MEASUREMENT OF π^+ – ARGON ABSORPTION AND CHARGE EXCHANGE INTERACTIONS USING PROTODUNE-SP

By

Jacob Calcutt

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Physics – Doctor of Philosophy

2021

ABSTRACT

MEASUREMENT OF π^+ – ARGON ABSORPTION AND CHARGE EXCHANGE INTERACTIONS USING PROTODUNE-SP

By

Jacob Calcutt

ProtoDUNE-SP is a prototype detector for the upcoming Deep Underground Neutrino Experiment (DUNE). It's a Liquid Argon Time Projection Chamber (LArTPC) with a similar configuration to DUNE's detector, and is designed to provide a test-bed for the future experiment. In addition to serving as a prototype, its 0.3 – 7 GeV/c charged particle beam line provided the ability to perform physics measurements of pions, protons, kaons, muons, and electrons. Importantly, the LArTPC allowed for the measurement of hadronic interactions on argon nuclei.

Pions are often present in the final state of neutrino interactions in the energy range of DUNE's neutrino beam. These particles can undergo various types of interactions with argon nuclei in the detector, and this can interfere with the characterization of neutrino interactions in DUNE's far detector. The rate of these so-called secondary interactions will be accounted for using Monte Carlo simulation of neutrino interactions. Measurements of secondary interaction rates provide necessary data which can be used to estimate and propagate uncertainties or provide tunes of the secondary interaction model used within DUNE's experimental simulation.

This analysis provides a simultaneous measurement of the π^+–Ar absorption and charge exchange cross sections using 1 GeV/c π^+ data taken by ProtoDUNE-SP during its initial run period in Fall 2018. This is one of the first hadronic interaction measurements provided by ProtoDUNE-SP. It is also the first π^+–Ar absorption measurement in 20 years and the first ever π^+–Ar charge exchange measurement.

Copyright by
JACOB CALCUTT
2021

ACKNOWLEDGEMENTS

I must mention a special set of people in my life, without whom, I would have never done this.

Firstly, my partner, Rae, we met in my first year of grad school, and so the journey through the last 6 years is essentially our journey together. You've offered me countless hours of support when I thought I couldn't keep going. You've shared joy in my successes. You've taught me so many things and exposed me to so many new ideas and ways of thinking about the world. I'm excited to continue on this journey together, and cannot wait to marry you next year.

To my sisters, Caitlin and Erin, you've been there since the beginning. I would never have even reached college, let alone get to this point, without you two in my life. Your guidance, love, and support is without equal. You've helped me get through heartbreak and loss, and gave me the motivation to keep going and even start on this course. Erin, when I first mentioned I wanted to study physics, you were so enthusiastic that you immediately bought me books on the subject. When I had a rough first year of undergrad and wanted to drop out, you convinced me to stay. Caitlin, you helped me find my first few jobs in high school and college and helped me stay afloat. Your sage wisdom and advice has helped me navigate through the toughest times of my life. I love geeking out about music and video games with you. I'm so proud of where you two are now, especially knowing the hard times we've had together and individually, and I know we'll all do great things.

Tyler, Greg, and Curtis, I'm so glad we all became friends in undergrad. That last year was the best time I had in college and it wouldn't have happened without you. From studying and doing homework together, to having some beers at Crunchy's, P.T.'s, Woody's (RIP), or What Up Dawg (also RIP), we had so many fun and memorable experiences together. Tyler, I'm so happy you moved back to MI after undergrad and we were able to stay close in-person over the years. Thanks for all the music you've introduced me to and the shows we went to. I'm glad you found someone like Mika to share your life with you, and thank you for bringing her into my life as a friend as well. Congrats on the new job! I'm happy we've all stayed in touch and look forward to when we can all meet up in person again.

Justin and Rob, you were two of the best roommates I could ever have hoped for. I loved hanging out, throwing parties, talking physics, and playing games at the house together. You were some of the nicest and funnest people in our cohort, and I was ecstatic when we all made plans to live together. I've been bad at keeping in touch, but I'll try to fix that. Talk to you soon.

To Andrew, Dan, and Luke, you three are all amazing officemates, scientists, and, above all, friends. Your camaraderie throughout the last few years is without equal. Thank you for all the fun times together like playing games together in the cramped office. Thank you even more for all the amazing help and insight that helped me through this analysis (especially Andrew, as this analysis is largely modeled after yours). I hope that collaboration continues in the future on DUNE. To Luke, a special thanks is in order. You've been an amazing role model to me. Your skills in programming have inspired me to keep improving my own (though they'll never be anywhere close to yours). I look up to you in so many ways, and I'm glad you joined Kendall's team when you did.

A huge thank you to everyone I met while working on ProtoDUNE: Milo, Casandra, Bryan, Chris, Aidan, Ajib, Heng-Ye, Peter, David, Richie; and a special thanks to Rob and Geoff for being my roommates for a few stays. The time at CERN would not have been the same without any of you there, and you all helped me make so many memories. The fun times we'd have while both working or exploring Geneva motivated me during those intense periods of work. I'm excited to keep working on DUNE with some of you. To the others, I wish you the best and hope to meet up again some time.

Francesca and Stefania, thank you for your close help on this analysis during the last few years. You are both amazing physicists and offered me so many useful ideas and examples of how to keep moving forward. Francesca, I'm glad we've become friends through doing this

analysis together, and hope we continue to work together in the future.

Clarence, thanks for being a friend since I spent the summer at Imperial. I'm glad you were there at Fermilab when I was there during my fellowship. I had so many fun times whether that be just hanging out, going to the climbing gym, or brewing beer. Looking forward to seeing you again.

To Kim and Brenda at MSU, and Maxine at DUNE, your administrative help throughout the years has always taken a giant load off my shoulders. Whether that be Kim's help figuring out all the little funding snafus and headaches, Brenda's work with scheduling and reimbursing all the conferences and trips, or Maxine's help finding a place to stay at CERN or helping me when I accidentally put diesel into a petrol car (oops..): you've all done so much work because of me, and I greatly appreciate it. Kim, you've been there since undergrad, and helped me keep going through some very rough times then and through grad school (all with a huge smile on your face). I'll always remember you.

To Tingjun, Leigh, and Tom, thank you all for hard work leading efforts at ProtoDUNE and providing guidance on this analysis and other work I've done. Leigh, thank you for your advice and help with the beam simulation. Tom, you're a master of statistics, and I have so much more to learn from you. Tingjun, thank you for working so closely with me on this analysis, and providing help and support through it. Additional thanks to you for all the letters of recommendation you've written for me. I'm glad to continue working with you all on DUNE.

Laura, thank you for being my fellowship advisor on the Geant4Reweight project. That project opened up so many opportunities for me, and helped me grow as a researcher. I couldn't have done it without you. I'm glad to have worked under someone like you, and you're a large inspiration for me. And of course, thank you for writing letters for me.

And last, but certainly not least, I'd like to thank my advisor Kendall. I'm so lucky that you took a chance on me during undergrad to let me research with you. That was such a momentous point in my life, as it opened up the road to getting to here. You've

helped me grow as a researcher and physicist, and this work would not have been possible without your direction. You've helped me through dark times, and have always been caring and understanding. You do so much work, and all while starting a family, it's absolutely unbelievable. I can tell you really care about me and other students working under you. I find it amazing and admirable the way you look for so many undergraduate students to work with you, as that is such an important part of becoming a physicist or, equally important, deciding you want to do something else. You are brilliant, and always looking for a new project to get your hands on to help the advancement of science and the experiments you work on. I'm lucky and proud to have you as a mentor.

TABLE OF CONTENTS

CHAPTER 1

EXECUTIVE SUMMARY

The field of study of neutrino oscillation has entered the precision era. Next-generation experiments – the Deep Underground Neutrino Experiment (DUNE) and Hyper-Kamiokande (Hyper-K) – will collect a large rate of accelerator-based neutrino-interaction events. This will provide researchers with the ability to answer remaining key questions within oscillation physics. DUNE, the physics program on which this thesis will focus, will attempt to answer the following questions:

1. What are the precise values of the Pontecorvo-Maki-Nakagawa-Sakata (PMNS) matrix mixing angles (θ_{12}, θ_{23}, θ_{13})? Specifically, is θ_{23} lower than, greater than, or even equal to 45° (known as maximal mixing)?

2. Does neutrino oscillation violate Charge-Parity symmetry (is δ_{CP} of the PMNS matrix non-zero)?

3. What is the ordering of the neutrino masses (what is the sign of Δm^2_{31})?

Great effort must be taken to reduce systematic uncertainty to a suitable level to achieve precise measurements related to the questions stated above. Necessary for this is the prototyping of DUNE's far detector with ProtoDUNE-SP. ProtoDUNE-SP serves as a test-bed for DUNE's detector components and event reconstruction, a first-attempt at calibration which will be employed at DUNE, and a source of physics measurements using its charged particle beam line which will serve as necessary inputs to DUNE's simulation. One of the particles provided by the beam line – π^+ – is important to study, as it is often found in the final state of neutrino interactions. As such, it has the ability to interfere with the reconstruction of the incident neutrino's energy or its flavor. For example, if a π^+ is absorbed by an argon nucleus nearby the primary neutrino interaction, the pion's energy could be missed in reconstruction of the neutrino's energy. Additionally, if a π^+ is produced in a neutral current ν_μ interaction it could instead undergo a charge exchange interaction, where it is converted into a π^0, nearby. The π^0 will promptly decay into two photons, which will produce showers in the detector. These showers could mimic an electron shower and could cause the ν_μ neutral current interaction to be misidentified as a ν_e charged current interaction. These errors will be accounted for in DUNE's oscillation analyses using Monte Carlo simulation of events in DUNE's far detector. However, if the rate of pion interactions are misestimated, DUNE's measurements could be biased. This thesis presents a measurement of the π^+–Ar absorption and charge exchange in order to reduce these systematic effects in DUNE's analyses.

This thesis will be organized as follows. Chapter 2 will describe the Standard Model of particle physics and how neutrinos fit (or rather *do not fit*) within this theory. It will also provide an overview of interactions of both neutrinos and charged pions on nucleons and nuclei. Chapter 3 will describe DUNE's physics program and detector design and provide motivation for this measurement. In Chapter 4, the ProtoDUNE-SP detector – the detector used for this measurement – will be discussed. Specifically, this chapter will focus on the design and components of the detector, the software used for reconstructing particle trajectories and interactions, and calibration of the detector. Chapter 5 will discuss the beam line which supplies the ProtoDUNE-SP detector with its test beam particles. Chapter 6 will describe the event selection used to characterize Monte Carlo events and data sets. Chapter 7 will discuss the strategy used to conduct this measurement including the strategy used to extract the cross section and the statistical fit used in the analysis. Chapter 8 will discuss the systematic uncertainties within the analysis. Chapter 9 will discuss validations of the statistical fit using fake data generated from Monte Carlo simulation. Finally, Chapter 10 will present the results of the measurement on real data.

CHAPTER 2

THEORY

As DUNE's physics program centers around neutrinos, this chapter will provide a description of our current understanding of these particles as part of the Standard Model of Particle Physics. Within this, an overview of the theory of neutrino–nucleus interactions will be given. The physics of charged pions will also be discussed, as these are often produced within neutrino–nucleus interactions.

2.1 The Standard Model

The Standard Model of Particle Physics represents the most up-to-date understanding of the universe at the subatomic level. It has provided immensely accurate descriptions of particle interactions (manifested as the electromagnetic and strong and weak nuclear forces) and successfully predicted the presence of multiple elementary particles. The Standard Model is rooted in the local symmetry group

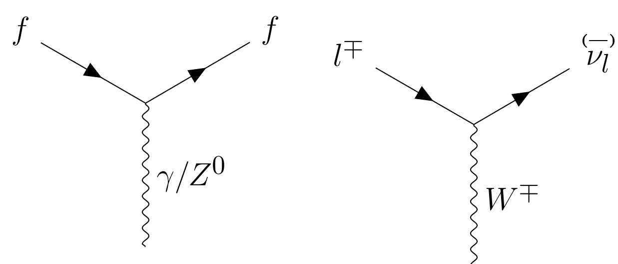
$$\mathrm{SU}(3) \times \mathrm{SU}(2)_L \times \mathrm{U}(1) \tag{2.1}$$

where the first term encompasses the strong interaction and the second and third terms give rise to the electroweak interaction. Here, the L subscript denotes this describes a "left-handed" chiral theory.

From the symmetry groups denoted in Equation 2.1, the interactions between matter and forces arise. In the development of the theory, a Lagrangian is constructed which describes a free fermion field and invariance under some local gauge transformation is enforced. If the fermion field is not invariant under that gauge transformation, an interaction with some vector field is introduced. Depending on the field and the gauge under consideration, these vector fields may also interact amongst themselves to ensure gauge invariance. The quanta of these vector fields are known as gauge bosons, and are modeled as being exchanged between interacting fermions. At this point, all gauge bosons and fermions have zero mass. However, this is wrong as known from experiment: three of the bosons and all fermions have mass. With the possible exception of neutrinos, these particles all gain mass due to the presence of the so-called Higgs field (the quantum of which is the scalar Higgs Boson, famously discovered in 2012 [1][2]). The bosons gain mass as the result of spontaneous symmetry breaking, and the fermions gain mass through coupling to the Higgs fields via Yukawa interactions [3]. The following sections describe the gauge bosons and the elementary fermions (quarks and leptons), as well as composite particles formed from quarks (hadrons).

2.1.1 Gauge Bosons

The most familiar of the gauge bosons is the photon, which is a massless, neutral particle that couples to electric charge and thus mediates the electromagnetic force. Figure 2.1a shows a Feynman diagram representing an elementary electromagnetic interaction vertex, where f is some charged fermion and the boson γ is the photon in the interaction.



(a) Electromagnetic and weak neutral current interactions.

(b) The weak charged current interaction.

Figure 2.1: Elementary interaction vertices of the electroweak interactions.

