F(Q^2) = \sum_{k=0}^{\infty} a_k z(Q^2)^k, \quad (81)$$

where  $a_k$  are dimensionless parameters that encode the nucleon structure.

Physical quantities can still be extracted from the general form of the form factors. For example, the electric charge radius is still defined by the slope of the electric form factor at  $Q^2 = 0$  and the axial mass can be redefined by the slope of the axial form factor at  $Q^2 = 0$ . Importantly, the net contribution to the proton spin from the individual quark spins,  $\Delta q$ , is the value of the axial form factor at  $Q^2 = 0$ , and specifically, the net contribution to the spin of the proton from the spin of the strange quarks in the nucleon is equal to the value of the strange axial form factor at  $Q^2 = 0$ .

### 2.6.2 Fits of $z$ Expansion Form Factors to Data

An approximation to the general form of the form factors can be made by fitting the coefficients,  $a_k$ , out to a value  $k_{max}$

$$F(Q^2) \approx \sum_{k=0}^{k_{max}} a_k z(Q^2)^k. \quad (82)$$

Obviously, the larger  $k_{max}$  is, the better the approximation will be. The limit of  $k_{max}$  is generally determined by the experimental data.

According to asymptotic scaling predictions in QCD [78], the vector and axial form factors should have a  $1/Q^4$  behavior at large values of  $Q^2$ . This can be encoded in the  $z$  expansion parameterization by enforcing four sum rules [79]

$$\left. \frac{d^n}{dz^n} F \right|_{z=1} = 0, \quad n = 0, 1, 2, 3. \quad (83)$$

Enforcing these sum rules means that there are fewer than  $k_{max}$  free parameters in the  $z$  expansion fit.

The optimal value for  $t_0$  can be chosen to minimize the maximum size of  $|z|$  [77]

$$t_0^{\text{optimal}}(Q^2) = t_{cut} \left( 1 - \sqrt{1 + Q_{\text{max}}^2/t_{cut}} \right), \quad (84)$$

where  $Q_{\text{max}}^2$  is the maximum  $Q^2$  in the data being fit to.

Fits of the electric and magnetic form factor  $z$  expansion coefficients to data have been performed in [80, 81, 79, 82]. We use the recent fit to electron scattering data performed in [82]. In it, the proton electric and magnetic form factors,  $G_E^p(Q^2)$  and  $G_M^p(Q^2)$ , are fit simultaneously up to  $k_{max} = 12$  (seven free parameters) using a previous unpolarized electron-proton scattering data and  $G_E^p/G_M^p$  ratios extracted from polarized electron-proton data. The neutron electric and magnetic form factors,  $G_E^n(Q^2)$  and  $G_M^n(Q^2)$ , are fit separately to up  $k_{max} = 10$  (five free parameters) using polarized and unpolarized electron-deuterium and electron-helium-3 scattering data.

Fits to the charged current axial form factor  $z$  expansion coefficients to neutrino data have been performed in [74, 83, 77] We use the fit to deuterium

bubble chamber neutrino data performed in [77]. The form factor,  $G_A(Q^2)$ , is fit up to  $k_{max} = 8$  (four free parameters) using accelerator neutrino-deuterium data from the deuterium bubble chamber experiments at Argonne National Lab (ANL) [84, 85, 86], BNL [87], and FNAL [88]. They found results similar to a dipole form but with larger, more realistic uncertainty estimates.

### 2.6.3 $z$ Expansion Fit to the Neutral Current Axial Form Factor

There are no existing fits of the neutral current axial form factor to data using the  $z$  expansion parameterization. In this analysis, we do a three-parameter fit of the strange part of the NC axial form factor to MicroBooNE neutral current elastic neutrino-proton scattering data and use the previous fit to the CC axial form factor in [77] for the up and down quark spin contributions.

$$G_A^{NC}(Q^2) = G_A(Q^2) \text{ (previous fit)} \\ + G_A^s(Q^2) \text{ (fit to MicroBooNE data)}. \quad (85)$$

The  $Q^2$  range of the NC elastic data that we will use for the fit in MicroBooNE is from  $0.1 \text{ GeV}^2$  to  $1.0 \text{ GeV}^2$ . Because the MicroBooNE NC elastic proton data set has relatively low statistics (see Sec. 5.3), we do a fit to the data with only three free parameters. This corresponds to  $k_{max} = 6$  with  $a_i^s$  for  $i = 3, 4, 5, 6$  determined by  $a_i^s$  for  $i = 1, 2, 3$ , and the sum rules in Eq. 83. The strange axial form factor is written as

$$G_A^s(z) = a_0^s + a_1^s z + a_2^s z^2 + a_3^s z^3 + a_4^s z^4 + a_5^s z^5 + a_6^s z^6, \quad (86)$$

with  $a_i^s$  ( $i = 0, 1, 2$ ) free parameters to fit to data, and

$$a_3^s = -20a_0^s - 10a_1^s - 4a_2^s, \quad (87)$$

$$a_4^s = +45a_0^s + 20a_1^s + 6a_2^s, \quad (88)$$

$$a_5^s = -36a_0^s - 15a_1^s - 4a_2^s, \quad (89)$$

$$a_6^s = +10a_0^s + 4a_1^s + a_2^s. \quad (90)$$

where  $a_i^s$  are the coefficients of the strange axial form factor that we are fitting to data, and  $z \equiv z(Q^2)$  from Eq. 80.

The net contribution of the strange quark spin to the spin of the proton,  $\Delta s$ , is defined as the value of the strange axial form factor at  $Q^2 = 0$ ,

$$\Delta s = a_0^s + a_1^s z_0 + a_2^s z_0^2 + a_3^s z_0^3 + a_4^s z_0^4 + a_5^s z_0^5 + a_6^s z_0^6, \quad (91)$$

where

$$z_0 = z(Q^2 = 0, t_{cut}, t_0) = \frac{\sqrt{t_{cut}} - \sqrt{t_{cut} - t_0}}{\sqrt{t_{cut}} + \sqrt{t_{cut} - t_0}}. \quad (92)$$

The specific fitting procedure and parameters are described in Sec. 6.



### 3 The MicroBooNE experiment

MicroBooNE is an accelerator neutrino experiment at Fermilab. The experiment measures neutrino physics properties by studying the interactions of neutrinos produced by Fermilab accelerators. The MicroBooNE detector is a liquid argon time projection chamber (LArTPC) with an additional light collection system. LArTPC technology is relatively new, and allows for high-resolution imaging of the neutrino interactions in the liquid argon. The production of the neutrino beam, the MicroBooNE LArTPC, and the light collection system are described in this section.

#### 3.1 MicroBooNE Physics Goals

MicroBooNE was designed to investigate the low-energy electron-like excess seen in MiniBooNE [20, 21]. MiniBooNE observed an excess of electron like tracks in the muon neutrino to electron neutrino oscillation search in the neutrino energy range below 0.5 GeV. Figure 11 shows the observed excess from the most recent analysis of the data in [21]. If the excess is due to the  $\nu_\mu$  to  $\nu_e$  oscillations, it cannot be described by the standard three neutrino oscillation model, and it may be due to new physics such a fourth generation of neutrinos that do not interact through any of the standard model forces. The MiniBooNE experiment uses an oil Cerenkov detector which is unable to distinguish between electrons and photons

at these low energies, and the excess might be due to a photon background. None of the known photon production mechanisms account for the excess of events seen.

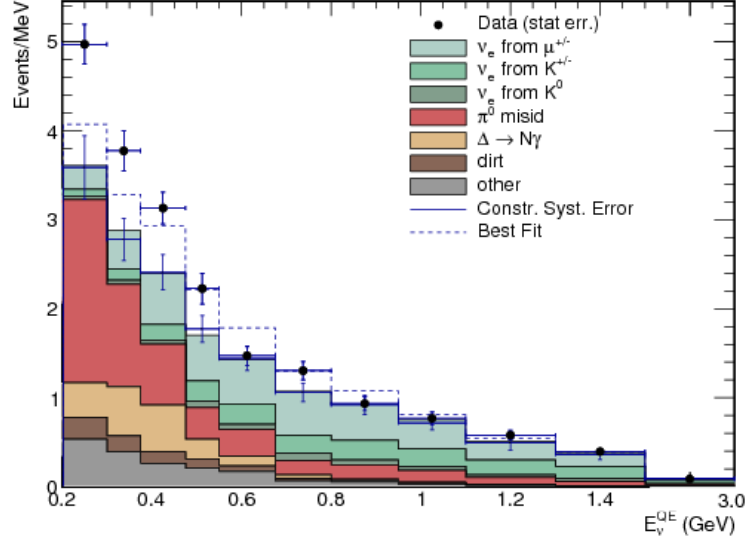


Figure 11: Observed low energy electron-like excess in MiniBooNE.

MicroBooNE was designed to be able to distinguish between low-energy electrons and photons and will be able to determine whether the excess is due to electrons or photons. One advantage of MicroBooNE is its high tracking resolution. This allows MicroBooNE to determine if a charged track started at the neutrino vertex or if there is a gap indicating that a neutral particle produced the charged one, like a photon pair producing into two electrons. Another advantage is MicroBooNE’s ability to determine the energy deposited along the length of the track. This allows MicroBooNE to determine if the energy deposited is consistent with one electron or two that would be produced by a photon. The details of

MicroBooNE's design are given in Sec. 3.3.

The design of MicroBooNE also allows it to study interesting regions of neutrino cross sections. Because of the high tracking and calorimetry resolution, the fine details of the neutrino interaction final states are visible [89]. This allows MicroBooNE to study unknown properties of scattering off nucleons in a nucleus. MicroBooNE's high resolution also allows it to search for exotic physics like neutrinos from supernova explosions and proton decay events. Lastly, MicroBooNE provides important research and development for future liquid argon TPC programs such as DUNE [90].

### **3.2 The Booster Neutrino Beam**

The neutrino beam used by the MicroBooNE experiment is produced using proton accelerators that were already in use at Fermilab. In fact, the proton beam that is extracted to go to the proton target used by MicroBooNE to produce neutrinos, is the same initial proton beam used by every other accelerator experiment at the lab. The Booster Neutrino experiments, which include MicroBooNE, are the first set of experiments on the proton accelerator line.

The first step in creating the Booster Neutrino Beamline (BNB) is to accelerate negatively ionized hydrogen through a linear accelerator, or Linac. The Linac is 500 ft long and accelerates the hydrogen ions to 400 MeV using radio frequency (RF) cavities. The hydrogen ions are then injected into the Booster accelerator.

During injection, the ions pass through a foil that strips both electrons leaving only the positive proton. The Booster [61] is a 474-meter-circumference, 15 Hz synchrotron that accelerates the protons from 400 MeV to 8 GeV. The proton beam leaving the Booster has a bunched structure. Each turn contains 81 proton bunches that are 2 ns wide and 19 ns apart. At this stage, a fraction of the 8 GeV protons are extracted to be sent to the Booster neutrino target. A kicker magnet is used to extract all 81 proton bunches in a turn which we refer to as a spill. One spill is 1.6  $\mu$ s long and contains up to  $5 \times 10^{12}$  protons.

The proton target is a beryllium cylinder 71.1 cm long and 0.51 cm in diameter that is aligned with the proton beam. The interactions of the protons in the beryllium produce secondary hadrons including charged and neutral pions, kaons, and nucleons. The decay of the charged pions, charged and neutral kaons, and muons from pion and kaon decays all contribute to the neutrino flux at the Micro-BooNE detector. The charged pions and kaons are focused by a pulsed toroidal electromagnet operating at a maximum frequency of 5 Hz which is the limiting factor in the beam spill frequency. In neutrino mode, the magnet focuses positive particles and defocuses negative particles. The positive pions and kaons then pass through an air-filled decay pipe where they quickly decay predominantly into positive muons (anti-muons) and muon neutrinos. In anti-neutrino mode, which is not used in this analysis, negative pions and kaons are focused which decay into negative muons and anti-neutrinos. The anti-muons and neutrinos next pass through

a beam stop made of steel and concrete to filter out remaining protons, pions, and kaons. Finally, the anti-muons and neutrinos pass through almost 500 meters of dirt where the anti-muons stop and decay, and a (mostly) pure neutrino beam remains. The predicted neutrino flux at MicroBooNE is shown in Fig. 12 per  $1 \times 10^6$  protons on the target (POT). The neutrino flux prediction is based on Geant4 [91] (describe in Sec. 4.1.3) based Monte Carlo simulation of protons interacting with the beryllium target and the magnetic focusing horn. The production rate of the secondary hadrons are determined from external proton-beryllium experimental data [92].

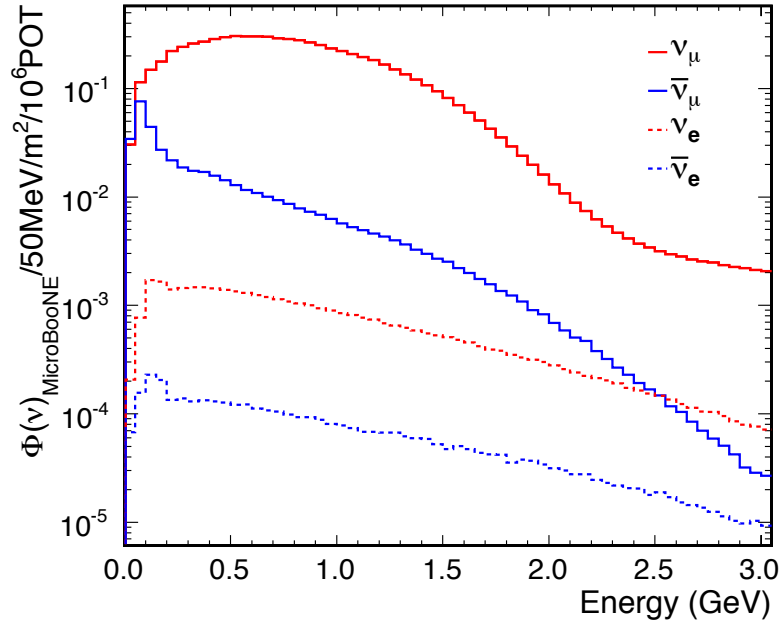


Figure 12: Predicted neutrino beam flux at MicroBooNE.

Neutrinos from the BNB can interact in the dirt upstream of the MicroBooNE detector. Most of these interactions have no effect in the detector because the

secondary particles do not travel far enough to pass through the dirt and enter the detector. However, secondary neutrons that are produced in the dirt near the detector can enter the detector at a significant rate. Figures 13 and 14 show the position of all of simulated neutrino interactions in the dirt surrounding MicroBooNE in gray and the subset of those interactions that produce a neutron which enters the MicroBooNE TPC in red from two different angles. The details of the simulation are described in Sec. 4. If one of these neutrons interacts in the MicroBooNE detector and scatters a single proton, the signal could look indistinguishable from a neutral-current elastic neutrino interaction that occurred in the detector.

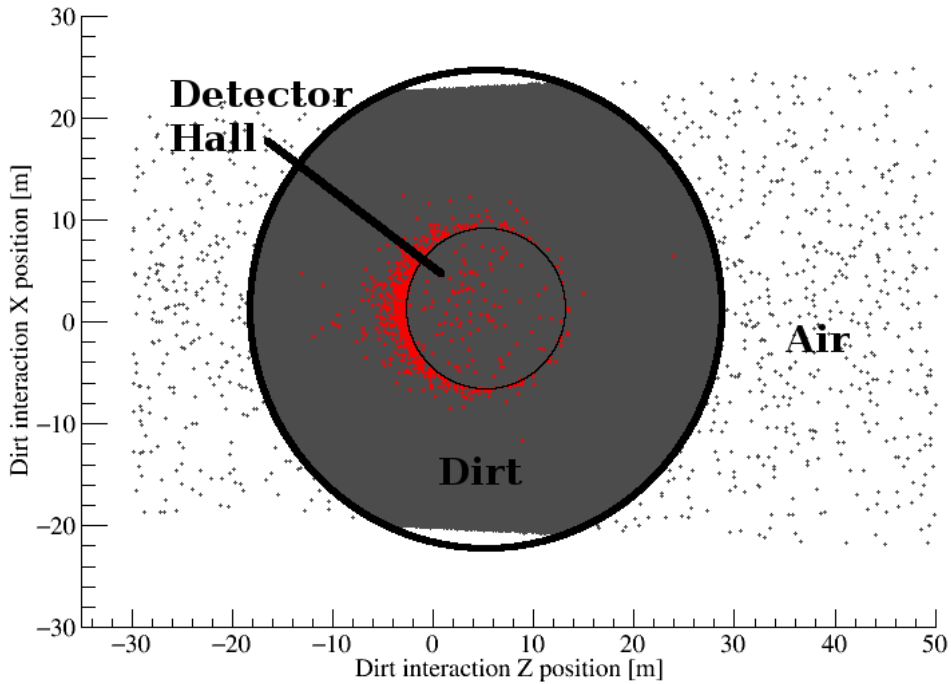


Figure 13: Top view of all simulated neutrino interaction positions.

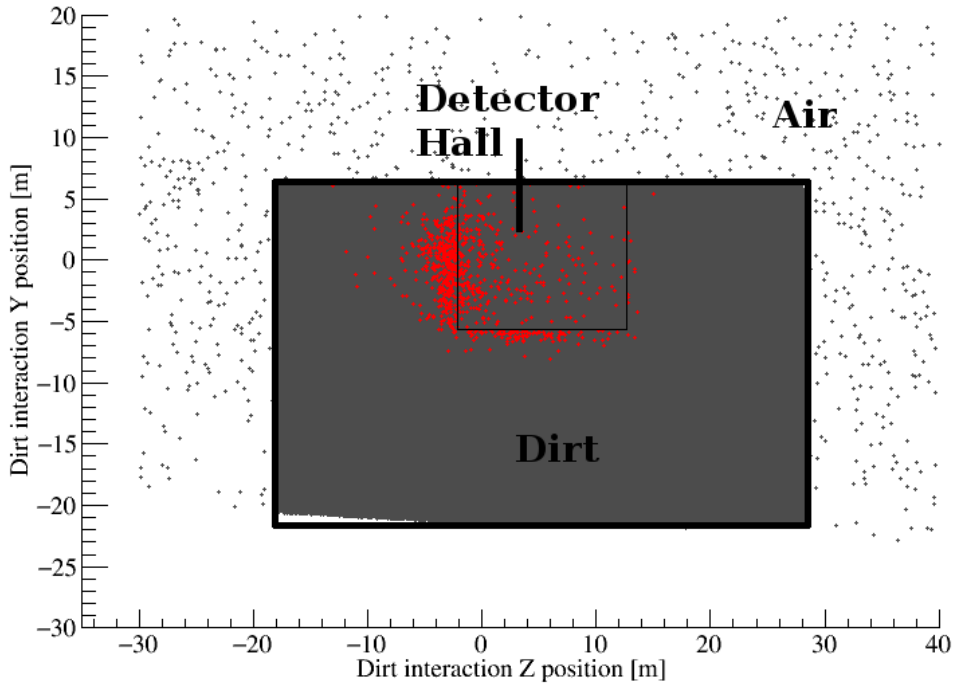


Figure 14: Side view of all simulated neutrino interaction positions.

### 3.3 MicroBooNE LArTPC

The MicroBooNE LArTPC [22] acts as both a target for the neutrino beam and a detector for the charged particles produced in the neutrino-argon interactions. The MicroBooNE TPC is submerged in 170 tons of liquid argon, of which 87 tons is contained inside the TPC. All of this is contained within a cylindrical cryostat.

Liquid argon is used as the detector material for several reasons. Like all noble liquids, argon produces both ionization charge and scintillation light when stimulated. The ionization electrons do not easily recombine, so they are able to pass through the liquid argon to be collected. Additionally, argon is transparent to its own scintillation light, making it detectable. Noble liquids also have good

dielectric properties allowing them to withstand high voltages without breaking down. Liquid argon is also relatively dense at  $1.4 \text{ g/cm}^3$ , making it a good target for neutrinos which interact extremely rarely. Lastly, argon is very abundant. It makes up 1% of Earth's atmosphere and is therefore an affordable option compared to heavier noble elements.

The MicroBooNE time projection chamber (TPC) uses a uniform electric field to guide the ionization electrons to an anode. The electrons are drifted horizontally, perpendicular to the beam direction. To produce this uniform, horizontal electric field, a  $10 \text{ m} \times 2.6 \text{ m}$  cathode plane makes up one face of the TPC. This face is vertical and parallel to the beam. A voltage of  $-70 \text{ kV}$  is applied to the anode plane which results in a  $273 \text{ V/cm}$  electric field across the  $2.3 \text{ m}$  TPC width. There are three wire planes at the anode, perpendicular to the electron drift direction and parallel to the cathode plane, that are strung with thousands of wires spaced  $3 \text{ mm}$  apart. The three planes are all parallel to each other and are also spaced  $3 \text{ mm}$  apart. The wires on the first plane have a small negative voltage ( $-200 \text{ V}$ ), the middle plane has no voltage, and the last plane has a small positive voltage ( $+440 \text{ V}$ ). When the electrons approach the first wire plane, they induce a signal on the closest wires, and are then attracted to the next plane because of its slightly less negative voltage. This is repeated at the second wire plane, and the electrons are terminated at the third plane that we refer to as the collection plane. The electronic signals from all 8256 wires are read out and



saved. The wires on each plane are at different angles to allow for an accurate 2-dimensional reconstruction of the event in the detector. The collection plane wires are vertical, and the wires on the two induction planes are at positive and negative 60 degrees from the beam direction. The third dimension, perpendicular to the anode planes, is reconstructed based on the time that the electrons arrived at the wires. A cartoon representation of this process is shown in Fig. 15. The wires are read out for 1.6 ms per event to ensure that the ionization electrons have time to traverse the entire width of the detector. One 1.6 ms readout window is referred to as a frame. In MicroBooNE the electron drift lifetime is greater than 6 ms, which is twice as long as what is required to perform analyses due to the high purity of the liquid argon [93].

The expected number of neutrino interactions in MicroBooNE as a function of protons on target (POT) is shown in Fig. 16. These were determined using a Monte Carlo simulation described in Sec. 4.1. The plots shows the expected number of events up to  $1 \times 10^{21}$  POT. MicroBooNE is approved to run for a total of  $1.3 \times 10^{21}$  POT. Figure 16b shows the expected number of the subset of neutral current interactions with NC elastic interactions in red. This line includes interactions on both proton and neutrons, so we expect the number of NC elastic proton events to be about half this. The integrated number of all expected NC elastic neutrino-proton events from  $1.3 \times 10^{21}$  POT is approximately 30,000.

The MicroBooNE detector is positioned just below the surface of the Earth,

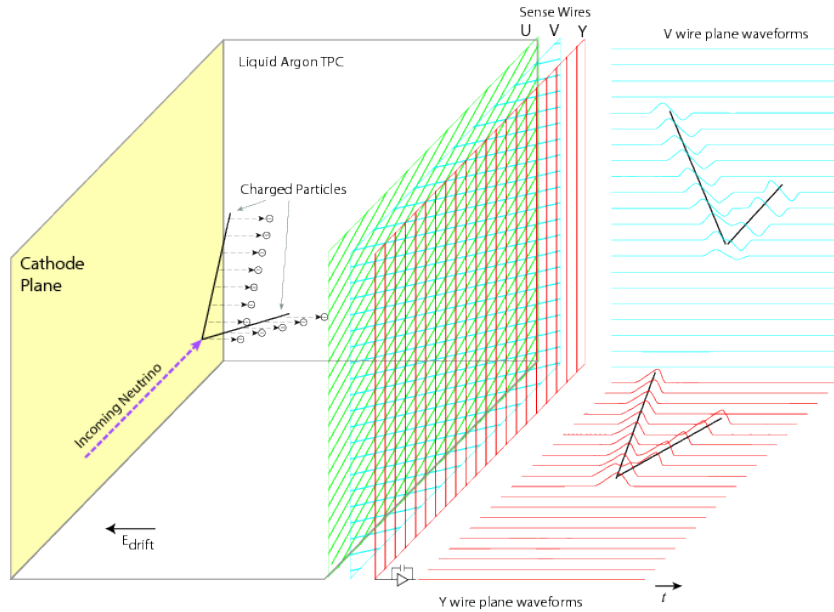
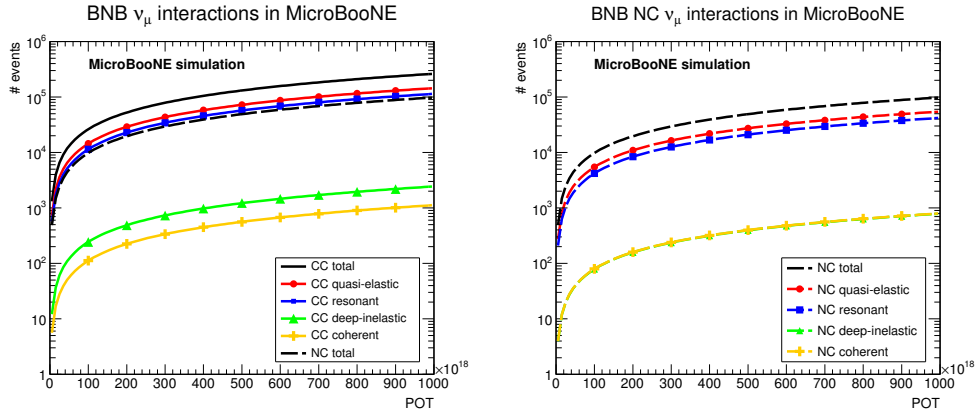


Figure 15: Representation of the operational principle of the MicroBooNE LArTPC.

with no substantial overburden for shielding cosmic rays. The rate of cosmic ray interactions in the detector is therefore large. Based on Fermilab’s elevation (226 m above sea level) and MicroBooNE’s geometry within the detector hall, the expected cosmic muon flux in MicroBooNE is  $160 \text{ m}^{-2}\text{s}^{-1}$  based on MicroBooNE simulation [7] using the CORSIKA Monte Carlo cosmic ray generator [94] (described in Sec. 4.1.2). This gives a rate of about 9 muons per 1.6 ms readout frame. We only expect one detectable neutrino interaction in the TPC for every 500 neutrino beam spills (about one every 10 s), which is much smaller than the



(a) The number of interactions is broken down by interaction type for CC interactions.

(b) The number of interactions is broken down by interaction type for NC interactions. NC elastic scattering off of both protons and neutrons is shown in red.

Figure 16: Expected number of muon neutrino interactions in the MicroBooNE TPC based on Monte Carlo simulation as a function of POT.

cosmic ray rate.

### 3.4 MicroBooNE light collection system and event trigger

Liquid argon also produces scintillation light. Tens of thousands of photons are produced per MeV of deposited energy [22]. The scintillation photons have a vacuum wavelength of 128 nm and do not re-interact with the argon assuming there is not a significant amount of contaminants. The light collection system in MicroBooNE consists of 32 photomultiplier tubes (PMTs) behind the TPC anode

wire planes. Each PMT is shielded by an acrylic plate coated in tetraphenylbutadiene (TPB) that shifts the ultraviolet photons to the visible spectrum before they interact with the PMT.

The light collection system in MicroBooNE is particularly useful for timing information. Since each TPC event is read out over several milliseconds, it is difficult to determine which events have activity coinciding with the neutrino beam spill and which consist of only cosmic background activity based on the TPC information alone. The PMTs have nanosecond timing resolution which allows us to only save events that have optical activity during a time window surrounding the  $1.6 \mu\text{s}$  neutrino beam spill. Figure 17 shows the fractional increase in flashes during the neutrino beam spill window.

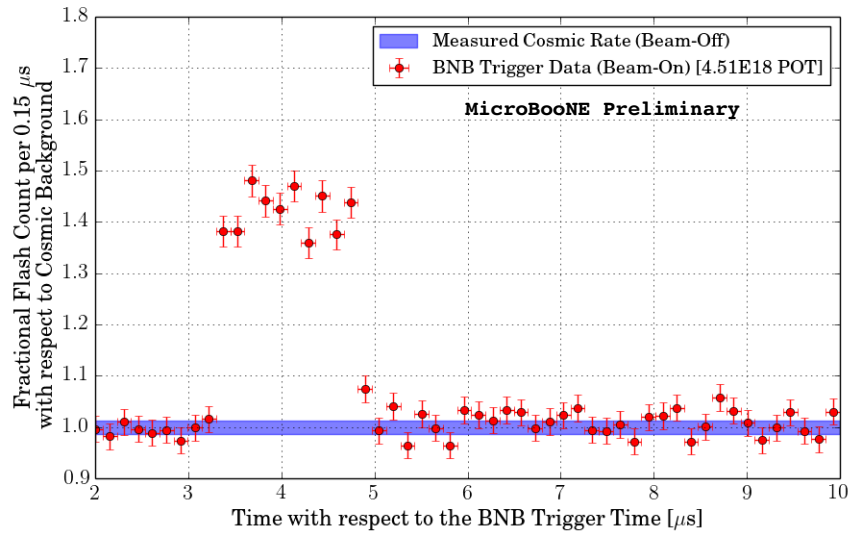


Figure 17: Fractional rate of flashes surrounding BNB spill window.

An event is saved when there is a coincidence between the accelerator beam signal and the PMT trigger implemented in the MicroBooNE data acquisition (DAQ) software. The PMT trigger is based on a  $23 \mu\text{s}$  window of PMT data which include the time of the neutrino beam spill. To form a trigger, first pulses are found on the thirty-two individual PMT signals using constant-fraction discriminators that open a 100 ns discriminator window, and a new window cannot be opened until the previous window has closed and there has been a 15 ns time period without the signal going above the discriminator threshold. The discriminator threshold in an individual PMT is 10 ADC counts which corresponds to approximately 0.5 photoelectrons (PE) in the PMT. The maximum of the PMT signal during the discriminator window is saved as the pulse height. Next, coincident pulses across the different PMTs are found and combined. Pulse windows that occur at the same time are combined and their pulse heights are summed. If the sum of the coincident pulse heights across all 32 PMTs is above 20 ADC counts, corresponding to approximately 6.5 PE, and the time of the coincident pulses is within the neutrino beam-spill window, the event is saved.

## 4 Simulation and reconstruction

In order to understand how different physics models would affect what we would see in the detector, large and complex Monte Carlo simulations of the beam and the detector are created. These simulations also allow us to develop and test algorithms that reconstruct the underlying neutrino interaction based on what particles are seen in the detector. The simulation and reconstruction algorithms used by MicroBooNE are described in this section.

### 4.1 Monte Carlo simulation

The entire experimental process from the neutrino interactions in and around the detector to electronic signal readout to particle identification is simulated in software. To interface the different software packages needed to simulate each step, a liquid argon software framework (LArSoft) was developed at Fermilab. Within the LArSoft framework, the simulation is divided into three steps: generation, propagation, and detector simulation. The output from the detector simulation stage is designed to match the real data output from the detector as closely as possible. Event reconstruction is also handled within LArSoft, and the same algorithms can be applied to both real data and simulated data in an identical way. This section describes all three stages of simulation and all of the reconstruction stages, including TPC particle track and optical flash reconstruction.

### 4.1.1 Cross section model

The initial neutrino interactions are simulated using the GENIE Neutrino Monte Carlo Generator [95, 96]. Assuming a given neutrino flux, GENIE simulates the interaction of the neutrino with the nucleons inside of the atoms in and around the detector. It also simulates the interactions that occur while the nucleons and pions from the initial neutrino-nucleon interaction traverse and exit the nucleus. Within GENIE the nuclear models and neutrino-nucleon cross sections are configurable by the user. This analysis uses the GENIE version v2.12.2 default settings with the addition of the empirical MEC cross section model [97]. The details of the physics models for each of the four processes (the nuclear physics model, the cross section model, the neutrino-induced hadron production model, and the intranuclear hadron transport model) within GENIE as well as the simulated cosmic ray generation are described in this section. The uncertainties in the analysis due to the models in this section are explored in Sec. 5.4.3.

First, the relativistic Fermi gas (RFG) nuclear model [98] is used for all processes. The mass density for the argon nucleus is taken from review articles [99] and the two-parameter Woods-Saxon density function is used [100]

$$\rho(r) = N_0 \frac{1}{1 + e^{(r-c)/z}}, \quad (93)$$

where  $\rho$  is the density,  $r$  is the distance from the center of the nucleus,  $c$  describes the size of the nucleus (approximately the radius where the density falls to half of

the central value), and  $z$  describes the thickness of the surface. For argon, GENIE uses  $c = 3.53$  fm and  $z = 0.54$  fm as default values. For elastic nucleon scattering in the nucleus, Pauli blocking is applied.

Next, in the case of elastic (and quasi-elastic) neutrino-nucleon scattering, the free-nucleon cross section is calculated. The neutrino-nucleon cross section and form factor models used are described in Secs. 2.6 and 2.5. The methods used to evaluate the Monte Carlo given a chosen cross section and form factor model are described in section 6.1.1.

The hadrons produced in neutrino-nucleon interactions in a nucleus are modeled separately from the nuclear model and the free-nucleon cross section model. This is because the hadron production models do not match the measured neutrino cross sections. The hadron discrepancy is corrected for in GENIE using the AGKY model [101] that was developed to account for the data seen in the MINOS [102] neutrino scattering experiment and tuned using bubble chamber experimental data.

Finally, the transport of the final state hadrons through the argon nucleus is modeled. The hadrons produced in the neutrino-nucleon interactions can rescatter as they exit the nucleus which changes the observable final state particle distributions. The intranuclear transport model is implemented by the GENIE subpackage called INTRANUKE. INTRANUKE uses a cascade model in which the hadron sees a nucleus of isolated nucleons. The interaction probability is calculated based



on the free nucleon cross section and the nucleon density in the nucleus [96]

$$\lambda(E, r) = \frac{1}{\sigma_{hN,tot}\rho(r)}, \quad (94)$$

where  $\lambda$  is the hadron interaction probability,  $E$  is the hadron energy,  $r$  is the distance from the center of the nucleus,  $\sigma_{hN,tot}$  is the total *free* nucleon cross section, and  $\rho$  is the density of the argon nucleus. The free nucleon cross section is different for protons, neutrons, and pions. The density of the argon nucleus is again determined by equation 93.

#### 4.1.2 Event generation

The GENIE neutrino events are generated in LArSoft based on the neutrino flux described in Sec. 3.2. The 1.6  $\mu$ s long spill containing  $5 \times 10^{12}$  protons on target is simulated. In addition to neutrino events, cosmic ray events are also simulated with the CORSIKA [94] cosmic ray generator. CORSIKA uses the FLUKA interaction and transport simulator [103] to model hadronic interactions in the atmosphere. To simulate real detector events both neutrino and cosmic ray interactions can be generated in the same simulation. We can also combine real cosmic data taken when the neutrino beam was turned off with simulated neutrino interactions in the same events to closely model the real detector data.

In the standard simulation neutrino interactions are only generated within the liquid argon filled cryostat, including interactions generated on the cryostat

itself. Additional special samples are made to study how secondary particles from neutrino interactions outside the cryostat contribute to our signal background. One special sample that we generate is a large *dirt* sample in which neutrino interactions that happen in the dirt and detector hall outside of the cryostat are generated. The neutrino interactions are allowed to happen anywhere in fifty feet of simulated dirt outside of the detector hall or anywhere in the detector hall outside of the cryostat. These are known as dirt events.

All of the Monte Carlo samples that are generated and used in this analysis are listed here.

1. Simulated neutrino interactions with simulated cosmic ray backgrounds. This sample is used to develop the particle identification and event selection algorithms.
2. Simulated neutrino interactions with real cosmic ray data backgrounds overlaid. This sample very closely represents the actual detector conditions when there is a neutrino interaction in MicroBooNE and is referred to as the overlay sample.
3. Simulated dirt interactions plus simulated cosmic interactions. This sample contains all known backgrounds to neutrino events in the TPC and is referred to as the dirt sample.

### 4.1.3 Detector simulation

The final state particles from GENIE are passed to the Geant4 [91] software package to be propagated through the simulated geometry. The entire MicroBooNE detector system, the detector hall, and fifty feet of dirt surrounding the detector hall are all included in the Geant4 simulation. Figure 18 shows the entire simulated geometry including the surrounding dirt. This includes the electric field and the detector electronics. The particles are stepped through the geometry and undergo a possible physics process at each step with a given probability. The particles are allowed to interact electromagnetically and hadronically with other particles and the detector system or decay through one of the physically possible decay modes. Additionally, the energy loss through ionization and scintillation is simulated for all particles traversing the detector geometry. In the case of the ionization of the liquid argon, the resulting electrons are propagated through the electric field to the wire readouts. For the scintillation of the argon, a photon library was generated for each position in the liquid cryostat. At each step a particle takes through the liquid argon, the resulting photons that would interact in the PMTs are determined from a look-up table that was generated in a previous full optical simulation. The full optical simulation of the scintillation photons includes the Rayleigh scattering of the photons in the liquid argon as well as the reflection and absorption of the photons at the surfaces in the detector.

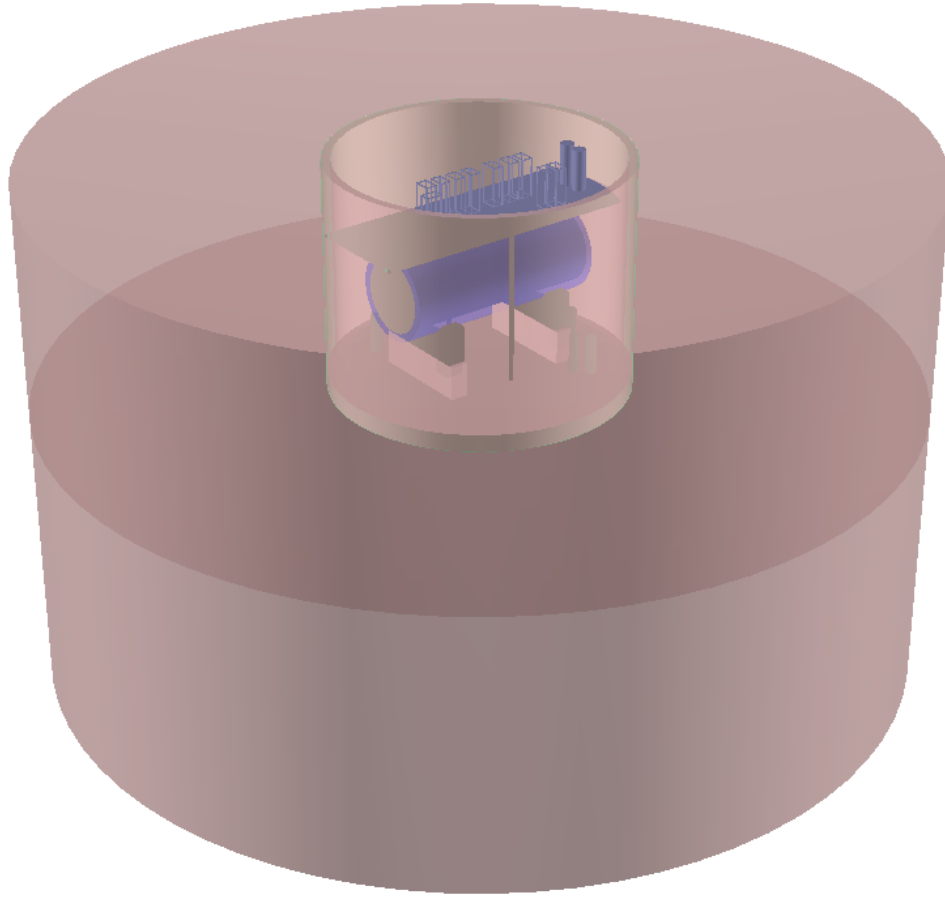


Figure 18: Rendering of the simulated MicroBooNE geometry.