The photon arises within the Standard Model as one of the vector fields necessary to achieve SU(2) × U(1) gauge invariance (which describes the electroweak interaction). The other fields introduced by requiring this invariance are the Z^0 and $W^{\pm}$ bosons which mediate the neutral and charged current weak interactions respectively. Elementary weak interactions are also shown in Figure 2.1, where the neutral current interaction is represented in Figure

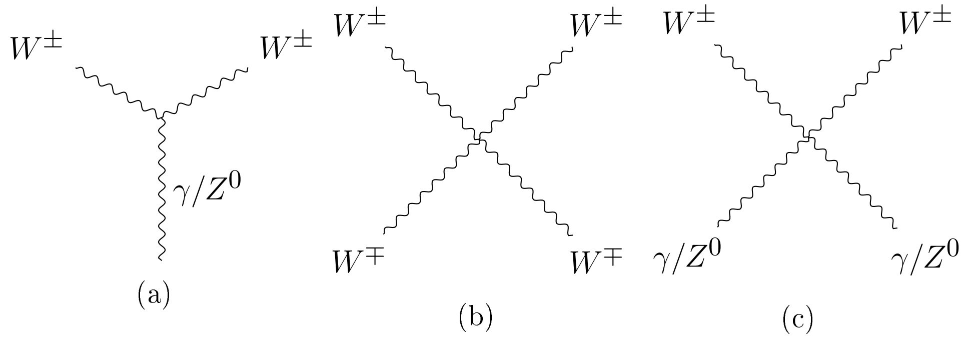


Figure 2.2: Electroweak self-interaction vertices.

2.1a with the Z^0 as the boson in the interaction and f is some fermion, and the charged current interaction is in Figure 2.1b. The $W^{\pm}$ and Z^0 bosons are the three gauge bosons mentioned above which gain their mass through the Higgs mechanism. Within electroweak theory, interactions between the W, Z, and photon also occur. These elementary interaction vertices are shown in Figure 2.2.

Finally, SU(3) invariance introduces eight gluons to facilitate the strong nuclear force. This force couples to a property known as color, which, like electric charge for the electromagnetic interaction, is common to all particles that experience the strong interaction (quarks and gluons themselves). Color differs from electric charge in that there are 3 colors (red, green, and blue) plus 3 anti-colors (anti-red, anti-green, and anti-blue) rather than just positive or negative electric charge. The gluons themselves carry color, and thus, due to color conservation, annihilate quarks of one color and create quarks of another color during interactions. In Figure 2.3a, the quarks entering and exiting the vertex are implied to have different colors. Similar to the electroweak bosons, interactions between the gluons arise as part of SU(3) invariance, giving rise to vertices such as Figures 2.3b and 2.3c.

2.1.2 Quarks

The elementary components of matter are a group of 12 fermions (and their antiparticle partners) separated into 6 quarks and 6 leptons. The quarks engage in all three forces

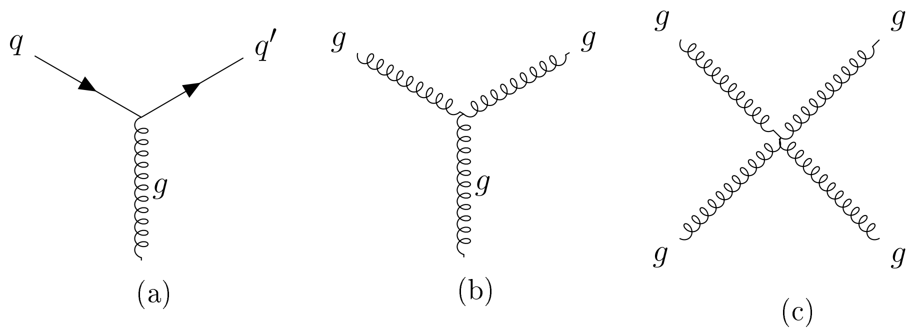


Figure 2.3: Interaction vertices of the strong force.

described previously, while the leptons feel only the electromagnetic and weak forces (and within the leptons, the neutrinos only engage in the weak forces as they are neutral).

The quarks come in six flavors: up, down, charm, strange, top, and bottom, which are separated into three generations as shown in Table 2.1. Each generation contains one quark with electric charge equal to $\frac{2}{3}e$ and the other with charge of $-\frac{1}{3}e$ (where e is the basic unit of electric charge). The first group includes u, c, and t quarks and are collectively known as "up-type" quarks, while the second group includes d, s, and b quarks and are known as the "down-type" quarks. As mentioned earlier, the quarks carry a property known as color which is similar to electric charge. In an analogy to primary colors of visible light, this comes in the form of three charges – (anti-)red, (anti-)green, or (anti-)blue. The analogy to the primary colors of light comes from the fact that the three color charges (or a color and its anticolor) add together to a net-0 or "white" color charge. A special property of color – color confinement – is that no isolated free particle can exist in a colored state. This leads to quarks forming bound states known as hadrons. Most often, these come in the form of mesons (bound states of a quark and an antiquark with one type of color and its anticolor) and baryons (bound states of three quarks each with one of the colors). Baryons and mesons will be discussed further in the next section.

In addition to the strong interaction, the quarks also take part in the weak interaction. In charged current weak interaction, the quarks transition between up-type and down-type flavors. Within the Standard Model Lagrangian, the quarks are described as two-component

Generation	Quark	Charge (e)	Mass (MeV/c^2)
I	u	2/3	2.4
	d	-1/3	4.8
II	c	2/3	1270
	s	-1/3	104
III	t	2/3	1.712 x 10^5
	b	-1/3	4200

TABLE 2.1: The quark generations along with the charge and mass of the individual quarks.

states, comprised of the up-type and down-type of each generation, as shown in Equation 2.2, which are operated upon in the weak interaction.

$$\Psi^q = \begin{pmatrix} \psi^u \\ \psi^d \end{pmatrix}, \begin{pmatrix} \psi^c \\ \psi^s \end{pmatrix}, \begin{pmatrix} \psi^t \\ \psi^b \end{pmatrix} \tag{2.2}$$

If these were eigenstates of the weak interaction, one would expect this process to transition between quarks only *within* generations (i.e. u would only transition to d, c to s, and t to b). This is not the case, and cross-generation transitions are allowed. This can represented by a set of quark states (d', s', b') that are linear combinations of the normal down-type quarks. The Cabibbo-Kobayashi-Maskawa (CKM) matrix mixes the normal down-type quarks into the special weak-eigenstate quarks as seen in Equation 2.3.

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} U_{ud} & U_{us} & U_{ub} \\ U_{cd} & U_{cs} & U_{cb} \\ U_{td} & U_{ts} & U_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} \tag{2.3}$$

2.1.3 Hadrons

As stated above, quarks cannot be observed in isolated states[1], and reside in composite particles known as hadrons. Most commonly, these are combinations of a quark and an antiquark

[1] That is, at energy scales relevant to this thesis. At high enough energies, a phase transition to a state known the quark-gluon-plasma occurs wherein quarks and gluons are not confined to hadrons [4]

Generation	Lepton	Charge (e)	Mass (MeV/c^2)
I	e	-1	.511
	ν_e	0	–
II	μ	-1	105.7
	ν_μ	0	–
III	τ	-1	1777
	ν_τ	0	–

TABLE 2.2: The lepton generations along with the charge and mass of the individual leptons. Note that the neutrinos are assumed to be massless within the Standard Model and so no masses are stated for these particles here.

(mesons) or combinations of three quarks/antiquarks (baryons). More exotic combinations such as tetraquarks (two quarks and two antiquarks) and pentaquarks (four quarks and one antiquark) have recently been discovered at the Large Hadron Collider [5].

The baryons include familiar particles like protons and neutrons (made of *uud* quarks and *udd* quarks respectively); higher energy resonances of the same sets of quarks such as Δ^+ and Δ^0; and particles including second or third generation quarks such as Λ^0, Λ_c^+, and Λ_b^0 (uds, udc, udb respectively). Many other combinations of quark flavors exist, as do similar combinations of quark flavors but with differing quantum numbers. For example, the Δ^+ and proton have the same flavors of quarks, but have total angular momentum of 3/2 and 1/2 respectively.

The mesons are similarly characterized by properties of their constituent quarks (flavor, angular momentum, etc.). The mesons include the charged and neutral pions (whose quark content is $\mathrm{u\bar{d}}$, $\mathrm{d\bar{u}}$, and $(\mathrm{u\bar{u}} - \mathrm{d\bar{d}})/\sqrt{2}$ for π^+, π^-, π^0 respectively), kaons (which include a strange quark), and various other combinations of quark flavors and angular momenta.

The pions are of particular interest within neutrino physics. They play an important role in nuclear dynamics, as they are the long-range mediator of the nuclear force according to Yukawa theory [4], and are discussed further in Section 2.4.

2.1.4 Leptons

Similar to the quarks, the 6 leptons are separated into 3 generations, each containing 1 charged lepton and its neutral partner. The charged leptons are the electron (e), muon (μ), and tau (τ), which each have a charge of -1e and masses as seen in Table 2.2. In each generation is also the neutral partner to the charged lepton: the electron neutrino (ν_e), muon neutrino (ν_μ), and tau neutrino (ν_τ). The leptons are all colorless particles, and thus do not feel the strong force. However, all left-handed leptons and right-handed antileptons engage in the weak interaction. To represent this, the leptons are given Lepton numbers: L_e, L_μ, and L_τ, which are equal to 1 (-1) for (anti-)leptons in the generation denoted by the subscript. These lepton numbers are absolutely conserved in the weak interaction.

2.2 Neutrinos: Not-so-standard Particles

2.2.1 Neutrino Oscillations

This overview of the Standard Model particles seems tidy, but there are some subtle peculiarities, especially with the neutrinos. This is hinted at in Table 2.2, where no masses are stated for the neutrinos. Within the Standard Model, the neutrinos are predicted to have no mass. However, it is now understood that at least two neutrinos in fact *do* have mass, as indicated by the presence of the process known as *neutrino oscillation*. This process was first theorized by Bruno Pontecorvo in an attempt to explain a deficit of observed electron neutrinos produced from nuclear reactions in the sun. In 1968, the Davis experiment [6] measured only about 1/3 of expected solar electron neutrinos. Pontecorvo suggested this deficit could be explained if electron neutrinos produced in the Sun transformed into muon or tau neutrinos (which the Davis experiment was unable to detect) before reaching Earth [7]. Takaaki Kajita from Super Kamiokande and Arthur B. McDonald from the Sudbury Neutrino Observatory were awarded the 2015 Nobel Prize in Physics for the discovery of oscillations [8][9].

2.2.2 Neutrino Mixing

Neutrino oscillations arise from the facts that a) neutrinos have nonzero mass and b) the flavor eigenstates are not equivalent to the mass-energy eigenstates. Similar to quark mixing and the CKM matrix, neutrino mixing is described by the unitary Pontecorvo-Maki-Nakagawa-Sakata (PMNS) matrix as shown in Equation 2.4, where $\nu_{1,2,3}$ are the mass-energy eigenstates.

$$\begin{pmatrix} \nu_e \\ \nu_\mu \\ \nu_\tau \end{pmatrix} = U_{PMNS} \begin{pmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \end{pmatrix} = \begin{pmatrix} U_{e1} & U_{e2} & U_{e3} \\ U_{\mu 1} & U_{\mu 2} & U_{\mu 3} \\ U_{\tau 1} & U_{\tau 2} & U_{\tau 3} \end{pmatrix} \begin{pmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \end{pmatrix} \tag{2.4}$$

When neutrinos are produced via the weak interaction, they are produced as definite flavor states. However, the propagation of the neutrinos is described by the time-evolution operator (equivalently the Hamiltonian operator). As such, the neutrinos travel as eigenstates of the Hamiltonian: the mass-energy states.

2.2.3 Oscillation Probability

The derivation of the oscillation probability is shown in the following example. Consider a neutrino that evolves in time, $|\nu(t)\rangle$. Suppose it begins as a muon neutrino, such that $|\nu(0)\rangle = |\nu_\mu\rangle$. In a vacuum, its evolution is described by the time-evolution operator (e^{-iHt}) as such:

$$|\nu(t)\rangle = e^{-iHt} |\nu(0)\rangle = e^{-iHt} |\nu_\mu\rangle \tag{2.5}$$

Because $|\nu_\mu\rangle$ is not an eigenstate of the Hamiltonian, this is expanded to the following:

$$\begin{aligned} |\nu(t)\rangle &= e^{-iHt} \left(U_{\mu 1} |\nu_1\rangle + U_{\mu 2} |\nu_2\rangle + U_{\mu 3} |\nu_3\rangle \right) \\ &= U_{\mu 1} e^{-iE_1 t} |\nu_1\rangle + U_{\mu 2} e^{-iE_2 t} |\nu_2\rangle + U_{\mu 3} e^{-iE_3 t} |\nu_3\rangle \end{aligned}$$

Suppose an experiment is attempting to measure the rate at which ν_μ oscillate to ν_e by detecting the ν_e[2]. The probability to detect the neutrino as ν_e at some point in time is related to the following matrix element

$$\langle \nu_e | \nu(t) \rangle = \sum_j U^*_{ej} U_{\mu j} e^{-iE_j t} \tag{2.6}$$

where the sum over j runs over the mass-energy states. The probability is then given by

$$P\left(\nu_\mu \to \nu_e\right) = |\langle \nu_e | \nu(t) \rangle|^2 = \sum_{j,k} U^*_{ej} U_{\mu j} U_{ek} U^*_{\mu k} e^{-i(E_j - E_k)t} \tag{2.7}$$

For a general pair of states a, b this probability is

$$P\left(\nu_a \to \nu_b\right) = \sum_{j,k} U^*_{bj} U_{aj} U_{bk} U^*_{ak} e^{-i(E_j - E_k)t} \tag{2.8}$$

Assuming the neutrino is ultrarelativistic, the energy can be expanded as such:

$$\begin{aligned} E_j &= \sqrt{|\vec{p}|^2 + m_j^2} = |\vec{p}|\sqrt{1 + \frac{m_j^2}{|\vec{p}|^2}} \\ &\approx |\vec{p}|\left(1 + \frac{m_j^2}{2|\vec{p}|^2}\right) = |\vec{p}| + \frac{m_j^2}{2|\vec{p}|} \\ &\approx E + \frac{m_j^2}{2E} \end{aligned}$$

Where the approximation $E \approx |\vec{p}|$ was used in the final step. The difference in the exponential term becomes

$$E_j - E_k \approx \frac{\Delta m^2_{jk}}{2E} \tag{2.9}$$

where $\Delta m^2_{jk} = m_j^2 - m_k^2$. There are three mass splittings, Δm^2_{21}, Δm^2_{32}, and Δm^2_{31}, two of which are independent[3]. Δm^2_{21} is known to be positive (this is discussed later), while the sign of Δm^2_{32} – also known as the *neutrino mass hierarchy* – is an open question in neutrino physics. This is highlighted in Figure 2.4, which shows the two possible orderings of the neutrino mass states.

[2]This is commonly known as an *electron neutrino appearance* analysis.
[3]$\Delta m^2_{31} = \Delta m^2_{32} + \Delta m^2_{21}$.

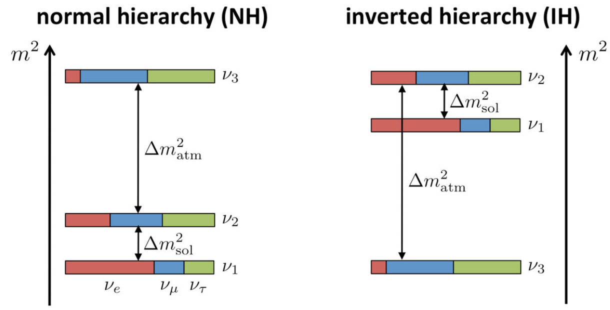


Figure 2.4: Representation of the two possible neutrino mass hierarchies. The left is the normal hierarchy with $m_3 > m_2 > m_1$. The right is the inverted hierarchy with $m_2 > m_1 > m_3$ [10].

The probability stated in 2.8 is often rewritten by splitting the real and imaginary components of the unitary PMNS matrix and approximating $t \approx L$ where L is the distance the neutrino has traveled:

$$\begin{aligned} P(\nu_a \rightarrow \nu_b) = \delta_{ab} & \\ & - 4 \sum_{j>k} \mathfrak{Re}\left[U^*_{bj} U_{aj} U_{bk} U^*_{ak}\right] \sin^2\left(\frac{\Delta m^2_{jk} L}{4E}\right) \\ & \pm 2 \sum_{j>k} \mathfrak{Im}\left[U^*_{bj} U_{aj} U_{bk} U^*_{ak}\right] \sin\left(\frac{\Delta m^2_{jk} L}{2E}\right) \end{aligned} \tag{2.10}$$

Here, the third term is positive (negative) for (anti-)neutrinos.

The PMNS matrix is parameterized by 3 mixing angles – θ_{12}, θ_{13}, and θ_{23} – and a phase factor – δ_{CP}– and is commonly factored into a product of three rotation matrices as such:

$$U_{PMNS} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix} \begin{pmatrix} c_{13} & 0 & s_{13} e^{-i\delta_{CP}} \\ 0 & 1 & 0 \\ -s_{13} e^{i\delta_{CP}} & 0 & c_{13} \end{pmatrix} \begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{2.11}$$

where c_{ij} and s_{ij} are $\cos\theta_{ij}$ and $\sin\theta_{ij}$. δ_{CP} is a phase factor that determines whether neutrino oscillations violate charge-parity (CP) symmetry. This symmetry and its violation are described in the next subsection, and the presence of CP symmetry violation within oscillations is one of the most important unanswered questions in neutrino physics.

2.2.4 CP Symmetry Violation

The violation of parity symmetry in the weak interaction was discovered in observations of cobalt 60 decays [7]. Under this symmetry, the weak interaction should behave the same under the complete reversal of coordinate system used to describe the system ($(\hat{x}, \hat{y}, \hat{z}) \rightarrow (-\hat{x}, -\hat{y}, -\hat{z})$). However, it was observed that, when the spins of cobalt atoms in a sample were aligned in a particular direction, the electrons resulting from neutron decays within the cobalt atoms (shown in Figure 2.5) came out opposite the cobalt spins. Under a parity transformation, the spin would not flip, but the direction of emission would. The electrons would then be emitted in the direction of the spins, thus violating the symmetry [7]. The parity symmetry violation in weak interactions is in fact maximal [7], leading to it being described by a left-handed chiral theory (hence the L in the subscript of 2.1). A result of the parity violation is that there are no interacting right-handed (left-handed) neutrinos (antineutrinos).

In addition to violating parity symmetry, the weak interaction also violates charge symmetry. For example, a left-handed neutrino would be transformed into a left-handed antineutrino under charge conjugation and undergo weak interactions under charge symmetry. This is not the case, however, and thus the weak interaction violates charge symmetry as well. The weak interaction in the lepton sector *does not* violate CP symmetry[4] under a simultaneous charge conjugation and parity transformation, left-handed neutrinos turn into right-handed antineutrinos which both undergo weak interactions.

Returning to the subject of neutrino oscillations, the CP-violating phase-factor δ_{CP} in the PMNS matrix (Equation 2.11) has the ability to introduce CP violation in the neutrino sector by producing an asymmetry between neutrinos and antineutrinos in the oscillation

[4]The weak interaction in the quark sector, however, does violate CP symmetry [7].

probability. If present (i.e. $\delta_{CP} \neq 0, \pi$), this oscillation asymmetry could be responsible the matter-antimatter asymmetry we observe in the universe [11]. The presence of a CP-violating value of δ_{CP} is one of the key unanswered questions in neutrino physics and is one focus of the upcoming oscillation experiments DUNE [12] and HyperK [13].

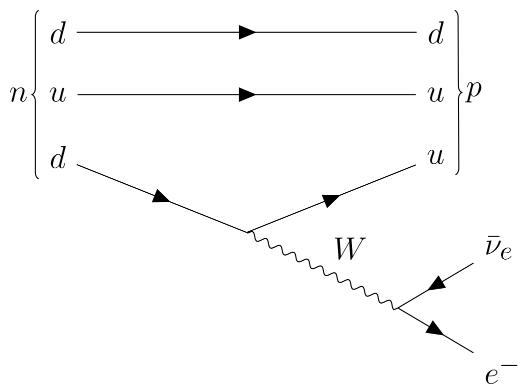


Figure 2.5: Feynman diagram of a neutron decaying to a proton.

2.2.5 MSW Effect

In matter, the oscillation probability is modified due to the presence of coherent forward elastic scattering of neutrinos by the surrounding matter. Specifically, the Charged Current process is available for ν_e, but not for $\bar{\nu}_e$, since only electrons and not positrons are present in normal matter (for example, the surrounding earth). A potential created by these processes is added to the Hamiltonian in Equation 2.5. This modifies the time evolution of the neutrino flavors, resulting in a modified oscillation probability and effective oscillation parameters in matter. This results in what is known as the MSW effect, wherein resonant behavior is exhibited in the effective mixing angles[5]. The resonance can only be present for either neutrinos or antineutrinos, depending on the ordering of the neutrino masses [14]. Thus, by measuring the asymmetry between neutrino and antineutrino oscillations, the MSW effect

[5]This is named after Mikheyev, Smirnov, and Wolfenstein. Wolfenstein first discovered that neutrinos were affected by the potential created by the surrounding matter. Mikheev and Smirnov discovered the resonant behavior.

can be exploited to determine the ordering of the neutrino masses. The sign of Δm^2_{21} was determined by analyzing solar neutrinos which are subject to the matter effect as they travel through the Sun [15]. The same principle can be used to determine the mass hierarchy (the sign of Δm^2_{32}) using accelerator-based neutrino experiments. Similar to the question of CP violation, this is an important, unanswered question in neutrino physics and will be explored by future experiments such as DUNE. A discussion of DUNE and its physics potential will be given in Chapter 3. In order to give context to that chapter, the following to sections discuss interactions of both neutrinos and pions with nucleons and nuclei.

2.3 Neutrino Interactions

Neutrino oscillation experiments rely on detecting neutrinos through identifying the particles produced by interactions on target nuclei. Through reconstructing these products, the flavor, sign (neutrino vs. antineutrino), and energy of the incident neutrino are inferred. This section provides a description of neutrino–nucleus interactions. In most types of neutrino–nucleus interactions[6], the neutrino interacts primarily with a constituent nucleon. These broadly fall into two categories: Charged Current (CC) and Neutral Current (NC). The CC interaction occurs with the exchange of a $W^{\pm}$ as shown by the vertex in Figure 2.1b while the NC interaction occurs with the exchange of a Z^0 shown in Figure 2.1a.

2.3.1 Quasielastic Scattering

The first major CC interaction is CC Quasielastic (CCQE). This interaction occurs in the forms given in Equation 2.12, and is represented by the Feynman diagrams shown in Figures 2.6 and 2.7 for neutrinos and antineutrinos respectively.

$$\begin{aligned} \nu_l + n &\rightarrow l^- + p \\ \bar{\nu}_l + p &\rightarrow l^+ + n \end{aligned} \tag{2.12}$$

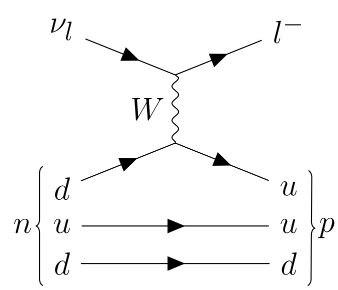


Figure 2.6: Feynman diagram for the neutrino CCQE interaction.

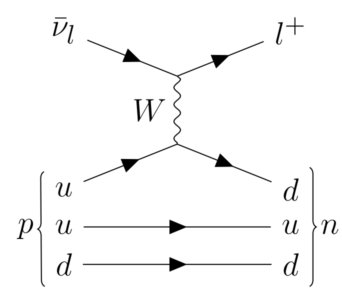


Figure 2.7: Feynman diagram for the antineutrino CCQE interaction.

The cross section for these processes depend on the vector form factors (F_1 and F_2) and axial form factors (F_A and F_P) of the CC interaction, which themselves depend only on the four-momentum transfer Q^2 between the neutrino and nucleon in the interaction [16]. F_1 and F_2 are related to the electromagnetic form factors of the nucleons, which are extracted from electron scattering data [14]. The pseudoscalar form factor F_P can either be neglected through approximation [14] or related to F_A [16], such that F_A is the only unknown portion of the CCQE cross section. The exact shape of the axial form factor is not described by theory, and a dipole approximation is generally used as shown in Equation 2.13.

$$F_A(Q^2) = \frac{g_A}{\left(1 + \frac{Q^2}{M_A}\right)^2} \tag{2.13}$$

Here, g_A is the axial-vector coupling constant of the weak charged current, which is obtained from neutron decay data, and M_A is the axial mass, which can be obtained from fitting to neutrino scattering data.

[6] Other than coherent neutrino–nucleus scattering.

2.3.2 Resonant Pion Production

The next set of major neutrino–nucleon interactions are resonant pion production. These interactions occur through both CC and NC channels, and result in a pion exiting the interaction along with the nucleon and final state lepton. In these interactions, the neutrino interacts inelastically with a nucleon and excites it into some resonance (i.e. a nucleon resonance or Δ resonance). The forms of the CC interactions (without specifying the intermediate resonance) are given in 2.14, while the NC interactions are given in 2.15. Multiple resonances contribute to the amplitudes of these processes, but at lower energies, the $\Delta(1232)$ resonance dominates [17]. The most commonly used model to describe the Δ resonance interaction is the Rein-Sehgal model [16][17]. The NC interaction resulting in a π^0 is important as a background to CCQE ν_e events, where the γ showers from the π^0 decay can be mistaken as an e shower during event reconstruction, and the rate of this background is important to constrain in ν_e appearance measurements.

$$
\begin{aligned}
\nu_l + p &\to l^- + p + \pi^+, & \bar{\nu}_l + p &\to l^+ + p + \pi^- \\
\nu_l + n &\to l^- + p + \pi^0, & \bar{\nu}_l + p &\to l^+ + n + \pi^0 \\
\nu_l + n &\to l^- + n + \pi^+, & \bar{\nu}_l + n &\to l^+ + n + \pi^-
\end{aligned}
\tag{2.14}
$$

$$
\begin{aligned}
\nu_l + p &\to \nu_l + p + \pi^0, & \bar{\nu}_l + p &\to \bar{\nu}_l + p + \pi^0 \\
\nu_l + p &\to \nu_l + n + \pi^+, & \bar{\nu}_l + p &\to \bar{\nu}_l + n + \pi^+ \\
\nu_l + n &\to \nu_l + n + \pi^0, & \bar{\nu}_l + n &\to \bar{\nu}_l + n + \pi^0 \\
\nu_l + n &\to \nu_l + p + \pi^-, & \bar{\nu}_l + n &\to \bar{\nu}_l + n + \pi^-
\end{aligned}
\tag{2.15}
$$

2.3.3 Deep Inelastic Scattering

Another important set of processes in oscillation experiments are the CC and NC Deep Inelastic Scattering (DIS) processes. In these, the neutrino and the intermediary gauge boson it exchanges with the nucleon are energetic enough to resolve the individual quark constituents of the nucleon. The nucleon is broken apart and a hadron shower is produced as a result of quark confinement. The form of these interactions is given in 2.16. Here, N is either nucleon and X is a set of hadrons.

$$\begin{aligned} &\nu_l + N \rightarrow l^- + X, && \bar{\nu}_l + N \rightarrow l^+ + X \\ &\nu_l + N \rightarrow \nu_l + X, && \bar{\nu}_l + N \rightarrow \bar{\nu}_l + X \end{aligned} \tag{2.16}$$

These processes dominate the total cross section at high energies ($E_\nu \gtrapprox 20$ GeV) [16]. The inclusive DIS cross section[7] is described by functions representing the structure of the nucleons known as parton distribution functions (PDFs) [16][17][14].

There exists a transition region between the resonance and DIS regimes called the Shallow Inelastic Scattering (SIS) region [16][18]. This region is not as well understood as the DIS-dominated region [16][18], and different simulation frameworks take a variety of approach to modeling this transition [16][18].

2.3.4 Neutrino–Nucleus Scattering

Neutrino oscillation experiments use nuclear targets for their detection medium. This complicates the relatively simple picture of neutrino–nucleon scatter in a few key ways. Firstly, coherent scattering become possible, wherein each component of the nucleus contributes to the interaction amplitude *coherently* and the nucleus is left in its ground state. An important type of coherent scattering is coherent pion production as shown in 2.17 (top: CC, bottom: NC). The NC process is an important background for ν_e appearance channels as the γ showers from the π^0 decay can mimic an e shower.

$$\begin{aligned} &\nu_l + A \rightarrow l^- + \pi^+ + A, && \bar{\nu}_l + A \rightarrow l^+ + \pi^- + A \\ &\nu_l + A \rightarrow \nu_l + \pi^0 + A, && \bar{\nu}_l + A \rightarrow \bar{\nu}_l + \pi^0 + A \end{aligned} \tag{2.17}$$

[7] Full expressions found in [16][17][14]

Additionally, the presence of the nuclear medium complicates the behavior of both the initial and final state. For CCQE interactions most models assume the Impulse Approximation, in which the neutrino scatters elastically off nucleons in the nuclear ground state, followed by quasifree ejection of the nucleons from the nucleus. The nuclear state (i.e. the kinematic distribution of nucleons within the nucleus) is commonly described by a Relativistic Fermi Gas model, where the nucleons are free particles subject to Fermi motion and populate states according the Pauli exclusion principle. Despite the fact it is commonly used, it poorly describes electron scattering data. Other models and approximations for the initial nuclear state have been utilized in recent years to overcome this limitation. An important development toward improved modeling of the initial nuclear state comes in the form of the inclusion of nucleon–nucleon correlations and meson exchange currents (MEC). These contribute to multinucleon excitation, and raise the cross section of events that produce no final state pion. In addition to these initial state effects, DIS interactions are further complicated through modifications of the nucleon PDFs by the nuclear medium [17][18].