After the simulated particles interact with the TPC or PMT system, the detector response is simulated. The detector-simulation stage includes the electronic responses of the sensitive detectors and reproduces the electronic signals from the TPC and PMT systems. First, the PMT signal is digitized and the PMT software trigger described in Sec. 3.4 is fully simulated. Events that do not pass the PMT trigger are dropped. The TPC electronics, including the electronic noise on the wires and unresponsive wires, are also included at this stage. At this point, the

simulated data resembles the actual raw detector data as closely as possible.

## **4.2 Event reconstruction**

Both the simulated waveforms and the raw detector waveforms need to be reconstructed into the initial neutrino interactions. A series of reconstruction algorithms for both PMT and TPC information exist in LArSoft. These algorithms start by identifying peaks in the waveforms and combine these peaks in stages to get to a 3-dimensional representation of the physics interaction.

### **4.2.1 Flash reconstruction**

The optical flash reconstruction algorithm is applied identically to detector data and the simulated data that is output from the detector simulation stage. The first step is to find pulses on the electronic signals read out from each of the 32 PMTs. This is done using a peak finding algorithm on the digitized signal. The time, amplitude, width, and area under the pulses are stored per pulse. Next, the flash reconstruction algorithm looks for coincident pulses across PMTs. The individual pulses are sorted by size, and all of the pulses that are within  $8 \mu\text{s}$  of the largest pulse are collected. If there are at least three pulses in that time window and the sum of the pulse areas is at least 6 PE above the noise background, it is considered a flash and saved. The peak time, width, position, and size are reconstructed and saved along with information about the pulses in the individual

PMTs that contributed to the flash. This process is repeated starting with the next largest remaining pulse until there are none left. An individual pulse can only contribute to one flash in an event.

#### 4.2.2 TPC event reconstruction

Reconstructing TPC events involves more steps than the PMTs since there are thousands of wires being read out for milliseconds resulting in approximately 30 MB of raw data per event. The reconstruction algorithms are again applied identically to detector data and the simulated data from the detector simulation stage. First, a noise-deconvolution filter is passed over each of the digitized wire signals. Then, similar to the optical reconstruction, pulses, or *hits*, are found on individual wires which are used as base building blocks for reconstructing 3D particle tracks across wires and wire planes.

The 1D hit finding algorithm starts by walking along a wire signal until the value is above a given threshold. The point where the signal goes above the threshold is considered the start of the pulse, and the end of the pulse is defined as the point where the signal goes back down below the threshold. Then, local minima and maxima are found between the start and end of the pulse which are used to determine where there are peaks within the pulse. Adjacent pulses are merged if they are close enough in time. Once the pulse and the number of peaks are established, the algorithm attempts to fit Gaussian peaks to the pulse. The

hypothesis signal is composed of one Gaussian per peak from the previous step. The mean and amplitude of each Gaussian is initially centered at the existing peaks and is allowed to float. If the residuals of the fit are sufficiently small, each Gaussian peak is saved as a 1D *hit* with an amplitude, width, and time given by the fit. The hit finding is repeating along the length of the wire for each wire on all three planes.

The 3D track finding algorithm is separated into two distinct parts. The first part attempts to reconstruct and tag as many cosmic-induced tracks. The hits associated with these tracks are then removed from the set of hits that are available to reconstruct neutrino-induced tracks in the second part.

The algorithm used to preferentially reconstruct cosmic tracks is called *PandoraCosmic*. It is implemented in the Pandora Software Development Kit [104] used by MicroBooNE. In PandoraCosmic, 1D hits are first clustered in 2D per wire plane. All of the reconstructed hits output by the hit finding algorithm on a given wire plane are used as input. The hits are clustered into unambiguous, continuous lines of hits. These initial clusters are meant to have a high purity, meaning all of the hits in the cluster were induced by the same true particle. The clusters in a 2D plane are then merged pairwise in an attempt to improve the completeness of the cluster, meaning most of the hits on that wire plane induced by a true particle are included in a single merged cluster. The 2D clusters are then matched across the three wire planes. The 3D track reconstruction algorithm

checks each plane for clusters which are likely to have come from the same true particle. All possible combinations of 2D cluster matching are considered by the algorithm. The most suitable set of cluster combinations are projected into 3D and saved as reconstructed track objects.

Next, the reconstructed tracks from PandoraCosmic are passed to a cosmic tagging stage. There are two algorithms used to identify cosmic-induced tracks. The first is a geometry tagger that looks for tracks that are not fully contained in the TPC during the event, and the second is a flash-matching algorithm that looks for tracks that are inconsistent with any flashes in the beam spill window. Both of the cosmic tagging algorithms try to remove as few neutrino-induced tracks as possible and only remove tracks that are very likely cosmic-induced.

The geometry tagger starts by locating any TPC wire hits that are reconstructed before or after the 1.6 millisecond readout frame. Any tracks that contain these hits are tagged as cosmic. Any tracks whose reconstructed start or end points are located within a given distance from the TPC boundary are also tagged. If both of the start and end points are near a TPC boundary, the track is given a cosmic score of 1. If only one of the start or end points is near a boundary, the track is given a cosmic score of 0.5.

The flash-matching algorithm creates a hypothesis flash for each reconstructed track based on its position, size, and energy deposited. It then compares the hypothesis flash to each of the reconstructed flashes in the beam spill window. If



a hypothesis flash is sufficiently incompatible with all true flashes in the window, the track is tagged as cosmic. After the cosmic tagging, all reconstructed hits associated with any of the tracks tagged as cosmic are removed from the set of possible hits that are used to reconstruct neutrino-induced tracks. This is referred to as the cosmic hit removal stage.

The remaining set of hits is used to reconstruct neutrino-induced tracks. First, two dimensional clusters of hits are formed on each of the three wire planes using the TrajCluster algorithm in LArSoft. TrajCluster creates line-like collections of hits and adds new hits to the cluster based on the 2D trajectory of the current set of hits. The algorithm stops when there are no additional hits along the trajectory of the cluster and is followed by an additional stage that combines clusters which start and end near each other. Next, 3D tracks are formed using the projection matching algorithm (PMA) [105] which was developed for the ICARUS experiment and implemented in LArSoft. PMA works by proposing nodes and lines connecting the nodes in 3D, projecting the 3D lines onto the 2D planes, and determining the most likely positions of the nodes in 3D based on the fit of the 2D projections to the existing 2D reconstructed clusters from the previous stage. The algorithm starts with a two-node hypothesis and adds nodes to the 3D line until a maximum number of nodes, which is based on the number of hits on the wires, has been reached.

Calorimetric information is extracted when the reconstructed track objects are

created. The calorimetric information that we calculate and use is related to the energy loss of the particle that created the track along its trajectory. At each point along the track, the difference in the total charge between the current point and the previous trajectory point is calculated. This gives us the change in charge as a function of distance along the track, which we label  $dQ/dx$ . The  $dQ/dx$  values along each track are found for each of the three wire planes. The  $dQ/dx$  values can be converted to energy loss per unit distance,  $dE/dx$ , by multiplying the  $dQ/dx$  by a measured conversion factor. This conversion factor depends on the strength of the electric field and the gain of the readout electronics and is determined empirically [106].

## 5 Particle Identification and Event Selection

This section describes the process for selecting events in data and simulation that we use to determine the strange axial form factor. First, proton tracks are identified using a gradient-boosted decision tree model based on reconstructed tracks features. Next, events are selected as neutral current elastic proton events using a logistic regression model based on reconstructed properties of the whole event including optical information. Comparisons of the selected NC elastic proton events in the neutrino beam data to the expectation from neutrino simulation and off-beam data are shown, and each of the remaining backgrounds is described in detail.

### 5.1 Particle Identification

After the particle tracks are reconstructed, we use a predictive model to classify proton tracks. The inputs to the model are the reconstructed physical variables, and the output is the probability that the track is from a proton vs. some other particle. There are many predictive models that we can use, each with advantages and disadvantages. We chose gradient-boosted decision trees for a few main reasons: they are easily interpretable, the inputs can be a mix of numeric and categorical variables, and boosted decision trees perform well at identifying a small signal in a large background. Each tree is essentially a series of cuts based on

physical variables which have been fine-tuned to increase the efficiency and purity of the final selected sample.

### 5.1.1 Reconstructed track features

The reconstructed features that are used as input to the classifier are listed below. Most of the features come directly from the track object, but some are created for this classifier. Each of the features used to identify protons either helps to separate neutrino-induced tracks from cosmic-induced tracks or to separate neutrino-induced proton tracks from other neutrino-induced particle types. For colorimetric information we only use information from the collection plane.

Below is a list and description of the features designed to separate neutrino-induced protons from other neutrino-induced particle types.

- **Number of hits:** This is the total number of hits on the collection plane that are associated with a track. When used in combination with track length and average energy deposited, this feature can be used to determine the hit and energy density of the track.
- **Straightness:** This is the ratio of distance between reconstructed end points (displacement) to reconstructed path length. It represents the amount of scattering a track undergoes. The value is always between zero and one with one being perfectly straight.

- **Cosmic score:** This is the geometry tagging cosmic score from Sec. 4.2.2. Tracks with a cosmic score of 1 have already been removed in the cosmic hit removal stage. So, this value is either 0 (fully contained within the TPC) or 0.5 (entering or exiting the TPC).
- **Length:** This is the reconstructed 3D track length found by stepping along the trajectory points.
- **Start  $dE/dx$ :** This is the total energy deposited on the collection plane in the first six non-zero hits along the track divided by the distance between hits to account for the angle with respect to the wire plane.
- **End to start  $dE/dx$  ratio:** This is the ratio of the total  $dE/dx$  from the last six non-zero hits along the track on the collection plane to the total  $dE/dx$  from the first six non-zero hits along the track on the collection plane.
- **Truncated total  $dE/dx$ :** This is the sum of the  $dE/dx$  of a truncated set of hits on the collection plane associated with a track. The truncated set includes all hits along the track with a  $dE/dx$  value within one standard deviation of the median  $dE/dx$  value of all hits along the track on the collection plane.
- **Truncated average  $dE/dx$ :** This is the truncated total  $dE/dx$  divided by the number of hits in the truncated hit set associated with a track.

Next is the list and description of the features designed to separate neutrino-induced tracks from cosmic-induced tracks.

- **Start and end positions:** These are the reconstructed  $x$ ,  $y$ , and  $z$  positions of start and end of the track. Tracks that start closer to a TPC boundary are more likely to be cosmic-induced.
- **$\theta$  and  $\phi$ :** These are the reconstructed polar and azimuthal angles with respect to the beam direction. Vertical tracks are much more likely to be cosmic-induced, while forward-going tracks are more likely to be from the neutrino beam.

Determining which end of a track is the beginning is difficult when a vertex is not observable. Since we are particularly interested in neutral-current elastic events with only a single proton, the direction of the track is a concern. A proton will deposit much more energy at the end of its track than at the beginning which can be used to determine the true direction. Since this correction is not currently implemented within the reconstruction, we take all reconstructed tracks that have a higher deposited energy at the beginning of the track than at the end of the track and flip them. The deposited energy at the beginning (ending) of the track is defined as the total  $dE/dx$  of the first (last) six non-zero hits along the track on the collection plane. This includes changing the saved start positions, end positions,  $\theta$ ,  $\phi$ , start  $dE/dx$ , end  $dE/dx$ , and the end to start  $dE/dx$  ratio.

### 5.1.2 Boosted decision trees

A decision tree can be thought of as a series of if/else statements that separate a data set into two or more classes as illustrated in Fig. 19. At each node of the tree, a split is chosen to maximize information gain until a set level of separation is reached. At the terminus of the series of splits, called a leaf, a class is assigned. The usual parameters that can be set when creating a decision tree are: the maximum depth of the tree (how many layers of nodes you will allow), the minimum split size (how many data points do you require to keep splitting), and minimum leaf size (how small does a leaf have to be before you stop).

A single tree can easily overfit a data set if it is at all complex, and its output is just a class label. Gradient-boosting addresses both of these issues by combining many weak classifiers into a strong one. Each weak classifier is built based on the error of the previous one. For a given training set, whenever a sample is classified incorrectly by a tree, that sample is given a higher importance when the next tree is being created. Mathematically, each tree is training on the gradient of the loss function. After all of the trees have been created, each tree is given a weight based on its ability to classify the training set, and the output of the gradient-boosted decision tree classifier is the probability that a sample is in a given class.

The gradient-boosted decision tree software package we use is XGBoost [107]. There are two types of classifiers we can use to separate protons from other tracks:

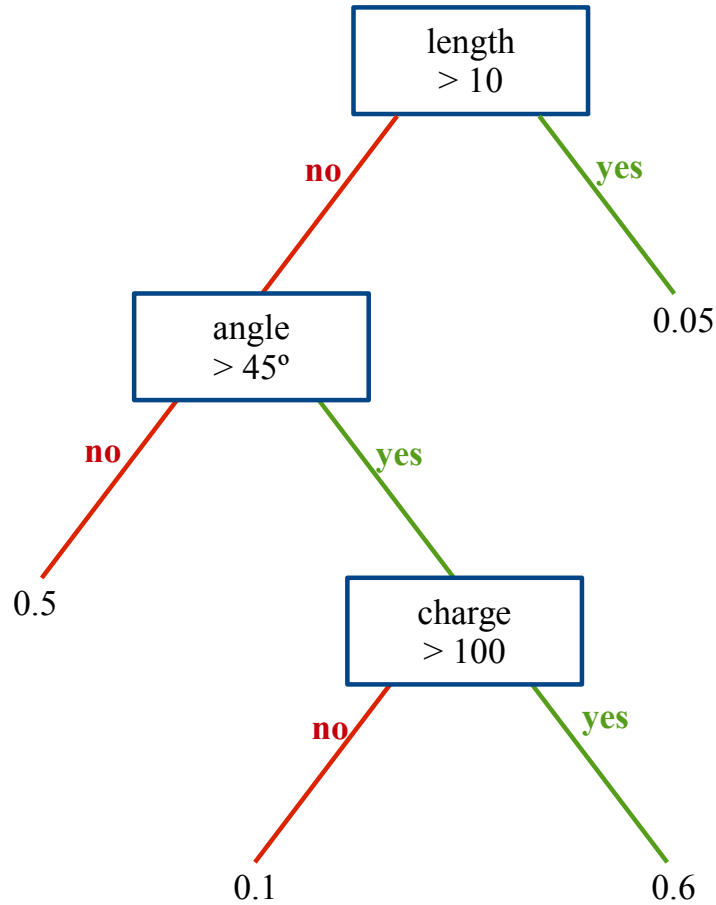


Figure 19: Graphical example of a decision tree.

binary and multiclass. Both classifiers are trained on all types of reconstructed tracks. A binary classifier classifies each track as either a proton or not a proton, and a multiclass classifier classifies a track as one of many types including a proton. We choose to use multiclass because the information about non-proton tracks is useful for selecting neutral current events. The five classes that we train the decision trees to classify are protons (both neutrino-induced and cosmic), neutrino-induced muons, neutrino-induced pions, neutrino-induced electrons/photons, and



all non-proton cosmics.

### 5.1.3 Training

Table 1: Breakdown by simulated particle type reconstructed tracks in the gradient-boosted decision tree training set.

	Protons	Muons	Pions	EM Showers	Non-proton Cosmics
No. of tracks	90,922	57,583	12,848	473,323	2,586,527
Fraction of set	0.028	0.018	0.004	0.147	0.803
Class weight	0.141	0.223	1.000	0.027	0.005

The gradient-boosted decision tree model was trained on 95,600 events with both simulated GENIE neutrino interactions and simulated CORSIKA cosmic interactions. Each track in every event was treated as a separate training sample. Table 1 shows the number of each type of track that was used for training. There were a total of 3,221,203 simulated training tracks.

Because the training set has unbalanced classes (there are different numbers of each particle type) each training sample is initially weighted so that the sum of weights is equal to the size of the smallest class, in this case pions.

$$N_s = \sum_{i=1}^{N_n} w_i^n,$$

where  $N_s$  is the number of samples in the smallest class,  $N_n$  is the number samples in the  $n^{th}$  class, and  $w_i^n$  is the weight given to the  $i^{th}$  sample in that class. The

same weight is used for each sample in a class, so the value of each positive weight is  $w^n = \frac{N_s}{N_n}$ . Balancing the training set prevents the classifier from only learning the most frequent classes. In our case, the classifier could achieve a high accuracy by classifying everything as a cosmic in the unbalanced set because over 80% of the tracks are cosmic-induced. One of our main goals is to have a proton ID efficiency, and since protons only make up 3% of the training set, giving them a higher weight makes it more important to the classifier that they are correctly classified.

The parameters used for training were chosen to both maximize classification accuracy and minimize overfitting to the training set. Overfitting occurs when the performance on the training set is more accurate than the performance on an external test set. The final training parameter settings are:

- **Objective: multiclass: softprob**

The learning objective. We want to classify five different track types and get a probability of each class.

- **Learning rate: 0.05**

The change in the factor that each incorrectly classified sample gets re-weighted by for the next tree. A smaller learning rate requires more trees but prevents overfitting.

- **Number of trees: 300**

The total number of trees in classifier.

- **Maximum depth: 6**

The maximum number of layers of nodes each tree can have.

- **Maximum sampled features: 1.0**

The fraction of total features that each tree can use to train. These are randomly sampled.

- **Maximum sampled observations: 0.75**

The fraction of total samples that each tree can use to train. These are randomly sampled.

#### 5.1.4 Performance on a Test Set

The performance of the gradient-boosted decision tree classifier was tested on a set of 3,200,000 reconstructed tracks from 96,200 events with simulated GENIE neutrino interactions and simulated CORSIKA cosmic interactions. This set of tracks was generated in the exact same way as the training set.

Figure 20 shows normalized histograms of the output proton score for every track in the test set. The proton score ranges from zero to one with zero being the least proton-like and one being the most. The blue histogram shows all simulated neutrino-induced and cosmic induced proton tracks normalized so that the area under the histogram is one. The orange histogram shows every other simulated

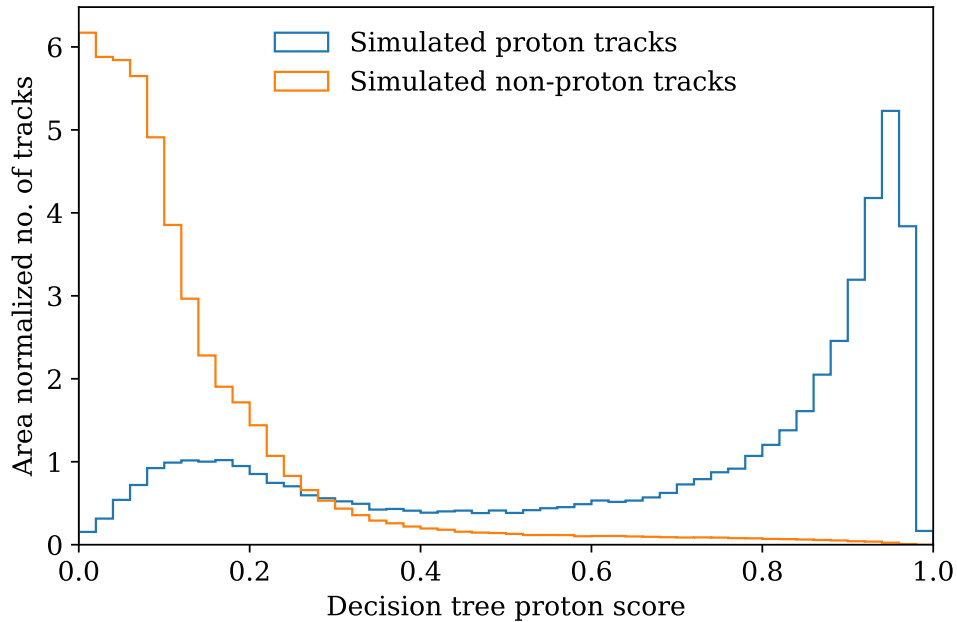


Figure 20: Area-normalized histograms of decision tree proton identification scores for simulated protons and other simulated proton tracks.

track type, also normalized so that the area under it is equal to one. Figure 21 shows the area-normalized histogram of proton scores for simulated proton tracks that were produced in neutral current elastic proton events.

Figure 22 shows the overall classification performance of the gradient-boosted decision tree model on the test set for each class. The  $x$  axis shows the true particle type and the  $y$  axis shows the particle classes. The numbers in the boxes are the fraction of the class that is made up of the given true particle type. The fraction of true protons in the set of tracks classified as protons is 0.71, the fraction of true muons in that set is 0.04, the fraction of true pions is 0.09, the fraction of

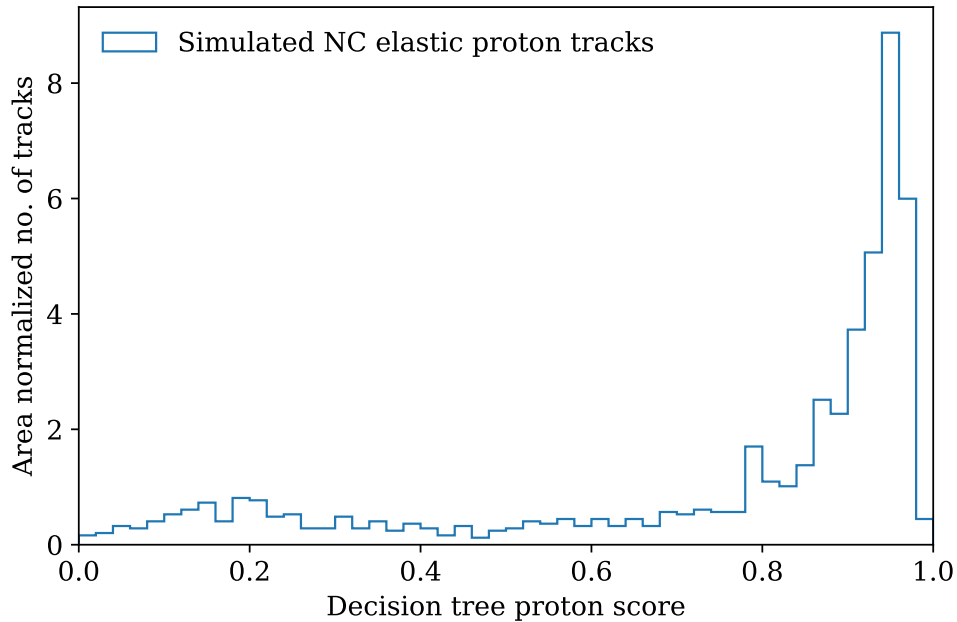


Figure 21: Area-normalized histogram of decision tree proton identification scores for simulated proton tracks from NC elastic proton interactions.

electromagnetic shower particles is 0.04, and the fraction of non-proton cosmics in the proton-classified set is 0.05. A track is labelled as a given class type in this plot if the particle’s decision tree score for that class is higher than its score for any of the other four classes. The numbers in this plot were calculated using equal numbers of each true particle type. In reality, there are far more non-proton cosmic tracks than there are true protons, and the fraction of true protons in the set classified as protons will be smaller.

Figure 23 shows the efficiency of the decision tree proton identification on simulated neutrino-induced protons as a function of true proton kinetic energy.

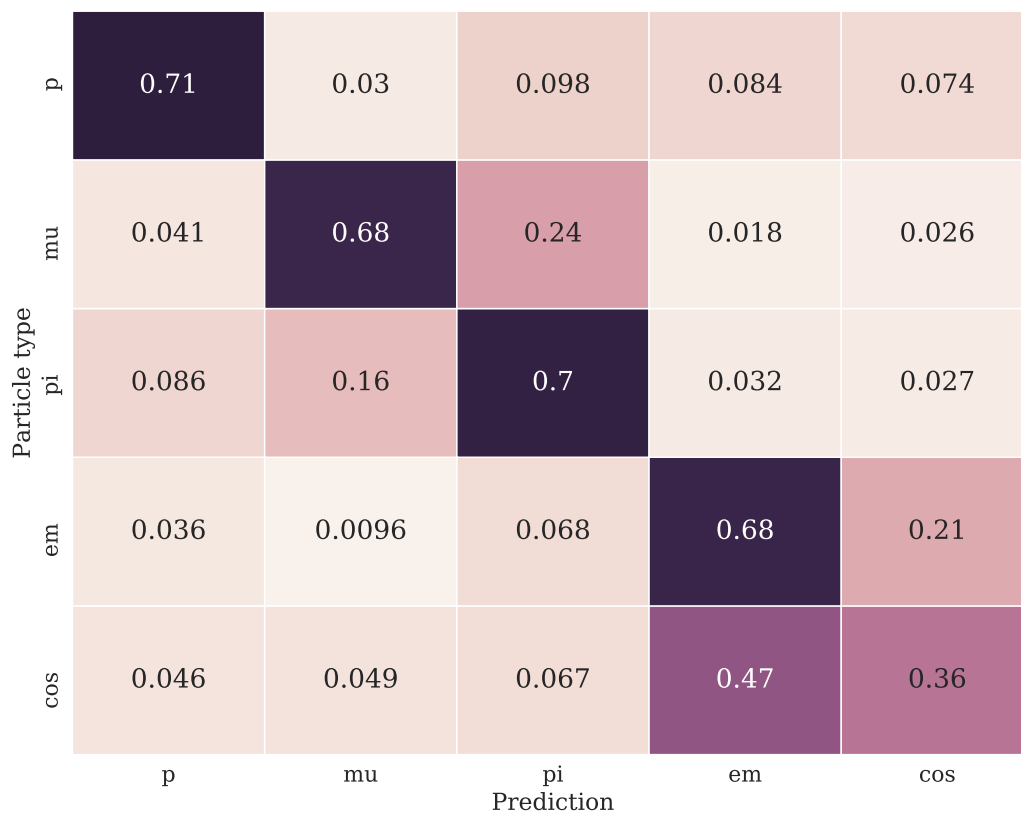
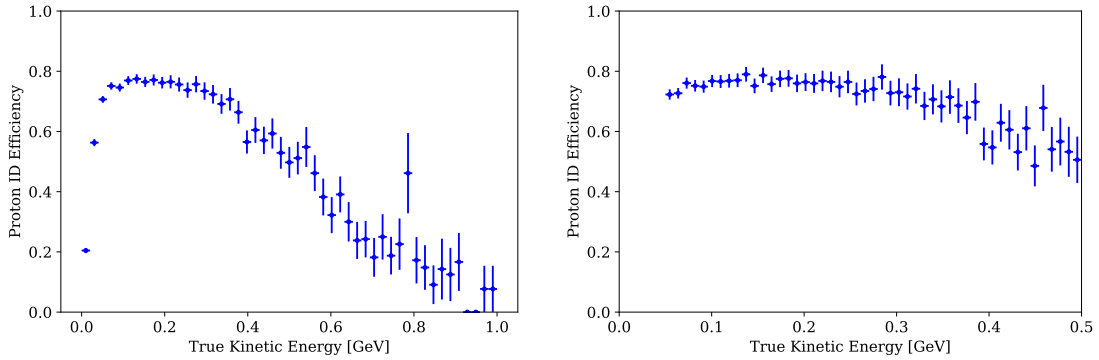


Figure 22: Heatmap showing the fraction of each class that is made up of a given particle type.

The left plot (23a) shows the full simulated range of true proton kinetic energy, and the right (23b) shows the range of interest to this analysis. In the interesting range of kinetic energies, the proton identification efficiency stays relatively flat between 0.6 and 0.8 efficiency, with an average efficiency of 0.71. A track is considered positively identified as a proton in these plots if its decision tree proton score is higher than 0.5, meaning it is more likely than not to be a proton.

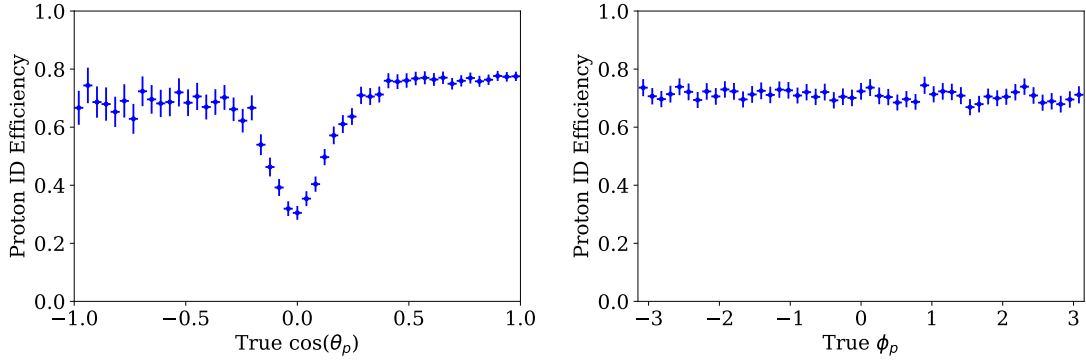
Figure 24 shows the efficiency of the decision tree proton identification on



(a) The full simulated kinetic energy range. (b) The kinetic energy range used in this analysis.

Figure 23: The efficiency of simulated neutrino-induced proton tracks correctly classified as protons as a function of true proton kinetic energy.

simulated neutrino-induced protons as a function of true proton angle. The efficiencies in these plots are calculated using only the simulated protons within the kinetic energy range of interest ( $0.05 \text{ GeV} \leq T_p \leq 0.5 \text{ GeV}$ ) used in Figure 23b. The left plot (24a) shows the efficiency as a function of  $\cos(\theta_p)$ , where  $\theta_p$  is the angle of the proton from the neutrino beam direction. At  $\cos(\theta_p) = 1$  the proton is parallel to the beam, at  $\cos(\theta_p) = -1$  the proton is anti-parallel to the beam, and at  $\cos(\theta_p) = 0$  the proton is perpendicular to the beam. When the proton is perpendicular to the beam it will not traverse more than one collection plane wire. A large contribution to the decrease in efficiency at  $\cos(\theta_p) = 0$  is the fact that the decision tree classifier only uses calorimetry information from the collection plane. The right plot(24b) shows the efficiency as a function of  $\phi_p$  which is the



(a) Efficiency as a function of the cosine of the true proton angle from the beam direction. (b) Efficiency as a function of the true proton angle around the beam direction.

Figure 24: The efficiency of simulated neutrino-induced proton tracks correctly classified as protons as a function of true proton angle.

angle around the neutrino beam direction. The flat efficiency is due to the fact that the neutrino-induced proton angle should be isotropic in  $\phi_p$ , and the angle around the beam direction has no effect on the angle with respect to the angle of the anode wires. Again, a track is considered positively identified as a proton in these plots if its decision tree proton score is higher than 0.5. Figure 25 shows the two-dimensional efficiency for true proton  $\cos(\theta_p)$  versus true proton kinetic energy. The kinetic energy range of interest to this analysis goes up to 0.5 GeV (the bottom half of the plot).



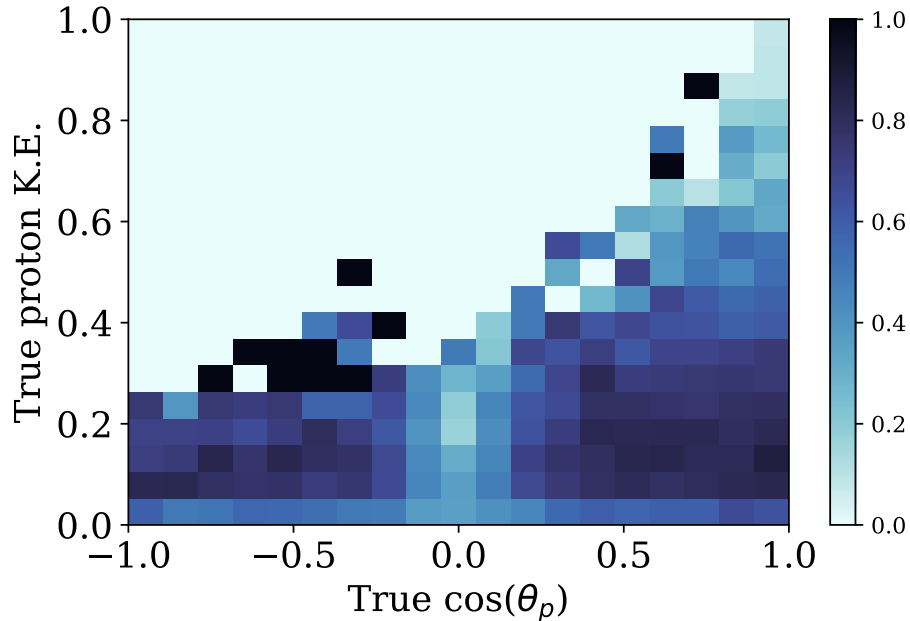


Figure 25: Two-dimensional efficiency for true proton  $\cos(\theta_p)$  versus true proton kinetic energy.

### 5.1.5 Performance on a Neutrino Data Subset

The gradient-boosted decision tree classifier was tested on a subset of MicroBooNE neutrino data corresponding to  $5 \times 10^{19}$  POT ( $< 5\%$  of the full MicroBooNE approved POT). The data set is taken entirely from MicroBooNE's first year of running (Run I). The results of the classifier on the neutrino data subset was compared to the results on a combination of neutrino and cosmic simulation and off-beam data. Each of the samples is scaled to  $5 \times 10^{19}$  POT. The samples used in these comparisons and the scaling factors are listed below.

1. Run I  $5 \times 10^{19}$  POT neutrino data subset

- Number of events: 171,603
- POT (`tor860_wcut`):  $4.41e19$
- Number of triggers (`E1DCNT_wcut`): 9,779,224
- Normalization factor: 1

2. Run I off-beam data subset

- Number of events: 189,226
- Number of triggers (`EXT`): 14,579,406
- Normalization factor: 0.7063

3. Neutrino Monte Carlo with cosmic data overlay set

- Number of events: 200,160
- POT:  $2.08e20$
- Flux correction factor: 1.029
- Normalization factor: 0.2429

4. Neutrino dirt Monte Carlo with simulated cosmic data set

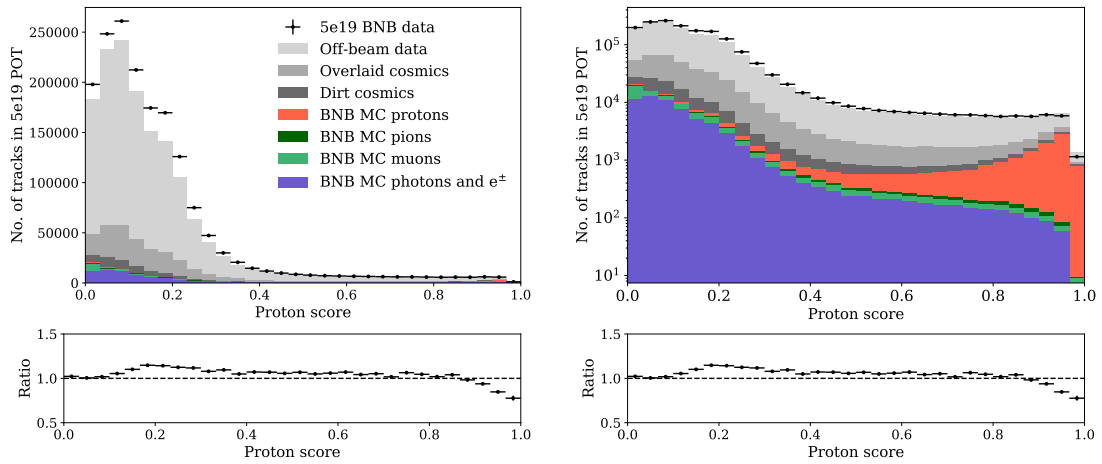
- Number of events: 105,214
- POT:  $4.66e20$
- POT normalization factor: 0.0947

- Data driven normalization factor: 0.5 (described in Sec. 5.3.1)
- Normalization factor: 0.0474

The number of triggers listed is the total number of beam spill triggers issued by the accelerator division and does not include the optical software trigger implemented in MicroBooNE. This is how we scale the off-beam data to the neutrino beam data. The number of events is the number of events left after the optical pre-selection described in section 5.2.1. The off-beam data is a direct measurement of the background to the neutrino beam data that is due to a cosmic interaction occurring in-time with the beam. The neutrino Monte Carlo simulates neutrino beam interactions that occur within the liquid argon cryostat, and the neutrino dirt Monte Carlo simulates the background neutrino beam interactions that occur outside of the cryostat.

Figure 26 shows the comparison of the decision tree proton score between the subset of MicroBooNE neutrino data and the MicroBooNE simulation. The top left plot is in linear scale and the top right plot is in log scale. The bottom plots are the same, and they show the ratio between the on-beam neutrino data and the combination of neutrino simulation and off-beam data.

In all of the figures in this section, the black points in the top plots show the subset of neutrino data. The horizontal bars represent the bin width, and the vertical bars represent the statistical uncertainty. The light gray filled histogram



(a) All reconstructed tracks.

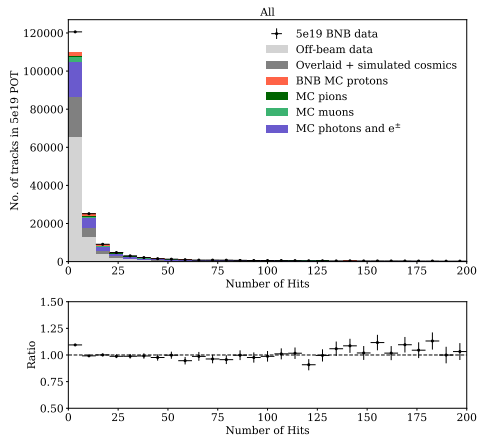
(b) All reconstructed tracks that are classified as protons.

Figure 26: Comparison of the decision tree proton scores between a subset of MicroBooNE neutrino data and a combination of MicroBooNE neutrino simulation and off-beam data.

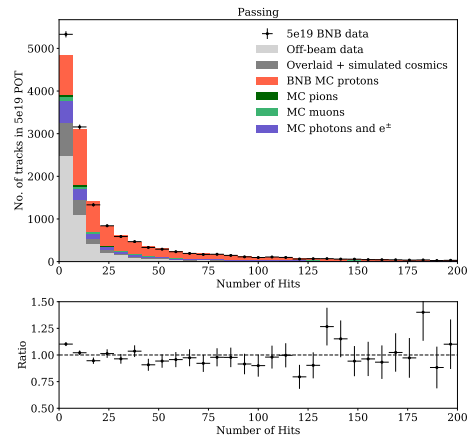
includes tracks from the off-beam data. These tracks represent the background of events where a cosmic interaction in the detector coincident with the beam time window triggered the event, and there was no actual neutrino interaction. The dark gray filled histograms include cosmic tracks that are in the background of events with actual neutrino interactions that triggered the event. For simulated neutrino interactions inside the detector, real data cosmic tracks are overlaid on the simulated event, and for simulated neutrino interaction outside the detector, the background cosmic tracks are from simulation. The color filled histograms include tracks from simulated neutrino interactions. The peach colored histograms

include simulated neutrino-induced proton tracks, the dark green includes simulated neutrino-induced pion tracks, the light green includes simulated neutrino-induced muon tracks, and the purple includes simulated neutrino-induced electromagnetic shower tracks. The fraction of proton tracks in the right plots (the tracks classified as protons) is much larger than in the left plots (all tracks), which is the goal of the classifier. The bottom plots in all of the figures show the ratio of the neutrino data points to the sum of all of the stacked, filled histograms. A ratio of one means perfect data to simulation agreement.