Finally, the presence of Final State Interactions (FSI) can modify the observable products of the primary interaction as they attempt to exit the nucleus. The hadronic products of each interaction (including pions in resonance interactions and hadron showers from DIS) can possibly reinteract as they travel through the nucleus. A common model for this is an intranuclear cascade, wherein the interaction products step through the medium and can undergo an interaction with the surrounding nucleons. These resulting particles are then added into the cascade process and can then go on to interact again. This goes on until all active particles exit the nucleus or are absorbed back into the nucleus. This results in a modified set of observable particles (i.e. with missing or additional particles, and/or with smeared kinematics). For resonance interactions, further complications arise from the fact that the surrounding nuclear medium modifies the properties of the intermittent Δ resonance. Processes such as those listed in 2.18 increase the width of the Δ within the nucleus [18].

$$\begin{aligned} \Delta + N &\to N + N \\ \Delta + N + N &\to N + N + N \\ \Delta + N &\to \pi + N + N \end{aligned} \tag{2.18}$$

2.4 Pions

The last section illustrated the complexities in neutrino scattering created by the nuclear environment. The pion is often produced in neutrino interactions, and as will be seen in Chapter 3, must be accounted for in neutrino oscillation analyses at DUNE. This section serves to describe the pion's role within the nucleus, and its interactions with nuclei.

The pion is the lightest meson, and is a spin-0, isospin-1 boson. It has three charge states (as evident by its isospin). These are described in Table 2.3. Yukawa predicted that a point-particle similar to the pion mediated the force between point-like nucleons within nuclei [7][4]. In fact, at ranges greater than 0.7 fm, intranuclear interactions are well described by this pion exchange picture [19]. At greater than 2 fm, one-pion exchange dominates, while two-pion exchange contributions become equal or greater than one-pion exchanges between 0.8 and 2 fm [20]. Below this, the point-like approximation of the pion and nucleons breaks down, and the quark-gluon degrees of freedom become important [4][20]. In the point-like approximation, the nucleon acts as a source of the pion field, resulting in a field of the form given by Equation 2.19. Here, τ_3 and σ are Pauli isospin and spin operators and f is a coupling constant. This has a striking similarity to the potential from a magnetic dipole, as shown in Equation 2.20 [20].

$$\phi_N(\vec{x}) = -\frac{f}{m_\pi}\tau_3\sigma \cdot \nabla_x \frac{e^{-m_\pi|\vec{x}-\vec{r}|}}{4\pi|\vec{x}-\vec{r}|} \tag{2.19}$$

$$\phi_M(\vec{x}) = -\mu \cdot \nabla_x \frac{1}{4\pi|\vec{x}-\vec{r}|} \tag{2.20}$$

Pion	Quark Content	Charge (e)	Mass (MeV/c^2)	I_3
π^+	$u\bar{d}$	+1	139.57	+1
π^0	$\frac{u\bar{u}+d\bar{d}}{\sqrt{2}}$	0	134.98	0
π^-	$d\bar{u}$	-1	139.57	-1

TABLE 2.3: The pion along with their quark coontent, charge, mass, and the third component of its isospin.

2.4.1 Pion–Nucleon Scattering

It is important to consider pion–nucleon scattering as a basis for pion–nucleus scattering. This interaction is purely elastic up to the threshold for the $\pi + N \rightarrow \pi + \pi + N$ process at $T_\pi \approx 170$ MeV [20]. When viewed in a partial wave analysis, the s- and p-wave (angular momentum $l = 0$, 1 respectively) contributions to the interaction dominate when compared to the d- and f-waves ($l = 2$, 3) [20]. Furthermore, the s-wave interactions are small compared to the p-wave interactions [21]. The dominant effect in the p-wave component, and thus the overall interaction, is the resonance appearing around pion kinetic energy $T_\pi \approx 180$ MeV. This is due to the coupling to the $\Delta(1232)$ spin 3/2, isospin 3/2 resonance [20][19][21]. This resonance can be seen in Figure 2.8.

Two processes are of particular interest for this thesis: single charge exchange and absorption. For absorption, additional particles must be involved in the process in order to conserve energy and momentum. As such, the absorption of pions by singular free nucleons is forbidden[8]. This will be discussed in Sections 2.4.2 and 2.4.3. Single charge exchange is free from this requirement and, for incident $\pi^\pm$, takes the forms in Equation 2.21. It too will be discussed in Section 2.4.3.

[8]This is approximately true for bound nucleons as well, as the interaction is suppressed due to the momentum that must be supplied by the nucleon, which is much larger than the Fermi momentum [21].

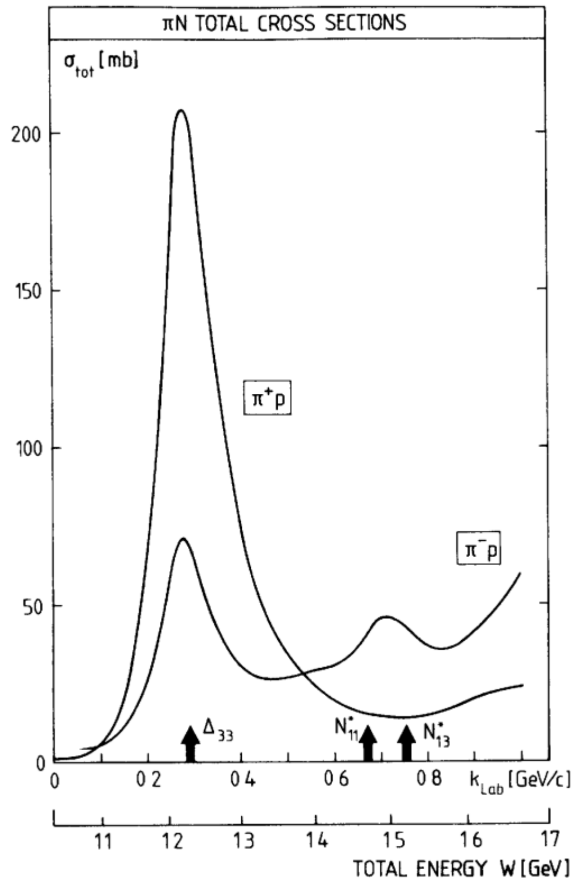


Figure 2.8: $\pi^{\pm} - p$ cross sections as functions of pion lab momentum k_{lab} and center-of-mass energy W [20].

$$\pi^{+} + n \rightarrow \pi^{0} + p \qquad (2.21)$$
$$\pi^{-} + p \rightarrow \pi^{0} + n$$

2.4.2 Pion–Deuteron Scattering

The scattering of pions by deuterons (a proton–neutron bound state) is the simplest extension of pion–nucleon scattering to multiple-body systems. The total cross section is comprised of contributions from elastic scattering, inelastic scattering (wherein the deuteron is broken up), absorption, and pion production at higher energies. The interaction can be well approximated by the sum of the $\pi - p$ and $\pi - n$ cross sections, as shown in Figure 2.9. However, the observed cross section is lower in the resonance region due to a broadening of the resonance caused by the motion of the nucleons within the deuteron as well as a shadowing effect of one nucleon by the other [20]. These effects are indicative of the complications that arise in the nuclear environment.

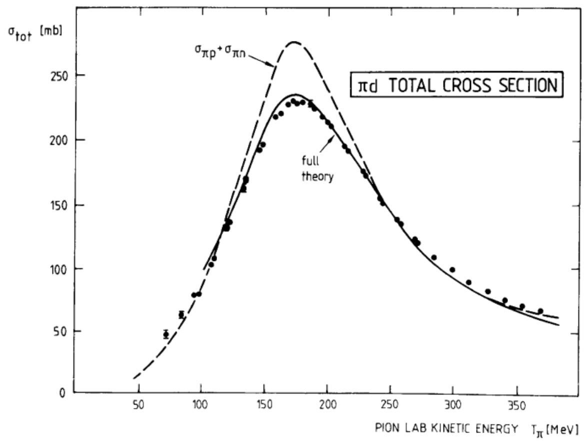


Figure 2.9: The $\pi - d$ total cross section. Black points are data, the dashed line is the sum of the $\pi - p$ and $\pi - n$ cross sections, and the solid line includes effects from nucleon motion and shadowing as described in the text [20].

As mentioned before, pion absorption on singular nucleons is forbidden, and multiple nucleons must contribute to the absorption process. As such, the $\pi - d$ absorption process is prototypical of this interaction in nuclei. It has been determined experimentally that two-nucleon absorption in nuclei is dominated by absorption on deuteron-like pairs in the $\Delta(1232)$ resonance region [19]. The "rescattering" model is an elementary model that gives a qualitative understanding of the physics of $\pi - d$ absorption. In this, a scatter on one nucleon is followed by absorption on the other. The leading terms in this theory are again the s- and p-wave contributions. The so-called s-wave rescattering consists of a the pion undergoing an s-wave scatter by the first nucleon, followed by p-wave absorption on the second nucleon. In p-wave rescattering, the pion strikes the first nucleon creating an intermediate Δ state. This Δ then interacts with the second nucleon creating the final dual-nucleon state [20].

Quantitatively, however, this description falls short, and a full three-body framework that treats $\pi - d$ absorption on equal footing with other $\pi - d$ scattering processes has been more successful in predicting experimental results [20].

2.4.3 Pion–Nucleus Scattering

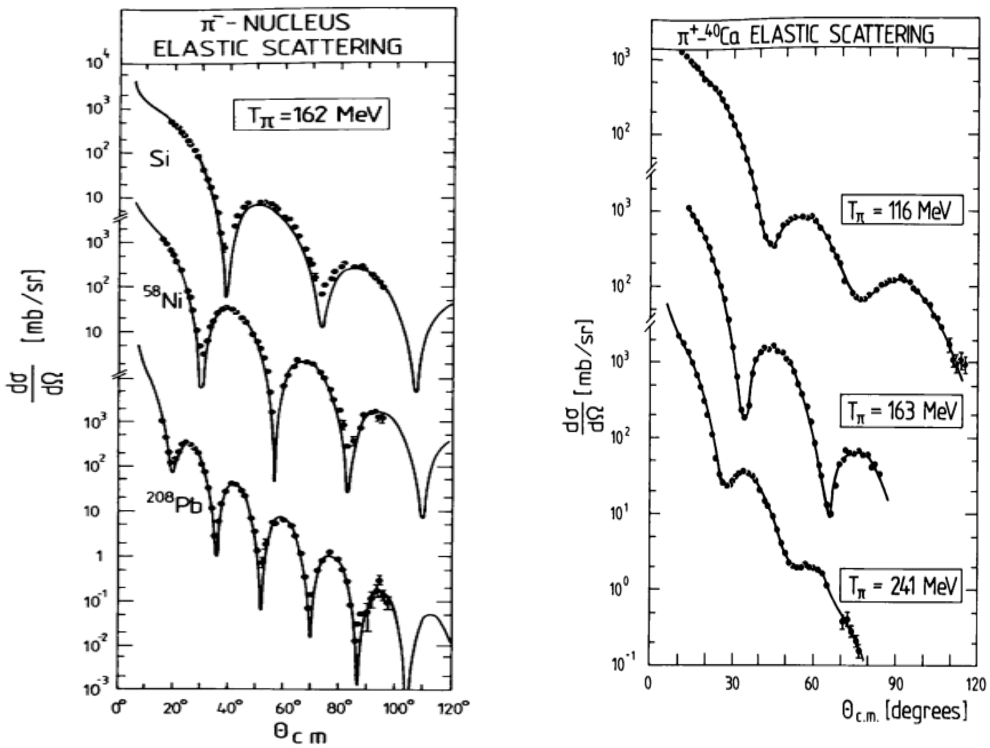
Similar to the extension of pion scattering from single-nucleon targets to the deuteron, the extension to the nucleus is complicated by the influence of additional nucleons on the interaction. The same basic processes (elastic and quasielastic scattering, single charge exchange, absorption on more than one nucleon) are present, and the $\Delta(1232)$ resonance still plays an important role. However, the dynamics are enriched by the nuclear environment. Recalling the dipole-like interaction of the pion with the nucleon, the nuclear environment acts as a polarizable and refractive medium for the pion, in analogy with the scattering of light by electromagnetic dipoles [20]. The Δ resonance is also influenced by the medium; its peak shifts lower and its width broadens as the nuclear mass increases [20][19].

2.4.3.1 Elastic Scattering

The analogy to light propagation is evident in elastic scattering off nuclei, whereby the nucleons diffract the incoming pion wave similar to light by atoms in an optically diffractive medium [20][19]. Within the resonance region, the imaginary part of the $\pi - N$ scattering amplitude becomes large, producing deep minima in the angular distribution of the scatters. These minima are still present, but become more shallow outside of the resonance region where the real part of the scattering amplitude becomes larger [20]. This can be seen in Figure 2.10.

2.4.3.2 Inelastic Scattering

As with elastic scattering, inelastic scattering is an extension of the $\pi - N$ interaction to the nuclear environment. The difference lies in the transition of the nucleus to excited states.



(a) (b)

Figure 2.10: Elastic pion scattering cross sections. a) Diffractive patterns are present within the resonance region. b) Diffractive patterns are suppressed outside of the resonance region [20].

Two broad regimes exist for this: 1) a low energy transfer region wherein the nucleus is excited to discrete states 2) a high energy transfer region in which the quasifree $\pi - N$ interaction dominates [20] and the struck nucleon is knocked out to continuum states [19]. This quasifree process is the leading contribution to the inelastic cross section [20], and is subject to in-medium effects that shift the location of the quasifree peak as seen in Figure 2.11.

2.4.3.3 Absorption

As previously stated, $\pi - d$ absorption is prototypical of absorption within nuclei, due to the suppression of single nucleon absorption by energy and momentum conservation. For nuclei with $A > 2$, absorption quasideuteron pairs ($I = 0$, pn pairs) remains the leading

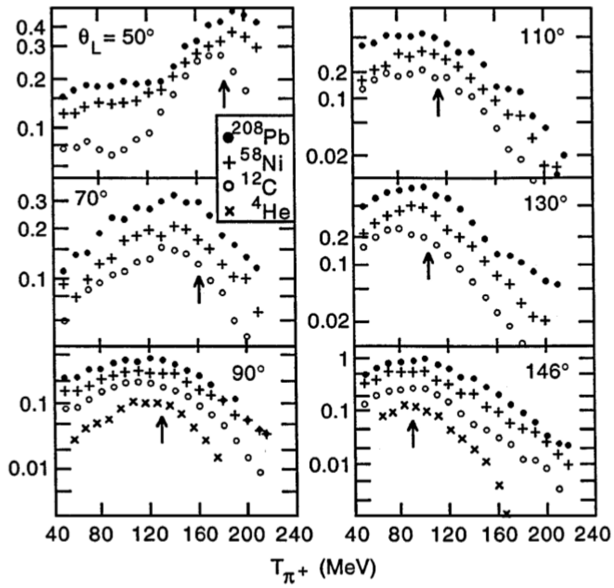


Figure 2.11: A collection of data representing spectra of pions relative to outgoing kinetic energy T_π and lab angles θ_L for inclusive inelastic scattering on various nuclei. The arrows represent quasifree peaks assuming no in-medium effects applied to the $\pi - N$ scattering amplitude [19].

contribution to 2–nucleon absorption [19]. However, the presence of more nucleons influences the interaction in a few ways. Firstly, direct interactions on multi-nucleon (N $>$ 2) groups contribute to the cross section. The absorption of π by 3 nucleons becomes significant even for ^{3}He targets [19]. These direct absorption processes provide insight into correlations between nucleons, and experiments such as LADS have detailed measurements of π–nuclear absorption relative to outgoing nucleon multiplicities and kinematics [22]. Additionally, multiple nucleons become involved in the absorption via Initial State Interactions (ISI – wherein the pion undergoes a quasifree scatter off a single nucleon and is absorbed later) and Final State Interactions (FSI – wherein the pion is absorbed by a set of nucleons and these nucleons go on to interact with other nucleons within the nucleus). Data has shown that the average number of nucleons substantially involved in the absorption process appears to be considerable and at least somewhat A-dependent [19]. The contribution of ISI to multinucleon absorption of pions in nuclei appears constant relative to A [19], while FSI contributes more as A increases [20]. Additionally, the energy spectra of exiting nucleons is similar to multistep, cascade processes [20].

2.4.3.4 Single and Double Charge Exchange

π–nuclear single charge exchange (the processes shown in Equation 2.21) is an extension of the quasifree inelastic interaction discussed above, but with a $\Delta I_3 = \pm 1$ transition of the nuclear isospin [20]. Charge exchange makes up roughly 10% of the π–nuclear reaction[9] in the resonance region [20]. Like the other interactions described so far, the process is complicated by the nuclear medium. Mainly, double charge exchange ($\pi^{\pm} + A \rightarrow \pi^{\mp} + A'$) can occur. Here, two subsequent single charge exchange interactions occur. The outgoing π^0 from the initial interaction exchanges charge with another nucleon, resulting in a $\Delta I_3 = \pm 2$ isospin transition and a flip of the pion's charge. Double charge exchange is a relatively rare process, with a cross section roughly 10% of the single charge exchange cross section [20].

2.4.4 Outlook

As highlighted by this section, pion–nucleus interactions contain complex dynamics. Particularly in heavy nuclear environments, these interactions become quite complicated. DUNE's nuclear target, argon, is no exception to this, and so care must be taken to model these processes within DUNE's experimental simulation. Currently, limited data exists for pion–Ar interactions, especially the exclusive interactions like absorption (a single measurement [22]) and charge exchange (no measurements). This thesis provides data for these interactions that can be used to validate and improve the pion interaction model used by DUNE.

[9] Absorption + Inelastic + Single Charge Exchange + Double Charge Exchange interactions

CHAPTER 3

THE DEEP UNDERGROUND NEUTRINO EXPERIMENT

As discussed in the previous chapter, there are a few questions in neutrino physics that remain unanswered. The Deep Underground Neutrino Experiment (DUNE) seeks to answer these questions once it begins to take data later this decade. Many experiments have already made enormous progress in getting us to the point where the answers to these questions are in reach. These experiments focused on neutrinos produced from several sources: neutrinos produced in nuclear reactors, neutrinos produced by cosmic ray interactions in the atmosphere, neutrinos produced within the Sun, and neutrinos produced from particle accelerators. A global fit to the data from these experiments has been performed to provide current estimates of the oscillation parameters [23]. These are presented in Table 3.1. Note that only the normal mass ordering is given here, as the analysis found the inverted ordering was disfavored at $\Delta\chi^2 = 4.7$ [23].

Though systematic uncertainties in previous-generation experiments have required great effort to overcome, the experiments were limited primarily by statistical uncertainty. DUNE

Parameter	Best-Fit $\pm 1\sigma$	3σ Range
$\sin^2\theta_{12}$	$0.310^{+0.013}_{-0.012}$	$0.275 - 0.350$
$\theta_{12}[^\circ]$	$33.82^{+0.78}_{-0.76}$	$31.61 - 36.27$
$\sin^2\theta_{23}$	$0.582^{+0.015}_{-0.019}$	$0.428 - 0.624$
$\theta_{23}[^\circ]$	$49.7^{+0.9}_{-1.1}$	$40.9 - 52.2$
$\sin^2\theta_{13}$	$0.022240^{+6.5\times10^{-4}}_{-6.6\times10^{-4}}$	$0.02044 - 0.02437$
$\theta_{13}[^\circ]$	$8.61^{+0.12}_{-0.13}$	$8.22 - 8.98$
$\delta_{CP}[^\circ]$	217^{+40}_{-28}	$135 - 366$
$\Delta m^2_{21}[10^{-5}\text{eV}^2]$	$7.39^{+0.21}_{-0.20}$	$6.79 - 8.01$
$\Delta m^2_{31}[10^{-3}\text{eV}^2]$	$2.525^{+0.033}_{-0.031}$	$2.431 - 2.622$

TABLE 3.1: Oscillation parameters as determined by the fit to global data in Reference [23]. Only the normal ordering of the mass hierarchy is shown here.

is a next-generation long baseline accelerator-based neutrino oscillation experiment, and will collect enough neutrino events to become limited primarily by systematic uncertainties. This chapter provides the motivation for the results of this thesis which will be used to meet DUNE's stringent systematic uncertainty requirement.

3.1 DUNE's Physics Program

The goals of DUNE's accelerator-based oscillation analyses will be to determine whether neutrino oscillations violate CP-symmetry, determine the neutrino mass hierarchy, and to determine precise values of the oscillation parameters. The DUNE Far Detector Technical Design Report [24] presents sensitivity studies which show DUNE's ability to achieve these goals. In these studies, simultaneous fits to $\nu_\mu \rightarrow \nu_\mu$, $\bar{\nu}_\mu \rightarrow \bar{\nu}_\mu$, $\nu_\mu \rightarrow \nu_e$, and $\bar{\nu}_\mu \rightarrow \bar{\nu}_e$ far detector samples were performed, with near detector samples included in order to introduce flux and cross section constraints. $\sin^2 2\theta_{13}$, θ_{12}, and Δm^2_{12} were all constrained with uncertainties derived from those shown in Table 3.1, while $\sin^2\theta_{23}$, Δm^2_{32}, and δ_{CP} were freely varied. More details on the fits can be found in [24].

The sensitivity studies show promise in DUNE's physics program. For 50% of true δ_{CP} values, DUNE can determine the presence of CP violation at the 5σ level after 10 years of its nominal run plan. If $\delta_{CP} = -\pi/2$ (which provides a maximal CP-violating effect), CP violation can be discovered after only seven years. For any value of δ_{CP}, the mass hierarchy can be determined after only two to three years. This reduces to only about one year if $\delta_{CP} = -\pi/2$. After about fifteen years of the nominal run plan, the resolution on the measurement of δ_{CP} approaches 5° for CP-conserving values and 15° for CP-violating values. After high exposure, the measurement of $\sin^2 2\theta_{13}$ approaches the precision of reactor experiments (which currently provide the main constraints on that angle), and the simultaneous measurement of all oscillation parameters without external constraint becomes possible.

3.2 The DUNE Detectors

DUNE seeks to achieve these goals as a long-baseline oscillation experiment, and, as such, is comprised of two sets of detectors: its far detector (FD) and near detector (ND) complexes. This is shown in Figure 3.1, which gives an overview of DUNE's facilities including the neutrino beam facility and the near detector complex located at Fermilab (Batavia, IL) and the far detector complex at Sanford Underground Research Facility (Lead, SD).

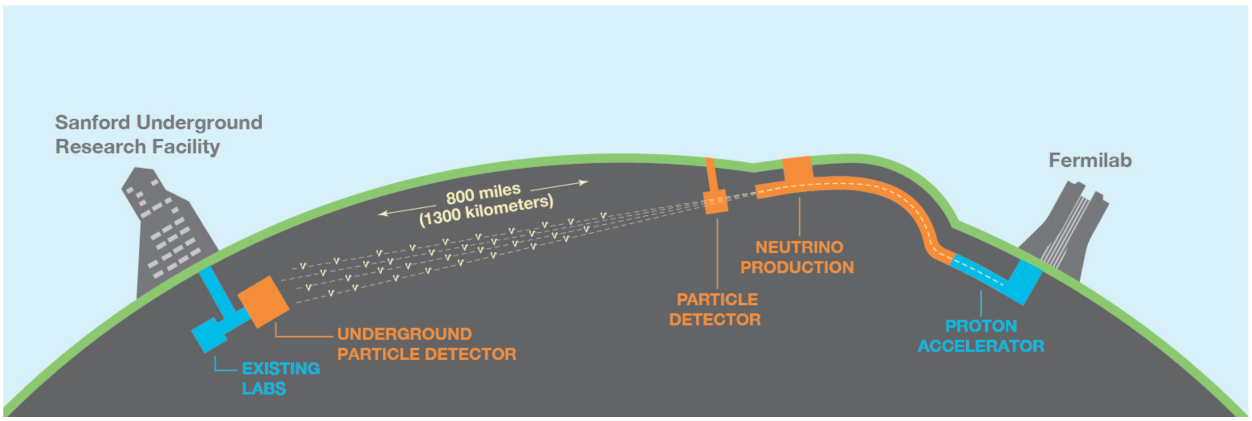


Figure 3.1: Overview of the future DUNE experiment. Toward the right is the neutrino beam facility and the near detector complex at Fermilab in Batavia, IL. Toward the left is far detector complex 1300km away at Sanfurd Underground Research Facility in Lead, SD [25].

The FD complex seeks to measure the number and flavor of neutrinos after they have had a chance to oscillate after traveling some distance. By measuring the rates at the FD, the oscillation probabilities (and more specifically, the parameters describing these probabilities) are probed directly. The ND, on the other hand, provides constraints on flux and neutrino cross section uncertainties within the models used for the oscillation analyses.

The planned DUNE FD will be comprised of four modules. Two of the modules (including the first to be installed) will be 10kt active volume single-phase (SP) Liquid Argon Time Projection Chambers (LArTPCs). This detector technology is the same as ProtoDUNE-SP and will be explained in detail in Chapter 4. One other module will be a dual-phase LArTPC (which is slightly different to the SP technology, but will not be explored here), while the final module's design is still to be determined. One common trait between all four modules is that their sensitive volumes will be 10kt of liquid argon. This argon will serve as both the target and detection medium for DUNE's neutrino beam.

The DUNE ND complex will be comprised of multiple detector subsystems. Included in these subsystems is a set of small, modular LArTPCs known as ArgonCube. It is necessary to have a portion of the near detector's target be argon in order to cancel neutrino interaction model uncertainties between the near and far detectors. This part of the detector will also be allowed to move lateral to the incident neutrino beam. Because the far detector is located at an angle of 0° with respect to the beam direction (it is "on-axis"), gathering data "off-axis" provides independent measurements of the neutrino beam. This off-axis data reduces systematic uncertainties surrounding the neutrino beam model. Other subsystems in the ND complex include a gaseous argon TPC downstream of the LArTPC portion (to help measure muons which punch through the back of the LArTPC), a fine-grained plastic scintillator detector (which remains on-axis to monitor the stability of the beam), and an electromagnetic calorimeter surrounding the previous two subsystems (which will assist in measuring all of the final state energy within the neutrino interactions).

3.3 The Role of Pion Interaction Systematic Uncertainties

To precisely measure the oscillation parameters, DUNE will attempt to discern the flavor and energy of the neutrinos interacting within the detector. Equation 2.10 (repeated here), shows the importance of successfully determining these quantities. Misidentification of the flavor will of course change the overall interaction rates of the various neutrino flavors, thus the extracted oscillation probability. Misestimation of the energy will change where in the energy distribution of interactions an event lies, thus distorting the energy spectrum of events, and further distorting the apparent oscillation probability.

$$
\begin{aligned}
P\left(\nu_a \rightarrow \nu_b\right) = \delta_{ab} \\
& - 4 \sum_{j>k} \mathfrak{Re}\left[U_{bj}^{*} U_{aj} U_{bk} U_{ak}^{*}\right] \sin^2\left(\frac{\Delta m_{jk}^2 L}{4E}\right) \\
& \pm 2 \sum_{j>k} \mathfrak{Im}\left[U_{bj}^{*} U_{aj} U_{bk} U_{ak}^{*}\right] \sin\left(\frac{\Delta m_{jk}^2 L}{2E}\right)
\end{aligned}
\tag{2.10}
$$

Both of these quantities are inferred from the final state particles resulting from the interaction. Figure 3.2 shows an example ν_μ interaction with multiple hadrons in the final state highlighting how complicated the final state of the interaction can be.

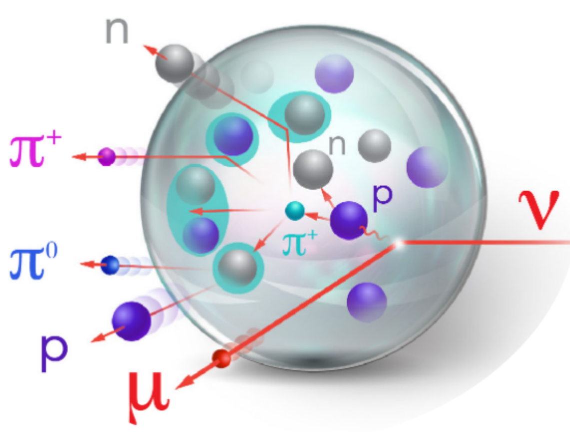


Figure 3.2: Cartoon of a ν_μ interaction with multiple hadrons in the final state [26].