Figures 27-39 show comparisons between the subset of MicroBooNE neutrino data and the MicroBooNE simulation for each of the input variables being used in the decision tree classifier. A description of each of these reconstructed track features is given in Sec. 5.1.1. The left plots show the histogram of the given variable for all tracks being input to the classifier, and the right plots show the histograms of the given variable for the tracks that were classified as protons. A track is considered classified as a proton if the decision tree proton score is greater than 0.5 for that track.

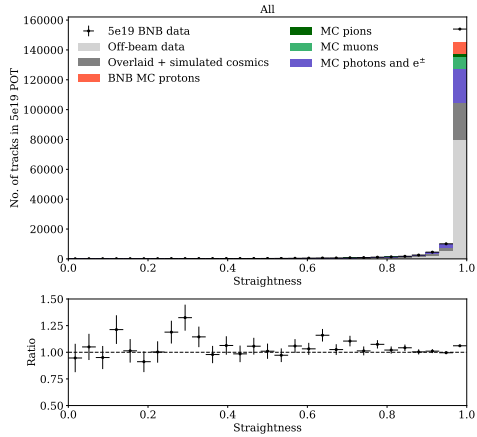


(a) All reconstructed tracks.

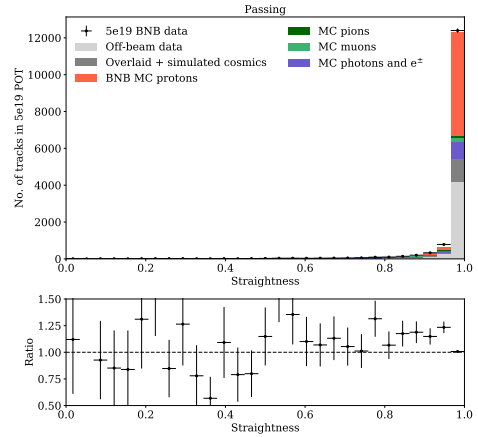


(b) All reconstructed tracks that are classified as protons.

Figure 27: Breakdown of the different particle track types in neutrino data and simulation as a function of the number of hits on the collection plane.

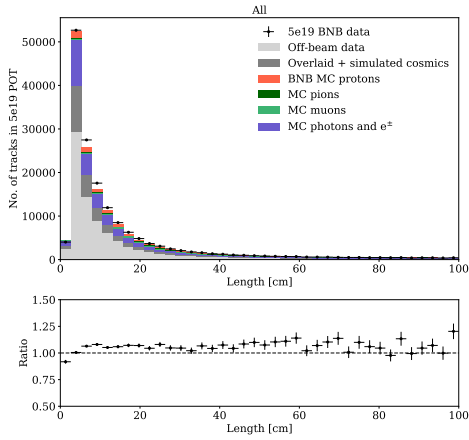


(a) All reconstructed tracks.

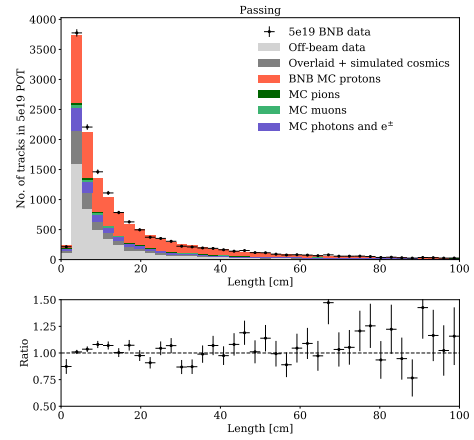


(b) All reconstructed tracks that are classified as protons.

Figure 28: Breakdown of the different particle track types in neutrino data and simulation as a function of the track straightness.

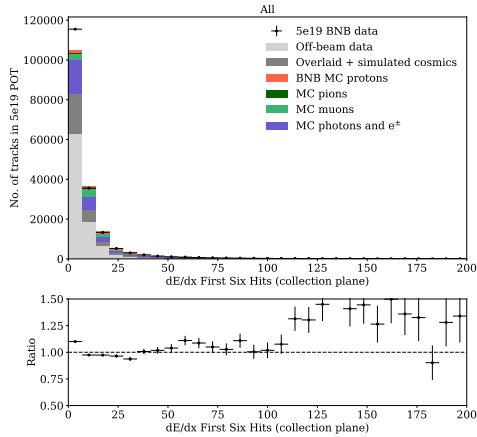


(a) All reconstructed tracks.

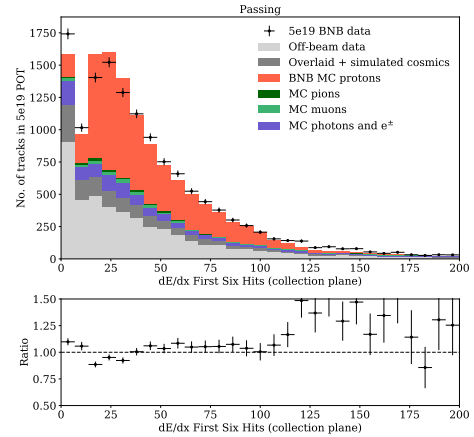


(b) All reconstructed tracks that are classified as protons.

Figure 29: Breakdown of the different particle track types in neutrino data and simulation as a function of the track length.

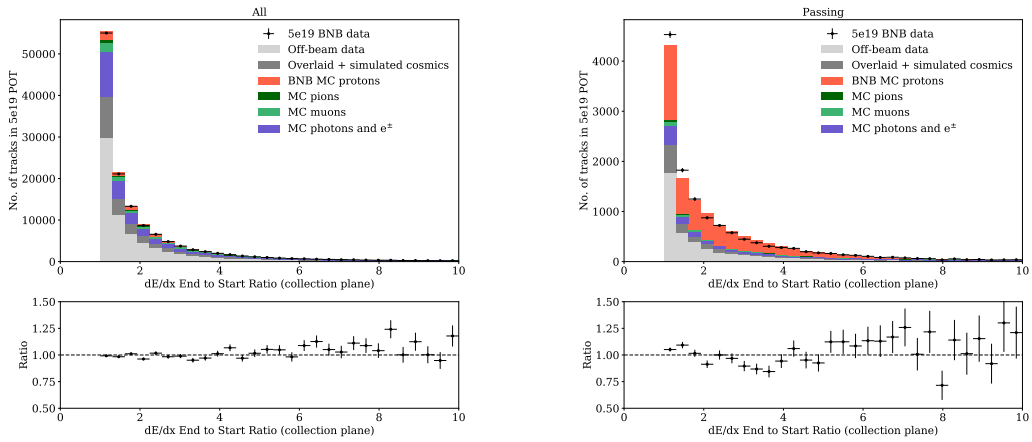


(a) All reconstructed tracks.



(b) All reconstructed tracks that are classified as protons.

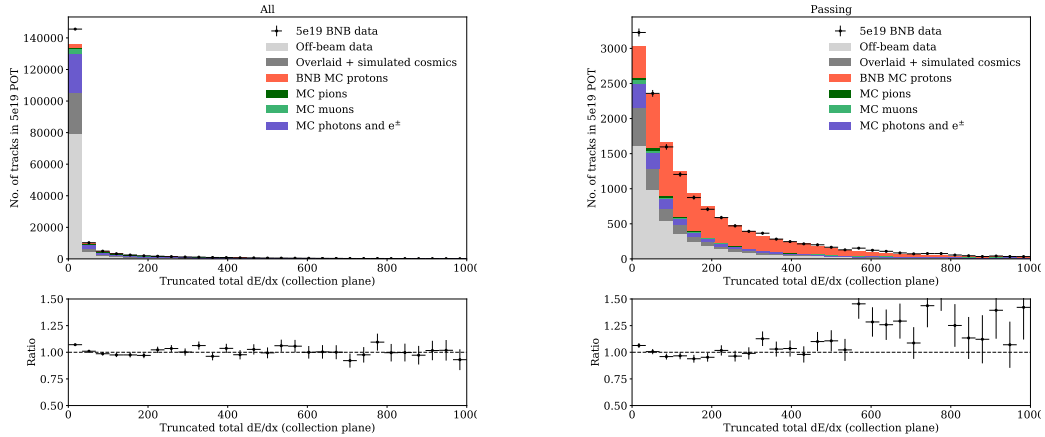
Figure 30: Breakdown of the different particle track types in neutrino data and simulation as a function of the track start  $dE/dx$ .



(a) All reconstructed tracks.

(b) All reconstructed tracks that are classified as protons.

Figure 31: Breakdown of the different particle track types in neutrino data and simulation as a function of the end to start  $dE/dx$  ratio.

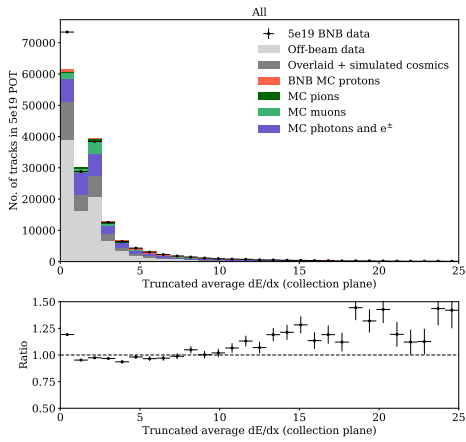


(a) All reconstructed tracks.

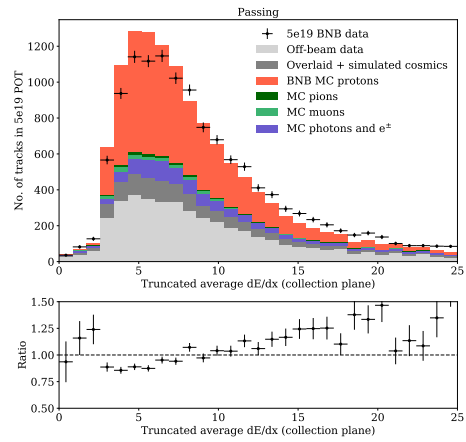
(b) All reconstructed tracks that are classified as protons.

Figure 32: Breakdown of the different particle track types in neutrino data and simulation as a function of the track truncated total  $dE/dx$ .



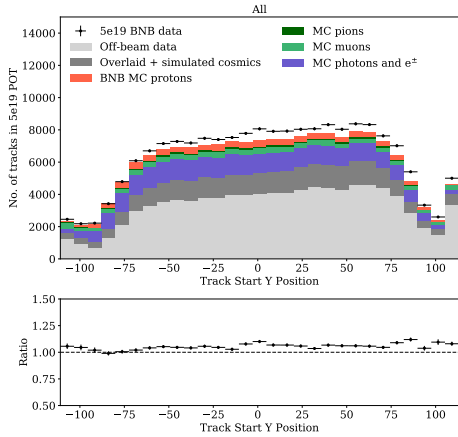


(a) All reconstructed tracks.

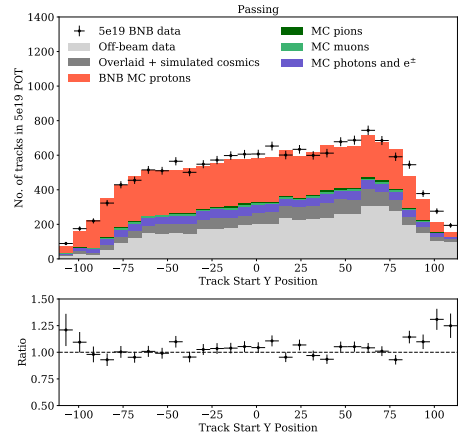


(b) All reconstructed tracks that are classified as protons.

Figure 33: Breakdown of the different particle track types in neutrino data and simulation as a function of the truncated average  $dE/dx$ .

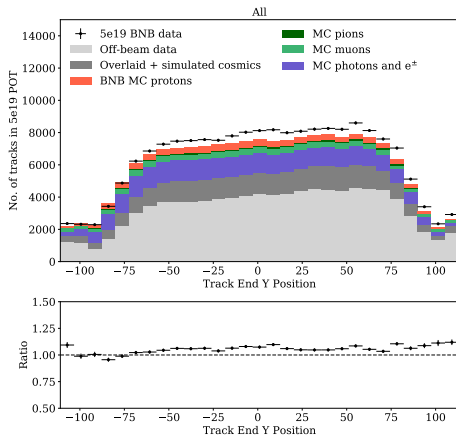


(a) All reconstructed tracks.

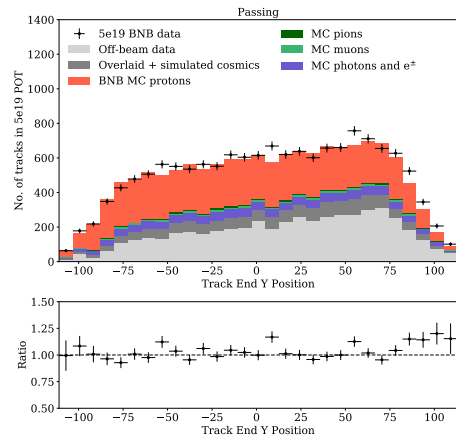


(b) All reconstructed tracks that are classified as protons.

Figure 34: Breakdown of the different particle track types in neutrino data and simulation as a function of the track starting  $y$  position.

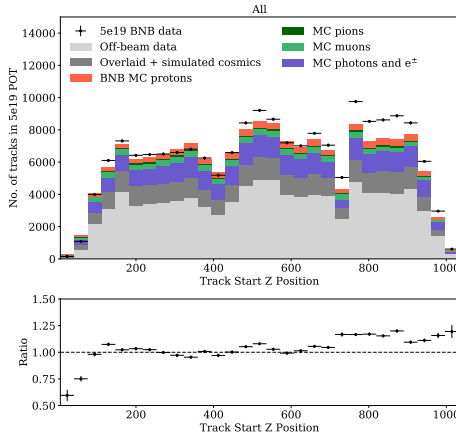


(a) All reconstructed tracks.

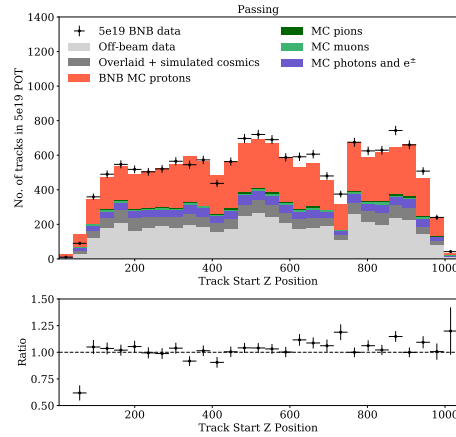


(b) All reconstructed tracks that are classified as protons.

Figure 35: Breakdown of the different particle track types in neutrino data and simulation as a function of the track ending  $y$  position.

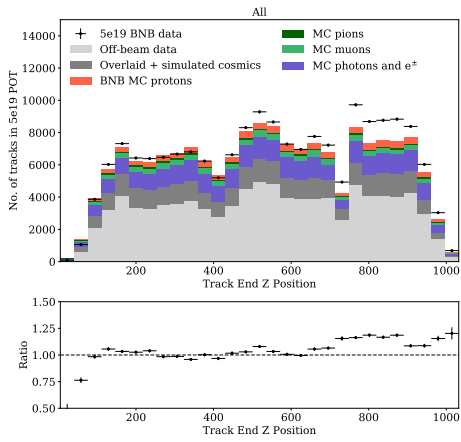


(a) All reconstructed tracks.

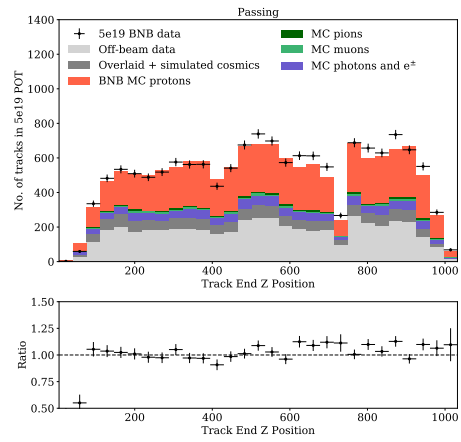


(b) All reconstructed tracks that are classified as protons.

Figure 36: Breakdown of the different particle track types in neutrino data and simulation as a function of the track starting  $z$  position.

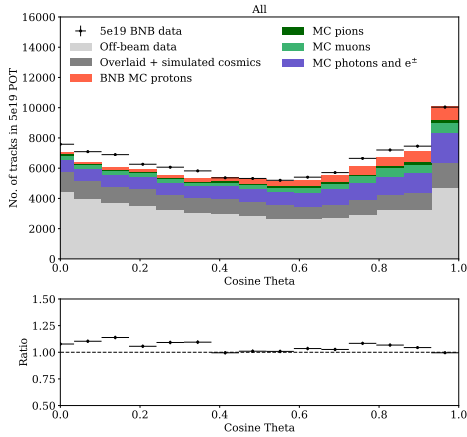


(a) All reconstructed tracks.

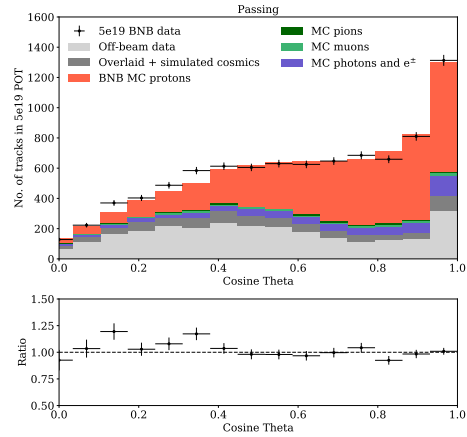


(b) All reconstructed tracks that are classified as protons.

Figure 37: Breakdown of the different particle track types in neutrino data and simulation as a function of the track ending  $z$  position.

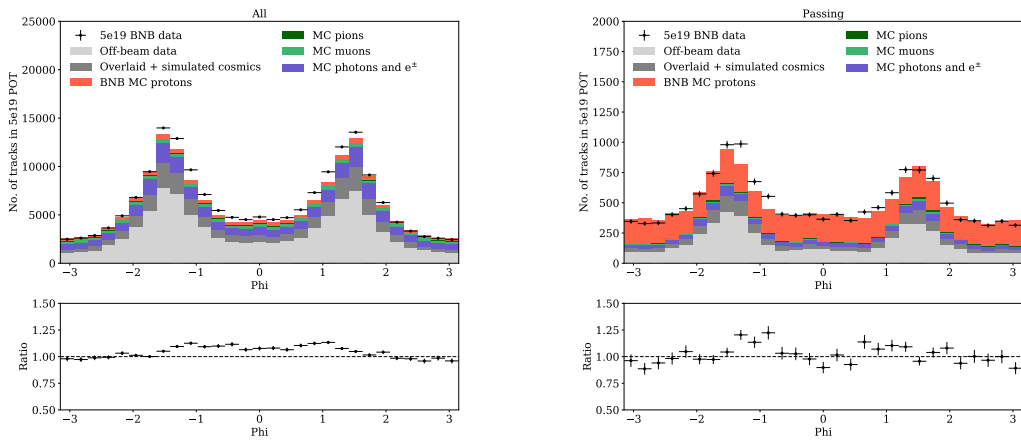


(a) All reconstructed tracks.



(b) All reconstructed tracks that are classified as protons.

Figure 38: Breakdown of the different particle track types in neutrino data and simulation as a function of the track  $\cos(\theta)$  angle.



(a) All reconstructed tracks.

(b) All reconstructed tracks that are classified as protons.

Figure 39: Breakdown of the different particle track types in neutrino data and simulation as a function of the track  $\phi$  angle.

## 5.2 Event Selection

The neutral current elastic proton event selection consists of some simple pre-selection cuts to remove events that are very unlike the signal and a final event selection using a logistic regression model based on event details. The proton identification in the previous section was designed to identify any proton induced track from any type of event. To select protons from NC elastic neutrino-proton interactions, we use optical timing and position information, information about activity surrounding the proton candidate, and information about unrelated tracks in the event that may be neutrino induced.

### 5.2.1 Optical Pre-selection

A common optical pre-filter is run over MicroBooNE data before any of the events are reconstructed. The common optical filter requires both that there is a sufficient optical flash within a  $2 \mu s$  window starting at the beginning of the  $1.6 \mu s$  beam window and that there is no such flash in a  $2 \mu s$  time window immediately preceding the beam window. This  $2 \mu s$  window before the beam is called the veto window.

First, each of the  $2 \mu s$  windows is sliced into 339 bins that are 94 ns wide. Then the total number of photoelectrons (PE) of any optical pulses that occur within a given bin are added to that time bin. If any of the time bins within the larger  $2 \mu s$  window add up to more than 20 PE, this is considered a sufficient flash to the

optical filter. If any one of these flashes occur within the beam window and none of the flashes occurs within the veto window the event is accepted. Otherwise, it is rejected.

Figure 40 shows the efficiency of the optical pre-selection on simulated NC elastic proton events as a function of true negative four-momentum squared. The overall efficiency in the range between  $Q^2 = 0.1 \text{ GeV}^2$  and  $Q^2 = 1.0 \text{ GeV}^2$  is 81%.

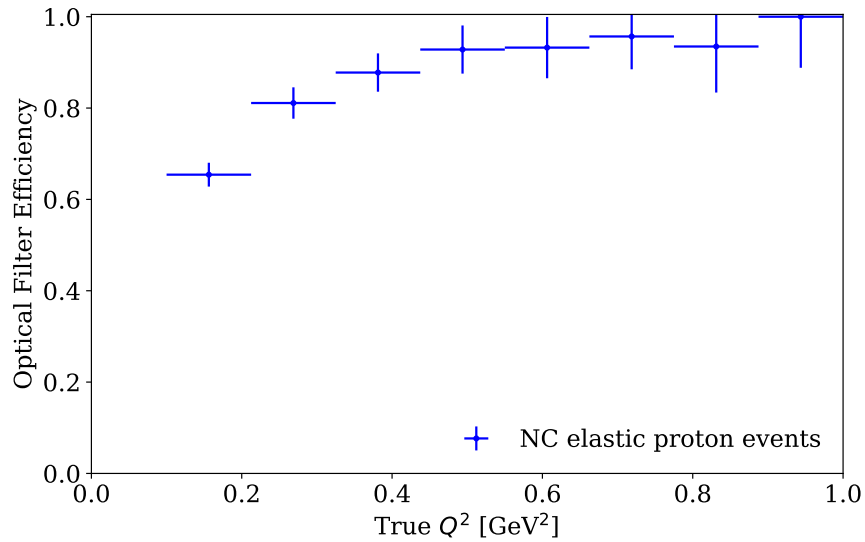


Figure 40: Efficiency of optical pre-selection on simulated NC elastic proton events.

### 5.2.2 NC Elastic Pre-selection

Before making a final neutral current elastic event selection, some simple cuts are made to reject a large number of background events that are very unlikely to be

Table 2: Number of remaining events in  $5 \times 10^{19}$  POT data set after each of the NC elastic pre-selection cuts.

Cut	Simulation	Off-Beam Data	Sim.+Off-Beam	Neutrino Data
Beam Flash	33743	71208	104951	108064
Containment	33742	71207	104949	108058
Length	33523	70426	103949	107123
Proton Score	16805	23491	40296	41867

NC elastic interactions. The result of each of these pre-cuts on the scaled data and simulation samples described in Sec. 5.1.5 are shown in Tab. 2. The first column shows the total number of simulated neutrino events that pass each cut scaled to  $5 \times 10^{19}$  POT, and the second column shows the total number of off-beam data events that pass each cut scaled to  $5 \times 10^{19}$  POT. The third column is the sum of the first two columns which is the expected number of on-beam events in  $5 \times 10^{19}$  POT. The last column shows the total number of on-beam events in the  $5 \times 10^{19}$  POT data after each cut. If the simulation were a perfect representation of reality, the last two columns would be the same. At each cut in the pre-selection the measured number of events in the on-beam neutrino data is within 5% of the expected number of events from simulation and off-beam data.

The first pre-selection requirement is that there is at least one reconstructed optical flash inside the  $1.6 \mu s$  neutrino beam time window. This cut is to reduce

the amount of cosmic background. The next two pre-selection requirements are that there is at least one reconstructed track that is at least 2.5 cm long and fully contained within a fiducial TPC volume. The fiducial volume is defined as being at least 10 cm away from either  $y$  boundary of the TPC active volume and at least 5 cm away from any of the  $x$  and  $z$  boundaries of the TPC active volume. The containment and length requirements don't reduce the number of events by a lot, but they reduce the number of tracks within each event that are considered when trying to select NC elastic proton tracks. The last pre-selection requirement is that there is at least one track in the event that has a proton score from the gradient-boosted decision tree classifier greater than 0.5. To summarize all of the cuts, the set of events after the pre-selection had an interaction in-time with the neutrino beam and at least one track that is likely to be a proton contained in the TPC fiducial volume. The efficiency and purity of simulated NC elastic proton events after each of the cuts is shown in Tab. 3. Figure 41 shows the efficiency of the pre-selection on simulated neutrino events as a function of true negative four-momentum squared.

Figure 42 shows the reconstructed track lengths of the longest remaining proton candidate track in the events remaining after the preselection. The black points in the top plot include the events in the  $5 \times 10^{19}$  POT neutrino data set with statistical uncertainty only. The gray filled histogram includes the off-beam data events scaled to  $5 \times 10^{19}$  POT, and the color filled histograms include the



Table 3: Efficiency and purity of simulated NC elastic proton events after each of the NC elastic pre-selection cuts.

Cut	Efficiency	Relative Efficiency	Purity
Optical Pre-Filter	0.81	0.81	0.004
Beam Flash	0.78	0.97	0.005
Reconstruction	0.62	0.79	0.005
Containment	0.53	0.86	0.005
Length	0.46	0.87	0.005
Proton Score	0.39	0.84	0.009

simulated neutrino events scaled to  $5 \times 10^{19}$  POT. The simulated NC elastic proton events in the TPC are in peach, the simulated charged current events in the TPC are in blue, the simulated neutral current background TPC events are in purple, the simulated events in which the neutrino interaction occurred in the liquid argon but outside of the TPC are in green, and the simulated events in which the neutrino interaction occurred outside of the liquid argon cryostat are in orange. Each of the remaining background types is described in more detail in section 5.3.

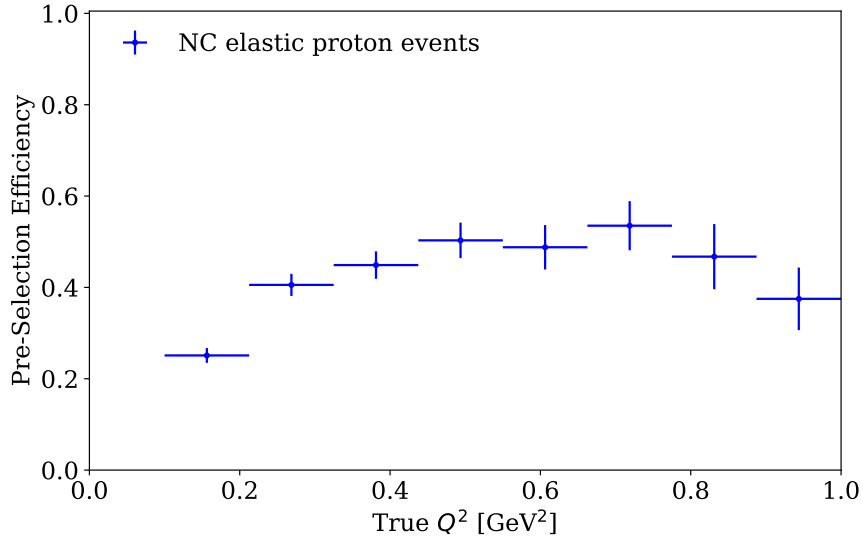


Figure 41: Efficiency of simulated NC elastic proton events as a function of true  $Q^2$  after the pre-selection cuts.

### 5.2.3 Selection Variables

To select NC elastic proton interactions we look for events with a track that is very likely to be a proton, the track is near the reconstructed beam flash, there are no other tracks near the proton candidate track, there are no tracks likely to be from charged current interactions near the reconstructed beam flash, and the proton candidate track is in the direction of the neutrino beam. The following seven variables are used to select these events,

1. the decision tree proton ID score,
2. the shortest distance between either reconstructed endpoint of the candidate track and the next closest endpoint of a different reconstructed track,

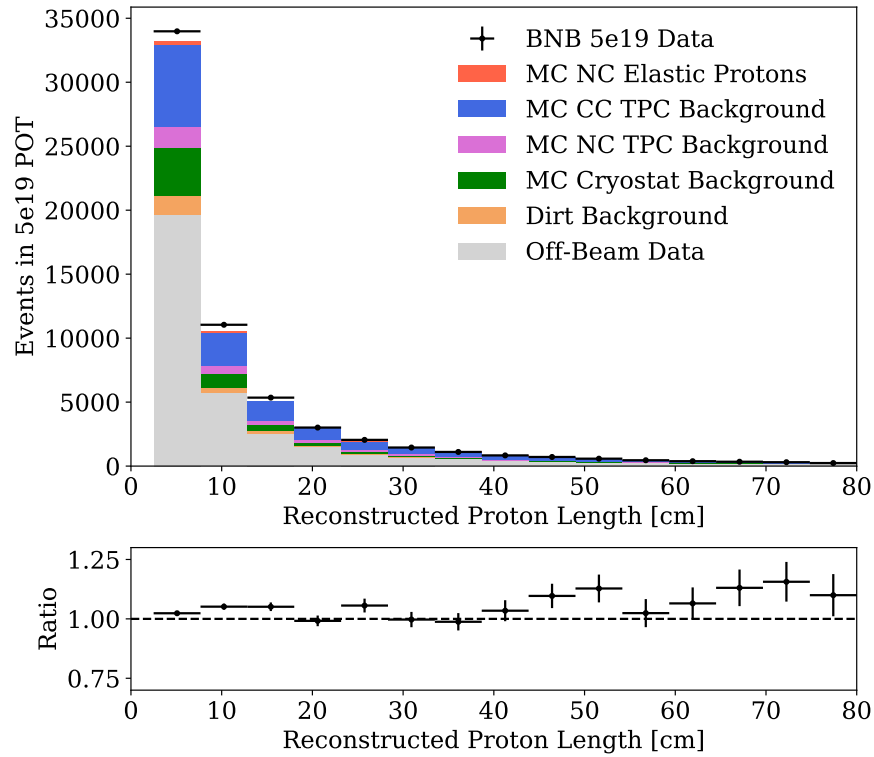


Figure 42: Reconstructed track length of the longest proton candidate track in the events after all pre-selection cuts.

3. the distance from the center of the reconstructed track to the center of the reconstructed beam flash in the  $z$  direction,
4. the distance from the center of the reconstructed track to the center of the reconstructed beam flash in the  $y$  direction,
5. whether or not the reconstructed track is in the neutrino beam direction,
6. the distance between any reconstructed tracks with a decision tree muon ID score greater than 0.5 and the reconstructed beam flash in the  $z$  direction,

7. the distance between any reconstructed tracks with a decision tree pion ID score greater than 0.5 and the reconstructed beam flash in the  $z$  direction.

The center of the reconstructed track is defined as the halfway point between the reconstructed track endpoints in the dimension of interest. The center of the reconstructed flash is the PE-weighted reconstructed center of the flash. The beam flash is defined as a flash whose peak amplitude occurs within the neutrino beam time window. A track is defined as in the beam direction if its reconstructed endpoint is downstream (higher in  $z$ ) than its reconstructed start point. If there are no tracks in the event identified as a muon (in item 6.) or a pion (in item 7.), the value is set to 999 cm, which is close to the maximum distance a track can be from a flash in the TPC. A comparison of each of these variables between data and simulation after the pre-selection is shown in Figs. 43-49.

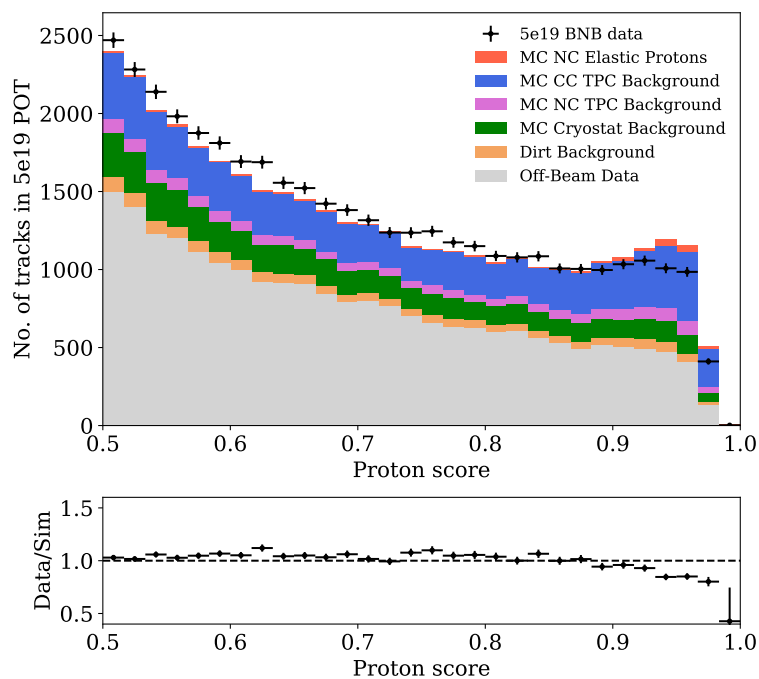


Figure 43: Decision tree proton score after the pre-selection cuts.

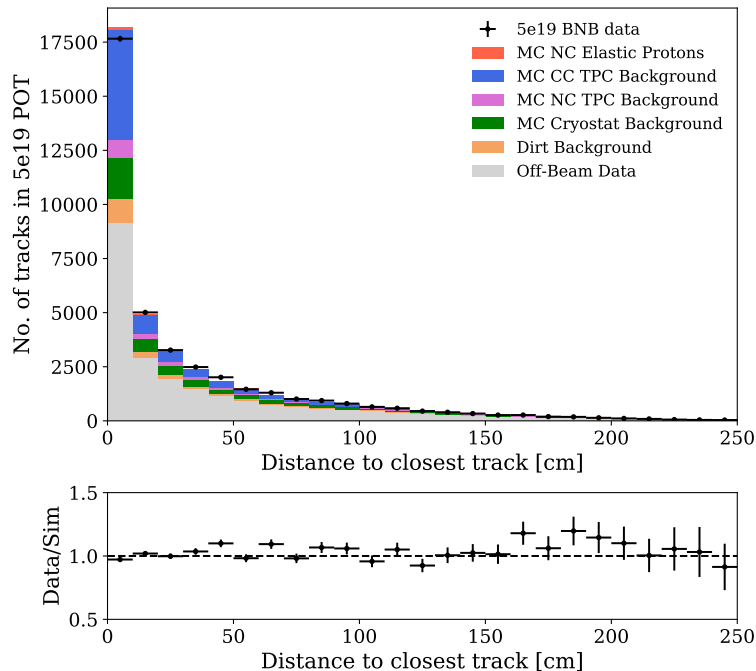


Figure 44: Distance to the next closest track after the pre-selection cuts.

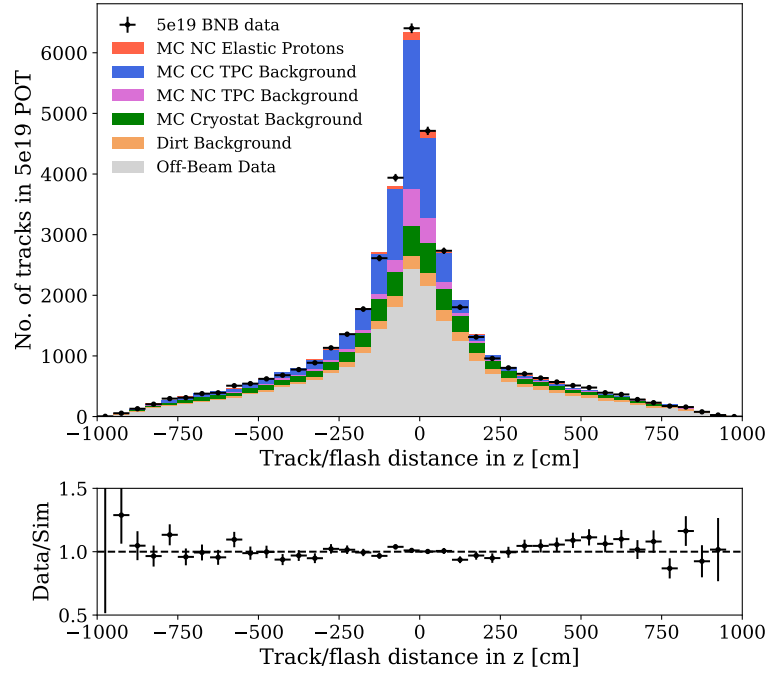


Figure 45: Distance to the beam flash in  $z$  after the pre-selection cuts.

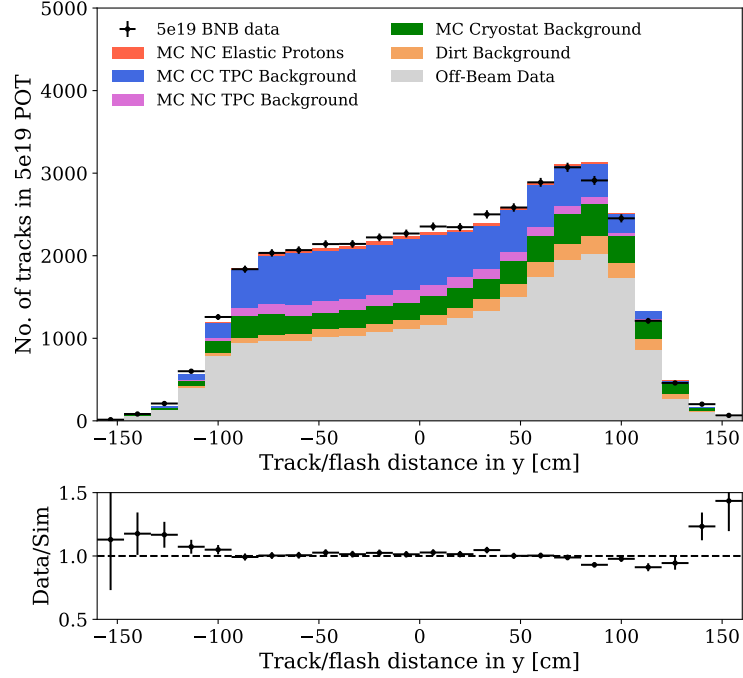


Figure 46: Distance to the beam flash in  $y$  after the pre-selection cuts.

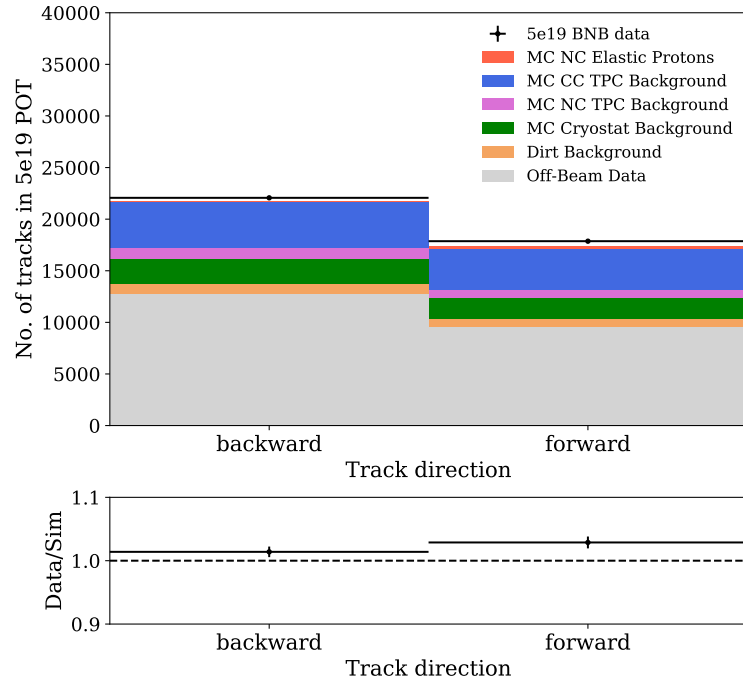


Figure 47: Whether or not the track is forward going after the pre-selection cuts.

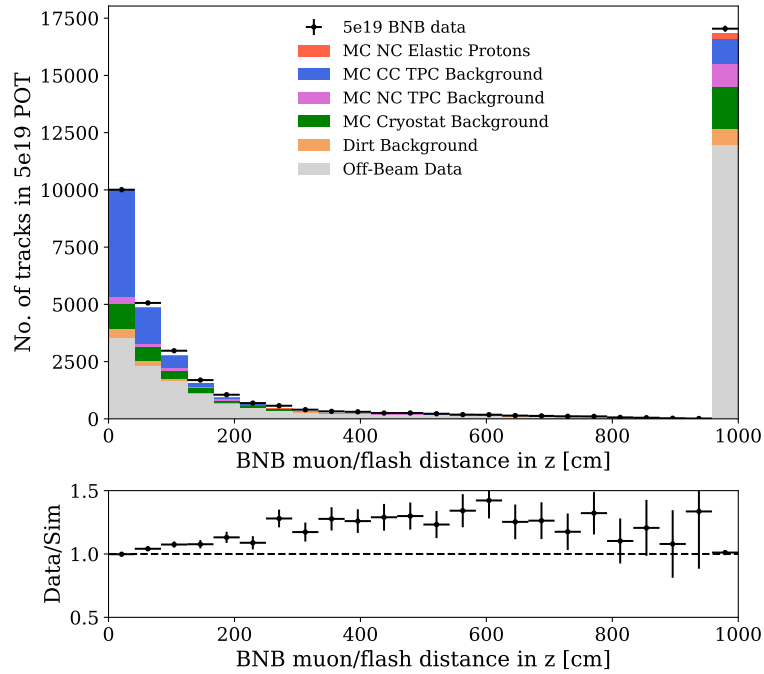


Figure 48: Distance between muon track and flash after the pre-selection cuts.

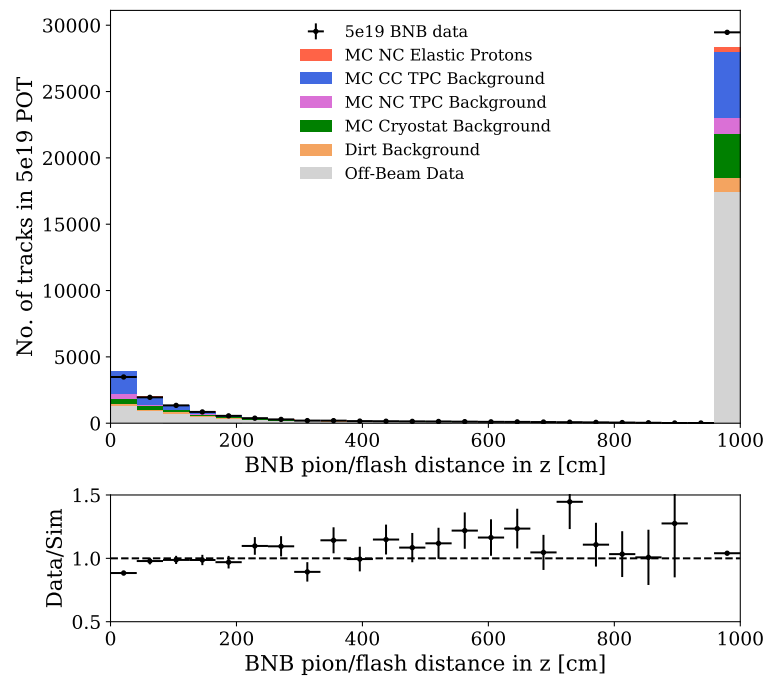


Figure 49: Distance between pion track and flash after the pre-selection cuts.



#### 5.2.4 Logistic Regression

To determine which events are NC elastic like based on these seven variables, we use them as input to a logistic regression model [108]. The difference between logistic regression and linear regression models is that a binary (or logistic) outcome is being modeled. Otherwise, the methods used to fit logistic and linear regression models follow the same principles. In our logistic regression a multi-dimensional sigmoid function is fit to the binary signal and background data. The data that the model is fit to binary (either background, 0, or signal, 1), but the model itself is continuous between zero and one. The output is a score that can be used to determine how signal-like a data point is.

$$S(g(\mathbf{x})) = \frac{e^{g(\mathbf{x})}}{1 + e^{g(\mathbf{x})}},$$

where  $g(\mathbf{x})$  is a linear combination of the selection variables,  $\mathbf{x}$ ,

$$g(\mathbf{x}) = w_0 + w_1x_1 + w_2x_2 + \dots + w_7x_7.$$

Here  $x_1$  is item 1 from the list in Sec. 5.2.3 (the proton ID score),  $x_2$  is item 2 (the distance to the next closest track), etc. The set of weights,  $w_0, \dots, w_7$  is determined from a fit to the data. We determined these weights using the StatsModels module [109] in Python to fit the model to a subset of the simulated neutrino events described in Sec. 5.1.4 along with a set of simulated cosmic events that produce an optical flash in-time with the neutrino beam. The simulated set

of in-time cosmics should match the off-beam data in a perfect simulation. For the fit, 998 simulated NC elastic proton events, 2000 simulated background neutrino events, and 1806 simulated background in-time cosmic events were used. The final set of weights used are

$$\begin{aligned}
w_0 &= -5.943956, \\
w_1 &= 5.388985, \\
w_2 &= 0.021189, \\
w_3 &= -0.016710, \\
w_4 &= -0.017510, \\
w_5 &= 0.592240, \\
w_6 &= 0.001084, \\
w_7 &= 0.000989.
\end{aligned}
\tag{95}$$

The  $\chi^2/DoF$  of the fit was  $8161/4804 = 1.70$ . The degrees of freedom in the fit are the 4804 simulated events used.

Figure 50 shows the output logistic regression score on the simulated neutrino events plus the off-beam data compared to the output on the  $5 \times 10^{19}$  POT neutrino beam data. Again, the black points in the top plot include the  $5 \times 10^{19}$  POT subset of neutrino data with statistical uncertainty only, the gray filled histogram includes off-beam data scaled to  $5 \times 10^{19}$  POT, and the color filled histograms include the simulated neutrino interactions overlaid with cosmic data backgrounds scaled to  $5 \times 10^{19}$  POT. The peach color is NC elastic proton events in the TPC,

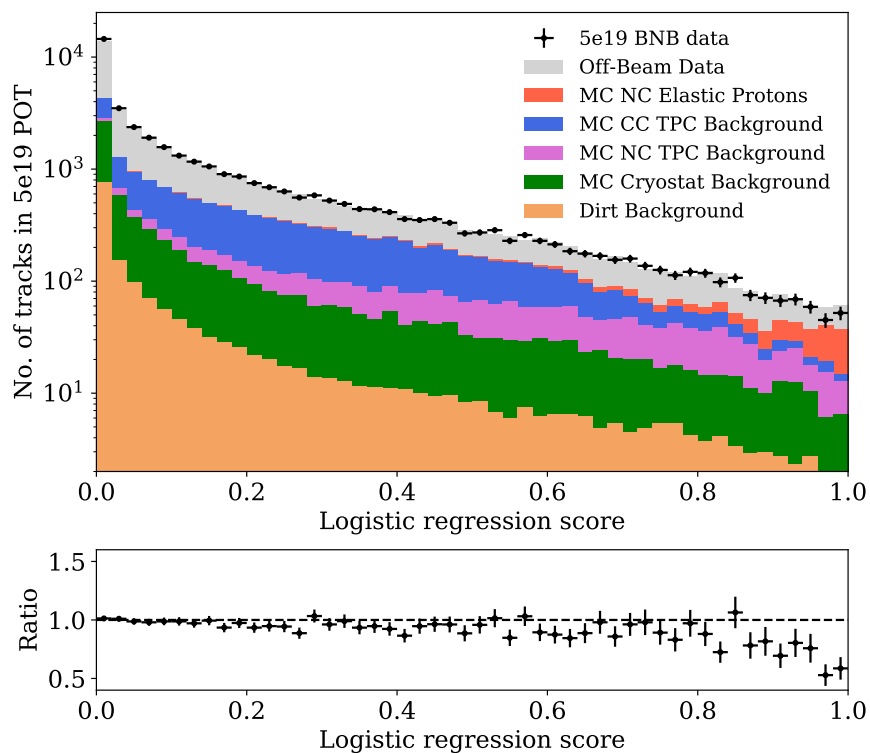
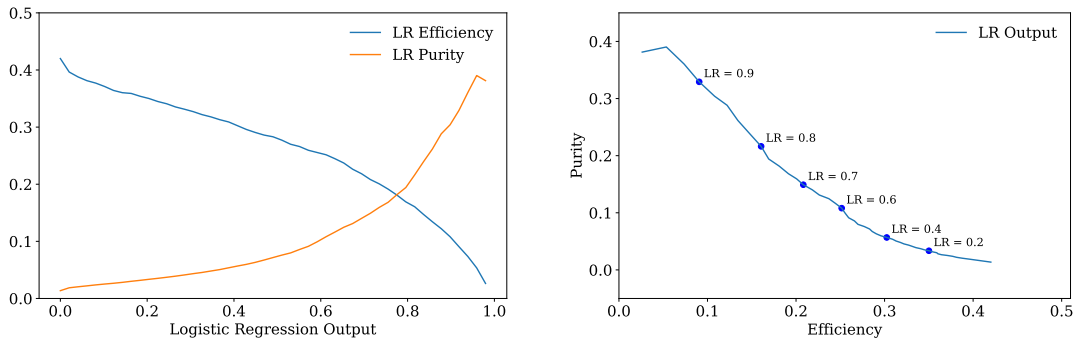


Figure 50: Logistic regression NC elastic events selection score.

blue is charged current background events in the TPC, purple is neutral current background events in the TPC, green is background from neutrino interactions in the liquid argon outside of the TPC, and yellow is background from neutrino interactions outside of the liquid argon cryostat. The higher the logistic regression score, the more likely the event is NC elastic proton. The bottom plot shows the ratio of the  $5 \times 10^{19}$  POT neutrino beam data to the sum of the simulated neutrino events and off-beam data.

Figure 51 shows the efficiency and purity of the NC elastic signal for cutting on given logistic regression scores. Figure 51a shows both the efficiency curve and

the purity curve of the NC elastic event selection on the simulated neutrino events and off-beam data as a function of the logistic regression score cut. This cut is imposed after the pre-selection cuts, which is why the efficiency only goes as high as 0.41. Figure 51b shows the purity on the  $y$ -axis as a function of efficiency on the  $x$ -axis for different logistic regression score cuts. Selected score cut values are labelled along the curve.



(a) Efficiency and purity curves overlaid.

(b) Purity as a function of efficiency.

Figure 51: Efficiency and purity of the NC elastic proton event selection given several different cut values on the logistic regression score.

In this analysis we choose a logistic regression score cut of 0.9 to minimize the backgrounds as much as reasonably possible. This gives us an overall NC elastic proton event selection efficiency of 0.11 and a purity of 0.30 based on simulated neutrino events and off-beam data. The efficiency of this selection on simulated NC elastic proton events is shown in Fig. 52 as a function of true  $Q^2$ . The overall shape of the NC elastic proton selection efficiency remains relatively flat across

the  $Q^2$  range of interest.

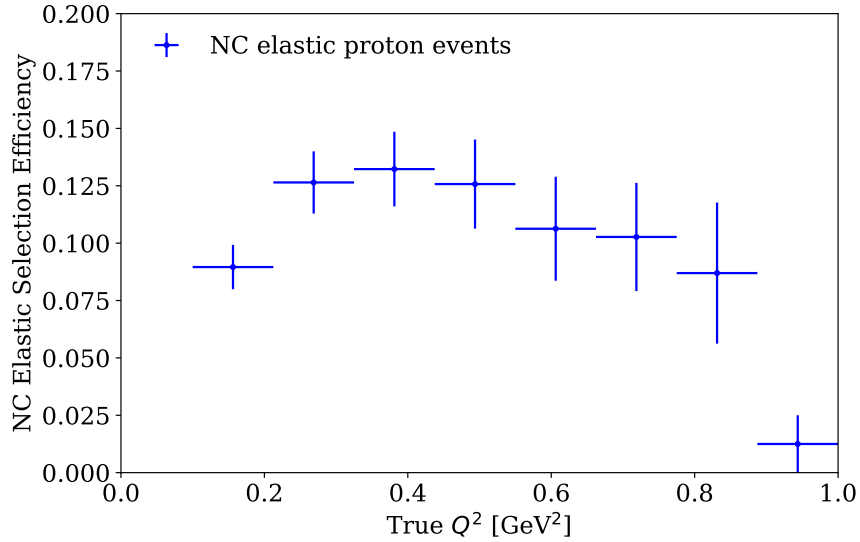


Figure 52: Efficiency of the NC elastic proton event selection on simulated NC elastic proton events as a function of true  $Q^2$ .

### 5.2.5 Comparison of Event Selection on Neutrino Beam Data to Expectation from Simulation and Off-Beam Data

Figures 54-58 show comparisons between the final selected events in the neutrino beam data and the neutrino simulation and off-beam data. The full Run I and simulated data sets are used in these comparisons. Each of the data sets is scaled to the size of the Run I data ( $1.6 \times 10^{20}$  POT). The data sets and the scaling factors are listed below.

1. Run I  $1.6 \times 10^{20}$  POT neutrino data set
  - Number of events: 621,231

- POT:  $1.6 \times 10^{20}$
- Number of triggers: 35,913,783
- Normalization factor: 1

2. Run I off-beam data subset

- Number of events: 941,584
- Number of triggers: 73,761,274
- Normalization factor: 0.4869

3. Neutrino Monte Carlo with cosmic data overlay set

- Number of events: 750,629
- POT:  $9.37 \times 10^{20}$
- Flux correction factor: 1.029
- Normalization factor: 0.1676

4. Neutrino dirt Monte Carlo with simulated cosmic data set

- Number of events: 105,214
- POT:  $3.73 \times 10^{20}$
- POT normalization factor: 0.4326
- Data driven normalization factor: 0.5 (described in Sec. 5.3.1)

- Normalization factor: 0.2163

Figure 53 shows the comparison between neutrino beam data and the expectation from simulation and off-beam data as a function of reconstructed negative four-momentum transfer squared, which is calculated from the proton kinetic energy,  $T_p$ , as derived in Sec. 1.3,

$$Q_p^2 = 2T_p M_p,$$

where  $M_p$  is the proton mass. This is the comparison that will be used to extract the strange axial form factor parameters (described in Sec. 6). In this and all of the following figures in this section, the black points in the top plot include the neutrino beam data with statistical uncertainty only, the gray filled histogram includes off-beam data, and the color filled histograms include simulated neutrino events with overlaid cosmic data backgrounds. The peach are NC elastic proton events, the blue are CC background events in the TPC, the purple are NC background events in the TPC, the green are background events from neutrino interactions in the liquid argon outside of the TPC, and the yellow are backgrounds from neutrino interactions outside of the liquid argon cryostat. Each of these backgrounds is described in more detail in Sec. 5.3. The bottom plot shows the ratio of the selected neutrino beam data events to the combination of selected simulated neutrino events and selected off-beam data events. If the simulation and data were exactly the same, the ratio would be equal to one.

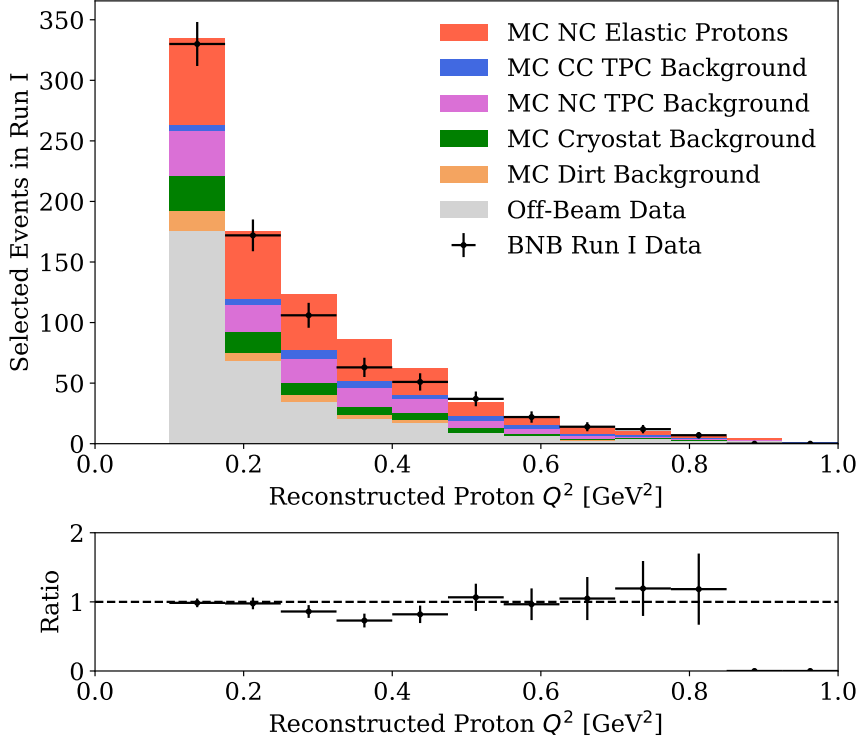


Figure 53: Selected NC elastic proton events as a function of reconstructed  $Q_p^2$ .

Figures 54 and 55 show the comparison between neutrino beam data and the expectation from neutrino simulation and off-beam data as a function of the reconstructed proton angle. Figure 54 shows this as a function of the cosine of the proton polar angle from the beam direction,  $\cos(\theta_p)$ , and Fig. 55 shows it as a function of the proton azimuthal angle around the beam direction,  $\phi_p$ . The  $\cos(\theta_p)$  distribution of the off-beam data appears to have a peak near 0.4 and another peak near 1.0. These peaks in the off-beam data are a combination of angle dependence of the cosmic ray flux and the reconstruction efficiency of the MicroBooNE TPC. The simulated neutrino-induced tracks tend to be more in the



direction of the neutrino beam. The  $\phi_p$  distribution has clear peaks in the off-beam data at  $\phi_p = \pm\pi/2$ . The  $-\pi/2$  peak corresponds to tracks that are reconstructed as vertically down-going in the TPC and the  $+\pi/2$  peak corresponds to tracks that are reconstructed as vertically up-going. Both peaks are due to down-going cosmic tracks, but some of these tracks are mis-reconstructed as going up through the detector. The simulated neutrino-induced tracks are relatively isotropic in  $\phi$ .

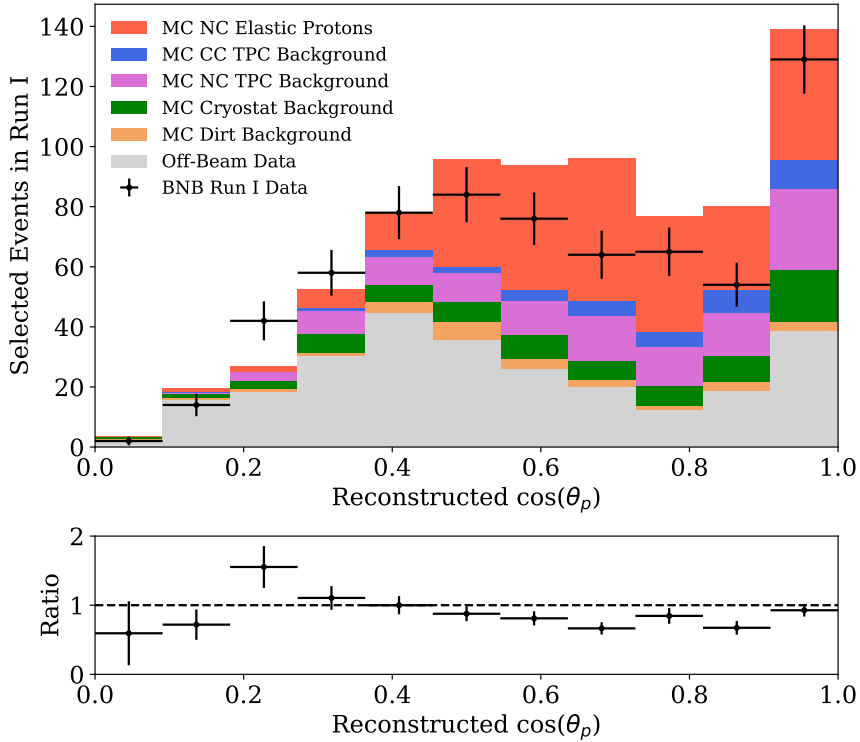


Figure 54: Selected NC elastic proton events as a function of reconstructed  $\cos(\theta_p)$ .

Figures 56, 57, and 58 show the comparison between neutrino beam data and the expectation from neutrino simulation and off-beam data as a function of the

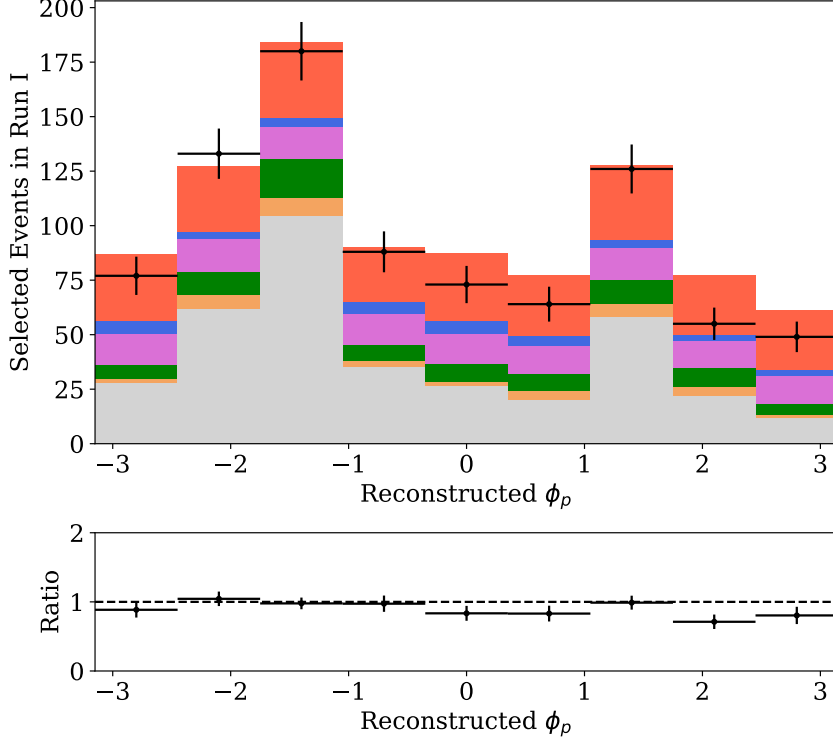


Figure 55: Selected NC elastic proton events as a function of reconstructed  $\phi_p$ .

start points of the reconstructed tracks in  $x$ ,  $y$ , and  $z$ , respectively. In the reference frame used by MicroBooNE, the  $x$  dimension is the dimension across the width of the detector with the anode near  $x = 0$  cm and the cathode near  $x = 250$  cm, the  $y$  dimension is the vertical dimension across the height of the detector with the bottom of the TPC near  $y = -100$  cm and the top near  $y = +100$  cm, and the  $z$  dimension is the dimension in the beam direction across the length of the detector with the upstream (front) end of the detector near  $z = 0$  cm and the downstream (back) end of the detector near  $z = 1000$  cm.

In Fig. 56, the higher number of selected events at low  $x$  is due to the greater

charge and light collection efficiency for particles that are near the anode since both the TPC readout wires and the PMTs are on the anode side of the detector at  $x = 0$ .

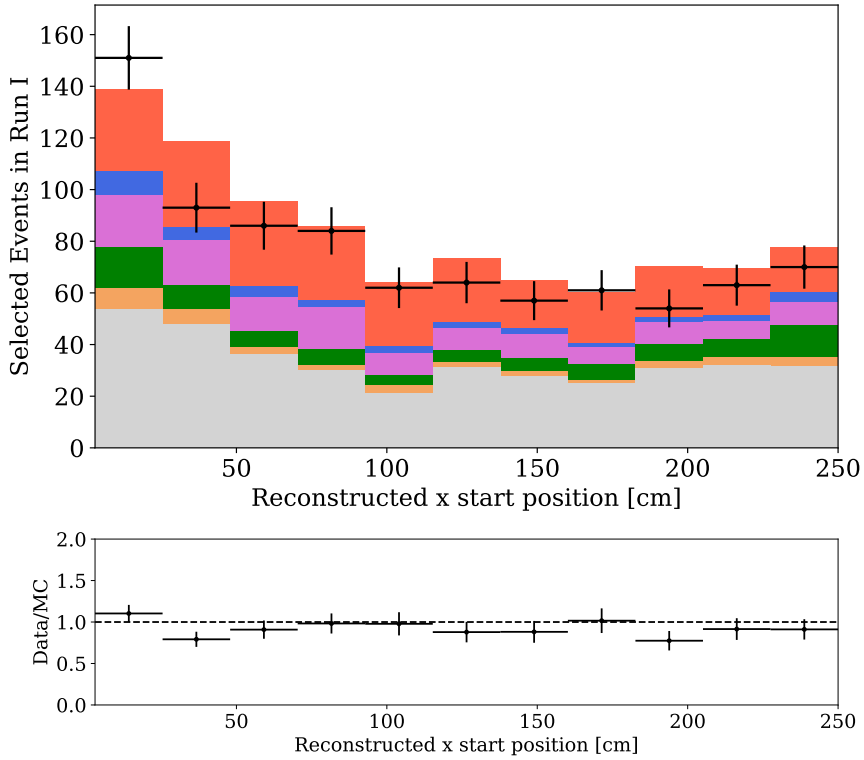


Figure 56: Selected NC elastic proton events as a function of reconstructed  $x$  position.

In Fig 57, the increase in the number of selected off-beam tracks from the bottom ( $y = -100$  cm) to the top ( $y = +100$  cm) of the TPC is due to the fact that most cosmic tracks enter through the top of the detector, and not all tracks make it all of the way through. The greater number of selected simulated neutrino

events near the center of the detector ( $y = 0$ ) might be due to the fact that tracks at an angle are more likely to fail the fiducial TPC containment cut when they are closer to the edges of the detector. This may be more visible in the  $y$  dimension than the  $x$  because there are no other competing effects.

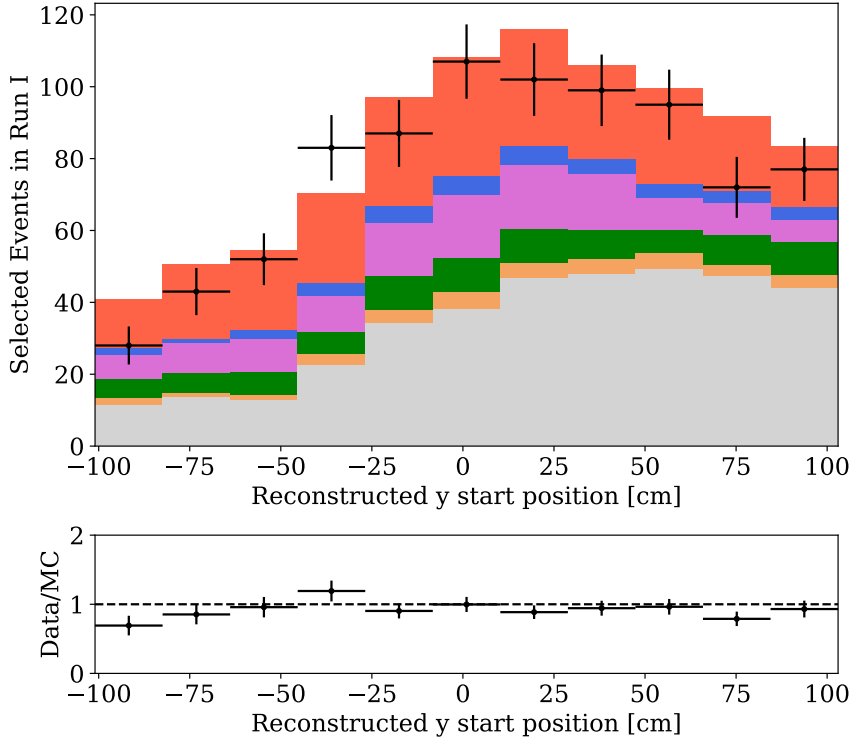


Figure 57: Selected NC elastic proton events as a function of reconstructed  $y$  position.

In Fig. 58, the selected off-beam and the selected neutrino events are fairly uniform in the  $z$  dimension. The small decrease in selected events near the very back of the TPC ( $z \approx 1000$  cm) is due to the fact that tracks are less likely to be contained in the TPC if they are produced so close to the back edge.

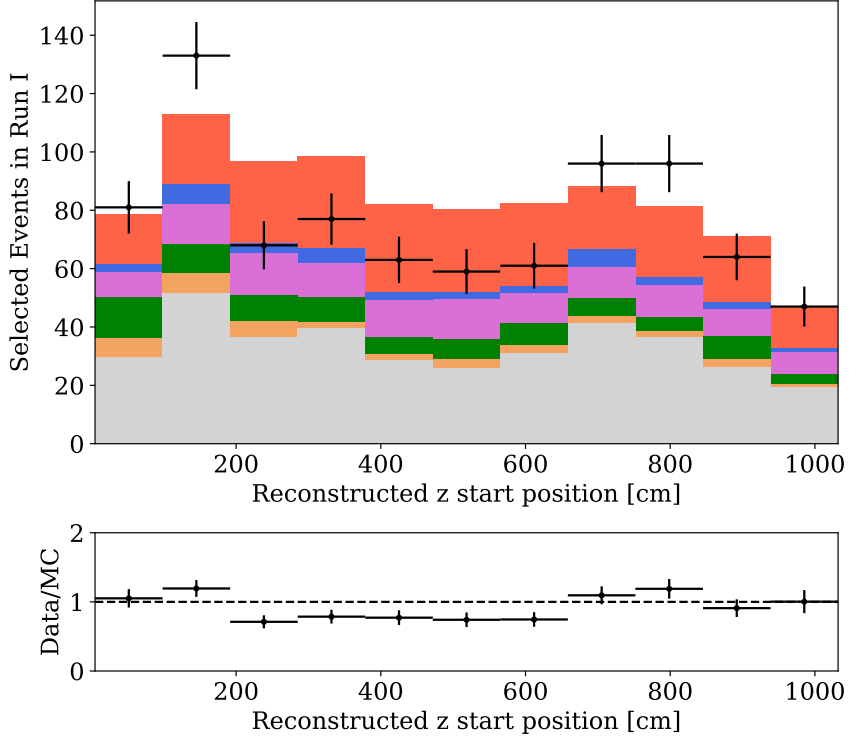


Figure 58: Selected NC elastic proton events as a function of reconstructed  $z$  position.

### 5.3 Remaining Backgrounds

The remaining backgrounds after the final neutral current elastic neutrino-proton event selection include events in which a neutrino interaction occurred outside of the liquid argon cryostat, events in which a neutrino interaction occurred in the liquid argon but outside of the active TPC volume, events in which a neutrino interaction that was not an NC elastic neutrino-proton interaction occurred within the active volume, and events in which a cosmic interaction occurred at the same time as the neutrino beam spill, and there was no neutrino interaction. In all

of these cases except the last one, tracks from the neutrino interaction produced scintillation light within the cryostat that triggered the event readout, but it could have been either a track from the neutrino event or cosmic-induced track that was tagged as the proton candidate.

### 5.3.1 Beam Induced Dirt Background

The dirt background, described in Sec. 3.2, consists of any selected events that are due to a neutrino interaction outside of the liquid argon cryostat. The majority of this background is caused by neutrons entering the detector and scattering protons. Sometimes the neutron is directly from the neutrino interaction in the dirt, and other times secondary particles from the neutrino interaction enter the cryostat and cause a flash, but an unrelated proton from a cosmic neutron happens to be spatially nearby. Figure 59 shows the total dirt background that passes the NC elastic proton event selection as a function of reconstructed  $Q^2$ . The filled yellow histogram includes all remaining dirt events, and the pale yellow box shows the statistical uncertainty on the number of events. An overall data-driven normalization has been applied and is described below. The hatched histograms overlaying the yellow filled histogram show what type of particle was selected as the NC elastic proton candidate. The blue hatched histograms include the events in which a proton was selected with hatching diagonal for cosmic protons and vertical for neutrino beam induced protons. The red hatched histograms include

the events in which a track from an electromagnetic shower was selected with diagonal hatching for cosmic showers and vertical hatching for neutrino beam induced showers. The purple hatched histograms include all other types of tracks.

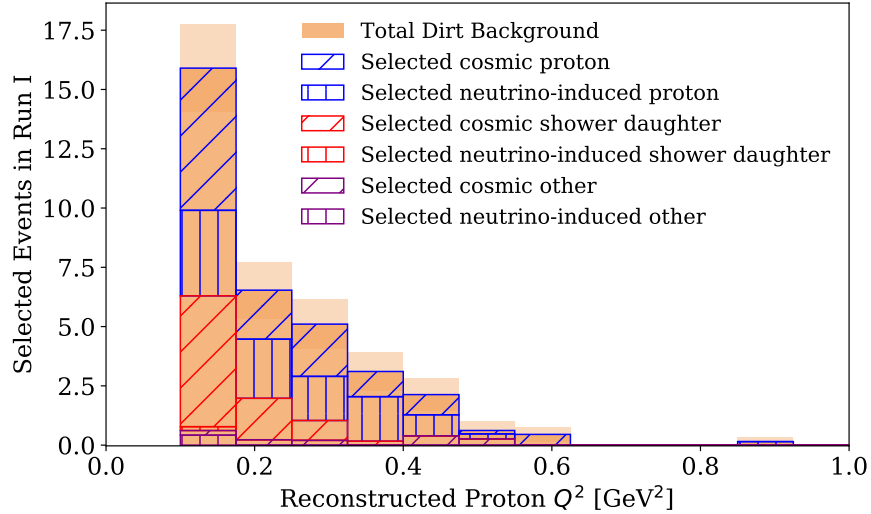


Figure 59: Remaining background dirt events after NC elastic proton event selection as a function of  $Q^2$ .

The dirt background is difficult to simulated for several reasons. First, the neutrino interactions occur outside of the detector in the surrounding infrastructure which is much more difficult to model perfectly than uniformly dense liquid argon. The composition and density distribution of the dirt outside of the detector hall is not known perfectly and changes throughout the year due to rain, freezing, and other weather. The neutrino interactions can also occur in the walls of the detector building, the electronics racks, the cryogenic plumbing, or anything else

that exists near the detector.

Additionally, the dirt sample is difficult to simulate because it depends on modeling the cosmic neutron flux, which is not well known at the energies that we are interested in [110]. The different cosmic ray flux models that were tested in MicroBooNE studies [7] vary by almost an order of magnitude in their predictions of the proton flux at proton kinetic energies below 100 MeV, shown in Fig. 60. The gray line in the figure represents the model used in MicroBooNE simulation, described in Sec. 4.1.2, which is the most consistent with measurements in Ref. [110].

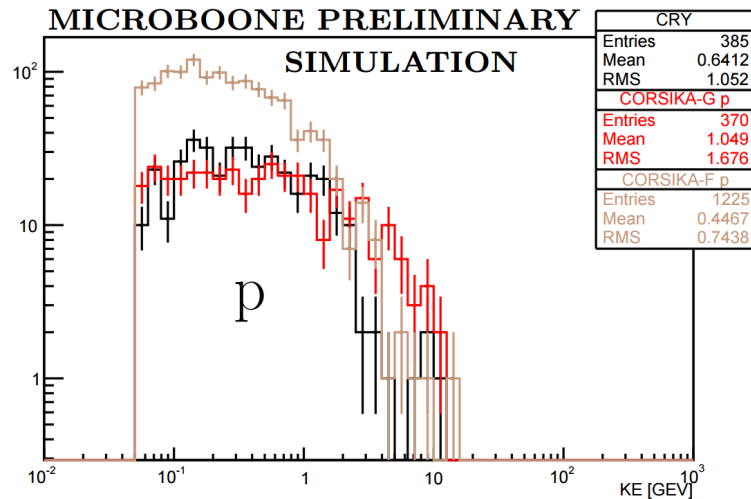


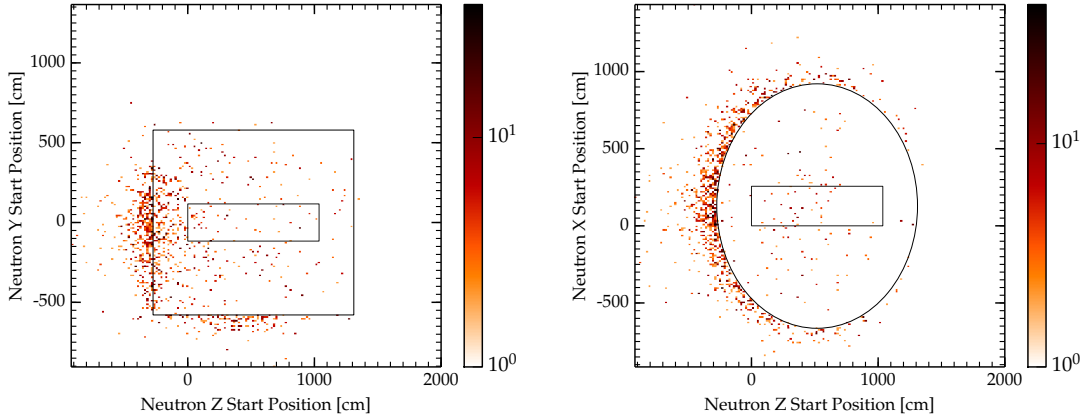
Figure 60: Comparison of the cosmic proton flux predictions from three different generator models in MicroBooNE [7].

Because of the difficulty on modeling the dirt background, we perform a data-driven normalization of the dirt sample. A similar scaling was performed in the



MiniBooNE neutral current elastic cross section measurement [5] using different observable variables. In MicroBooNE, the contribution from dirt events can be seen most clearly in the  $z$  distribution of reconstructed flashes since most of the dirt interactions that produce a signal in the detector occur upstream as shown in Figure 61. Figure 61a shows the view from the side of the detector hall with the neutrino beam going from left to right. The inner box shows the dimensions of the MicroBooNE TPC and the outer box shows the dimensions of the outer walls of the building. Figure 61b shows the top view of the detector with the beam again going from left to right. The inner box shows the dimension of the MicroBooNE TPC and the outer circle shows the dimensions of the outer wall of the building. Figure 62 shows a comparison of this distribution between the  $5 \times 10^{19}$  POT subset of neutrino-beam data and the expectation from neutrino simulation and off-beam data *without* the contribution from dirt interactions. There is a clear excess in neutrino beam data events in the upstream end of the detector ( $z < 300$  cm). Figure 63a shows the difference between the two distribution in Fig. 62, and Fig 63b shows the  $z$  distribution of reconstructed flashes from simulated dirt events. The shapes are similar, but the scale is not. We use these two distributions to extract an overall normalization factor for the dirt simulation.