For determining the flavor, reconstruction software attempts to identify the outgoing leptons from CC interactions ($\mu^\pm$ and $e^\pm$ from muon and electron neutrinos respectively). Specifically for ν_e CC events, the e^- will produce an electromagnetic shower at the interaction vertex. A background to this interaction is a ν_μ NC interaction with a π^+ in the final state. This π^+ can potentially strike a nearby Ar nucleus and create a π^0 in a charge exchange interaction. This π^0 will promptly decay into two photons which will shower similarly to the e^-. This could cause this event to be wrongly selected as a ν_e CC event. Corrections for this type of background is taken from simulation, and any uncertainty on the rate of π^+-Ar charge exchange interactions will translate to an uncertainty on the true number of ν_e events (and thus limit the precision of the oscillation measurements).

Similarly, smearing between true and reconstructed neutrino energy will be influenced on the modeling of π^+ interactions. In DUNE, the neutrino energy is estimated by the energy of the final state particles using a calorimetric energy reconstruction given in Equation 3.1. Here, E^l is the energy of the outgoing lepton, T_i^{Nucleon} is the kinetic energy of any final state protons or neutrons, and E_i^π is the total energy of pions in the final state[1]. The rest mass of any pion must be included since some of the incident neutrino energy must be used to produce the pion. A π^+ in the final state could undergo an absorption interaction on a nearby nucleus and produce a proton. The reconstruction software could fail to identify that there was a pion in the final state and the rest mass of the pion could be lost in Equation 3.1. Again, simulation is used to account for this type of effect (and similar other effects), and any uncertainty in the rate of π^+-Ar absorption limits the resolution of oscillation measurements.

$$E_{\text{Reco}}^{\nu} = E^l + \sum_i T_i^{\text{Nucleon}} + \sum_i E_i^{\pi} \tag{3.1}$$

These two examples are not an exhaustive list of regions where uncertainty on the rates of these interactions will add to DUNE's total systematic uncertainty. Rather, they are illustrative of the goal of the analysis presented in this thesis. Measurements of these interaction cross sections will provide constraints within DUNE's oscillation analyses, and will reduce DUNE's systematic uncertainty. This is an important task, as DUNE's systematic uncertainty budget is limited to 2% in order to achieve the physics goals laid out in this chapter [24]. An example of how pion scattering data can be used for the benefit of neutrino experiments is given by T2K's use of world π^+ scattering data to constrain the nuclear model used within their neutrino interaction simulation [27].

[1]Other particles such as kaons have been ignored for this example, but, in general, could be present in the final state.

CHAPTER 4

THE PROTODUNE-SP DETECTOR

Currently, the single-phase ProtoDUNE detector (ProtoDUNE-SP) is the world's largest active Liquid Argon Time Projection Chamber (LArTPC). This detector is designed to be a prototype of DUNE's single phase far detector, located in CERN's North Area. Detector installation and integration began in 2017 and finished Summer 2018. This was followed by a commissioning phase (including its charged particle beam line commissioning) in the late Summer & early Fall of 2018. After commissioning, cosmic ray data and beam line data was taken up to the CERN long shut down[1]. Since then, cosmic ray data taking has been ongoing.

Section 4.1 describes the general operation principles of LArTPCs. Section 4.2 describes the specific design of ProtoDUNE-SP. Section 4.3 provides a description of the characterization of data taken by the TPC. Section 4.4 describes the reconstruction of events in the TPC. Section 4.5 highlights the calibration of the detector. Finally, Section 4.6 describes the Monte Carlo simulation of events within the detector. The beam line will be described separately in Chapter 5.

4.1 LArTPC Principles

The detection principles of LArTPCs are based on the detection of ionization electrons and scintillation light produced by charged particles passing through the liquid argon (LAr). The argon sits between a set of anode wires and a cathode, which create a (nominally) uniform electric field. The ionization electrons drift along the electric field toward the anode wires. These wires are instrumented with electronics and detect signals produced by the drifting ionization. A configuration can be achieved such that several planes of wires can

[1]During this time, the Super Proton Synchrotron, from which the ProtoDUNE-SP beam line originates, was shut down to allow for upgrades.

measure the ionization. In such a configuration, two wire planes sit in front of a third. The electric field lines terminate on the third plane, meaning the drifting electrons ultimately deposit onto this wire. Thus this plane is called the "collection plane." Before collection, the electrons drift past the two other "induction planes." Bipolar signals are induced on these planes as the electrons drift first toward and then away from these wires on their way toward the collection plane. If these are oriented in different directions to the collection plane, the combination of signals provide a 2D projection of the charged particle's position as it traversed the LAr. The third dimension is given by the time which the ionization took to finish drifting. As the drift velocity is constant and known from the electric field, one can measure the initial position as:

$$x = t_{\text{drift}} * v_{\text{drift}} = (t_f - t_0) * v_{\text{drift}} \tag{4.1}$$

where x is the lateral position of the track, t_f is the readout time, t_0 is the time at the start of charge drift, and v_{drift} is the known drift velocity. These principles are shown in Figure 4.1. This shows a neutrino interaction producing two charged particles. These go on to ionize the LAr, and the ionization electrons drift against the electric field created by the anode wires and cathode plane. Signals are produced on the wires: the plane labeled 'V' shows bipolar signals created by induction; the plane labeled 'Y' shows unipolar collection signals.

By ionizing the LAr as it travels through the TPC, the charged particle loses energy. Thus, by measuring the amount of ionization (the size of the signals produced on the wires), one can measure the energy lost by the particle during its traversal of the LAr. This allows LArTPCs to provide calorimetric energy measurements of the particles it detects. The ProtoDUNE-SP event display shown in Figure 4.2 highlights this capability. In this, a beam particle enters the TPC from the left of the figure. It travels through the LAr until it undergoes an interaction with an Ar nucleus, producing two visible particle tracks. The strength of the signals is shown by the color of the tracks. The incident beam particle

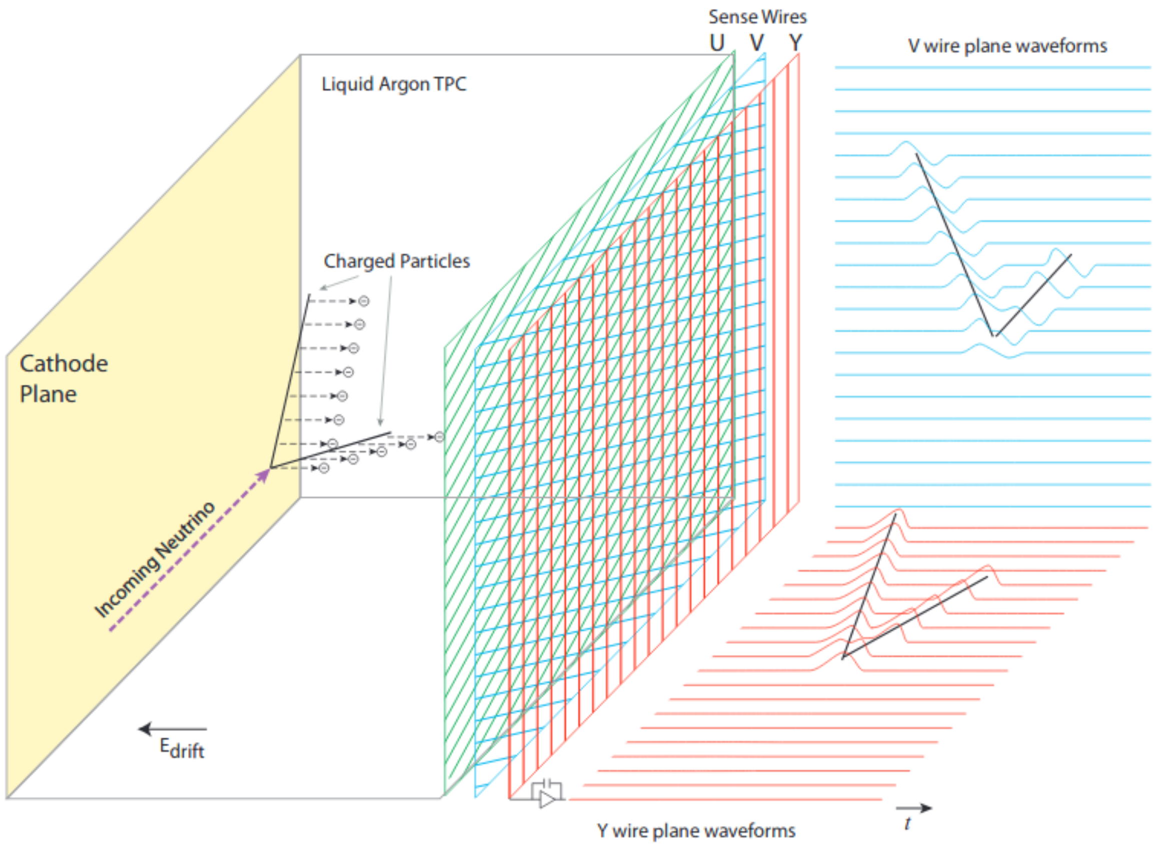


Figure 4.1: Design and operating principles of a LArTPC[28].

is a beam π^+ candidate, and deposits considerably less energy per unit distance than the products of its interaction.

Detecting the deposited energy provides Particle Identification (PID) of the charged particles because the mean rate of energy loss is well described by the Bethe formula[29] shown in Equation 4.2.

$$\left\langle -\frac{dE}{dx} \right\rangle = Kz^2 \frac{Z}{A} \frac{1}{\beta^2} \left(\frac{1}{2} \ln \frac{2m_e c^2 \beta^2 \gamma^2 W_{max}}{I^2} - \beta^2 - \frac{\delta(\beta\gamma)}{2} \right) \tag{4.2}$$

Here, z is the charge number of the incident particle, Z and A are the atomic number and mass of the material through which the particle is traveling, and I is the mean excitation of the material. $K = 4\pi N_A r_e^2 m_e c^2$ where N_A is Avogadro's number, and m_e and r_e are the mass and classical radius of the electron. W_{max} is the maximum energy transfer to an

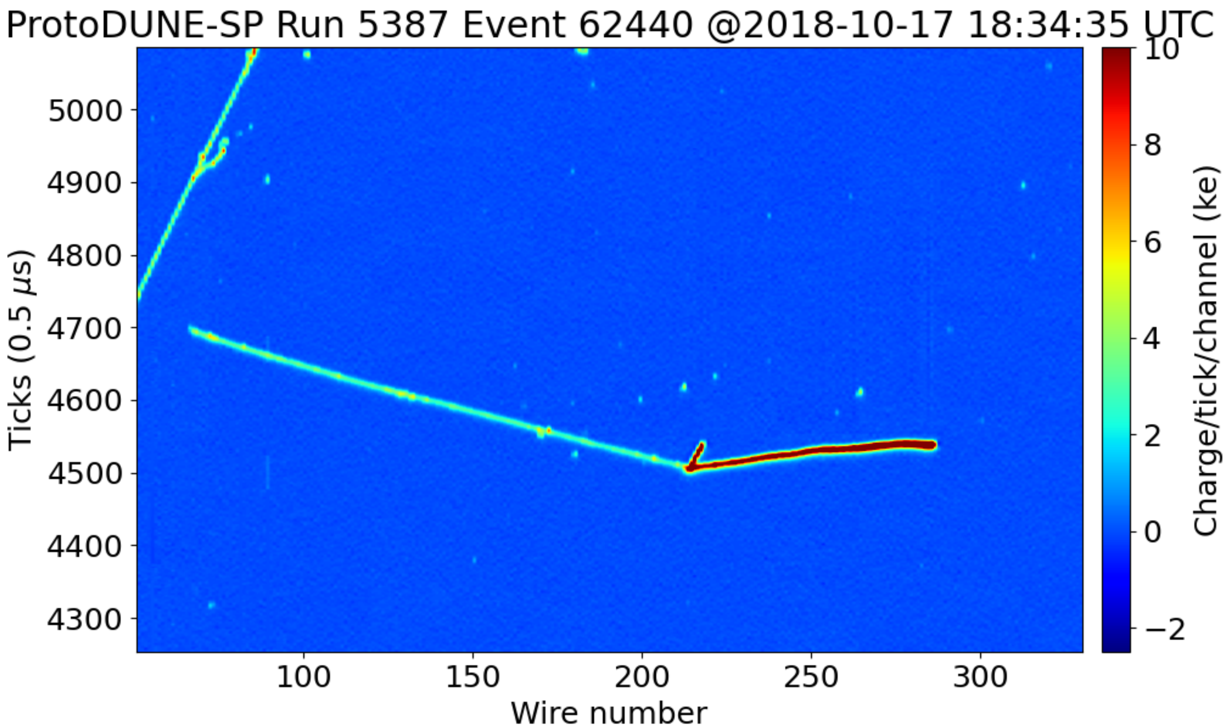


Figure 4.2: ProtoDUNE-SP event display showing a candidate beam π^+ entering from the left and undergoing an interaction with an Ar nucleus.

atomic electron for a single collision by a particle of mass M and is given by Equation 4.3.

$$W_{max} = \frac{2m_e c^2 \beta^2 \gamma^2}{1 + 2\gamma m_e/M + (m_e/M)^2} \tag{4.3}$$

The main dependence on the incident particles comes from the factor $1/\beta^2$[29]. For particles at the same energy, heavier particles will have a smaller β and thus will deposit a larger amount of energy per unit length. PID can be performed on particles by observing how much energy they deposited along their travel through the LAr. Thus, the products of the interaction in Figure 4.2 appear to be protons, making this event a candidate for π^+ absorption. The exact technique used in this analysis to separate $\pi^\pm$ from protons will be described in Chapter 7.

In addition to ionization, the charged particles create scintillation light by the excitation and subsequent radiative decay of argon excimers. Scintillation light from LAr is produced isotropically in a narrow band around 128 nm and has a large yield of 24,000 photons per

MeV deposited at a drift field of 500 V/cm (ProtoDUNE-SP's operating drift field value)[30]. It is produced in both a fast (∼5ns) and slow (∼1.3 – 1.4 μs) component at a ratio of 1:3[30]. The LAr is transparent to its own scintillation light[28], allowing photon detectors within the LAr to collect the light produced by the charged particles. This provides important timing and triggering capabilities for neutrino experiments.

This section provided an overview of LArTPC detection principles. However, certain complications arise in normal operating situations. The next subsections will provide overviews of the following effects in LArTPCs which modify this simple interpretation of tracking: recombination of ionization electrons(4.1.1); attenuation of ionization electrons due to LAr impurities (4.1.2); the Space Charge Effect created by the accumulation of positive Ar in the bulk of the LAr(4.1.3).

4.1.1 Recombination

Recombination is the effect by which ionized electrons thermalize with the Ar and then quickly attach to the positive Ar^{1+} ions created by the charged particle. This modifies the charge observed by the wires and must be accounted for in calibration in order to accurately measure the energy deposited by charged particles. There has been limitation in theoretical treatments of recombination to provide a global description of data [31]. The preferred models for LArTPCs with similar electric fields to ProtoDUNE-SP are the Birks and Box models. Both models are based off of the principle recombination effect arising from ionization electrons attaching to other Ar^{1+} ions created by the charged particle (as opposed to reattaching to its original atom) and both depend on the electric field. These models differ in that the Box model neglects electron diffusion and ion mobility during recombination and uses "Box model" boundary conditions rather than Birks' cylindrical assumptions [31]. The ICARUS experiment found good agreement to fits of the Birks model [32], while ArgoNeuT achieved good agreement with a "modified Box model" [31] which enabled another parameter to vary in order to achieve agreement to the Birks model at low dE/dx [31]. ProtoDUNE-SP also adopted this modified Box model.

4.1.2 Ionization Attenuation

Impurities in the LAr, such as water and O_2, can capture ionization electrons as they drift toward the anode plane. This reduces the final amount of collected charge, and is simply modeled as an exponential decay as in Equation 4.4. Here, Q_C is the collected charge, Q_0 is the initial charge deposited, t_d is the drift time, and τ is the "drift electron lifetime." This lifetime is lowered by the presence of impurities.

$$Q_C = Q_0 e^{-t_d/\tau} \tag{4.4}$$

4.1.3 Space Charge Effect

The Ar^{1+} ions created by the charged particles drift toward the cathode, but at a much slower velocity. As such, if enough positive ions are created, positive charge can build up in the bulk of the LArTPC. This accumulated charge can distort the electric fields, causing the so-called Space Charge Effect (SCE). This is especially the case for LArTPCs on Earth's surface such as ProtoDUNE-SP. These surface detectors are subject to a large cosmic ray flux, which constantly replenishes the positive charge. This large accumulation of charge causes the field lines to be bend toward the center of the TPC, resulting in distorted particle tracks. Through modifying the electric field, the SCE also changes recombination. The specifics of SCE in ProtoDUNE-SP will be discussed in Section 4.5.1.

4.2 The ProtoDUNE-SP Detector

With a total of 770 tonnes of LAr (420 tonnes are within the instrumented volume), ProtoDUNE-SP is the largest LArTPC ever constructed [30]. It provided a test bed for many components and engineering challenges of the single phase technology that will comprise the first DUNE far detector. It was designed to satisfy stringent requirements and achieve improved levels of LArTPC performance required by DUNE, and it surpassed these in many cases [30]. This section describes the design of the ProtoDUNE-SP components including the following: the cryostat surrounding the TPC and the LAr purification system (4.2.1); the TPC components (4.2.2); the Cold Electronics (CE) used to readout the TPC signals (4.2.3); the photon detector system used to readout scintillation light (4.2.4); the cosmic ray tagger (4.2.5); the Data Acquisition (DAQ), timing, and triggering systems (4.2.6). The detector will be described in terms of a right-handed coordinate system with y as the vertical axis pointing up, z horizontal and pointing approximately along the beam axis, and x horizontal and pointing along the electric field.

4.2.1 Cryostat and Purification

The cryostat, cryogenics, and purification system serve the role of keeping the argon in a liquid state with as little impurities as possible in order to avoid signal attenuation as described in Section 4.1.2. The TPC is encased in a membrane cryostat, which is formed of a corrugated membrane that holds the liquid and gaseous (from boil-off) argon with insulation, fireproofing, and supports outside of this [33]. The internal dimensions are 8.5m x 7.9m x 8.5m, making this the largest LAr cryostat ever constructed [30]. The membrane contains several openings to allow installation of detector elements, electrical/signal feedthroughs, the support structure for the TPC (which is suspended within the membrane), and cryogenic systems [34].

Due to ProtoDUNE-SP's larger drift distance (3.6m), a higher purity of LAr had to be achieved in order to limit the attenuation of ionization during drift. The purification systems used for ProtoDUNE-SP were built off those developed for ICARUS, MicroBooNE (another LArTPC neutrino experiment at Fermilab), and a LAr purity demonstrator based at Fermilab [30]. This purification system is the largest to date, and, along with the rate of recirculation and avoidance of leaks in the cryostat, reached an equivalent oxygen contamination of a few parts per trillion (ppt) [30]. This is in line with DUNE's requirement for

<100 ppt contamination in its single phase far detector [35].

4.2.2 TPC

The TPC of ProtoDUNE-SP is an active volume of 7.2m x 6.0m x 6.9m separated by a cathode at $x = 0$ into two drift volumes each of drift distance 3.6m and a drift field of 500 V/cm. The cathode is formed of six Cathode Plane Assemblies (CPAs) biased at -180 kV. Each side contains three Anode Plane Assemblies (APAs) opposite the cathode which contain the instrumentation wires and CE used to readout the wires. Surrounding the top and bottom and sides parallel to the drift field is the Field Cage (FC) that provides (in addition to the APAs and CPAs) electrostatic boundary conditions to achieve the intended drift field. Penetrating into the $x < 0$ drift volume (henceforth called the "beam side") is the Beam Plug which minimizes the energy loss and interactions of beam particles with inactive material. This layout is shown in Figure 4.3.

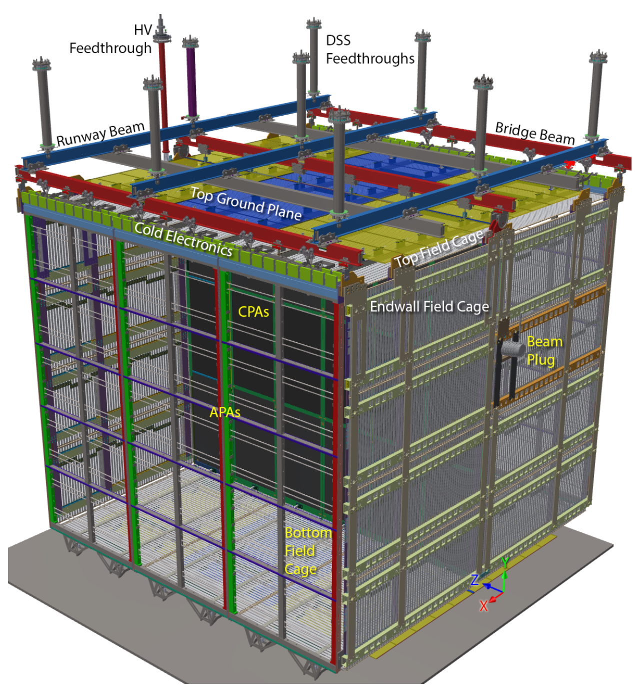


Figure 4.3: Diagram of the ProtoDUNE-SP TPC components [34].

The CPAs are 1.15m wide and 6.1m and consist of three vertically stacked cathode panels.

In order to avoid an electrical breakdown of the TPC which could destroy the CE, the CPAs are constructed of heavily resistive materials which give them a very long discharge time. The panels are constructed from a fire-retardant fiberglass-epoxy composite and are laminated on both sides with a Kapton film [30].

The APAs are formed of a rectangular stainless steel frame 6.1m high, 2.3m wide, and 76mm thick. Bonded directly over each side of the frame is a bronze wire mesh with 85% optical transparency that provides a grounded shield plane for four sets of wires on each side of the frame. Each successive wire plane is 4.75mm above the previous, with the inner most plane also 4.75mm above the mesh. The inner most plane is the X plane and is oriented vertically. Above the X plane is the V layer oriented at -35.7° from vertical, proceeded by the U plane oriented at +35.7° from vertical. Finally, the Grid (G) plane lies above the U plane and is oriented vertically. The G plane serves as a protective shield against electrostatic discharge and is not read out. The rest of the wires are connected to front-end CE and serve as the main instrumentation wires. The voltages of the wire planes ($V_{\mathrm{G}=-665}$ V, $V_{\mathrm{U}=-370}$ V, $V_{\mathrm{V}=0}$ V, $V_{\mathrm{X}=+820}$ V) are chosen such that the field lines terminate on the X plane, thus designating the X plane as the collection plane. The V and U planes are thus the induction planes. The X and G planes both have a wire pitch of 4.79mm, but are staggered from each other by half a wire pitch (meaning the G plane wires sit above and between two X plane wires). The V and U planes both have a wire pitch of 4.67mm. Each side has separate X and G planes, while the V and U planes are wrapped once around the APA. The angle of the V and U planes is such that 1) each wire crosses only a given collection wire on each side only once and 2) an integral number of CE boards reads out one APA. The first point serves to reduce ambiguities in track reconstruction. A diagram of an APA with a limited number of wires displayed is shown in Figure 4.4. Additionally, electron diverters were installed between the APAs on the beam. These were formed of two vertical electrode strips mounted on insulating board that, with voltages applied between the electrodes, modified the local drift field such that electrons drifted away from the gaps and toward the active area. During operation, high currents were drawn from the diverters' power supplies due electrical shorts in the cold volume. They were therefore left unpowered during operation, and, due to a resistive path to ground, the outer electrode was grounded. This was not the intended voltage, as it then collected charge near the gaps between APAs and distorted tracks crossing between APAs [30].

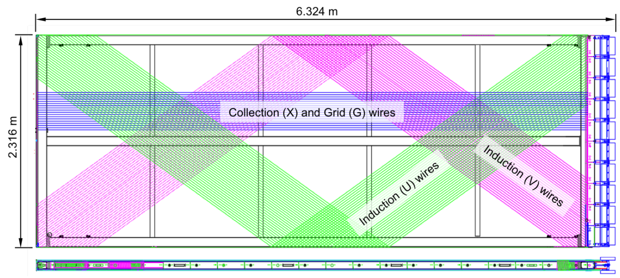


Figure 4.4: Diagram of an APA with its wire planes labeled. The bronze wire mesh is not shown. As it is shown, it is oriented on its side. The right side of the figure is the top of the APA when it is oriented vertically. The connections to the front-end CE boxes can be seen on the right side [34].

The Field Cage covers the remaining four sides of the drift volumes not covered by the APAs or cathode plane. It provides the remaining electrostatic boundary conditions to create uniform electric fields in the drift volumes. The top and bottom are comprised of six FC assemblies each, while four end wall panels each consisting of four assemblies oriented parallel to the x direction (the nominal drift direction). The assemblies are made of parallel metal profiles connected to each other by a resistive divider chain to provide the voltage gradient, I-beams that form an insulating support structure, and ground planes for the top and bottom assemblies. The ground planes prevent (at the top) a high electric field entering the gaseous argon and (at the bottom) a high electric field reaching the cryostat floor and cryogenic services [34].

On the beam side FC wall closest to the beam ($z \approx 0$), a beam plug is installed. This plug displaces the LAr and reduces the mass through which the beam particles must travel before reaching the TPC. This then reduces the energy loss and interactions upstream of the active volume. It is formed of a series of alternating fiberglass and stainless steel rings, forming a cylinder capped by low mass fiberglass plates. It extends about 5 cm inside the field cage boundary. A printed circuit board acting as a mini field cage covers the inside face of the plug in order to reduce drift field distortions. It is filled with nitrogen at a pressure of 1.3 bar to balance against the hydrostatic pressure of LAr at its positioned height. The beam plug can be seen on the right side of Figure 4.3. In addition to the beam plug, the cryostat warm structure and insulation are modified to further reduce upstream interactions [30].

4.2.3 Cold Electronics

Each APA has a total of 2560 sense wires, resulting in a total of 15,360 channels to be read out. 20 Front End Mother Boards (FEMBs) are located directly on top of each APA and within the LAr to read out the sense wires. By being placed close to the wires, the capacitance of each channel is reduced, thus reducing the noise recorded by the electronics. The CE collect the signals from the APA wires, then amplify, shape, and digitize them before transmitting them to Warm Interface Boards (WIBs). These interface electronics then handle transmitting these signals to the DAQ.

The FEMBs consist of an analog motherboard containing eight 16-channel analog Front-End (FE) ASICs that provide the amplification and shaping of the signals, and eight 16-channel Analog to Digital Converter (ADC) ASICs. These ASICs are both custom circuits designed by Brookhaven National Laboratory (BNL) [36]. In addition to the analog motherboard is a mezzanine card containing a commercial Altera Cyclone IV FPGA which provides clock and control signals to the two sets of ASICs. The FEMB layout can be seen in Figure 4.5.

The FE ASICs provided amplification with a programmable gain of 4.7, 7.8, 14, and 25 mV/fC and a 5th-order anti-aliasing shaper with programmable peaking time of 0.5, 1, 2, and 3 μs. It also included options for enabling AC coupling, selectable baseline adjustment for operating at 200 mV for unipolar pulses on the collection plane or 900mV for bipolar pulses on the induction planes, and a selectable pre-amplified leakage current of either 100, 500, 1000,

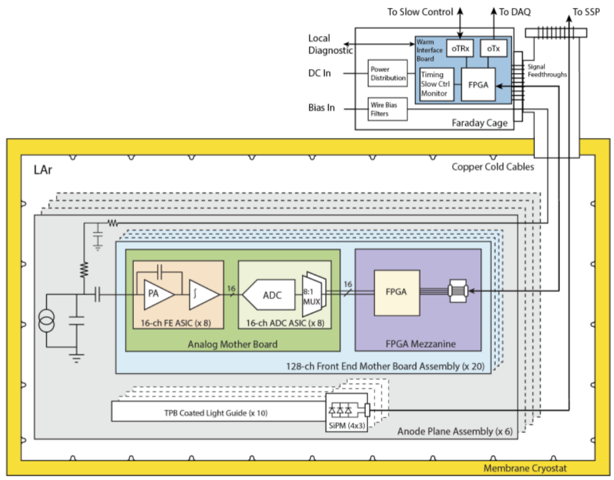


Figure 4.5: Diagram of the Cold Electronics in ProtoDUNE-SP [36].

or 5000 pA [36]. These ASICs also contained an internal, programmable pulse generator for electronics calibration. At normal running conditions, the FE ASIC gain is set to 14 mV/fC and the peaking time is set to 2 μs for all channels [30]. At cryogenic temperature, the FE ASIC packaging puts stress on the ASIC chip causing a channel dependent non-uniform lowering by up to 150 mV of the 200 mV collection mode baseline [36]. In addition to this, large input charge caused the FE ASICs to saturate. Due to this, the baselines for both collection and induction plane channels were set to 900 mV [30].