The normalization factor is determined by performing a weighted-least-squares regression on the  $z$  distribution of the reconstructed flashes. The value that we



(a) Position in the  $z - y$  plane (looking at the side of the detector hall). (b) Position in the  $z - x$  plane (looking down at the detector hall).

Figure 61: Position of simulated neutrino interactions which produce a neutron that enters and interacts in the MicroBooNE TPC.

minimized is

$$WLS = \sum_{i=0}^{N_{\text{bins}}} w_i r_i^2,$$

where the sum is over the bins in the flash position distribution,  $w_i$  is the variance of the excess of measured events over the expectation in bin  $i$ , and

$$r_i = x_i^{(\text{data exc.})} - a \cdot y_i^{(\text{dirt})}.$$

Here  $x_i$  is the number of the data excess events in bin  $i$ ,  $y_i$  is the number of simulated dirt events in bin  $i$ , and  $a$  is the dirt normalization factor that we are trying to determine. Using the Python SciPy [111] Optimize library, we found that  $WLS$  is minimized at  $a = 0.5$ . The estimated uncertainty on this normalization factor is 50%, which covers the statistical variations in the first peak in the  $z$  flash dis-

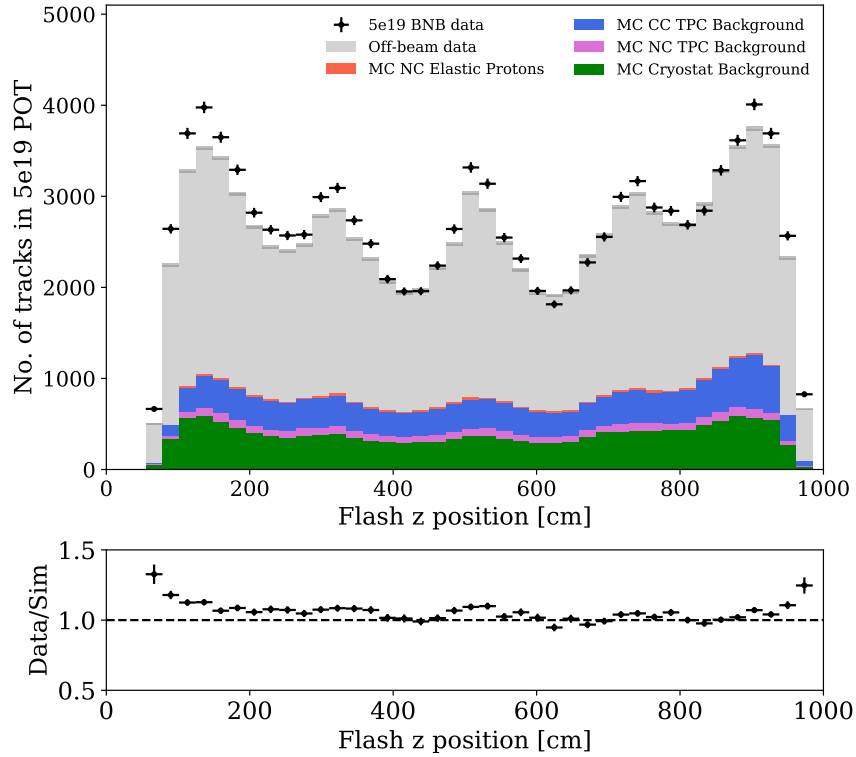
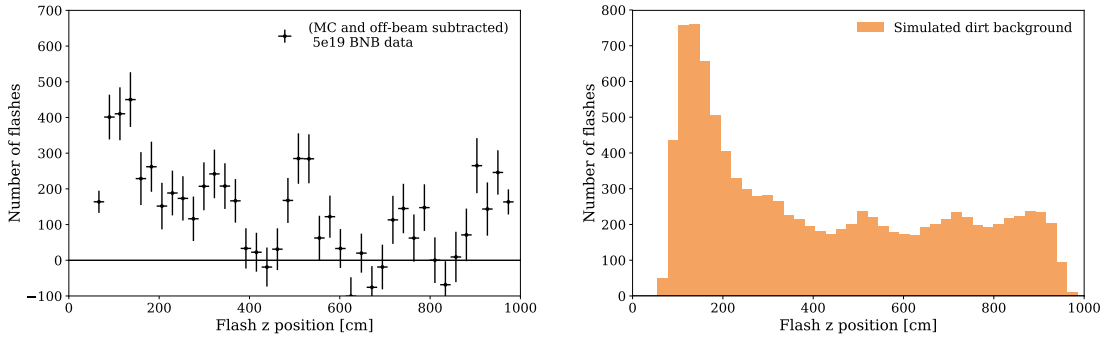


Figure 62: Comparison between neutrino beam data and the expectation from simulation and off-beam data of the  $z$  distribution of reconstructed flashes *without* the contribution from neutrino interactions in the dirt.

tribution. Figure 64 shows the distribution of the  $z$  position of the reconstructed flashes in the neutrino beam data compared to the expectation from the neutrino simulation and off-beam data including the normalized simulated dirt events. The brown bars show the 50% uncertainty on the dirt normalization factor.

With the normalization, the total number of background dirt events remaining after the NC elastic proton event selection (scaled to  $1.6 \times 10^{20}$  POT) is

$$N_{\text{dirt}} = 34 \pm 6 (\text{stat.}) \pm 17 (\text{syst.}).$$



(a) Difference in the number of flashes between neutrino beam data and the expectation from simulation and off-beam data *without* the contribution from neutrino interactions in the dirt. (b) Distribution of flashes from dirt simulation.

Figure 63: Distribution of the  $z$  position of reconstructed flashes in MicroBooNE.

The normalized distribution of remaining events is the one shown in Fig. 59. The effect on the overall systematic uncertainty on the NC elastic proton selection due to the uncertainty on the number of dirt events is shown in Sec. 5.4.3.

### 5.3.2 Beam Induced TPC and Cryostat Background

The beam induced TPC and cryostat background includes all background events in which there was a neutrino beam interaction inside the liquid argon cryostat. This background can be separated into cryostat background events, charged current TPC background events, and neutral current TPC background events.

Like the dirt background, the neutrino interactions in the cryostat background

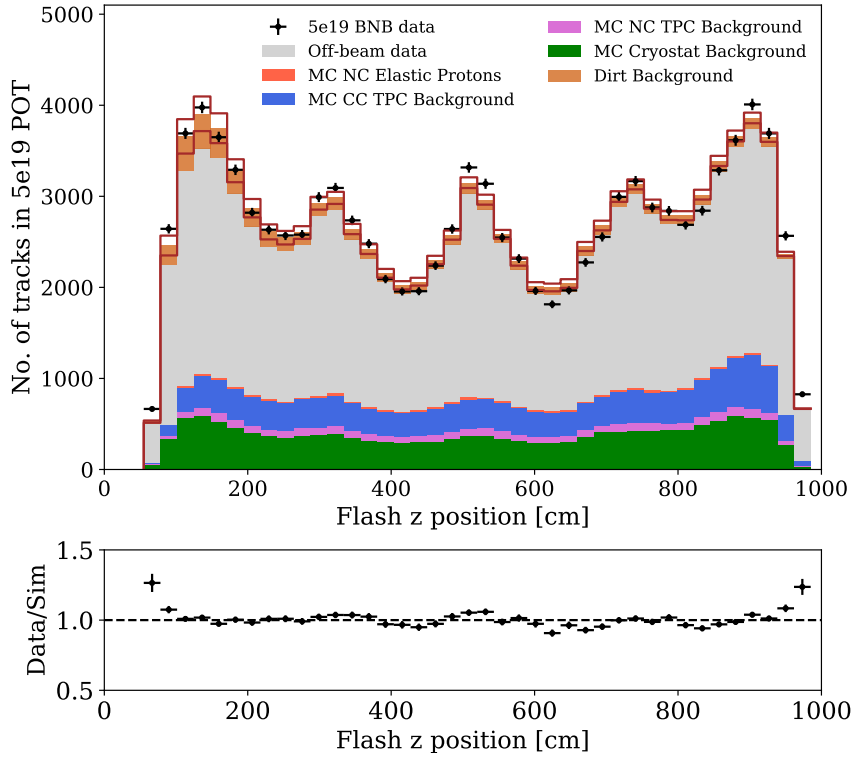


Figure 64: Comparison between neutrino beam data and the expectation from simulation and off-beam data of the  $z$  distribution of reconstructed flashes *with* the contribution from neutrino interactions in the dirt.

occur outside of the active TPC volume, but they do occur inside the liquid argon cryostat (or in the wall of cryostat vessel itself), making them easier to simulate for the reasons described in Sec. 5.3.1. Like the other two background categories in this section, the cryostat background is included in the simulated neutrino data set with an overlaid cosmic data background. Figure 65 shows the remaining cryostat background events in the final NC elastic event selection as a function of reconstructed  $Q^2$ . The green filled histogram includes all remaining cryostat

background events, and the pale green box represents the statistical uncertainty on the number of events. The hatched histograms overlaid on the green filled histogram show what type of particle was selected as the NC elastic proton candidate. The purple diagonally hatched histogram includes events in which an overlaid cosmic track was selected, the blue hatched histogram includes events in which a neutrino-induced proton was selected, the red hatched histogram includes events in which a track from a neutrino-induced electromagnetic shower was selected, and the purple vertically hatched histogram includes events in which any other neutrino-induced track was selected. Again, the majority of the tracks selected as the proton candidate were actual protons. Cosmic tracks are the next most common track selected as the proton candidate in these events. Since real cosmic data are used as the background, we don't know how many of the selected cosmic tracks were actual protons. The other background tracks are negligible. The number of remaining cryostat background events after the NC elastic proton event selection (scaled to  $1.6 \times 10^{20}$  POT) is  $N_{\text{cryo}} = 78 \pm 10$ .

The CC TPC background includes all events in which a charged current neutrino interaction occurred in the active TPC volume. Figure 66 shows the remaining CC TPC background after the NC elastic proton selection as a function of reconstructed  $Q^2$ . The blue filled histogram includes all remaining CC TPC events, and the pale blue box represents the statistical uncertainty on the number of events. The type of particle selected as the proton candidate is shown

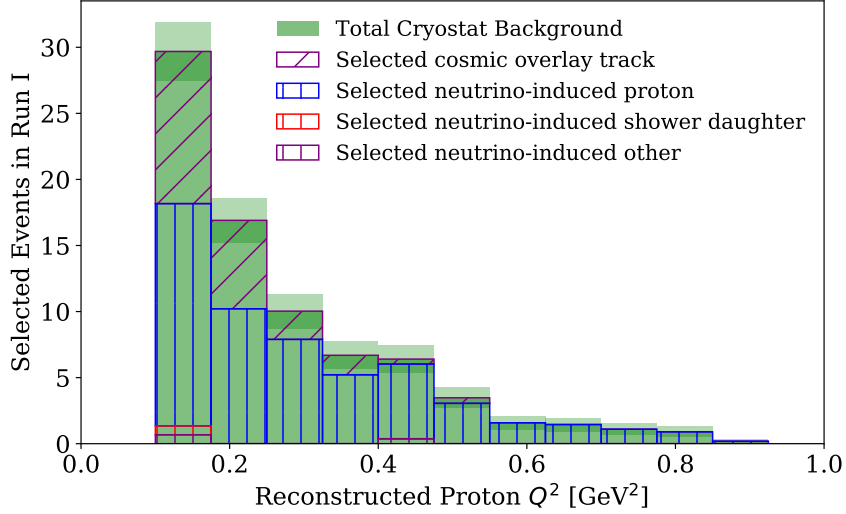


Figure 65: Remaining background cryostat events after NC elastic proton event selection as a function of  $Q^2$ .

in the overlaid hatched histograms with the same color and hatching scheme as in Fig. 65. A large majority of the tracks selected as the proton candidate are actual neutrino-induced protons. These background proton tracks can be from a charged current interaction in which a neutron is ejected from the nucleus during the initial interaction which later scatters a proton some distance from the interaction vertex. If the muon from the charged current interaction is not correctly identified as a neutrino-induced muon by the gradient-boosted decision tree classifier the proton can be tagged as NC elastic. The proton tracks could also be primary particles from a charged current interaction, but the muon from the interaction is either not reconstructed or reconstructed poorly enough that it is not associated with the same vertex as the proton and also not identified correctly by

the decision tree classifier. The next most common tracks selected as the proton candidate are cosmic overlay tracks. Again, we don't know how many of these tracks are actual protons. In these events, the charged current interaction was either not reconstructed or poorly reconstructed, and a nearby unrelated cosmic track was selected. The number of remaining CC TPC background events after the NC elastic proton event selection (scaled to  $1.6 \times 10^{20}$  POT) is  $N_{\text{CC}} = 38 \pm 8$ .

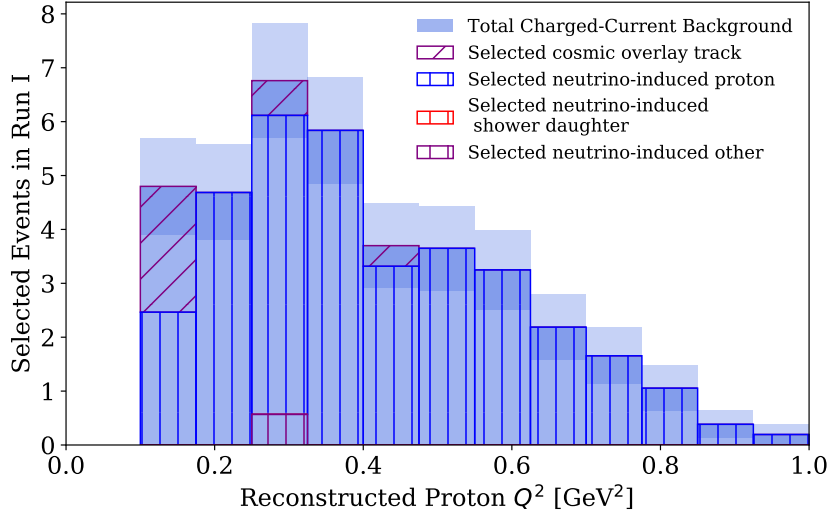


Figure 66: Remaining CC TPC events background after NC elastic proton event selection as a function of  $Q^2$ .

The NC TPC background includes all events in which a neutral current neutrino interaction occurred in the active TPC volume, but the interaction was not a neutral current elastic neutrino-proton interaction. Figure 67 shows the remaining NC TPC background after the NC elastic proton selection as a function of



reconstructed  $Q^2$ . The purple filled histogram includes all remaining NC TPC background events, and the pale purple box represents the systematic uncertainty on the number of events. The overlaid hatched histograms show which type of track was selected as the proton candidate. In the NC TPC background events the selected track was almost always a neutrino-induced proton track.

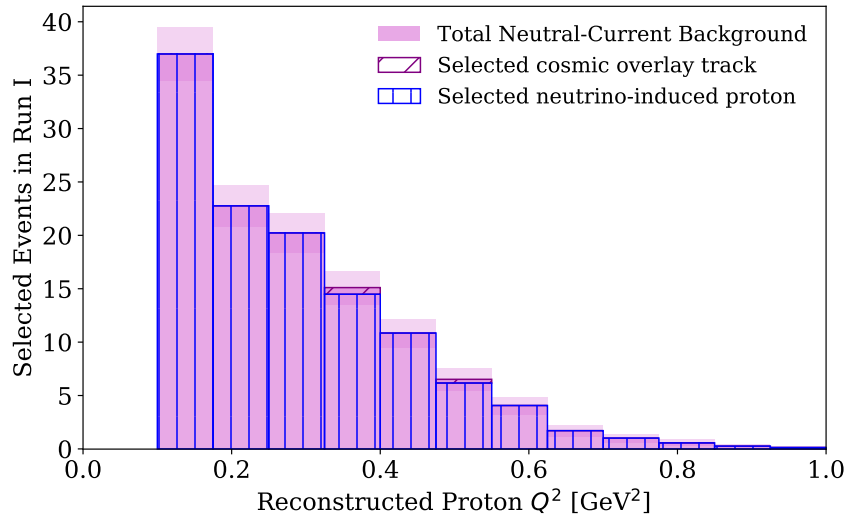


Figure 67: Remaining NC TPC background events after NC elastic proton event selection as a function of  $Q^2$ .

The NC TPC background includes NC elastic neutrino-neutron interactions, NC elastic interactions with a correlated neutron-proton pair in the nucleus, also referred to as meson-exchange-current (MEC) interactions, and all other nonelastic NC neutrino interactions. The NC elastic neutrino-neutron and NC MEC interactions are the largest NC TPC backgrounds as shown in Fig. 68. The darkest purple histogram includes all NC elastic neutrino-neutron selected events, the

medium purple histograms includes all NC MEC selected events, and the pale purple includes all other background events. The NC elastic neutrino-neutron cross section is a difficult background because both of the final state particles in the initial interaction, the neutrino and the neutron are neutral making them difficult to detect in the TPC. The neutron can then scatter a single proton in the liquid argon. These neutrino-neutron NC elastic events also depends on the strange axial form factor, but in the opposite way from the NC elastic neutrino-proton cross section. If there were exactly as many neutrino-neutron as neutron-proton NC elastic events selected, the sensitivity to  $G_A^s$  would disappear.

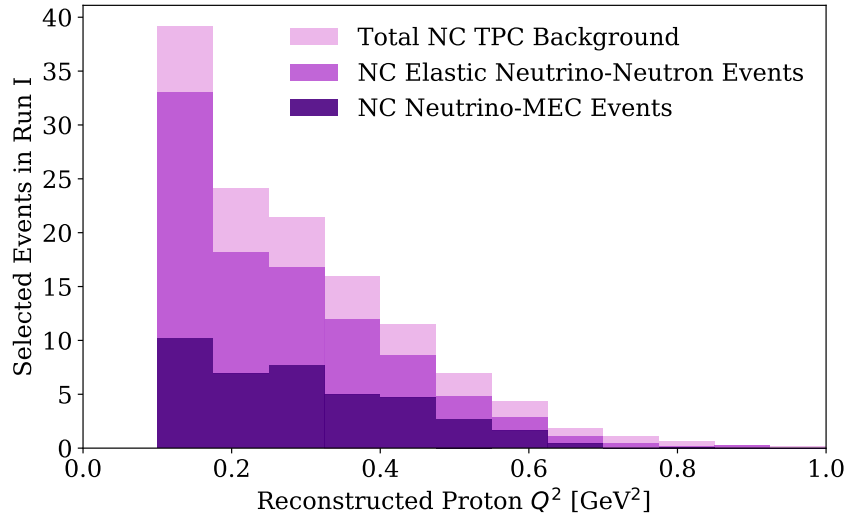


Figure 68: Remaining background NC TPC background events broken down by interaction type after NC elastic proton event selection as a function of  $Q^2$ .

The NC MEC events are difficult to distinguish from a signal event because the final state particles are usually a neutrino, a neutron, and a proton, with

the proton being the only easily detectable particle. NC MEC events have not been very well studied because of very little data, and there is a large uncertainty and variation between nuclear MEC models even in the charged current sector where there are more experimental data. In the version of the GENIE Monte Carlo neutrino generator that was used for this analysis, only one model for NC MEC events is available referred to as the “empirical MEC model.” However, for CC MEC interactions there are two models available: the empirical model and the Valencia model [112, 113]. We can look at the difference between the two CC MEC models in MicroBooNE simulation to estimate the correct scale and uncertainty on the number of NC MEC events. Figure 69 shows the number of CC MEC events in the MicroBooNE simulation as a function of true  $Q^2$  for each of the two models scaled to  $5 \times 10^{19}$  POT. The two models were also simulated with different nuclear and FSI models, which have a much smaller effect on the number of events than the MEC model.

To cover the difference in the MEC models in this analysis, we multiply all MEC events by a correction factor so that the total number is the average of the two models, and apply an uncertainty so that the prediction of either model is one standard deviation from the scaled number of events. This is shown in Fig. 70 with the simulated events for the two models the same as in Fig. 69 and the black points showing the scaled events with one standard deviation systematic uncertainty bars. All MEC events above  $Q^2 = 0.5 \text{ GeV}^2$  are given a correction

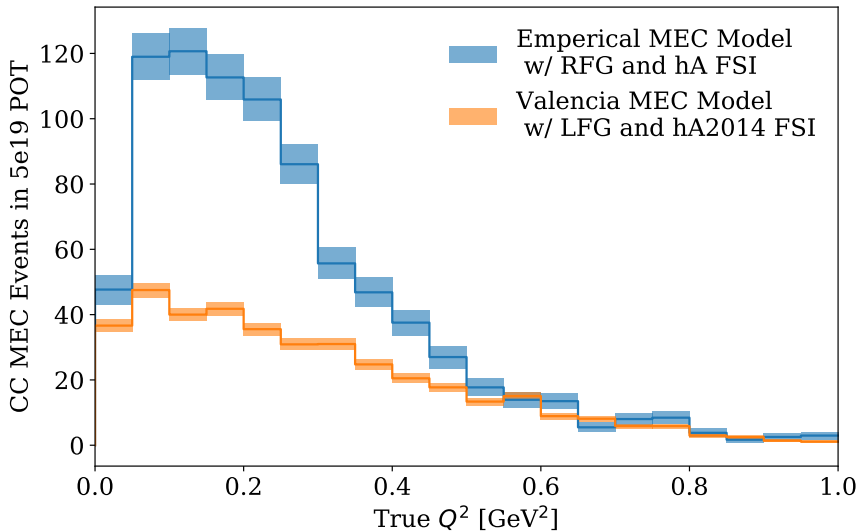


Figure 69: Comparison of the number of CC MEC events in MicroBooNE simulation between the two models as a function of true  $Q^2$ .

factor of one. The MEC correction factor and uncertainty is applied to both NC and CC MEC events. The effect on the overall systematic uncertainty of the NC elastic proton selection due to the MEC uncertainty is shown in Sec. 5.4.3. Including the MEC event correction factor, the total number of remaining NC TPC background events after the NC elastic proton event selection (scaled to  $1.6 \times 10^{20}$  POT) is

$$N_{\text{NC}} = 120 \pm 13 \text{ (stat.)} \pm 10 \text{ (syst.)}. \quad (96)$$

### 5.3.3 Cosmic Background

The cosmic background comes from both cosmic interactions that occur within the  $1.6 \mu\text{s}$  neutrino beam spill window triggering the event readout and from

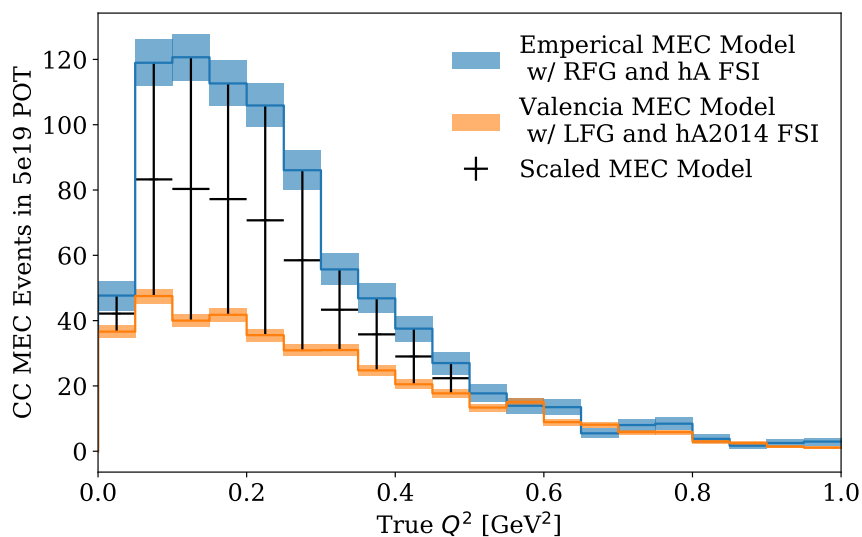


Figure 70: Corrected number of CC MEC events as a function of true  $Q^2$  with one standard deviation systematic uncertainty bars.

cosmic interactions that occur within the 1.6 ms TPC readout window that was triggered by a neutrino beam interaction. The latter is included in the previous sections because they cosmic tracks are part of the neutrino triggered events. The rate of cosmic interactions that trigger the readout, referred to as in-time cosmic interactions, are determined from the off-beam data stream. The exact same trigger logic is run in the on-beam and off-beam data streams, but with the off-beam trigger being offset in time from the neutrino beam spill. This allows us the determine the cosmic in-time background almost exactly. Figure 71 shows the remaining off-beam events passing the NC elastic proton selection after scaling to the correct number of triggers in the on-beam data set as shown in Sec. 5.2.5. The pale gray box represents the statistical uncertainty on the number of events.

The number of remaining cosmic in-time background events after the NC elastic proton event selection (scaled to  $1.6 \times 10^{20}$  POT) is  $N_{\text{cosmic}} = 340 \pm 34$ .

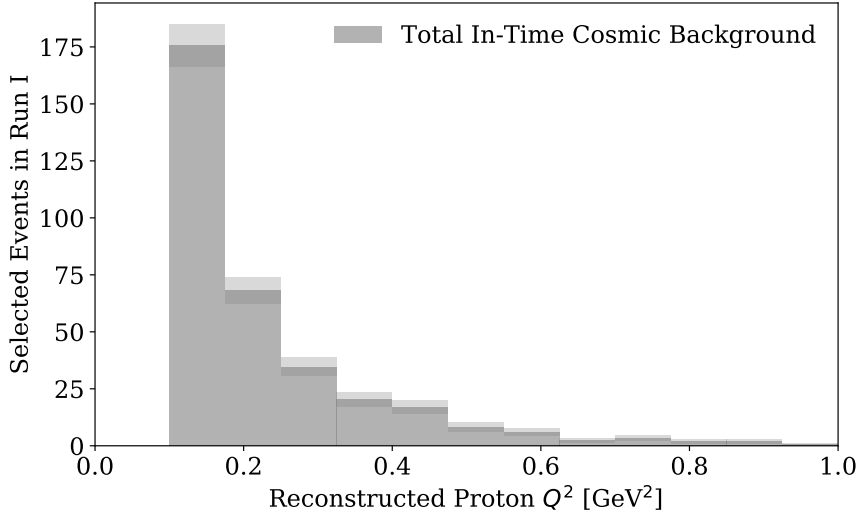


Figure 71: Remaining in-time cosmic background events after NC elastic proton event selection as a function of  $Q^2$ .

#### 5.4 Estimation of Systematic Uncertainty

The estimation of the systematic uncertainty due to the neutrino beam flux, the detector physics modeling, and the nuclear and cross section physics models are described in this section. Since the measured neutrino beam data that passes the selection are compared directly to the expectation from simulation and off-beam data, the systematic uncertainty is estimated for and applied to the simulation only, and the data are unaltered.

### 5.4.1 Neutrino Beam Flux Uncertainty

The systematic uncertainties on the neutrino beam flux prediction at MicroBooNE have been evaluated in detail in Ref. [8]. Figure 72 shows the calculated neutrino flux with the associated uncertainty. The region of neutrino energy with the largest uncertainty is also the region with a low elastic neutrino cross section. Figure 73 shows the uncertainty on the number of simulated neutrino interactions in  $5 \times 10^{19}$  POT of MicroBooNE data due to the flux uncertainty as a function of true neutrino energy. Figure 73a includes all simulated neutrino interactions and Fig. 73b only includes simulated NC elastic neutrino-proton interactions.

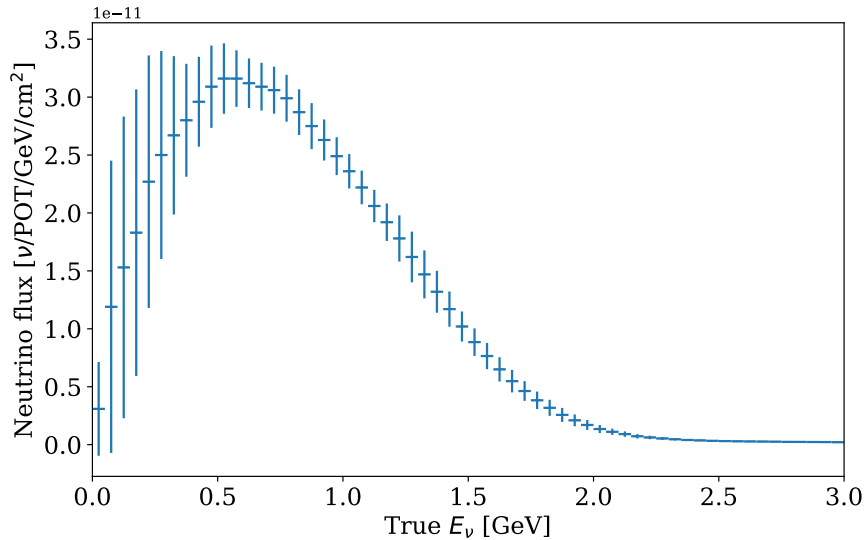
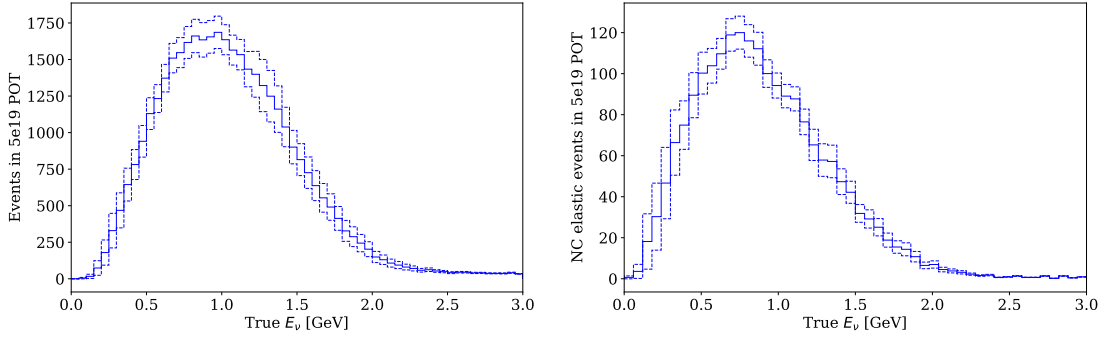


Figure 72: Neutrino flux prediction with systematic uncertainty at MicroBooNE [8].

Since we are interested in a specific range of negative four-momentum transfer squared, it is useful to know the effect of the neutrino flux uncertainty as a function



(a) All simulated neutrino interactions. (b) Simulated NC elastic neutrino-proton interactions.

Figure 73: The uncertainty on the number of simulated events due to the neutrino beam flux uncertainty as a function of  $E_\nu$ .

of  $Q^2$ . Figure 74 shows the relationship between the true four-momentum transfer squared and the true neutrino energy in simulated NC elastic neutrino-proton interactions in MicroBooNE. The color scale represents the number of events in arbitrary units with yellow being the highest and dark blue being the lowest. The dashed white lines represent the  $Q^2$  range of interest in this analysis. The interactions with the lowest neutrino energy and the highest fractional uncertainty have a four-momentum transfer squared below the range of interest. Figure 75 shows the fractional uncertainty on the number of simulated NC elastic neutrino-proton interactions due to the neutrino flux uncertainty as a function of true  $Q^2$ . The dashed blue line represents the lower end of the  $Q^2$  range of interest.

Figure 76 shows the uncertainty on the number of simulated neutrino interac-



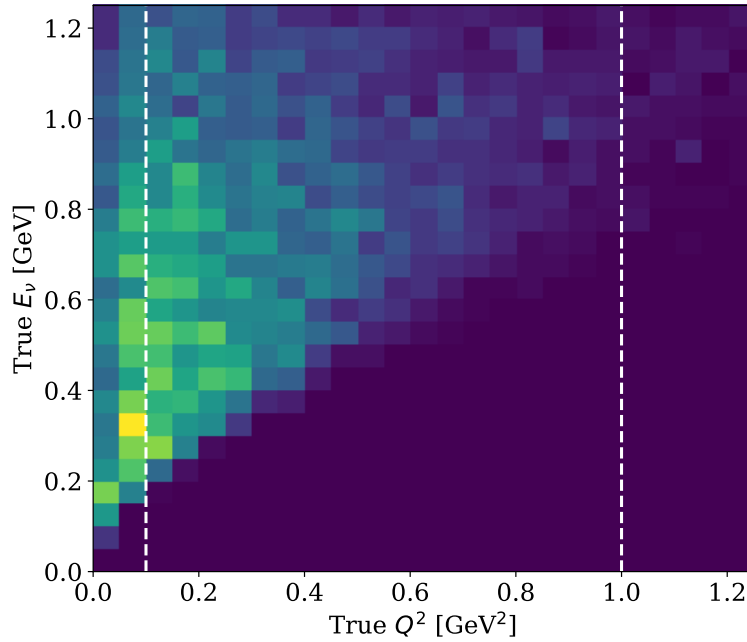


Figure 74: Two-dimensional histogram showing the relationship between true  $Q^2$  and true  $E_\nu$  in simulated NC elastic neutrino-proton interactions in MicroBooNE.

tions in  $5 \times 10^{19}$  POT of MicroBooNE data due to the flux uncertainty as a function of true neutrino energy. Figure 76a includes all simulated neutrino interactions and Fig. 76b only includes simulated NC elastic neutrino-proton interactions.

Figure 77 shows the NC elastic proton selection on simulation and off-beam data with the statistical and systematic uncertainty due to the uncertainty on the neutrino beam flux as a function of reconstructed  $Q^2$ . In this figure, and all of the similar figures in this section, the color filled histograms include simulated neutrino data with true NC elastic proton events in peach and the gray filled histogram includes off-beam data events, all scaled to the expected Run I  $1.6 \times 10^{20}$  POT of

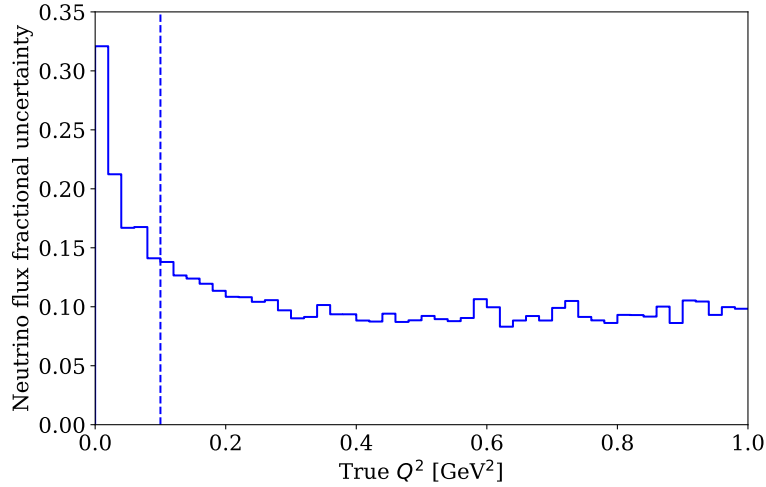
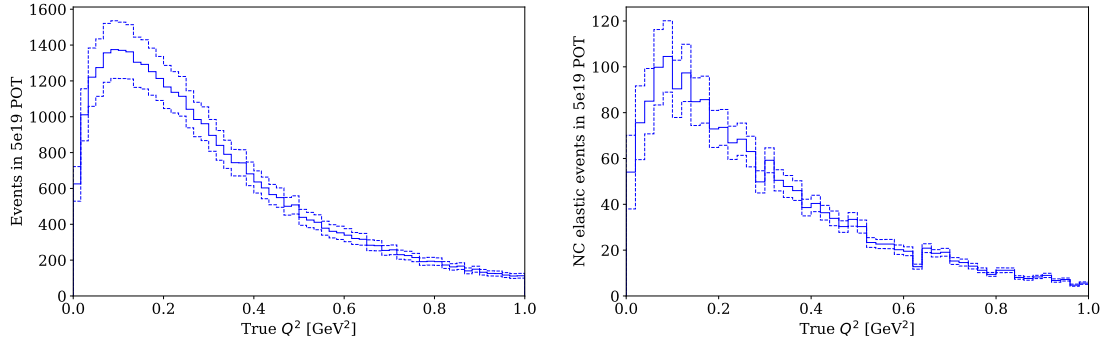


Figure 75: Fractional uncertainty on the number of simulated NC elastic neutrino-proton interactions due to the beam flux uncertainty as a function of  $Q^2$ .

data. The inner error bar represents the statistical uncertainty on the number of events, and the outer error bar represents the combined statistical and systematic uncertainty due to the flux uncertainty.

#### 5.4.2 Detector Physics Uncertainty

The uncertainty on the expected number of events due to mis-modeling of the detector physics in the simulation is estimated by generating simulated events with variations in the detector physics parameters. First, a set of events are generated, propagated through the detector geometry, and reconstructed all with the default MicroBooNE simulation parameters. Then, the same generated events are propagated through the detector geometry with a given parameter varied by one standard deviation of the estimated uncertainty on the parameter. This is



(a) All simulated neutrino interactions. (b) Simulated NC elastic neutrino-proton interactions.

Figure 76: The uncertainty on the number of simulated events due to the neutrino beam flux uncertainty as a function of  $Q^2$ .

done for all detector physics parameters that are expected to have a large effect on the final event selection. The events simulated with the default parameters can be directly compared to the events simulated with the varied parameters to determine the effect.

The detector physics effects that are expected to have the largest effect on the NC elastic proton selection in simulation are the dynamic induced charge on the wires, the space charge effect, and the single PE background rate, as described below. Full MicroBooNE simulation samples were generated to test each of these three effects.

In the default MicroBooNE simulation, charge induction is only simulated on the wire closest to the drifting charge on each of the wire planes [114]. An esti-

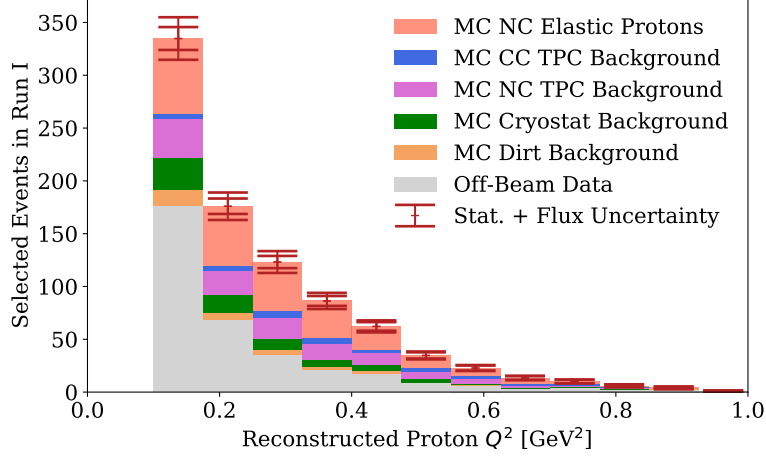
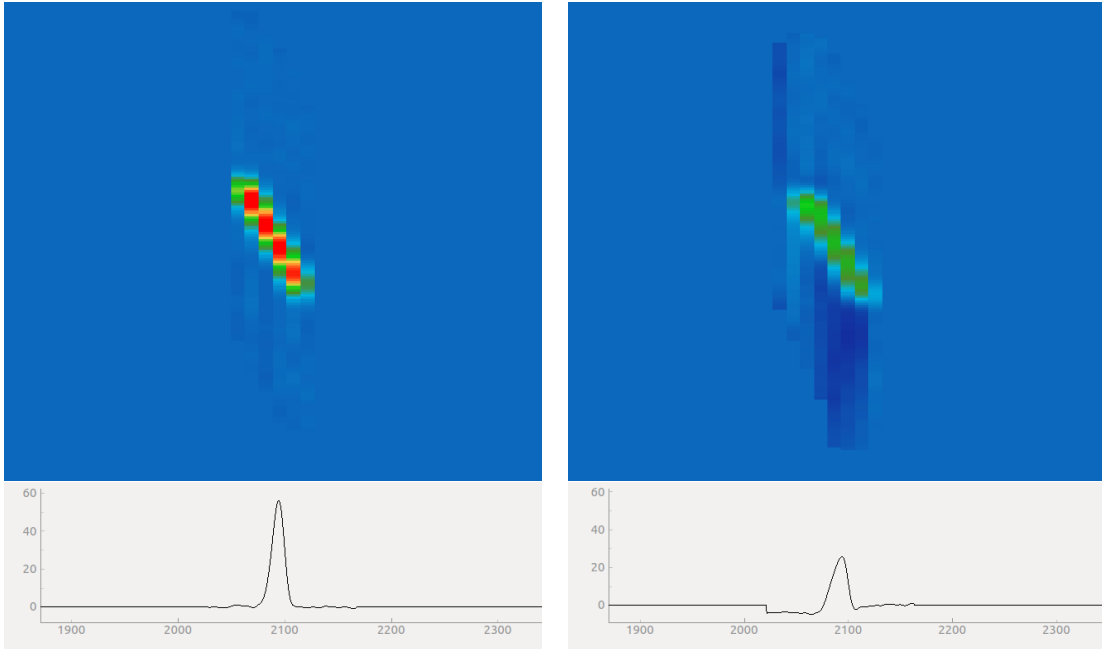


Figure 77: The NC elastic proton selection on simulation and off-beam data with statistical and systematic uncertainty due to the uncertainty on the neutrino beam flux.

mated induced charged is added to the ten closest wires to the drifting charge in a detector variation sample. The estimation is not a full 2D simulation of effect of the drifting charge on the wires, which will be added to a later version of MicroBooNE simulations. Figure 78 shows event displays of a simulated proton on the second induction plane event using the default settings without the dynamic induced charge estimation, Fig 78a, and with the dynamic induced charge estimation in the alternative sample, Fig. 78b. At the bottom of both events display, an individual wire signal is shown from one wire in the middle of the track. The track was selected as an NC elastic proton in the default simulation, but not in the alternative simulation.

Figure 79 shows the total number of selected events as a function of recon-



(a) Using the default detector physics parameters. (b) With the estimated dynamic induced charge on the wires.

Figure 78: Comparison of an event display of the second induction plane using the default detector physics parameters and using the estimated dynamic induced charge on the wires.

structured  $Q^2$  in the default detector physics sample and the dynamic induced charge estimation sample. The effect is much larger in the first bin (a 40% change) than the higher  $Q^2$  bins (a 25% change). Figure 79a includes all simulated events passing the NC elastic proton selection, and Fig. 79b only includes simulated true NC elastic proton events passing the NC elastic proton selection. The percent change as a function of  $Q^2$  bin is similar between the two. To account for the difference

due to not having a full dynamic induced charge simulation, a 40% uncertainty in the negative direction is added to the first reconstructed  $Q^2$  bin, and 25% uncertainty in the negative direction to each of the higher reconstructed  $Q^2$  bins. No uncertainty in the positive direction is attributed to the dynamic induced charge model because the default simulation includes no model at all, and adding induced charge isn't expected to increase the number of events selected.

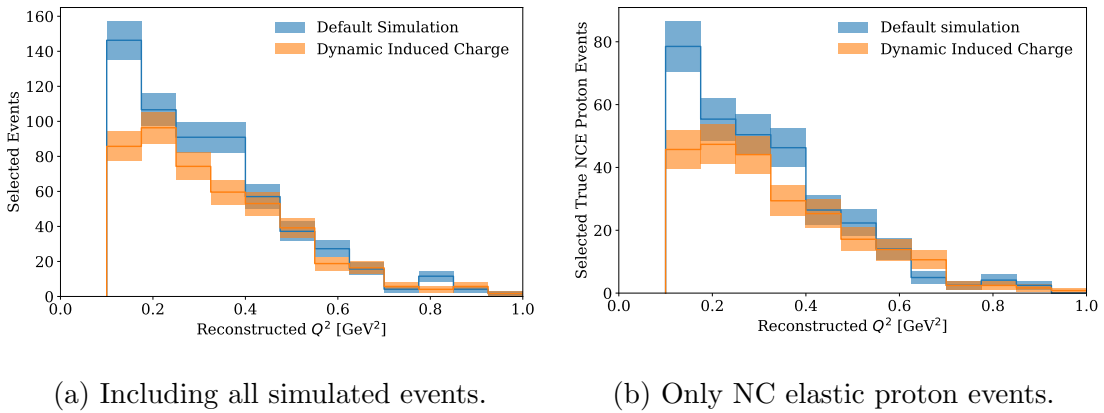


Figure 79: Change in the number of simulated events passing the NC elastic proton selection as a function of reconstructed  $Q^2$  between the default detector physics and with the estimated dynamic induced charge on the wires.

Figure 80 shows the NC elastic proton selection on simulation and off-beam data with the statistical and systematic uncertainty due to the dynamic induced charge on the wires as a function of reconstructed  $Q^2$ . The inner error bar represents the statistical uncertainty on the number of selected events and the outer error bar represents the combined statistic and systematic uncertainty due to the

induced charge.

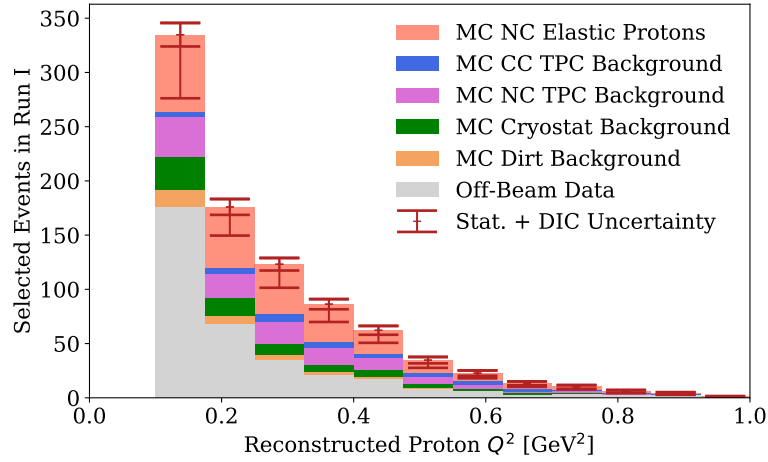
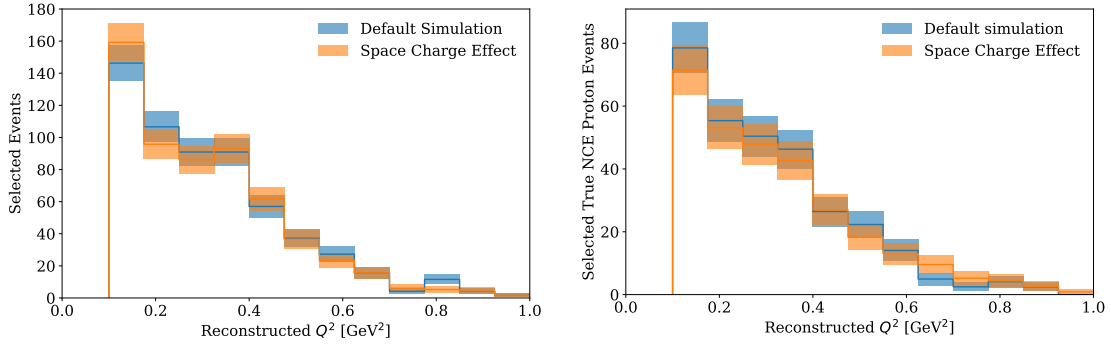


Figure 80: The NC elastic proton selection on simulation and off-beam data with statistical and systematic uncertainty due to the dynamic induced charge on the wires.

The space charge effect is due to the build up of positive ions in areas of the detector that lead to a distortion of the electric field. The effect in MicroBooNE is described in detail in Ref. [115]. In the detector physics variation sample, a data-driven correction is applied to the electric field distortion, which will be included in future MicroBooNE simulations. Figure 81 shows the change in the number of simulated events passing the NC elastic proton selection as a function of reconstructed  $Q^2$  between the default detector physics sample and the space charge correction sample. Figure 81a includes all passing simulated events, and Fig. 81b only includes true NC elastic proton events. The change in the number of selected events between the two samples is small.



(a) Including all simulated events.

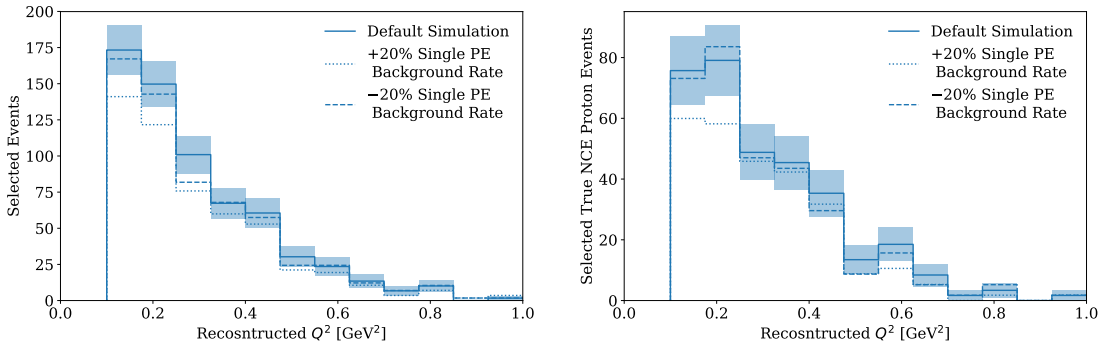
(b) Only NC elastic proton events.

Figure 81: Change in the number of simulated events passing the NC elastic proton selection as a function of reconstructed  $Q^2$  between the default detector physics and with the space charge correction.

Another source of uncertainty in MicroBooNE is the rate of the single PE optical background. The single PE background rate is a single PE rate with an unknown source of about 250 kHz per PMT. To test the effect of the rate of the single photoelectron background in the PMTs, a simulation sample was generated which varies the single PE rate by  $\pm 20\%$ . The flash produced by short single proton tracks can be small and near the PE threshold for optical pre-selection. Additionally, the NC elastic proton selection is dependent on the reconstructed position of the beam flash which can be affected by the background PEs in the PMTs, especially for smaller flashes. This is because the reconstructed position of the flash is weighted by the number of PEs in each PMT, and the single PE background is uniform in MicroBooNE. Figure 82 shows the change in the number



of simulated events passing the NC elastic proton selection as a function of reconstructed  $Q^2$  when the single PE background rate is varied by  $\pm 20\%$ . Figure 82a includes all passing simulated events, and Fig. 82b only includes true NC elastic proton events. The small reduction in the number of selected events when the single PE rate is decreased by 20% (dashed line) is mainly due to events with flashes near the PE threshold being cut in the optical pre-selection. The large reduction when the single PE rate is increased by 20% is mainly due to the distance between the center of the reconstructed flash and track increasing. Even



(a) Including all simulated events.

(b) Only NC elastic proton events.

Figure 82: Change in the number of simulated events passing the NC elastic proton selection as a function of reconstructed  $Q^2$  when the single PE background rate is varied.

though increasing and decreasing the single PE rate by 20% both have a negative effect on the number of events, it is due to different effects, and a smaller variation could increase the number of selected events. To account for this we

apply a symmetric uncertainty of 20% which is the percent change in the number of selected events when the single PE rate was increased. Figure 83 shows the NC elastic proton selection on simulation and off-beam data with the statistical and systematic uncertainty due to the single PE background rate as a function of reconstructed  $Q^2$ . The inner error bar represents the statistical uncertainty on the number of selected events and the outer error bar represents the combined statistic and systematic uncertainty due to the single PE rate. Additional sources

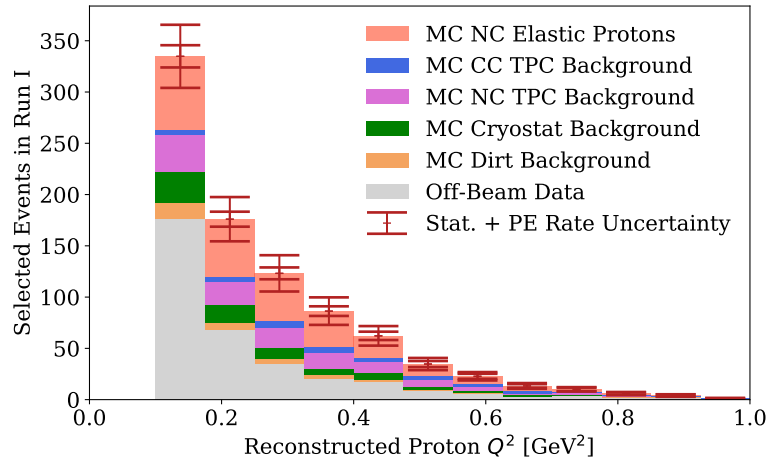


Figure 83: The NC elastic proton selection on simulation and off-beam data with statistical and systematic uncertainty due to the single PE background rate.

of disagreement between the optical simulation and the light detected in MicroBooNE are from Cerenkov radiation and a time-dependence of the light collection efficiency. Both of these effects are known to occur in the detector, but are not included in the MicroBooNE simulation.

### 5.4.3 Model Uncertainty

There is systematic uncertainty on the number of simulated events passing the NC elastic proton selection due to the nuclear models and the neutrino-nucleon cross section models. The uncertainty on the cross section model is due to the uncertainty on the electromagnetic and charged current axial form factors used in the model. The uncertainty on the nuclear models is much larger and comes from several different effects.

The estimate of the uncertainty due to the MEC model was discussed in Sec. 5.3.2. The effect of this uncertainty on the NC elastic proton selection is shown in Fig. 84 as a function of  $Q^2$ . The color filled histograms include the simulated neutrino events with NC elastic neutrino-proton events in peach, and the gray filled histogram includes the off-beam data all scaled to the expected selection in the  $1.6 \times 10^{20}$  POT Run I data. The inner error bars represent the statistical uncertainty on the number of events, and the outer error bar represents the combined statistical and systematic uncertainty due to the estimated uncertainty on the MEC model. Although the uncertainty on the MEC model is large, as shown in Fig. 70, the number of selected MEC events is small enough that the effect on the overall selection uncertainty is small.

The GENIE version used in this analysis implements a relativistic Fermi gas (RFG) nuclear momentum model [116]. The model assumes that all nucleon

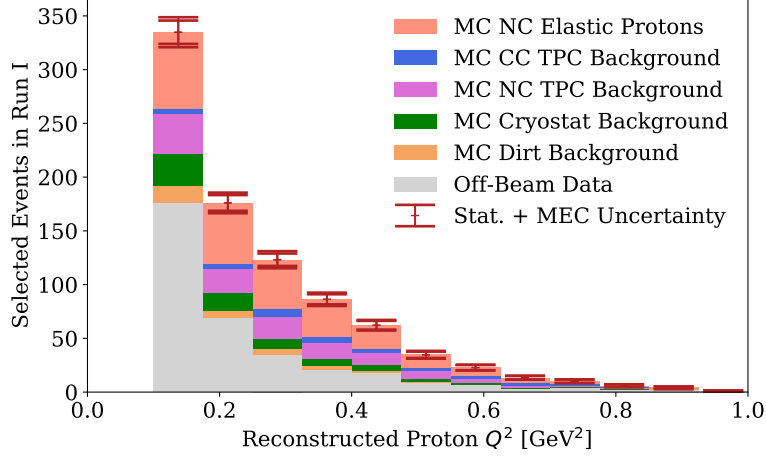


Figure 84: The NC elastic proton selection on simulation and off-beam data with statistical and systematic uncertainty due to the uncertainty on the MEC model.

momentum states up to the Fermi momentum,  $k_F$ , are equally likely. Neutrino-nucleon interactions are only allowed if the outgoing nucleon momentum is greater than  $k_F$ . This effect is referred to as Pauli blocking (or Pauli suppression). In the GENIE Reweighting tool, the uncertainty in the amount of Pauli blocking is accounted for by varying the Fermi momentum by  $\pm 35\%$  [96]. However, this is only implemented in the GENIE Reweighting tool for charged current quasi-elastic (CCQE) interactions. Since Pauli blocking is a nuclear effect, we assume that the uncertainty due to Pauli blocking as a function of true  $Q^2$  is the same for NC elastic and CCQE interactions in argon and apply the uncertainty found for CCQE interactions to NC elastic interactions, as well. Figure 85 shows the uncertainty due to Pauli blocking on CCQE interactions as a function of true  $Q^2$  for an arbitrary number of simulated interactions. The solid line shows the number

of simulated CCQE interactions using the default values of  $k_F = 0.242$  GeV for protons and  $k_F = 0.259$  GeV for neutrons, the upper dashed line shows the number of simulated CCQE interactions using  $(1 - 0.35) \cdot k_F$ , and the lower dashed line shows the number of simulated CCQE interactions using  $(1 + 0.35) \cdot k_F$  (a higher value of  $k_F$  corresponds to greater suppression). The effect is largest at true  $Q^2$  values less than  $0.1$  GeV<sup>2</sup>.

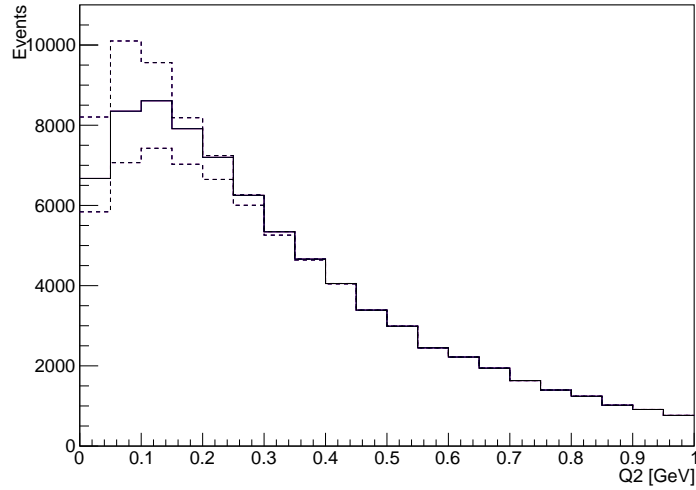


Figure 85: The systematic uncertainty due to Pauli blocking on simulated CCQE events as a function of true  $Q^2$  using the GENIE Reweighting tool.

Figure 86 shows the effect on the simulated NC elastic proton event selection as a function of reconstructed  $Q^2$  due to applying the Pauli blocking uncertainty to NC elastic and CCQE events. The inner error bar represents the statistical uncertainty on the number of events, and the outer error bar represents the combined statistical and systematic uncertainty due to Pauli blocking.

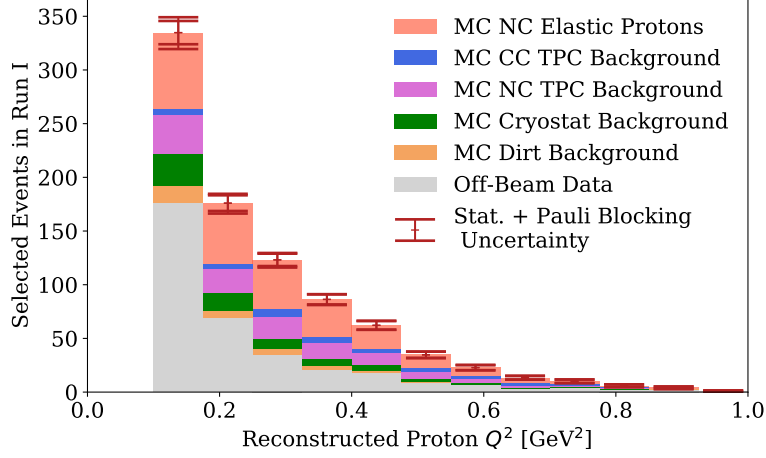


Figure 86: The NC elastic proton selection on simulation and off-beam data with statistical and systematic uncertainty due to Pauli blocking.

The GENIE Reweighting package also includes tools to determine the uncertainty on the simulation due to the probability of the nucleon re-interacting in the nucleus. This is done for by varying the mean free path of the nucleon,  $N_{\text{mfp}}$ , in the argon nucleus which is proportional to the total rescattering probability. Figure 87 shows the systematic uncertainty on the number of simulated events due to varying  $N_{\text{mfp}}$  by  $\pm 20\%$  for all events using the GENIE Reweighting tool. The inner error bars represent the statistical uncertainty on the number of selected events, and the outer error bars represent the combined statistical and systematic uncertainty due to the FSI probability.

To determine the uncertainty due to the choice of the nuclear model itself, simulation samples were generated with alternative nuclear momentum and FSI models. The nuclear momentum model is changed from the default RFG model to

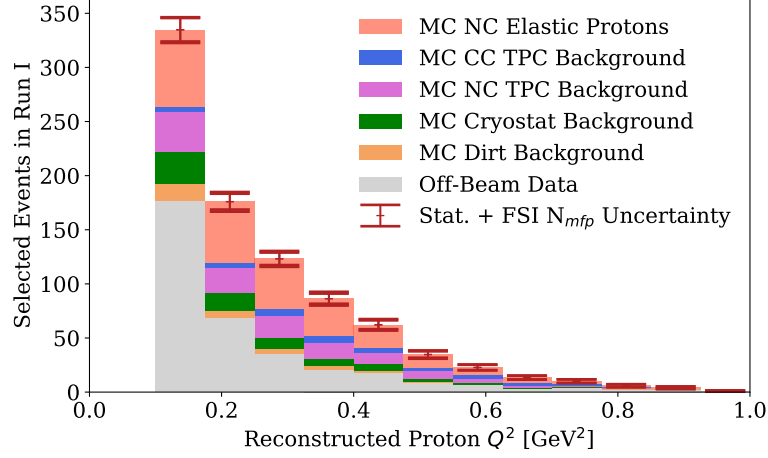
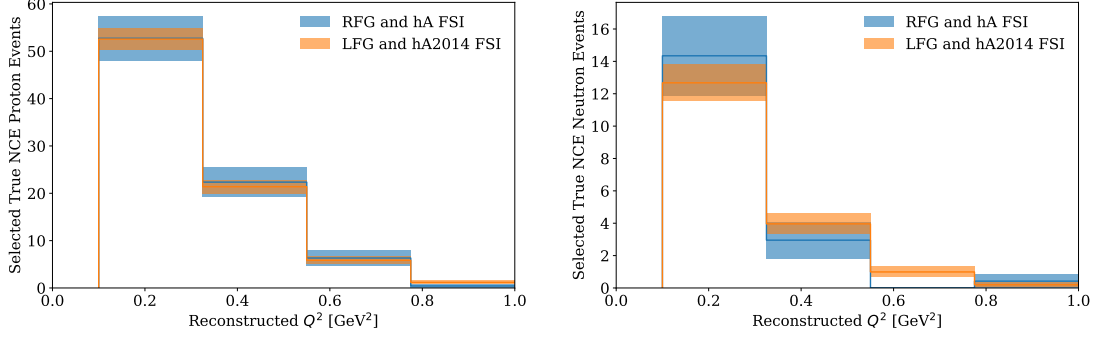


Figure 87: The NC elastic proton selection on simulation and off-beam data with statistical and systematic uncertainty due to the nucleon rescattering probability in the argon nucleus.

a local Fermi gas (LFG) model [117]. The FSI model changed from the hA GENIE model to the hA2014 GENIE model [97]. The NC elastic neutrino-nucleon cross section is unchanged in the alternative simulation sample, so the effect of changing the Fermi gas and FSI models was tested in NC elastic interactions. Figure 88 shows the change in the number of simulated NC elastic events in  $5 \times 10^{19}$  POT of data for the two different simulation sets as a function of reconstructed  $Q^2$ . The blue lines show the events simulated with the RFG nuclear momentum model and the hA FSI model, and the orange lines show the events simulated with the LFG nuclear momentum model and the hA2014 FSI model. The corresponding colored boxes show the statistical uncertainty on the number of events. Figure 88a only includes the NC elastic *proton* events, and Fig. 88b includes the NC elastic

*neutron* events. For both interaction types, the change in the number of events between models is within statistical uncertainty. The overall change in the number of events in the NC elastic proton selection is about 1%.



(a) NC elastic neutrino-proton events. (b) NC elastic neutrino-neutron events.

Figure 88: The change in the number of NC elastic events with different nuclear momentum and FSI models as a function of reconstructed  $Q^2$ .

The estimation of the uncertainty on the number of dirt interactions in Micro-BooNE was described in Sec. 5.3.1. Figure 89 shows the systematic uncertainty on the simulated NC elastic proton selection due to the dirt uncertainty. The inner error bars represent the statistical uncertainty on the number of events, and the outer error bars represent the combined statistical and systematic uncertainty due to the number of dirt events.

The uncertainty on the NC elastic proton selection due to the electric, magnetic, and charged current axial form factor  $z$  expansion models was found to be small. The uncertainty on the models were taken directly from the Refs. [82, 77]



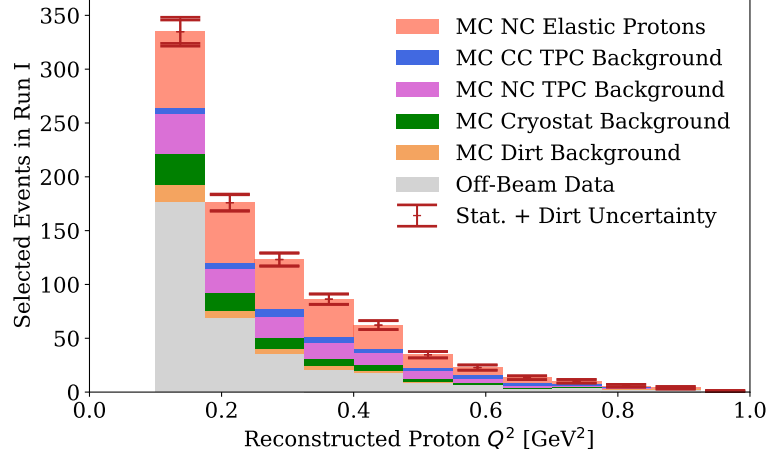


Figure 89: The NC elastic proton selection on simulation and off-beam data with statistical and systematic uncertainty due to the number of dirt interactions.

and extrapolated to the NC elastic proton selection. The total uncertainty on the selection due to the form factor models was found to be 0.5%.

Table 4 shows the contribution from each of the systematic effects to the overall uncertainty on the number of expected events in the NC elastic proton selection. If an asymmetric systematic uncertainty is applied, the larger of the two is shown in the table.

Figure 90 shows the total combined systematic uncertainty on the simulated NC elastic proton selection from all of the effects described in this section. The inner error bars represent the statistical uncertainty on the number of events, and the outer error bars represent the combined statistical and systematic uncertainty.

Table 4: Contributions to the systematic uncertainty on the number of simulated events passing the NC elastic proton selection.

Source	Uncertainty %
Neutrino Beam Flux	9.8
Dynamic Induced Charged	27.5
Space Charge	5.0
Single PE Rate	20.0
MEC Model	5.9
Pauli Blocking	4.9
Fermi Gas and FSI Models	1.0
FSI Interaction Probability	5.5
Dirt Events	3.2
Form Factor Model	1.7

## 6 Analysis

This section describes the analysis tools used to extract the strange axial form factor parameters from the number of events in data and simulation that pass the NC elastic proton selection. First, the method for determining the expected number of events in data given different model and systematic parameter values is discussed. Next, the formula for calculating the likelihood of the observed data

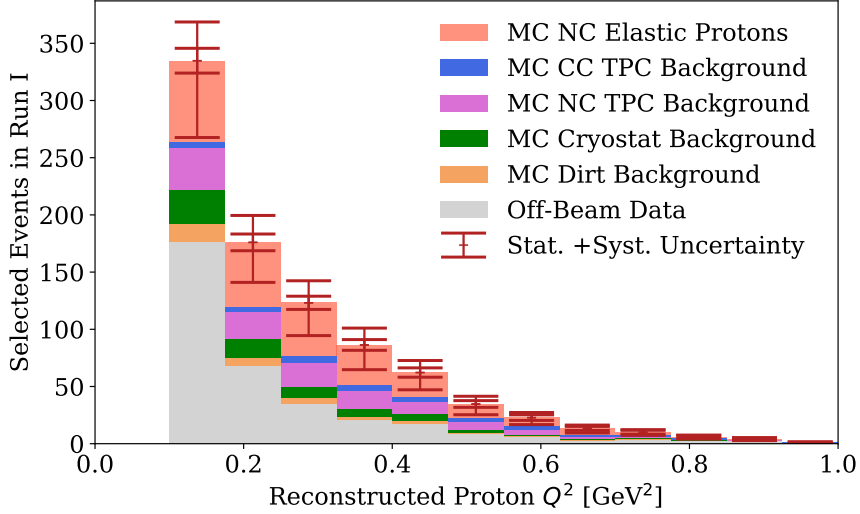


Figure 90: The NC elastic proton selection on simulation and off-beam data with statistical and systematic uncertainty.

given a set of parameter values, and the technique used to sample the probability distribution of the parameters is described. Last, the results of the measured probability distributions of the strange axial form factor parameters are shown.

### 6.1 Comparison of Data to Simulation

To determine what underlying true physics values cause the data that we measure, we compare the data to Monte Carlo simulation. The physics model of the data is a combination of many different models including nuclear physics models, neutrino cross section models, nucleon structure models, cosmic ray models, which would be very complicated to calculate directly. Instead, we simulate the data we expect to see in the detector given a set of physical parameters and models

and compare that directly to the actual data we measured. We do this for many possible values of the parameters and calculate the likelihood for each set of parameter values. In addition to varying the parameters that we want to measure in the simulation, we vary the physical parameters whose true values aren't well constrained which might have a large effect on the final data. This allows us to quantify the uncertainty due to the unknown quantities.

### 6.1.1 Event Reweighting

Full Monte Carlo simulations are both time and compute intensive. Instead of re-running the simulation for each possible parameter value that we are interested in, we can calculate the ratio of the probabilities of each interaction occurring given the new parameter values to the probabilities of each interaction occurring for the original simulated parameter values. We refer to this ratio as an event weight:

$$w = \frac{P(\text{event}|\theta')}{P(\text{event}|\theta)}, \quad (97)$$

where  $w$  is the event weight for a given event,  $P(\text{event}|\theta)$  is the probability of the simulated event given the set of original parameters used in the simulation,  $\theta$ , and  $P(\text{event}|\theta')$  is the probability of the simulated event given a new set of parameters,  $\theta'$ .

The four contributions to the GENIE cross section model (the nuclear physics model, the neutrino-nucleon cross section model, the hadron production model,

and the intranuclear hadron transport model from Sec. 4.1.1) are treated independently in NC elastic events [95]. This allows us to factor the event weight,

$$w = w_{\text{nuclear}} \times w_{\text{neutrino-nucleon}} \times w_{\text{hadron prod.}} \times w_{\text{intranuclear}}, \quad (98)$$

where  $w_{\text{nuclear}}$  is the nuclear physics model weight,  $w_{\text{neutrino-nucleon}}$  is the neutrino-nucleon cross section model weight,  $w_{\text{hadron prod.}}$  is the hadron production model weight, and  $w_{\text{intranuclear}}$  is the intranuclear hadron transport model weight. Only the NC elastic cross section probability ratio needs to be calculated to see the effect of  $\Delta s$  or another cross section parameter of interest on reconstructed  $Q^2$  of selected events.

The probability of a neutrino interaction is proportional to the interaction cross section, so the weight,  $w_{\text{neutrino-nucleon}}$  is simply a ratio of the cross sections

$$w_{\text{neutrino-nucleon}} = w_{\sigma} = \frac{d^n \sigma'_{\nu}/dK^n}{d^n \sigma_{\nu}/dK^n}, \quad (99)$$

where  $d^n \sigma/dK^n$  is the differential cross section for the initially simulated neutrino-nucleon interaction, and  $d^n \sigma'/dK^n$  is the differential cross section with the modified parameters evaluated at the kinematical phase space  $\{K^n\}^3$ . The differential cross section is a function of the neutrino energy in the rest frame of the scattered nucleon,  $E_{\nu}^{(NRF)}$ , the interaction four-momentum transfer,  $Q^2$ , and the physics model, including the model parameters.

To determine the effect of the strange axial form factor parameters  $\Delta s$  and  $M_A^s$  (or  $a_0^s$ ,  $a_1^s$ , and  $a_2^s$ ) on the data, we calculate the NC elastic cross section given

these new parameters and the cross section given the initial simulation parameters for each NC elastic event in the Monte Carlo simulation. To get an accurate weight, the denominator needs to be calculated exactly as the cross section was calculated in the initial GENIE simulation. However, there is no reason that a different model can't be used for the numerator. For the numerator in elastic interactions, we use the Llewellyn-Smith neutrino-nucleon elastic cross section parameterization described in Sec. 2.5 with the  $z$  expansion vector and strange axial form factors described in Sec. 2.6.1.

To determine the effect of a set of NC elastic cross section parameters given our parameterization, we calculate the neutrino-nucleon weight for each NC elastic event

$$w_\sigma = \frac{\left(\frac{d\sigma}{dQ^2}(a_0^s, a_1^s, a_2^s)\right)_{\text{L.S.}}}{\left(\frac{d\sigma}{dQ^2}\right)_{\text{GENIE}}}, \quad (100)$$

where the L.S. subscript represents the cross section calculated using the Llewellyn-Smith parameterization with the  $z$  expansion form factors, and the GENIE subscript represents the default cross section in the simulation.

We also calculate weights to determine the effect on the simulated data of each of the sources of systematic uncertainty described in Sec. 5.4. We can sample the probability space of each of these “nuisance” parameters and calculate a weight based on that value. If the nuisance parameters are independent of each other, multiplying the nuisance parameter weights together is equivalent to sampling the

combined,  $N$ -dimensional probability space of the systematic parameters, where  $N$  is the number of parameters. Sampling each of the relevant systematic parameters from Sec. 5.4 gives

$$w_{syst} = w_{\text{flux}} \times w_{\text{D.I.C.}} \times w_{\text{S.P.E.}} \times w_{\text{MEC}} \times w_{\text{P.B.}} \times w_{\text{dirt}} . \quad (101)$$

Combining Eqns. 98, 100, and 101 gives an event weight which combines the effect due to a sample from the model parameter distribution and the values of the strange axial form factor parameters that we want to measure

$$w = w_{\sigma} \times w_{syst} . \quad (102)$$

### 6.1.2 Likelihood calculation

To compare the simulation and the weight calculations directly to the data, we sum the weights for each reconstructed  $Q^2$  bin in the distribution of events passing the NC elastic proton selection

$$N_{NCE}(Q_i^2) = \sum_{j \in NC_i} w_j , \quad (103)$$

where  $i$  is the  $Q^2$  bin,  $NC_i$  is the set of events selected as NC elastic neutrino-proton events in the  $i^{\text{th}}$   $Q^2$  bin, and  $w = w_j$  is the calculated weight of the event. If the event is not a true simulated elastic event, then  $w_{\sigma} = 1$  giving  $w_j = w_{syst}$ . At this point, the Monte Carlo simulated data is directly comparable to the detector data.

To evaluate how well the model represents the data, we calculate the probability of the observed data given a model and set of parameters. This probability is called the likelihood. We assume that the measured data is normally distributed, and the likelihood is

$$\mathcal{L} = P(D^{obs}|\theta) = \prod_{i \in I_{Q^2}} P(D_i^{obs}|\theta) = \frac{1}{\sqrt{2\pi\sigma_i^2}} e^{-\frac{1}{2}(D_i^{obs} - D_i^{exp}(\theta))^2/\sigma_i^2}, \quad (104)$$

where  $D^{obs}$  is the measured NC elastic neutrino-proton event selection distribution in data,  $I_{Q^2}$  is the set of  $Q^2$  bins,  $D_i^{obs}$  is the measured number of selected events of that bin, and  $\sigma_i$  is the statistical uncertainty of the difference between the observed and expected events in the measured bin value. The expected value of the event selection distribution,  $D_i^{exp}(\theta)$ , is a function of the model and the set of model parameters,  $\theta$ . It is equal to  $N_{NCE}(Q_i^2)$  in Eqn. 103. The set of model parameters,  $\theta$ , contains the systematic model parameters described in Sec. 5.4 and the strange axial form factor parameters,  $a_0^s$ ,  $a_1^s$ , and  $a_2^s$  described in Sec. 2.6.3. In this analysis, only the statistical uncertainty is contained in  $\sigma_i$ , and the systematic uncertainty is handled by sampling the systematic parameter values contained in  $\theta$  as described in Sec. 6.2.2.

When computing and comparing the likelihood values, the natural log of the likelihood (log-likelihood) is actually used to avoid rounding errors in the computation. Taking the log of Eq. 104 gives

$$\log \mathcal{L} = -\frac{N_{bins}}{2} \log(2\pi) \sum_{i \in I_{Q^2}} \left[ \log(\sigma_i) \frac{1}{2} (D_i^{obs} - D_i^{exp}(\theta))^2 / \sigma_i^2 \right] \quad (105)$$



The likelihood calculation depends on nine different parameters, the three parameters of the strange axial form factor and the six systematic parameters described in Sec. 5.4. Calculating likelihood distributions for nine parameters would be computationally very time consuming. Figure 91 shows the likelihood distributions calculated on a grid over the three strange axial form factor parameters,  $a_0^s$ ,  $a_1^s$ , and  $a_2^s$ , while setting all of the systematic uncertainties to zero. Figures 91a and 91b show the grids of likelihood values for  $a_0^s$  and  $a_1^s$  with  $a_2^s$  held at  $-20$  and  $10$ , respectively. Figures 91c and 91d show the grids of likelihood values for  $a_0^s$  and  $a_2^s$  with  $a_1^s$  held to  $-3$  and  $2$ , respectively. Figures 91e and 91f show the grids of likelihood values for  $a_1^s$  and  $a_2^s$  with  $a_0^s$  held to  $0$  and  $1$ , respectively. Even ignoring the systematic parameters and just varying three parameters is computationally difficult and can only be done for discrete values of the parameters. The sampling method described in Sec. 6.2.2 makes calculating the distributions in larger dimensions achievable.

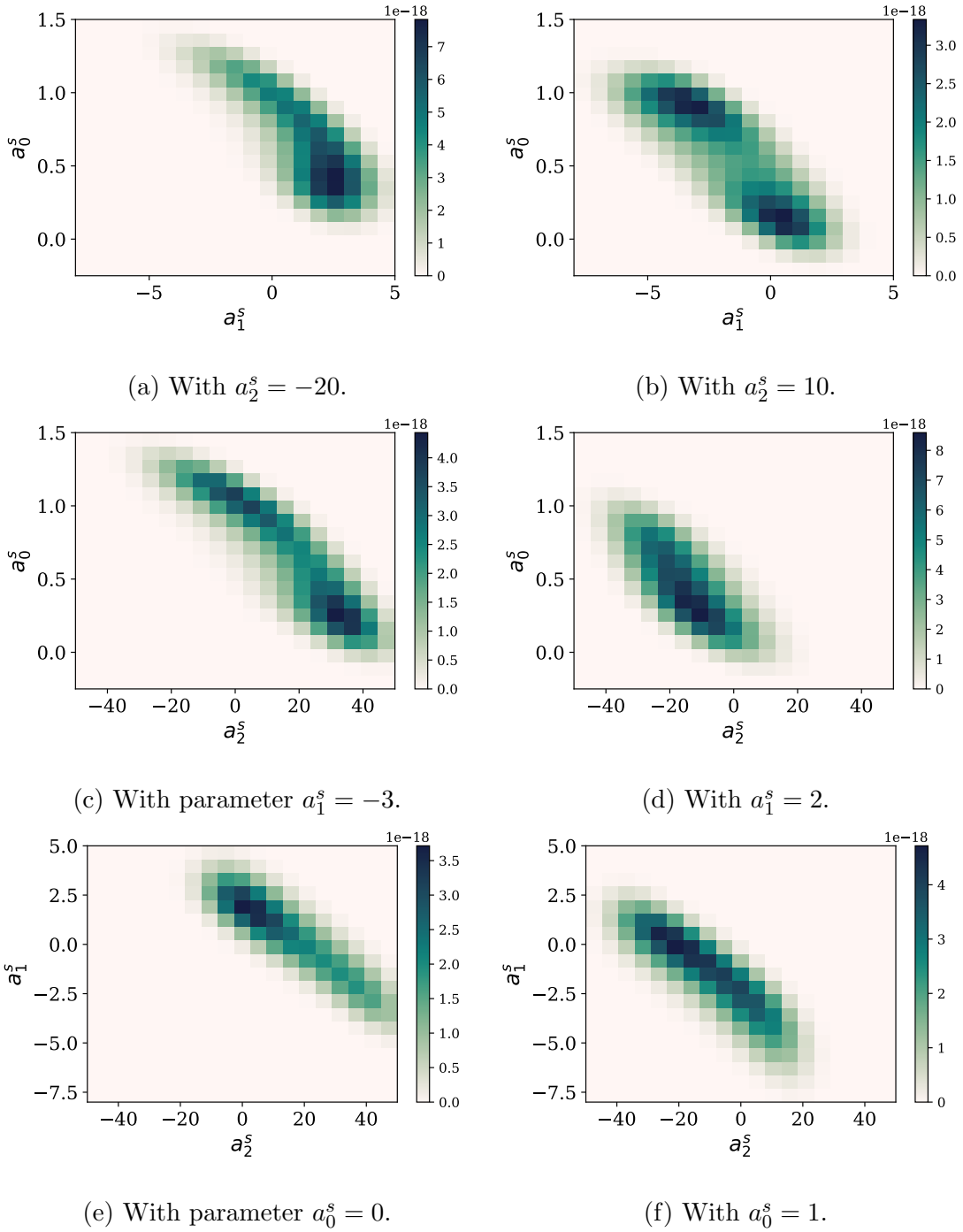


Figure 91: Calculated likelihood values at discrete  $a_0^s$ ,  $a_1^s$ , and  $a_2^s$  with zero systematic uncertainty.

## 6.2 Strange axial form factor parameter estimation

The likelihood gives the probability of the observed data given a model and a set of parameters. What we are really interested in is the probability of the parameters  $a_0^2$ ,  $a_1^s$ , and  $a_2^s$  given the observed data and our model. We can determine this probability distribution using Bayesian inference and probability sampling methods.

### 6.2.1 Bayesian inference

Bayes' theorem is used to convert between the likelihood and the probability of the model given the data

$$P(\theta|D^{obs}) = \frac{P(D^{obs}|\theta)P(\theta)}{P(D^{obs})}. \quad (106)$$

The likelihood,  $P(D^{obs}|\theta)$ , was described in detail in Sec. 6.1.2 and is defined in Eqn. 104. The other three factors in Bayes' theorem deserve some explanation.

First, the probability of  $\theta$  given the observed data,  $P(\theta|D^{obs})$ , is the probability distribution that we ultimately want to determine. It is referred to as the posterior distribution. Implicit in the notation is the model. If we wanted to write it explicitly, it would be:

$$P(\theta|D^{obs}) \equiv P(\theta|D^{obs}, \mathcal{M}), \quad (107)$$

where  $\mathcal{M}$  represents the physics model. The parameter set  $\theta$  is still the set containing the strange axial form factor parameters, and the systematic parameters

from Sec. 6.1.1.

Next, the probability of the parameters,  $P(\theta)$ , is referred to as the prior distribution. It is also implicitly conditional on the model,  $\mathcal{M}$ . This is where we include prior information that we know to be true. It is impossible not to include some prior information in inference. For example, using a uniform prior on  $\Delta s$  is the same as saying that  $\Delta s$  has the same probability of being zero as it does of being infinite. The prior should be used to exclude unphysical parameter values, like negative mass. In a good model with adequate data the posterior distribution should be robust to the choice of a prior. It is always necessary to evaluate the effect of the choice of priors on the posterior. In Sec. 6.3 we show the effect of different priors, including a uniform prior, on the posterior distribution.

Last, the marginal probability of the observed data  $P(D^{obs})$  is integrated over all  $\theta$  values. It too is implicitly conditional on  $\mathcal{M}$  and is referred to as the evidence of the model. Explicitly, it can be written as

$$P(R^{obs}) \equiv P(D^{obs}|\mathcal{M}) = \int_{\theta} P(D^{obs}|\theta)P(\theta|\mathcal{M})d\theta. \quad (108)$$

### 6.2.2 Markov Chain Monte Carlo

There is no way to calculate an exact solution to our posterior distribution analytically, but it can be estimated numerically by sampling. Most sampling techniques would be computationally impossible due to the fact that our posterior distribution is nine-dimensional ( $a_0^g$ ,  $a_1^s$ ,  $a_2^s$ , and the six systematic nuisance parameters),

and each likelihood calculation requires thousands of event weights to be calculated. Markov chain Monte Carlo (MCMC) is a class of methods for sampling multi-dimensional posterior distributions. Two of the most common MCMC algorithms are the Metropolis algorithm [118] and Gibbs sampling [119]. Both are used in this analysis.

The Metropolis algorithm is a random walk in the  $N$ -dimensional parameter space with a rule to either accept or reject each step in the walk. Each proposed step is drawn from a proposal distribution. In this analysis, we use a multivariate normal proposal distribution centered at the current position. The step is accepted if the value of the posterior distribution at the proposed position is greater than at the current position. If the value at the proposed distribution is smaller than the current value, the proposed step is accepted with a probability equal to the ratio of the value at the proposed position to the value at the current position. The decision to accept or reject a proposed step can be determined entirely by calculating the *ratio* of the posterior values at the proposed and current positions. Since the evidence,  $P(D^{obs})$ , doesn't depend on  $\theta$ , it never needs to be calculated. At every proposed step only the likelihood,  $P(D^{obs}|\theta)$ , and the prior,  $P(\theta)$ , need to be calculated.

Gibbs sampling as used in this analysis is also a random walk in the  $N$ -dimensional parameter space, but it samples the posterior directly. We cannot sample the posterior directly for the strange axial form factor parameters that

we want to infer from the data, but we can make an approximation that allows us to use Gibbs sampling for our nuisance parameters. At each Gibbs sampling iteration, subvectors of parameters,  $\theta_i \in \theta$ , are cycled through and sampled conditionally on the position of the previous subvector  $P(\theta_i^t | \theta_{-i}^{t-1}, D^{obs})$ , where  $t$  is the current iteration,  $\theta_{-i}^{t-1}$  represents the current position all of the other subvectors (excluding  $\theta_i$ ),  $\theta_{-i}^{t-1} = (\theta_0^t, \dots, \theta_{i-1}^t, \theta_{i+1}^{t-1}, \dots, \theta_d^{t-1})$ , and  $d$  is the number of subvectors in  $\theta$ . Since the posterior is being sampled directly, the sample is also conditional on the data,  $D^{obs}$ . This is where the approximation comes in. If we assume that the nuisance parameters are independent of the data, then

$$P(\theta_i | \theta_{-i}, D^{obs}) \approx P(\theta_i | \theta_{-i}). \quad (109)$$

The assumption being made here is that these parameter distributions, the position and variance of the systematic nuisance parameters, would not change very much based on the data we are comparing to,  $D^{obs}$ . In other words we would not be able to tell from the observed data events passing the NC elastic proton selection what the true mean and variance of the systematic nuisance parameters are. For example, we don't believe that we can measure the single PE background rate and its uncertainty by looking at the NC elastic event selection in the data. If that assumption is true, then the posterior distributions of the systematic parameters are equal to the prior distributions, which were estimated in Sec. 5.4.

The Metropolis and Gibbs sampling methods are combined for the parameter

estimation. First, a Metropolis step is proposed for new  $a_0^s$ ,  $a_1^s$  and  $a_2^s$  values. The likelihood value at the proposed position is evaluated conditionally on the current position of the systematic nuisance parameters. Then, a Gibbs step is taken in all of the systematic parameters at once ( $\theta_i$  is all of the nuisance parameters), and the likelihood is evaluated conditionally on the current position of  $a_0^s$ ,  $a_1^s$ , and  $a_2^s$ . The proposed positions of each of the six parameters in the step are independent of each other and of the current parameter values. This is repeated iteratively until the target distribution has been reached. Metrics for determining how many samples are required are discussed in Sec. 6.3.2.

### 6.3 Results

Using the reweighting method, the likelihood calculation, and the Metropolis and Gibbs MCMC sampling, the strange axial form factor parameters can now be estimated. First, the tools are tested by sampling the strange axial form factor parameter posterior distributions without including the systematic nuisance parameters. Then, the systematic parameters are included, the convergence of the MCMC sampling is evaluated, and the posterior distributions in terms of  $\Delta s$  and  $M_A^s$  are shown.

### 6.3.1 MCMC Sampling without Systematic Nuisance Parameters

The three strange axial form factor parameters are sampled first assuming no systematic uncertainty. The main reason for doing this is to see if we get something reasonable as a proof of concept. The extracted parameters don't inform us about the true strange axial form factor when the systematic uncertainty isn't accounted for.

To sample without the systematic uncertainty, the proposal distribution for each systematic nuisance parameter is set to zero. The proposal distribution for the strange axial form factors, described below, was determined empirically to maximize the coverage of the allowed parameter space. The final proposal distribution used for the strange axial form factor parameters is a multivariate normal distribution with a small amount of covariance between each of the parameters. The covariance is included because of the strong anti-correlation in the parameter likelihood distributions in Sec. 6.1.2. If the covariance were not included, the final sampled distributions would eventually be the same, but it would take longer to cover the parameter space. The metric used to determine the sampling performance is the acceptance rate. If the proposed step is too large, most moves will be rejected (a low acceptance rate), and it will take a long time to cover the parameter space. If the proposed step is too small, most moves will be accepted (a high acceptance rate), and it will take a long time to cover more than just a



small portion of the parameter space [120].

The proposal distribution used for the strange axial form factor parameters as each MCMC iteration is a multivariate normal centered at the current position of the MCMC chain at that iteration with covariance:

$$\Sigma = \begin{bmatrix} 0.075 & -0.1 & -0.2 \\ -0.1 & 1.5 & -2.0 \\ -0.2 & -2.0 & 30 \end{bmatrix}, \quad (110)$$

which gives a acceptance rate when ignoring systematic parameters of 0.55. Figure 92 shows the covariance between the three parameters in the proposal distribution. The initial position of the MCMC chain is sampled from the proposal distribution centered at the zero for each of the parameters.

The prior distributions on the strange axial form factor parameters are used to encode bounds on the coefficients  $a_k^s$  for  $k = 0, \dots, 6$  introduced in Sec. 2.6.3. Following the  $z$  expansion form factor analyses in Refs. [80, 74], we enforce  $|a_k^s| \leq 10$ . Recall that  $a_3^s, \dots, a_6^s$  are linear combinations of  $a_0^s$ ,  $a_1^s$ , and  $a_2^s$ , making the coefficient bounds:

$$|a_0^s| \leq 10, \quad |a_1^s| \leq 10, \quad |a_2^s| \leq 10, \quad (111)$$

$$| -20a_0^s - 10a_1^s - 4a_2^s | \leq 10, \quad (112)$$

$$| 45a_0^s + 20a_1^s + 6a_2^s | \leq 10, \quad (113)$$

$$| -36a_0^s - 15a_1^s - 4a_2^s | \leq 10, \quad (114)$$

$$| 10a_0^s + 4a_1^s + a_2^s | \leq 10. \quad (115)$$

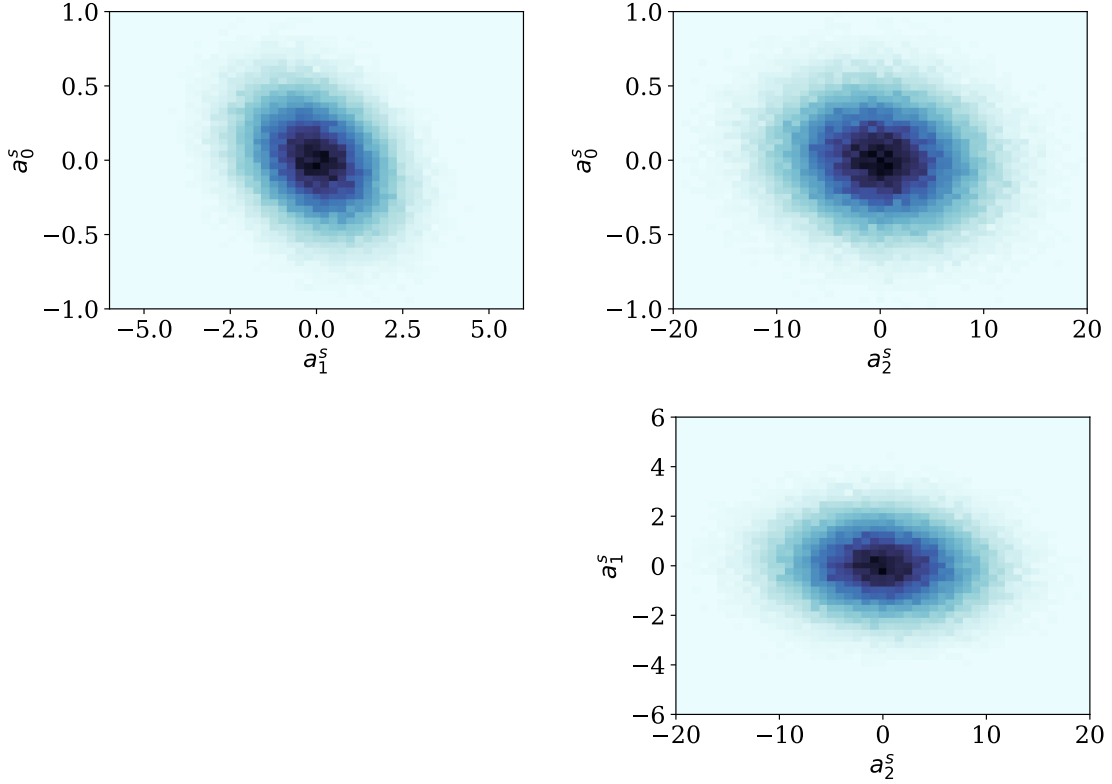


Figure 92: Two-dimensional views of the multivariate normal proposal distribution of the strange axial form factor parameters.

The optimal value of  $t_0$  from Eq. 84 is used for the mapping of  $Q^2$  to  $z(Q^2)$ ,

$$\begin{aligned}
 t_0^{\text{optimal}}(Q^2) &= t_{\text{cut}} \left( 1 - \sqrt{1 + Q_{\text{max}}^2/t_{\text{cut}}} \right) \\
 &= 9m_\pi^2 \left( 1 - \sqrt{1 + 1.0 \text{ GeV}^2/(9m_\pi^2)} \right) \quad (116) \\
 &= -0.28 \text{ GeV}.
 \end{aligned}$$

Figures 93-96 show the results of 50,000 MCMC steps in  $a_0^s$ ,  $a_1^s$ , and  $a_2^s$  with the systematic parameters held at zero. Figure 93 shows the chains of MCMC samples for  $a_0^s$ ,  $a_1^s$ , and  $a_2^s$ . Figure 94 shows the 50,000 samples of the posterior distributions for each of the two-parameter pairs of the strange axial form factor

parameters. The correlation between the  $a_0^s$  and  $a_1^s$  posterior distributions and the correlation between the  $a_1^s$  and  $a_2^s$  posterior distributions are similar to the correlations in the corresponding likelihood distributions even after introducing the coefficient bounds and allowing all three parameters to vary simultaneously. The calculated Pearson correlation coefficient between  $a_0^s$  and  $a_1^s$  is  $r_{01} = -0.77$ , and the coefficient between  $a_1^s$  and  $a_2^s$  is  $r_{12} = -0.93$ . There doesn't appear to be a negative correlation between the  $a_0^s$  and  $a_2^s$  posterior distributions, however, even though one exists between the likelihood distributions. In fact, the correlation coefficient between  $a_0^s$  and  $a_2^s$  is positive,  $r_{02} = 0.51$ .

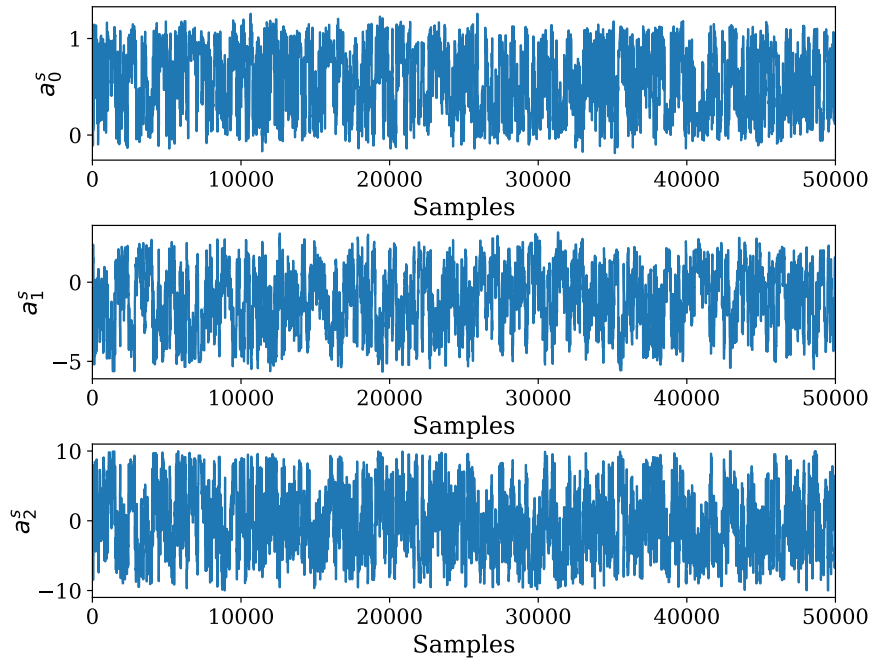


Figure 93: MCMC chains of the strange axial form factor parameters after 50,000 steps with the systematic parameters held to zero.

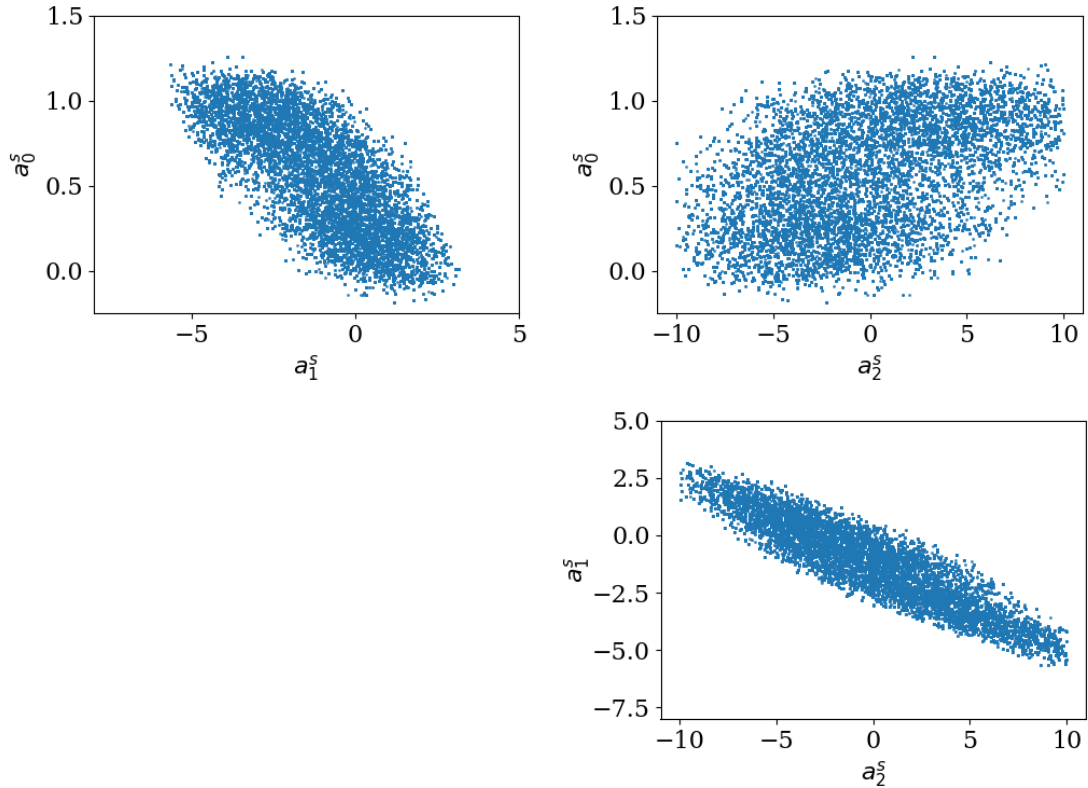


Figure 94: Two-dimensional views of the 50,000 MCMC samples of the posterior distributions of the strange axial form factor parameters with the systematic parameters held to zero.

### 6.3.2 MCMC Sampling with Systematic Nuisance Parameters

The prior distributions of the systematic nuisance parameters are described in Sec. 5.4. Since we don't expect the NC elastic proton selection in data to increase our knowledge of the mean and variance of the systematic parameters significantly, the posterior distributions of the systematic parameters should look similar to the prior distributions. The best proposal distributions of the parameters are therefore

their prior distributions since they are our best estimate of their posteriors which we are not actually trying to measure.

To account for the systematic uncertainty from the sources described in Sec. 5.4, we sample directly from the distributions described in that section in the Gibbs sampling step. For each of the systematic proposals, we sample a standard Gaussian with a mean of zero and a variance of one and translate the sampled value to the effect of that parameter on the number of simulated events passing the NC elastic proton selection. In the case of the dynamic induced charge the negative of the absolute value is used because the uncertainty is only in one direction. Each simulated event passing the NC elastic proton selection is multiplied by a weight equal to the sampled parameter value times the effect of the parameter on the event for each parameter. These weights are samples from the uncertainty distributions that were described for each systematic parameter in Sec. 5.4. We assume there is no correlation between the parameters and that there is a correlation of one between bins for a given parameter. For the single PE background rate the effect is the same for every event, and the weight is

$$w_{\text{S.P.E}} = 1 + 0.2 \cdot r_{\text{S.P.E}}, \quad (117)$$

where  $r_{\text{S.P.E}}$  is the randomly sampled single PE value, and the factor of 0.2 is the 20% uncertainty on the number of simulated events passing the NC elastic proton selection due to the uncertainty on the single PE rate. The effect of the dirt

normalization uncertainty is the same for all dirt events and zero for all non-dirt events,

$$w_{\text{dirt}} = \begin{cases} 1 & \text{non-dirt events} \\ 1 + 0.5 \cdot r_{\text{dirt}} & \text{dirt events} \end{cases}, \quad (118)$$

where  $r_{\text{dirt}}$  is the randomly sampled dirt parameter value, and the factor of 0.5 is the 50% uncertainty on the normalization of the simulated dirt events. The effect of the dynamic induced charge is applied to all simulated events and depending on the reconstructed  $Q^2$  value of the event,

$$w_{\text{D.I.C.}} = \begin{cases} 1 - 0.4 \cdot |r_{\text{D.I.C.}}| & Q_{\text{reco}}^2 < 0.175 \text{ GeV}^2 \\ 1 - 0.25 \cdot |r_{\text{D.I.C.}}| & Q_{\text{reco}}^2 \geq 0.175 \text{ GeV}^2 \end{cases}, \quad (119)$$

where  $r_{\text{D.I.C.}}$  is the randomly sampled induced charged parameter value, and the factors of 0.4 and 0.25 are the 40% uncertainty on the lowest reconstructed  $Q^2$  simulated events and the 25% uncertainty on all other simulated events, respectively. The effect of the neutrino flux uncertainty is applied to all simulated events depending on the true neutrino energy of the event as shown in Fig. 72

$$w_{\text{flux}} = 1 + f_{\text{flux}}(E_\nu) \cdot r_{\text{flux}}, \quad (120)$$

where  $f_{\text{flux}}(E_\nu)$  is the fractional uncertainty on the simulated flux for the given true neutrino energy, and  $r_{\text{flux}}$  is the randomly sampled flux parameter value. The MEC uncertainty weight is applied only to simulated true MEC events and is dependent on the true  $Q^2$  value of the event as shown in Fig. 70,

$$w_{\text{MEC}} = \begin{cases} 1 & \text{non-MEC events} \\ 1 + f_{\text{MEC}}(Q_{\text{true}}^2) \cdot r_{\text{MEC}} & \text{MEC events} \end{cases}, \quad (121)$$

where  $f_{\text{MEC}}(Q_{\text{true}}^2)$  is the fractional uncertainty on the number of simulated MEC events for the given true  $Q^2$ , and  $r_{\text{MEC}}$  is the randomly sampled MEC parameter value. Finally, the Pauli blocking uncertainty weight is applied only to all simulated events depending on the true  $Q^2$  value of the event as shown in Fig 85,

$$w_{\text{P.B.}} = 1 + f_{\text{P.B.}}(Q_{\text{true}}^2) \cdot r_{\text{P.B.}} , \quad (122)$$

where  $f_{\text{P.B.}}(Q_{\text{true}}^2)$  is the fractional uncertainty on the simulated events for the given true  $Q^2$ , and  $r_{\text{P.B.}}$  is the randomly sampled Pauli blocking parameter value. The asymmetry in the uncertainty is ignored to make it easier to evaluate the probability distribution. This is expected to be inconsequential because the asymmetry is small, the overall uncertainty is small relative to the other sources of uncertainty, and the larger of the negative or positive uncertainties is used at each true  $Q^2$  bin which is the conservative choice.

The strange axial form factor parameters are treated the same as when the systematic parameters were held to zero. The proposal distribution is the same multivariate normal in Eq. 110 centered at the current position, and the prior distributions are the same in Eq. 111 used to enforce the coefficient bounds.

Figures 95-96 show the results of 200,000 MCMC steps in  $a_0^s$ ,  $a_1^s$ , and  $a_2^s$  with the systematic parameters included. Figure 95 shows the chains of MCMC samples for  $a_0^s$ ,  $a_1^s$ , and  $a_2^s$ . Figure 96 shows the 200,000 samples of the posterior distributions for each of the two-parameter pairs of the strange axial form factor

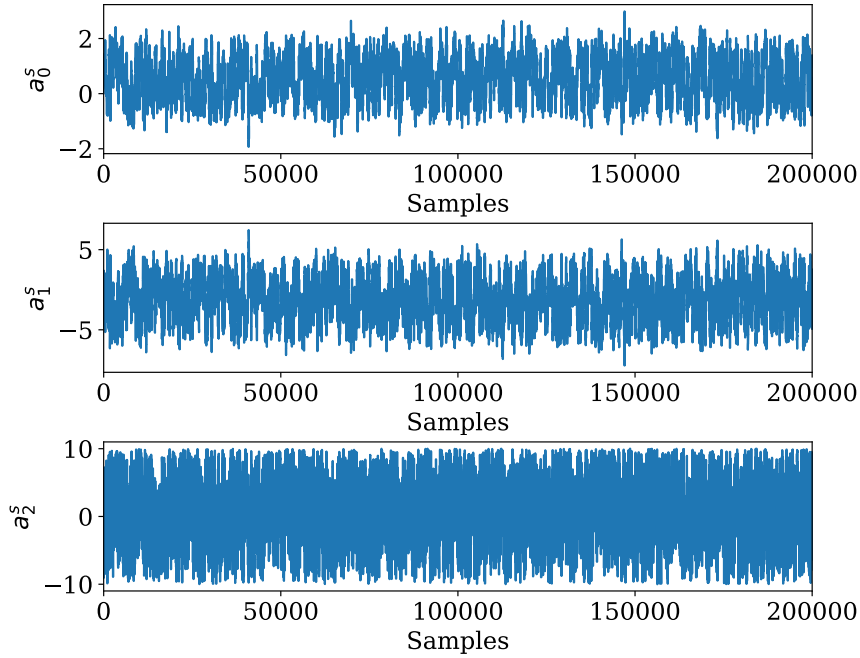


Figure 95: MCMC chains of the strange axial form factor parameters after 200,000 steps with the systematic parameters included.

parameters. The correlations between the parameters are similar to the case with no systematic uncertainty, but with larger variance. The correlations between the parameters are similar to the correlations when the systematic parameters are not included. The Pearson correlation coefficient between  $a_0^s$  and  $a_1^s$  is  $r_{01} = -0.84$ , the coefficient between  $a_1^s$  and  $a_2^s$  is  $r_{12} = -0.80$ , and the coefficient between  $a_0^s$  and  $a_2^s$  is again positive at  $r_{02} = 0.36$ .

A bimodal shape in the  $a_0^s$  and  $a_1^s$  posterior distributions is slightly visible in Figs. 95 and 96 and becomes much more apparent when we look at the two-dimensional histogram of the samples in  $a_1^s$  and  $a_0^s$  shown in Fig. 97. Figure 98



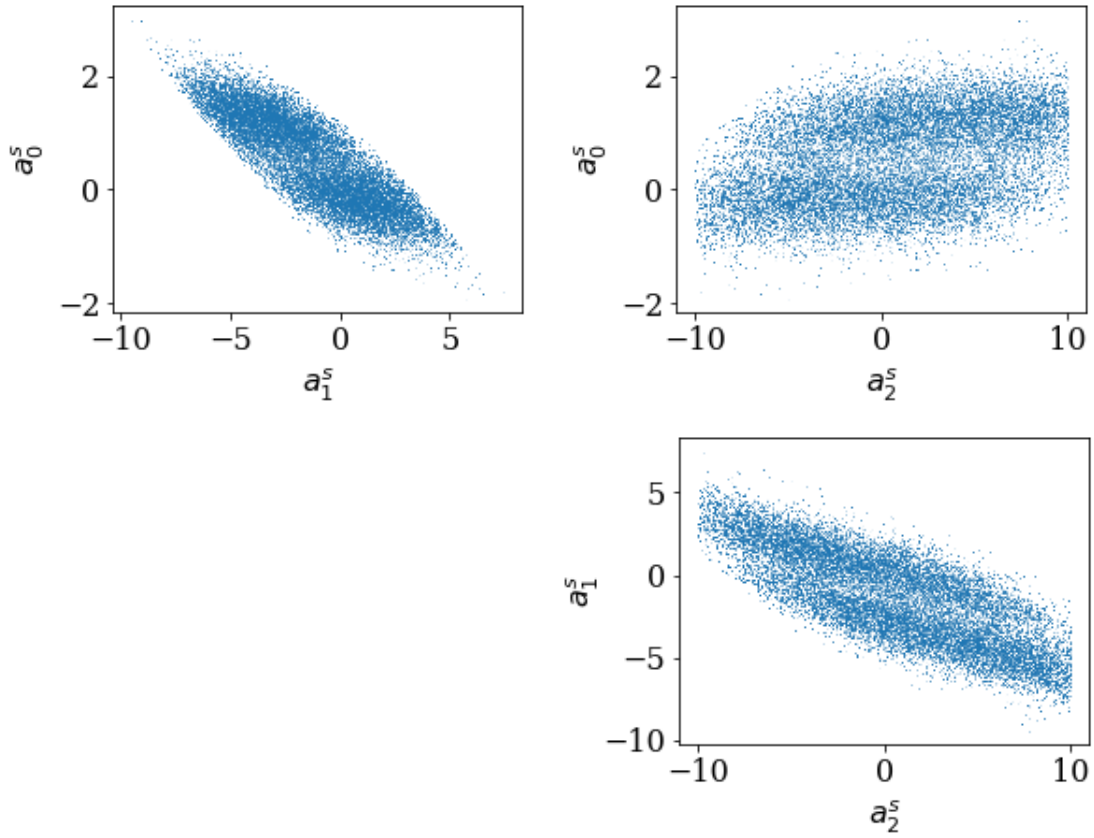


Figure 96: Two-dimensional views of the 200,000 MCMC samples of the posterior distributions of the strange axial form factor parameters with the systematic parameters included.

shows the one-dimensional histograms of the samples of  $a_0^s$ ,  $a_1^s$ , and  $a_2^s$ . The bimodal nature of the  $a_0^s$  posterior distribution is very obvious in one dimension with one clear mode at  $a_0^s = -0.1$  and another at  $a_0^s = 1.3$ . The  $a_1^s$  posterior distribution also shows a slight bimodal shape in one dimension, but it is not nearly as pronounced as in  $a_0^s$ .

Figure 99 shows the posterior predictive distribution of the number of expected

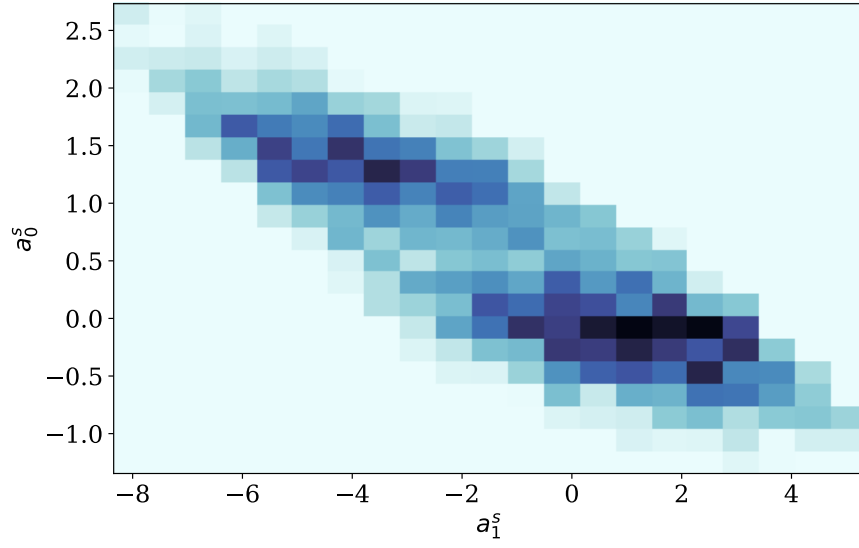


Figure 97: Two-dimension histogram of the MCMC samples of the posterior distribution for the  $a_0^s$  and  $a_1^s$  parameters.

events passing the NC elastic proton selection in Run I. The posterior predictive distribution is a prediction of what the data would look like based on the posterior distribution of the parameters. It is used to check how well the model fits the data. If the posterior predictive distribution is very different than the observed data, it shows that the model is unable to fit the data. In Fig. 99, the blue distributions at each value of reconstructed  $Q^2$  are the posterior predictive distributions, and the black points are the observed neutrino-beam data events in Run I that pass the NC elastic proton selection and the corresponding statistical uncertainty.

To understand the different  $a_0^s$  and  $a_1^s$  modes better, the posterior was divided into two samples: one with  $a_0^s$  sample values less than 0.5 and one with  $a_0^s$  sample values greater than 0.5. The posterior predictive distributions were determined

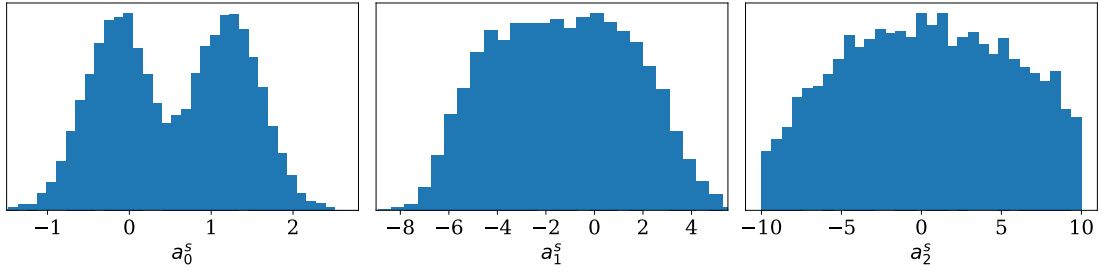


Figure 98: One-dimensional histograms of the strange axial form factor parameter samples.

for each sample and the median of each sample is plotted in Fig. 99 in red and yellow. Mode 1, in red, is the sample with  $a_0^s$  sample values less than 0.5, and Mode 2, in yellow, is the sample with  $a_0^s$  values greater than 0.5. The results are what would be expected. The Mode 1 sample, with low  $a_0^s$  values and high  $a_1^s$  corresponds to a larger number of selected events at low  $Q^2$  and a steeper slope in that region. The strange axial form factor contributes negatively to the NC elastic neutrino-proton cross section, as shown in Sec. 2.4, so a large negative value for the  $a_0^s$  term in the  $z$  expansion corresponds to a larger cross section at low  $Q^2$ . The Mode 2 sample, with more positive values of  $a_0^s$ , corresponds to a smaller number of selected events at low  $Q^2$  and a shallower slope in that region. It isn't clear from the posterior predictive distribution why there are two distinct modes, however.

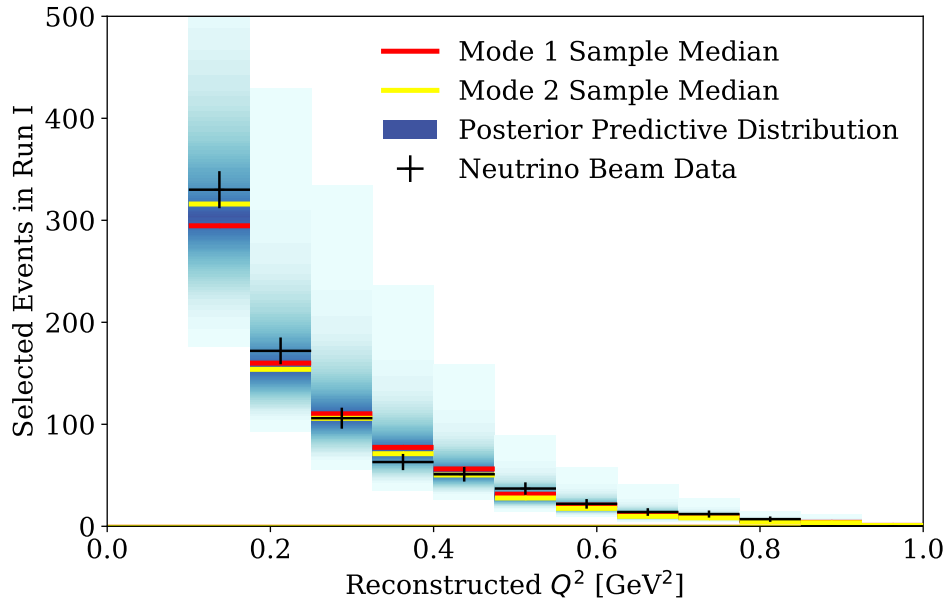


Figure 99: The posterior predictive distribution of the number of selected events compared to Run I neutrino-beam data as a function of reconstructed  $Q^2$ .

### 6.3.3 MCMC Test Metrics

There is no perfect test of whether an MCMC chain has sampled to entire target distribution [120], but there are some metrics to check to avoid common problems. The obvious problems to check for are whether the step size is too small and the MCMC chain is simply performing a random walk, whether the step size is too large and the MCMC chain isn't moving, whether the MCMC chain is stuck in a local maximum, called pseudo-convergence, or if the MCMC chain simply hasn't taken enough samples.

To test whether the step size in the proposal distribution is too large or small, the simplest and most effective metrics are the acceptance rate and plotting the

trace of the MCMC chain itself. As discussed briefly in Sec. 6.3.1, if the acceptance rate is close to one, the MCMC chain is performing a random walk, and if the acceptance rate is close to zero, the MCMC chain is stationary. It has been shown that the ideal acceptance rate for efficient Metropolis MCMC sampling with one parameter is 0.5 decreasing to 0.23 for many parameters [121]. The acceptance rate of the MCMC sampling chain was consistently 0.55, which is slightly higher than the theoretical ideal, but still far from one and zero. The trace of the MCMC chains are shown in Fig. 95. The chains do not look to be taking random walks and do not have large changes over time.

Another straight-forward test is to change the prior and proposal distributions and check that the posterior distribution is as expected. The MCMC sampling was run for 50,000 iterations without any priors on the strange axial form factor parameters and compared to the results with the priors. Recall that the priors on the strange axial form factor parameters are used to enforce the coefficient bounds described in Sec. 6.3.1 and are otherwise uniform. The mean values of the samples of the  $a_0^s$ ,  $a_1^s$ ,  $a_2^s$  did not change significantly when the coefficient bounds were removed. In fact, the mean values of each parameter were within a standard deviation of the mean values when the coefficient bounds were enforced. However, the variance of the samples greatly increased. To test the validity of the proposal distribution the MCMC was run for 50,000 iterations with a proposal distribution with twice the variance on each parameter half of the correlation between each

pair of parameters. The means and variances of the MCMC samples did not change. These simple checks show that the sampling is fairly robust to the choice of prior and proposal distributions.

The Gelman-Rubin test [122] compares the variance between separate MCMC chains to the variance within an MCMC chain to test the potential reduction in variance if the chain were to be run longer. The  $\hat{R}$  value is the measure of the potential scale reduction, and an ideal value is close to one. A result of  $\hat{R} < 1.1$  is a sign that sampling variability is negligible [123]. Ten separate MCMC chains were run for 10,000 iterations each to calculate  $\hat{R}$  for each strange axial form factor parameter

$$\begin{aligned}\hat{R}_{a_0^s} &= 1.02, \\ \hat{R}_{a_1^s} &= 1.02, \\ \hat{R}_{a_2^s} &= 1.01.\end{aligned}\tag{123}$$

These Gelman-Rubin test scores represent the potential scale reduction after 10,000 iterations. Since the MCMC sampling used for the final results here was run for 200,000 iterations, we can be reasonably assured that the final posterior distributions would not change drastically with further running.

### 6.3.4 Distributions of $\Delta s$ and $M_A^s$

To transform the strange axial form factor  $z$  expansion coefficients into  $\Delta s$ , we use the relationship derived in Sec. 2.6.3 between  $\Delta s$  and the coefficients,

$$\Delta s = a_0^s + a_1^s z_0 + a_2^s z_0^2 + a_3^s z_0^3 + a_4^s z_0^4 + a_5^s z_0^5 + a_6^s z_0^6. \quad (124)$$

Recall that the coefficients  $a_3^s, \dots, a_6^s$  are linear combinations of the first three coefficients as shown in Eq. 87. Additionally, the strange axial mass can be redefined in terms of the slope of the strange axial form factor at  $Q^2 = 0$ ,

$$\begin{aligned} M_A^s &= \sqrt{2 \frac{G_A^s(Q^2 = 0)}{G_A^{sl}(Q^2 = 0)}} \\ &= \sqrt{2\Delta s / (a_1^s + 2a_2^s z_0 + 3a_3^s z_0^2 + 4a_4^s z_0^3 + 5a_5^s z_0^4 + 6a_6^s z_0^5)}. \end{aligned} \quad (125)$$

The dependence of  $M_A^s$  on the square root of  $\Delta s$  leads to an interesting looking relationship in the posterior distribution between the two parameters. Figure 100 shows the two-dimensional histogram of the samples from the posterior distribution in terms of  $M_A^s$  and  $\Delta s$ . There is a clear difference in the shape of the distribution for  $\Delta s$  above or below zero. When  $\Delta s$  is positive, there is a positive correlation between  $\Delta s$  and  $M_A^s$ , and when  $\Delta s$  is negative, there is a negative correlation between the two.

Figures 101 and 102 show the one-dimensional histograms of the samples MCMC of the posterior distribution in terms of  $\Delta s$ , in Fig. 101, and  $M_A^s$ , in Fig. 102. The one-dimensional distributions make it clear how large the variance of the parameters is. The mode, and 68% and 95% credible intervals are shown for

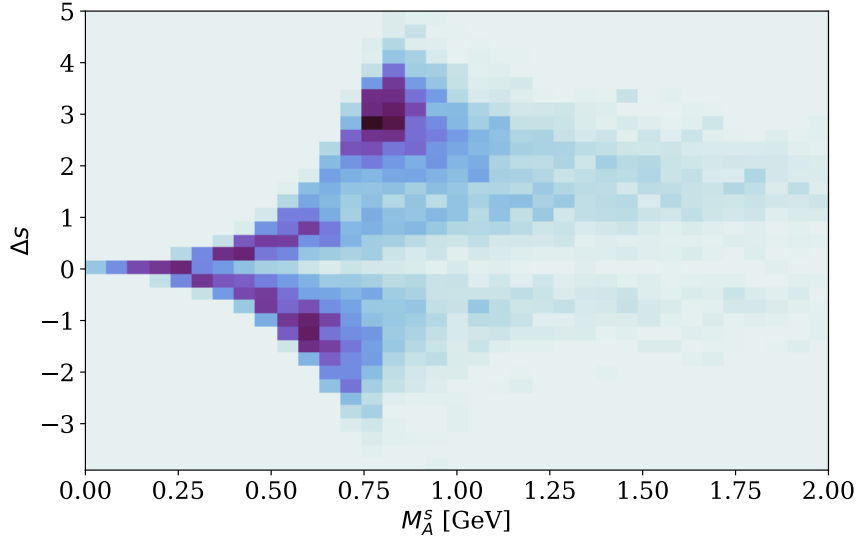


Figure 100: Two-dimensional histogram of the MCMC samples transformed to  $\Delta s$  and  $M_A^s$ .

both parameters. The credible intervals used here are the highest posterior density intervals (HPDs). The HPD is defined as the narrowest interval that contains the given percentage of the data.

For  $\Delta s$ , the mode of the samples is at  $\Delta s = -0.4$  with a 95% credible interval of  $-1.8 < \Delta s < 3.8$  and a 68% credible interval of  $-1.0 < \Delta s < 2.6$ , as shown in Fig. 101. For  $M_A^s$ , the mode of the samples is at  $M_A^s = 0.8$  GeV with a 95% credible interval of  $0 \text{ GeV} < M_A^s < 2.1 \text{ GeV}$  and a 68% credible interval of  $0.4 \text{ GeV} < M_A^s < 1.0 \text{ GeV}$ .



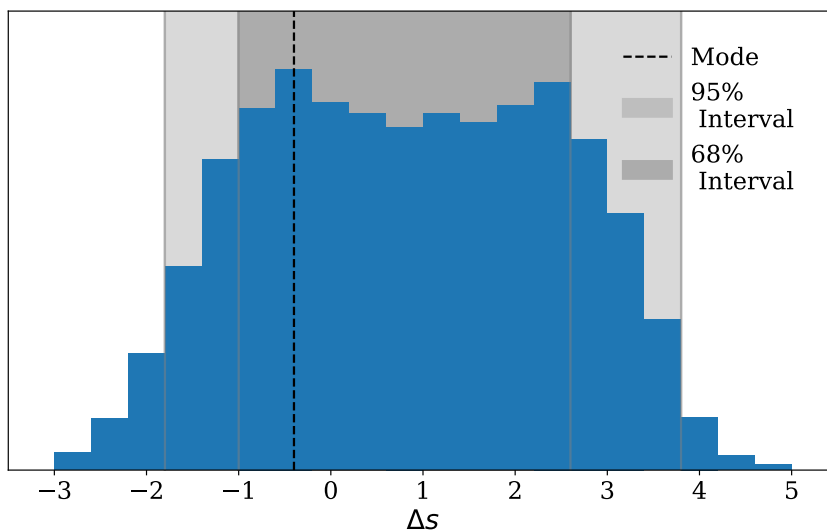


Figure 101: One-dimensional histogram of the MCMC samples transformed to  $\Delta s$  with 68% and 95% credible intervals.

#### 6.4 Conclusions and Prospects for a Future MicroBooNE $\Delta s$ Measurement

The measurement of the strange quark spin structure in the nucleon through NC elastic neutrino-proton scattering remains an intriguing and important cross check to measurements performed through deep inelastic scattering of polarized charged-leptons and nucleons. MicroBooNE remains a good prospect for carrying out this measurement because of its high resolution and neutrino energy range. The machinery to identify protons in liquid argon TPCs and select NC elastic events has been developed and is in a promising state. The limiting factor of the measurement of  $\Delta s$  in MicroBooNE is currently the understanding of the detector physics that effect both the energy reconstruction of TPC objects and

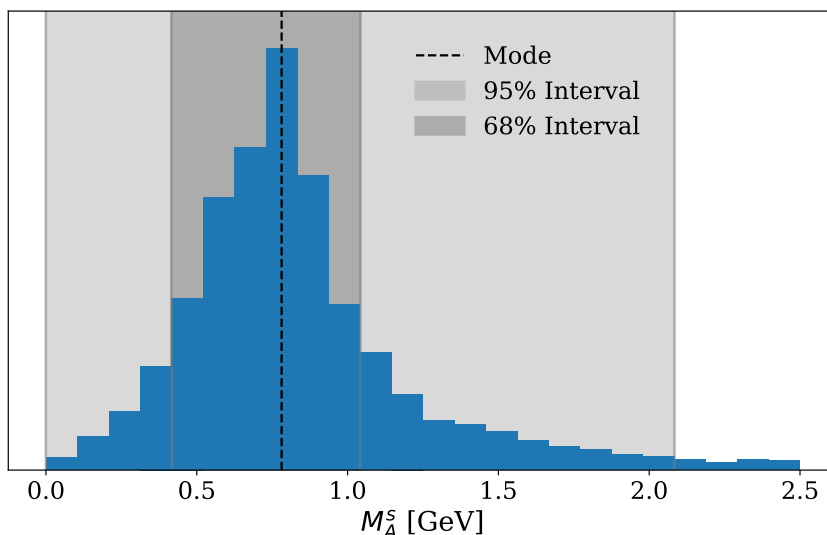


Figure 102: One-dimensional histogram of the MCMC samples transformed to  $M_A^s$  with 68% and 95% credible intervals.

the reconstruction of optical flashes in the PMTs.

The 68% credible interval around  $\Delta s$  is larger than the previous two NC elastic neutrino measurements of  $\Delta s$  described in Sec. 1.3. If we ignore the second mode in the  $\Delta s$  posterior distribution, the 68% interval ranges from less than negative one to greater than positive one. However, the potential to drastically reduce the uncertainty on the  $\Delta s$  measurement exists. Due to the fact that liquid argon TPCs are such a new technology, the ability to identify protons, select low- $Q^2$  events with a reasonable efficiency, and understand the relationship between simulated events in MicroBooNE and measured data events are all hindered by the understanding of the detector physics.

The largest source of systematic uncertainty on the expected number of events

is the dynamic induced charge, described in Sec. 5.4.2. A full, two-dimensional simulation of the induced charge on each of the wire is currently being incorporated into the MicroBooNE simulation, which should have a huge effect on the uncertainty in the simulation. The dynamic induced charge and other sources of uncertainty on the reconstruction of TPC objects effect the use of calorimetry to identify particles, as well. A lot of work has gone in to calibrating the energy deposited in the detector and the charge collected on the wires, but the track reconstruction needs to be precise enough to precisely determine the shape of the energy deposited along the distance of the track. Improvements in our understanding of the TPC physics will lead to reduced systematic uncertainty on the NC elastic proton selection, as well as a higher selection efficiency because we will be able to make more precise cuts on reconstructed variables.

The next largest effect is due to the optical model. The uncertainty on the simulation of the light in MicroBooNE effects several aspects of this analysis. As discussed in Sec. 5.4.2, the rate of a single PE background shifts the center of the reconstructed flashes causing an increase or reduction in the number of events selected. Additionally, because the light is not well modeled in MicroBooNE, the PE threshold for events to be reconstructed was raised to reduce the disagreement between the measured data and the expectation from simulation. Since single, low energy protons produce flashes near the threshold, this has a large effect on the efficiency of the selection. Future iterations of MicroBooNE simulation will

have improvements to the optical model including a treatment of Cerenkov light production and the time dependence of the single PE rate leading to an increase in the selection efficiency and reduction of systematic uncertainty.

The large variance in the posterior distribution of  $\Delta s$  is mainly due to a disagreement between the observed data and the expectation from simulation that cannot be accounted for easily in the model. Once the detector model has improved, a more precise measurement of  $\Delta s$  will be achievable. Several improvements are currently being made, and a better measurement should be possible in the next iteration of MicroBooNE the simulation model.

## REFERENCES

- [1] User: MissMJ, “The standard model,” *Wikimedia Commons*.
- [2] J. Ashman *et al.*, “An Investigation of the Spin Structure of the Proton in Deep Inelastic Scattering of Polarized Muons on Polarized Protons,” *Nucl. Phys.*, vol. B328, p. 1, 1989.
- [3] L. A. Ahrens *et al.*, “Measurement of Neutrino - Proton and anti-neutrino - Proton Elastic Scattering,” *Phys. Rev.*, vol. D35, p. 785, 1987.
- [4] G. Cheng *et al.*, “Dual baseline search for muon antineutrino disappearance at  $0.1\text{eV}^2 < \Delta m^2 < 100\text{eV}^2$ ,” *Phys. Rev.*, vol. D86, p. 052009, 2012.
- [5] A. A. Aguilar-Arevalo *et al.*, “Measurement of the Neutrino Neutral-Current Elastic Differential Cross Section on Mineral Oil at  $E_\nu \sim 1 \text{ GeV}$ ,” *Phys. Rev.*, vol. D82, p. 092005, 2010.
- [6] R. S. Sufian, K.-F. Liu, and D. G. Richards, “Weak Neutral Current Axial Form Factor and (Anti)Neutrino-Nucleon Scattering,” 2018.
- [7] T. M. Collaboration, “Cosmic Shielding Studies at MicroBooNE,” *Micro-BooNE Public Note*, vol. MICROBOONE-NOTE-1005-PUB, 2016.
- [8] T. M. Collaboration, “Booster Neutrino Flux Prediction at MicroBooNE,” *MicroBooNE Public Note*, vol. MICROBOONE-NOTE-1031-PUB, 2018.
- [9] E. Rutherford, “Bakerian Lecture. Nuclear Constitution of Atoms,” *Proc.Roy.Soc.Lond.*, vol. A97, p. 374, 1920.
- [10] F. Reines and C. L. Cowan, “The neutrino,” *Nature*, vol. 178, pp. 446–449, 1956.
- [11] M. Gell-Mann, “A Schematic Model of Baryons and Mesons,” *Phys. Lett.*, vol. 8, pp. 214–215, 1964.
- [12] G. Zweig, “An SU(3) model for strong interaction symmetry and its breaking. Version 1,” 1964.
- [13] G. Zweig, “An SU(3) model for strong interaction symmetry and its breaking. Version 2,” in *DEVELOPMENTS IN THE QUARK THEORY OF HADRONS. VOL. 1. 1964 - 1978* (D. Lichtenberg and S. P. Rosen, eds.), pp. 22–101, 1964.
- [14] J. I. Friedman and H. W. Kendall, “Deep inelastic electron scattering,” *Ann. Rev. Nucl. Part. Sci.*, vol. 22, pp. 203–254, 1972.

- [15] R. Davis, Jr., D. S. Harmer, and K. C. Hoffman, “Search for neutrinos from the sun,” *Phys. Rev. Lett.*, vol. 20, pp. 1205–1209, 1968.
- [16] J. N. Bahcall, N. A. Bahcall, and G. Shaviv, “Present status of the theoretical predictions for the Cl-36 solar neutrino experiment,” *Phys. Rev. Lett.*, vol. 20, pp. 1209–1212, 1968. [45(1968)].
- [17] Y. Fukuda *et al.*, “Evidence for oscillation of atmospheric neutrinos,” *Phys. Rev. Lett.*, vol. 81, pp. 1562–1567, 1998.
- [18] J. Ashman *et al.*, “A Measurement of the Spin Asymmetry and Determination of the Structure Function  $g(1)$  in Deep Inelastic Muon-Proton Scattering,” *Phys. Lett.*, vol. B206, p. 364, 1988.
- [19] A. Aguilar-Arevalo *et al.*, “Evidence for neutrino oscillations from the observation of anti-neutrino(electron) appearance in a anti-neutrino(muon) beam,” *Phys. Rev.*, vol. D64, p. 112007, 2001.
- [20] A. A. Aguilar-Arevalo *et al.*, “Event Excess in the MiniBooNE Search for  $\bar{\nu}_\mu \rightarrow \bar{\nu}_e$  Oscillations,” *Phys. Rev. Lett.*, vol. 105, p. 181801, 2010.
- [21] A. A. Aguilar-Arevalo *et al.*, “Observation of a Significant Excess of Electron-Like Events in the MiniBooNE Short-Baseline Neutrino Experiment,” 2018.
- [22] R. Acciarri *et al.*, “Design and Construction of the MicroBooNE Detector,” *JINST*, vol. 12, no. 02, p. P02017, 2017.
- [23] R. P. Feynman, “The behavior of hadron collisions at extreme energies,” *Conf. Proc.*, vol. C690905, pp. 237–258, 1969.
- [24] A. W. Thomas and W. Weise, *The Structure of the Nucleon*. 2001.
- [25] S. D. Bass, *The Spin structure of the proton*. 2007.
- [26] J. R. Ellis and R. L. Jaffe, “A Sum Rule for Deep Inelastic Electroproduction from Polarized Protons,” *Phys. Rev.*, vol. D9, p. 1444, 1974. [Erratum: *Phys. Rev.*D10,1669(1974)].
- [27] D. Dubbers, “Particle physics with cold neutrons,” *Prog. Part. Nucl. Phys.*, vol. 26, pp. 173–252, 1991.
- [28] C. A. Aidala, S. D. Bass, D. Hasch, and G. K. Mallot, “The Spin Structure of the Nucleon,” *Rev. Mod. Phys.*, vol. 85, pp. 655–691, 2013.

- [29] G. Baum *et al.*, “A New Measurement of Deep Inelastic e p Asymmetries,” *Phys. Rev. Lett.*, vol. 51, p. 1135, 1983.
- [30] P. L. Anthony *et al.*, “Deep inelastic scattering of polarized electrons by polarized He-3 and the study of the neutron spin structure,” *Phys. Rev.*, vol. D54, pp. 6620–6650, 1996.
- [31] K. Abe *et al.*, “Measurements of the proton and deuteron spin structure functions  $g(1)$  and  $g(2)$ ,” *Phys. Rev.*, vol. D58, p. 112003, 1998.
- [32] B. Adeva *et al.*, “Measurement of the spin dependent structure function  $g_1(x)$  of the deuteron,” *Phys. Lett.*, vol. B302, pp. 533–539, 1993.
- [33] B. Adeva *et al.*, “Spin asymmetries  $A(1)$  and structure functions  $g_1$  of the proton and the deuteron from polarized high-energy muon scattering,” *Phys. Rev.*, vol. D58, p. 112001, 1998.
- [34] V. Yu. Alexakhin *et al.*, “The Deuteron Spin-dependent Structure Function  $g_1(d)$  and its First Moment,” *Phys. Lett.*, vol. B647, pp. 8–17, 2007.
- [35] K. Ackerstaff *et al.*, “Measurement of the neutron spin structure function  $g_1(n)$  with a polarized He-3 internal target,” *Phys. Lett.*, vol. B404, pp. 383–389, 1997.
- [36] K. Ackerstaff *et al.*, “Flavor decomposition of the polarized quark distributions in the nucleon from inclusive and semiinclusive deep inelastic scattering,” *Phys. Lett.*, vol. B464, pp. 123–134, 1999.
- [37] A. Airapetian *et al.*, “Precise determination of the spin structure function  $g_1$  of the proton, deuteron and neutron,” *Phys. Rev.*, vol. D75, p. 012007, 2007.
- [38] M. Alekseev *et al.*, “Flavour Separation of Helicity Distributions from Deep Inelastic Muon-Deuteron Scattering,” *Phys. Lett.*, vol. B680, pp. 217–224, 2009.
- [39] M. G. Alekseev *et al.*, “Quark helicity distributions from longitudinal spin asymmetries in muon-proton and muon-deuteron scattering,” *Phys. Lett.*, vol. B693, pp. 227–235, 2010.
- [40] A. Airapetian *et al.*, “Flavor decomposition of the sea quark helicity distributions in the nucleon from semiinclusive deep inelastic scattering,” *Phys. Rev. Lett.*, vol. 92, p. 012005, 2004.

- [41] A. Airapetian *et al.*, “Quark helicity distributions in the nucleon for up, down, and strange quarks from semi-inclusive deep-inelastic scattering,” *Phys. Rev.*, vol. D71, p. 012003, 2005.
- [42] A. Airapetian *et al.*, “Measurement of Parton Distributions of Strange Quarks in the Nucleon from Charged-Kaon Production in Deep-Inelastic Scattering on the Deuteron,” *Phys. Lett.*, vol. B666, pp. 446–450, 2008.
- [43] D. de Florian, R. Sassot, M. Stratmann, and W. Vogelsang, “Global Analysis of Helicity Parton Densities and Their Uncertainties,” *Phys. Rev. Lett.*, vol. 101, p. 072001, 2008.
- [44] D. de Florian, R. Sassot, M. Stratmann, and W. Vogelsang, “Extraction of Spin-Dependent Parton Densities and Their Uncertainties,” *Phys. Rev.*, vol. D80, p. 034030, 2009.
- [45] J. Blumlein and H. Bottcher, “QCD Analysis of Polarized Deep Inelastic Scattering Data,” *Nucl. Phys.*, vol. B841, pp. 205–230, 2010.
- [46] E. R. Nocera, R. D. Ball, S. Forte, G. Ridolfi, and J. Rojo, “A first unbiased global determination of polarized PDFs and their uncertainties,” *Nucl. Phys.*, vol. B887, pp. 276–308, 2014.
- [47] E. Leader, A. V. Sidorov, and D. B. Stamenov, “New analysis concerning the strange quark polarization puzzle,” *Phys. Rev.*, vol. D91, no. 5, p. 054017, 2015.
- [48] M. J. Savage and J. Walden, “SU(3) breaking in neutral current axial matrix elements and the spin content of the nucleon,” *Phys. Rev.*, vol. D55, pp. 5376–5384, 1997.
- [49] S. J. Dong, J. F. Lagae, and K. F. Liu, “Flavor singlet  $g(A)$  from lattice QCD,” *Phys. Rev. Lett.*, vol. 75, pp. 2096–2099, 1995.
- [50] G. S. Bali *et al.*, “Strangeness Contribution to the Proton Spin from Lattice QCD,” *Phys. Rev. Lett.*, vol. 108, p. 222001, 2012.
- [51] M. Engelhardt, “Strange quark contributions to nucleon mass and spin from lattice QCD,” *Phys. Rev.*, vol. D86, p. 114510, 2012.
- [52] A. Abdel-Rehim, C. Alexandrou, M. Constantinou, K. Hadjiyiannakou, K. Jansen, C. Kallidonis, G. Koutsou, and A. V. Avils-Casco, “Disconnected quark loop contributions to nucleon observables using  $N_f = 2$  twisted clover fermions at the physical value of the light quark mass,” *PoS*, vol. LATTICE2015, p. 136, 2016.



- [53] A. J. Chambers *et al.*, “Disconnected contributions to the spin of the nucleon,” *Phys. Rev.*, vol. D92, no. 11, p. 114517, 2015.
- [54] V. Lyubushkin *et al.*, “A Study of quasi-elastic muon neutrino and antineutrino scattering in the NOMAD experiment,” *Eur. Phys. J.*, vol. C63, pp. 355–381, 2009.
- [55] J. A. Formaggio and G. P. Zeller, “From eV to EeV: Neutrino Cross Sections Across Energy Scales,” *Rev. Mod. Phys.*, vol. 84, pp. 1307–1341, 2012.
- [56] G. T. Garvey, W. C. Louis, and D. H. White, “Determination of proton strange form-factors from neutrino p elastic scattering,” *Phys. Rev.*, vol. C48, pp. 761–765, 1993.
- [57] W. M. Alberico, M. B. Barbaro, S. M. Bilenky, J. A. Caballero, C. Giunti, C. Maieron, E. Moya de Guerra, and J. M. Udias, “Strange form-factors of the proton: A New analysis of the neutrino (anti-neutrino) data of the BNL-734 experiment,” *Nucl. Phys.*, vol. A651, pp. 277–286, 1999.
- [58] S. F. Pate, D. W. McKee, and V. Papavassiliou, “Strange Quark Contribution to the Vector and Axial Form Factors of the Nucleon: Combined Analysis of G0, HAPPEX, and Brookhaven E734 Data,” *Phys. Rev.*, vol. C78, p. 015207, 2008.
- [59] D. S. Armstrong *et al.*, “Strange quark contributions to parity-violating asymmetries in the forward G0 electron-proton scattering experiment,” *Phys. Rev. Lett.*, vol. 95, p. 092001, 2005.
- [60] K. A. Aniol *et al.*, “Parity violating electroweak asymmetry in polarized-e p scattering,” *Phys. Rev.*, vol. C69, p. 065501, 2004.
- [61] A. A. Aguilar-Arevalo *et al.*, “The Neutrino Flux prediction at MiniBooNE,” *Phys. Rev.*, vol. D79, p. 072002, 2009.
- [62] T. Golan, K. M. Graczyk, C. Juszczak, and J. T. Sobczyk, “Extraction of Axial Mass and Strangeness Values from the MiniBooNE Neutral Current Elastic Cross Section Measurement,” *Phys. Rev.*, vol. C88, p. 024612, 2013.
- [63] T. Golan, C. Juszczak, and J. T. Sobczyk, “Final State Interactions Effects in Neutrino-Nucleus Interactions,” *Phys. Rev.*, vol. C86, p. 015505, 2012.
- [64] D. Casper, “The Nuance neutrino physics simulation, and the future,” *Nucl. Phys. Proc. Suppl.*, vol. 112, pp. 161–170, 2002.

- [65] I. J. R. Aitchison and A. J. G. Hey, *Gauge theories in particle physics: A practical introduction. Vol. 2: Non-Abelian gauge theories: QCD and the electroweak theory.* Bristol, UK: CRC Press, 2012.
- [66] E. A. Paschos, *Electroweak theory.* Cambridge University Press, 2010.
- [67] W. M. Alberico, S. M. Bilenky, and C. Maieron, “Strangeness in the nucleon: Neutrino - nucleon and polarized electron - nucleon scattering,” *Phys. Rept.*, vol. 358, pp. 227–308, 2002.
- [68] J. C. Ward, “An Identity in Quantum Electrodynamics,” *Phys. Rev.*, vol. 78, p. 182, 1950.
- [69] W. Gordon, “Der strom der Diracschen elektronentheorie,” *Zeitschrift für Physik*, vol. 50, pp. 631–632, 1928.
- [70] S. Gerstein and Y. Zeldovich, “Meson Corrections in the Theory of Beta Decay,” *JETP*, vol. 2, p. 576, 1950.
- [71] R. P. Feynman and M. Gell-Mann, “Theory of Fermi interaction,” *Phys. Rev.*, vol. 109, pp. 193–198, 1958.
- [72] C. H. Llewellyn Smith, “Neutrino Reactions at Accelerator Energies,” *Phys. Rept.*, vol. 3, pp. 261–379, 1972.
- [73] C. F. Perdrisat, V. Punjabi, and M. Vanderhaeghen, “Nucleon Electromagnetic Form Factors,” *Prog. Part. Nucl. Phys.*, vol. 59, pp. 694–764, 2007.
- [74] B. Bhattacharya, R. J. Hill, and G. Paz, “Model independent determination of the axial mass parameter in quasielastic neutrino-nucleon scattering,” *Phys. Rev.*, vol. D84, p. 073006, 2011.
- [75] C. G. Boyd and M. J. Savage, “Analyticity, shapes of semileptonic form-factors, and anti-B  $\rightarrow$  pi lepton anti-neutrino,” *Phys. Rev.*, vol. D56, pp. 303–311, 1997.
- [76] P. Federbush, M. L. Goldberger, and S. B. Treiman, “Electromagnetic Structure of the Nucleon,” *Phys. Rev.*, vol. 112, pp. 642–665, 1958.
- [77] A. S. Meyer, M. Betancourt, R. Gran, and R. J. Hill, “Deuterium target data for precision neutrino-nucleus cross sections,” *Phys. Rev.*, vol. D93, no. 11, p. 113015, 2016.
- [78] G. P. Lepage and S. J. Rodsky, “Exclusive Processes in Perturbative Quantum Chromodynamics,” *Phys. Rev.*, vol. D22, p. 2157, 1980.

- [79] G. Lee, J. R. Arrington, and R. J. Hill, “Extraction of the proton radius from electron-proton scattering data,” *Phys. Rev.*, vol. D92, no. 1, p. 013013, 2015.
- [80] R. J. Hill and G. Paz, “Model independent extraction of the proton charge radius from electron scattering,” *Phys. Rev.*, vol. D82, p. 113005, 2010.
- [81] Z. Epstein, G. Paz, and J. Roy, “Model independent extraction of the proton magnetic radius from electron scattering,” *Phys. Rev.*, vol. D90, no. 7, p. 074027, 2014.
- [82] Z. Ye, J. Arrington, R. J. Hill, and G. Lee, “Proton and Neutron Electromagnetic Form Factors and Uncertainties,” *Phys. Lett.*, vol. B777, pp. 8–15, 2018.
- [83] B. Bhattacharya, G. Paz, and A. J. Tropicano, “Model-independent determination of the axial mass parameter in quasielastic antineutrino-nucleon scattering,” *Phys. Rev.*, vol. D92, no. 11, p. 113011, 2015.
- [84] W. A. Mann *et al.*, “Study of the reaction  $\nu n \rightarrow \mu^- p$ ,” *Phys. Rev. Lett.*, vol. 31, pp. 844–847, 1973.
- [85] S. J. Barish *et al.*, “Study of Neutrino Interactions in Hydrogen and Deuterium. 1. Description of the Experiment and Study of the Reaction  $\text{Neutrino } d \rightarrow \mu^- p p(s)$ ,” *Phys. Rev.*, vol. D16, p. 3103, 1977.
- [86] K. L. Miller *et al.*, “Study of the Reaction  $\mu\text{-neutrino } D \rightarrow \mu^- p p(S)$ ,” *Phys. Rev.*, vol. D26, pp. 537–542, 1982.
- [87] N. J. Baker, A. M. Cnops, P. L. Connolly, S. A. Kahn, H. G. Kirk, M. J. Murtagh, R. B. Palmer, N. P. Samios, and M. Tanaka, “Quasielastic Neutrino Scattering: A Measurement of the Weak Nucleon Axial Vector Form-Factor,” *Phys. Rev.*, vol. D23, pp. 2499–2505, 1981.
- [88] T. Kitagaki *et al.*, “High-Energy Quasielastic Muon-neutrino  $n \rightarrow \mu^- p$  Scattering in Deuterium,” *Phys. Rev.*, vol. D28, pp. 436–442, 1983.
- [89] C. Adams *et al.*, “Comparison of  $\nu_\mu$ -Ar multiplicity distributions observed by MicroBooNE to GENIE model predictions,” *Submitted to: Phys. Rev. D*, 2018.
- [90] M. V. Diwan, V. Galymov, X. Qian, and A. Rubbia, “Long-Baseline Neutrino Experiments,” *Ann. Rev. Nucl. Part. Sci.*, vol. 66, pp. 47–71, 2016.

- [91] S. Agostinelli *et al.*, “GEANT4: A Simulation toolkit,” *Nucl. Instrum. Meth.*, vol. A506, pp. 250–303, 2003.
- [92] M. G. Catanesi *et al.*, “Measurement of the production cross-section of positive pions in the collision of 8.9-GeV/c protons on beryllium,” *Eur. Phys. J.*, vol. C52, pp. 29–53, 2007.
- [93] T. M. Collaboration, “Measurement of the Electronegative Contaminants and Drift Electron Lifetime in the MicroBooNE Experiment,” *MicroBooNE Public Note*, vol. MICROBOONE-NOTE-1003-PUB, 2016.
- [94] D. Heck, J. Knapp, J. N. Capdevielle, G. Schatz, and T. Thouw, “CORSIKA: A Monte Carlo code to simulate extensive air showers,” 1998.
- [95] C. Andreopoulos *et al.*, “The GENIE Neutrino Monte Carlo Generator,” *Nucl. Instrum. Meth.*, vol. A614, pp. 87–104, 2010.
- [96] C. Andreopoulos, C. Barry, S. Dytman, H. Gallagher, T. Golan, R. Hatcher, G. Perdue, and J. Yarba, “The GENIE Neutrino Monte Carlo Generator: Physics and User Manual,” 2015.
- [97] M. Alam *et al.*, “GENIE Production Release 2.10.0,” 2015.
- [98] R. A. Smith and E. J. Moniz, “NEUTRINO REACTIONS ON NUCLEAR TARGETS,” *Nucl. Phys.*, vol. B43, p. 605, 1972. [Erratum: *Nucl. Phys.*B101,547(1975)].
- [99] H. De Vries, C. W. De Jager, and C. De Vries, “Nuclear charge and magnetization density distribution parameters from elastic electron scattering,” *Atom. Data Nucl. Data Tabl.*, vol. 36, pp. 495–536, 1987.
- [100] R. D. Woods and D. S. Saxon, “Diffuse Surface Optical Model for Nucleon-Nuclei Scattering,” *Phys. Rev.*, vol. 95, pp. 577–578, 1954.
- [101] T. Yang, C. Andreopoulos, H. Gallagher, K. Hoffmann, and P. Kehayias, “A Hadronization Model for Few-GeV Neutrino Interactions,” *Eur. Phys. J.*, vol. C63, pp. 1–10, 2009.
- [102] P. Adamson *et al.*, “A Study of Muon Neutrino Disappearance Using the Fermilab Main Injector Neutrino Beam,” *Phys. Rev.*, vol. D77, p. 072002, 2008.
- [103] G. Battistoni, A. Ferrari, P. Sala, C. Forti, J. Ranft, and S. Roesler, “The production of residual nuclei in peripheral hadron nucleus and nucleus nucleus collisions,” *Nucl. Phys. Proc. Suppl.*, vol. 52B, pp. 120–122, 1997.

- [104] R. Acciarri *et al.*, “The Pandora multi-algorithm approach to automated pattern recognition of cosmic-ray muon and neutrino events in the MicroBooNE detector,” *Eur. Phys. J.*, vol. C78, no. 1, p. 82, 2018.
- [105] M. Antonello *et al.*, “Precise 3D track reconstruction algorithm for the ICARUS T600 liquid argon time projection chamber detector,” *Adv. High Energy Phys.*, vol. 2013, p. 260820, 2013.
- [106] T. M. Collaboration, “Detector calibration using through going and stopping muons in the MicroBooNE LArTPC,” *MicroBooNE Public Note*, vol. MICROBOONE-NOTE-1048-PUB, 2016.
- [107] T. Chen and C. Guestrin, “XGBoost: A Scalable Tree Boosting System,” 2016.
- [108] D. W. Hosmer, S. Lemeshow, and R. X. Sturdivant, *Applied Logistic Regression*, vol. 3. Wiley, 2013.
- [109] S. Seabold and J. Perktold, “Statsmodels: Econometric and statistical modeling with python,” *9th Python in Science Conference*, 2010.
- [110] A. Bhadra, S. K. Ghosh, P. S. Joarder, A. Mukherjee, and S. Raha, “Study of low energy hadronic interaction models based on BESS observed cosmic ray proton and antiproton spectra at medium high altitude,” *Phys. Rev.*, vol. D79, p. 114027, 2009.
- [111] E. Jones, T. Oliphant, P. Peterson, *et al.*, “SciPy: Open source scientific tools for Python,” 2001. [Online; accessed 2018-10-05].
- [112] J. Nieves, I. Ruiz Simo, and M. J. Vicente Vacas, “Inclusive Charged-Current Neutrino-Nucleus Reactions,” *Phys. Rev.*, vol. C83, p. 045501, 2011.
- [113] J. Nieves, I. Ruiz Simo, and M. J. Vicente Vacas, “The nucleon axial mass and the MiniBooNE Quasielastic Neutrino-Nucleus Scattering problem,” *Phys. Lett.*, vol. B707, pp. 72–75, 2012.
- [114] C. Adams *et al.*, “Ionization electron signal processing in single phase LArTPCs. Part I. Algorithm Description and quantitative evaluation with MicroBooNE simulation,” *JINST*, vol. 13, no. 07, p. P07006, 2018.
- [115] T. M. Collaboration, “Study of Space Charge Effects in MicroBooNE,” *MicroBooNE Public Note*, vol. MICROBOONE-NOTE-1018-PUB, 2016.

- [116] A. Bodek and J. L. Ritchie, “Further Studies of Fermi Motion Effects in Lepton Scattering from Nuclear Targets,” *Phys. Rev.*, vol. D24, p. 1400, 1981.
- [117] T. Leitner, O. Buss, L. Alvarez-Ruso, and U. Mosel, “Electron- and neutrino-nucleus scattering from the quasielastic to the resonance region,” *Phys. Rev.*, vol. C79, p. 034601, 2009.
- [118] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, “Equation of state calculations by fast computing machines,” *J. Chem. Phys.*, vol. 21, pp. 1087–1092, 1953.
- [119] S. Geman and D. Geman, “Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-6, no. 6, pp. 721–741, 1984.
- [120] S. Brooks, A. Gelman, G. Jones, and X.-L. Meng, *Handbook of Markov Chain Monte Carlo*. 2011.
- [121] G. Roberts, A. Gelman, and W. Gilks, “Weak Convergence and Optical Scaling of Random Walk Metropolis Algorithms,” *Ann. Appl. Probab.*, vol. 7, no. 1, pp. 110–120, 1997.
- [122] A. Gelman and D. B. Rubin, “Inference from Iterative Simulation Using Multiple Sequences,” *Statist. Sci.*, vol. 7, pp. 457–472, 1992.
- [123] S. P. Brooks and A. Gelman, “General Methods for Monitoring Convergence of Iterative Simulations,” *J. of Comp. and Graph. Stats.*, vol. 7, pp. 434–455, 1997.