The ADC ASICs have 16 12-bit digitizers operating at speeds up to 2 MHz and an 8:1 multiplexing stage resulting in a pair of parallel serial readout lines that send output signals to the FPGA. At cryogenic temperature, the ADC ASIC suffered from an issue caused by failures in transistor matching. This effect is hard to simulate at LAr temperatures and is not present at room temperature. The mismatch between transistors effected the transition between the six most significant bits and six least significant bits in the ADC's “domino” architecture, causing the ADC output to prefer 0 and multiples of 63 in the dynamic range of the ADC [36]. This issue, referred to as “sticky codes,” was corrected for after data-taking and will be further discussed in Section 4.3.

The signals from the ADC ASICs are collected by the FEMB's FPGA, which further serializes the 16 pairs of data streams into four 1.25 Gbps links to the WIBs. The FPGA also provides a calibration pulse to the FE ASICs as a cross-check for electronics calibration [36].

The WIBs serve as the interface between the CE and DAQ, and are each controlled by an Altera Arria V GT FPGA. Data cables from all FEMBs on a given APA feed through a signal flange to a Warm Interface Electronics Crate (WIEC). The WIECs each contain one Power and Timing Card (PTC) which is connected to both a 48 V power supply and the detector timing system via a bidirectional fiber optical link. The PTC is connected to a Power and Timing Backplane (PTB) also housed in the WIEC. The PTB steps down the power and then fans out the power and clock signals from the PTC to five WIBs also contained in the WIEC. Each WIB distributes power to and controls up to 4 FEMBs. The WIB FPGA reorganizes and transmits FEMB data over fiber optical links to the DAQ. It also includes a real-time digital diagnostic readout on a Gb Ethernet link and an on-board component that can provide independent clocking to the FEMBs. These two components allowed for installation and checkout tests to be performed on the FEMBs before they were connected to the timing system and DAQ [36].

4.2.4 Photon Detectors

To collect scintillation light produced by charged particles in the LAr, 10 bar-shaped photon detectors 8.6 cm in height, 2.2 m in length, and 0.6 cm thick were embedded in each APA frame. Three different designs of photon detection technology were used in order to test options for use in DUNE's far detector modules. In each, the ∼128 nm scintillation photons were converted into visible light using wavelength shifters. This visible light is trapped within the photon detectors and eventually collected by an array of silicon photomultipliers [30].

4.2.5 Cosmic Ray Tagger

Located upstream and downstream (relative to z/beam direction) of the ProtoDUNE-SP cryostat is a cosmic ray tagger (CRT) used to provide triggers from cosmogenic muons. The

CRT is formed of scintillation counters recycled from the outer veto of the Double Chooz experiment that coarsely measure the x and y position of cosmic muons which pass through it. Coincidence hits registered upstream and downstream of the detector can be used to form tracks that can then be matched to reconstructed tracks in the TPC and provide calibration [30].

4.2.6 Data Acquisition, Timing, Triggering

The DAQ reads in data from the TPC, photon detectors, and CRT. Two readout solutions were employed for the TPC as tests for the DUNE far detector readout: RCE [37] and FELIX [38]. During the beam run, one APA (located on the $x > 0$ side) used FELIX, while the other 5 APAs used RCE. artDAQ [39] was used as the software framework that controlled the data-flow including event building, configuration, and writing of data to disk [30].

The timing system provides a 50 MHz clock to all subsystems of the detector. It also serves to distribute triggers created by the Central Trigger Board (CTB). The CTB is a hardware triggering system that forms trigger words based on the status of individual subsystems (CRT, photon detectors, beam instrumentation). These words are sent to the timing system which ultimately makes the readout decisions. Various trigger conditions can be created by creating requirements of subsystem statuses (active vs. inactive). When these requirements are met, the CTB sends off its trigger words to the timing system, which then determines if an event should be formed. If so, it issues the trigger to the DAQ and the various readout systems [30]. Importantly, the CTB can create beam-on and beam-off triggers based on if the beam instrumentation recorded a particle passing through the beam line. This way, TPC events containing a beam particle can be easily identified and used for analysis.

Each triggered readout of the detector, also known as an “event,” consists of 3ms of data taking: 6000 consecutive samples taken at a rate of 2MHz from each ADC. The event is built from data taken in by the DAQ starting 250 μs before the trigger time. This collects signals from charge deposited in the detector before the trigger, but that arrive within the time of the event. Coinciding data from the photon detectors and CRT are saved in the output stream as well, and matching to beam instrumentation (which will be described in Chapter 5) is done after data taking.

4.3 TPC Characterization

Before performing analysis, several data preparation steps are required to convert the waveforms in units of ADC to units of charge, as well as to mitigate readout issues. The first step is to determine the pedestal of each channel, as voltage offsets are introduced at the input of the front end amplifiers and these vary on a channel-by-channel basis. Additionally, for a given channel, the pedestal varies from one TPC event to the next. As such, the pedestal is evaluated separately for each channel and each event [30]. The pedestals are determined by finding the mean of all (typically) 6000 samples in an event for each channel.

For each channel in the event, its pedestal is subtracted from all ADC samples in the waveform. This difference is then multiplied by the channel's gain. This gain g is determined by using the 6-bit DAC included in the FE ASIC to inject a known amount of charge Q. For the collection plane, the integral of the ADC signal over the pulse A is related to the input charge as $Q = gA$. Special runs were taken where the DAC injected known amounts of charge. For each charge setting, the mean of the ADC integral of the resulting waveforms were determined. A line constrained to pass through 0 was fit to a set of these mean values near charge inputs typical to operation (up to several overlapping Minimum Ionizing Particles). The slope of this line is proportional to g for that channel [30]. Figure 4.6 shows an example of this.

In addition to the gain calibration, readout issues are identified and mitigated. The first readout issue is the aforementioned sticky codes issue. ADC values subject to sticking as well as the channels which exhibit the issue were initially identified by scanning a few waveforms and the pedestal histograms for every channel. The channels with particularly prevalent sticky codes are identified, and the list of known sticky codes is used to mitigate the

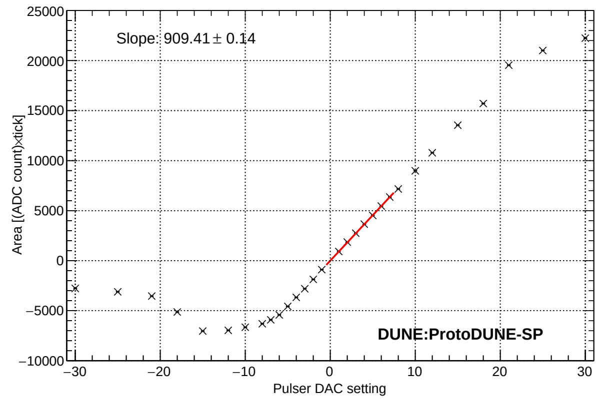


Figure 4.6: Example of gain determination for one channel in PDSP. The slope of the line divided by the charge level of each DAC step (Q_s = 3.43fC = .4 ke) gives the gain of the channel [30].

issue in less problematic channels. The mitigation works on these channels known to exhibit sticky code issues by replacing any ADC sample at a sticky code with a value taken from interpolation between the nearest non-sticky neighbors [30]. An example of this is shown in Figure 4.7.

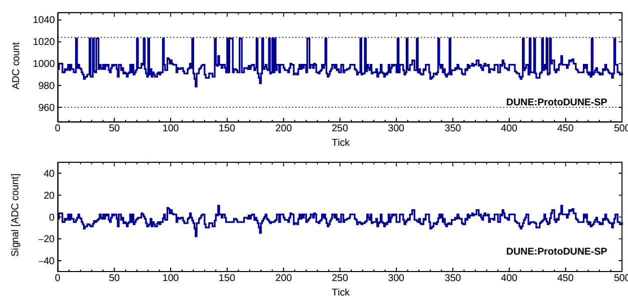


Figure 4.7: Example of ADC waveform before (top) and after (bottom) pedestal subtraction and sticky code mitigation. The spikes are samples which have stuck to the code represented by the upper horizontal dashed line. They are removed and replaced by interpolating to the nearest non-sticky neighbors [30].

In addition to the sticky code mitigation, preparations to remove tails resulting from AC coupling in the CE and correlated noise are also performed [30].

4.4 Event Reconstruction

After events are recorded, and an amount of data preparation is performed (described previously in Section 4.3), reconstruction software builds up a description of what happened during the event. This is done in a two step process: 1) hit finding which identifies localized charge deposits on wires and 2) pattern recognition which separates collections of hits into objects representing particle tracks and showers, and which also attempts to associate particles together in a hierarchy representing series of interactions.

Ideally, charge depositions on the wires should form (possibly overlapping) Gaussian-shaped signals when read out by the electronics. Thus, a hit-finding algorithm attempts to identify these separate depositions of charge by fitting Gaussian peaks to the waveform in a given wire. Each Gaussian peak thus represents one reconstructed hit or, in other terms, a localized deposition of charge in the detector. An example of this is given in Figure 4.8, where three hits have been reconstructed to the shown waveform.

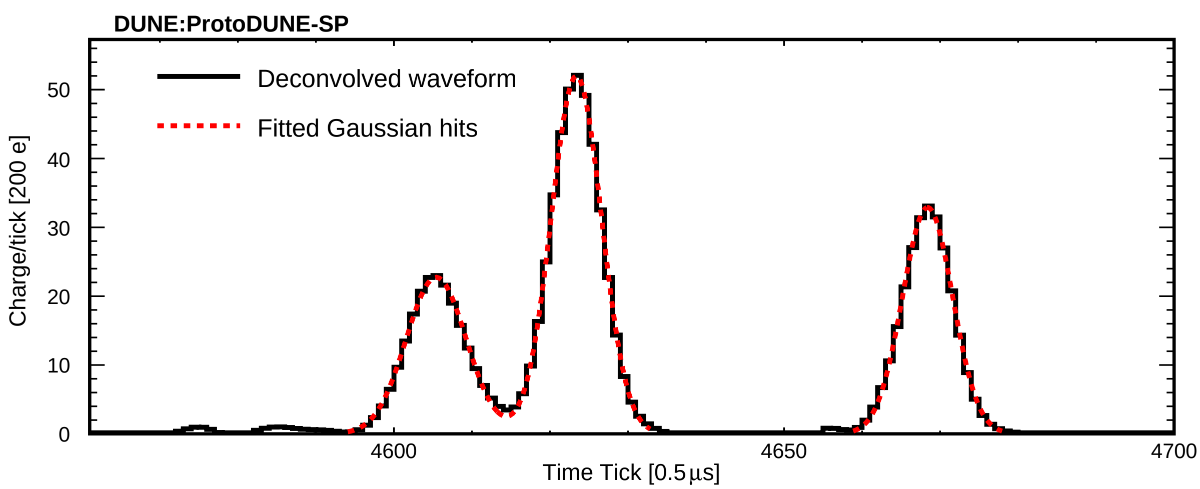


Figure 4.8: Example of three hits reconstructed to a single wire's waveform [30].

Because induction wires are wrapped around the APA, charge on either side of the APA can create a signal on a given wire. Thus, a disambiguation must be performed to determine which side of the APA the signal came from. Collections one wire from each plane are formed by identifying signals that arrived within a narrow time window. Sometimes, multiple pairs of induction wires can be matched to the collection plain. To determine which ones were truly paired, the algorithm attempts to minimize the difference in charge between that on the collection plane wire and on the induction plane wires. Simulation shows that this assigns >99% of hits to correct wire segments [30].

The second part of reconstruction is pattern recognition, which is performed by the Pandora framework [40]. This software has been successfully used in other LArTPC experiments such as MicroBooNE [41]. The first step in the pattern recognition is to preform two dimensional clustering of the reconstructed hits in each view. It then attempts to match sets of 2D clusters between the views, with care taken to resolve ambiguities. Afterwards, 3D hits are created. Then, in order to provide a detailed description of events, particle interaction hierarchies are created. Pandora then attempts to pick out particles originating from the beam line. All clusters are reconstructed first under a cosmic ray hypothesis. Clear cosmic ray candidates are then identified and removed. After these cosmic particles are removed, Pandora attempts to divide the detector into 3D regions containing all hits produced by a given particle interaction. These regions could contain cosmic rays that were not previously identified as such or particles that originated from the beam. Parallel reconstruction chains (one for cosmic rays and the other for test-beam particles) are then performed on these detector regions. The reconstruction for the beam particles is intended to resolve intricate hierarchies of particles such as from hadronic interactions. After the dual reconstruction is performed, a boosted-decision-tree algorithm tries to identify which region (if any) originated from the incoming beam [30]. The full reconstructed hierarchy (links between parent and child particles) in the beam region are made available for analysis. Further pattern recognition tries to identify whether the reconstructed particles were track-like (such as pions, protons, muons) or shower-like (electromagnetic showers from electrons or photons). Reconstructed track and shower objects are created for the corresponding particles. These provide information such as track length or shower direction (depending on the object) to users.

4.4.1 Hit Classification Using Machine Learning

In addition to the track/shower discrimination from Pandora, a machine learning-based classification was developed. A convolutional neural network (CNN) was trained to classify hits into track-like, shower-like, empty, or Michel-like categories. The track-like category represents hits coming from particles like pions and muons, the shower-like category represents hits from electron or photon showers, and the empty category represents hits resulting from noise. The Michel-like category represents electrons which originate from the decay of muons in the LAr, which appear different to electron showers. The output of the network is a set of scores representing how similar to each category the hit appears. The Michel-like category can overlap with the track-like and shower-like categories, and was not used for this analysis. The other three category scores are constrained to sum to one such that the hit can be classified as only one of these categories (that with the highest score).

The network uses as input 48x48 pixel^2 images created from wire readout data with the hit in question at the center. Each pixel is filled with the ADC value from the readout data. One axis of the image represents the wire which recorded the hit, while the other axis is the time coordinate (which has been downsampled by taking an average over time samples). The readout data used as input has been prepared according to the procedure described in 4.3. MC simulation was used to train the network by identifying whether the hit was due to charge deposited by a particle (or was created by noise) and what type of particle created the hit. Further information on network architecture and training can be found in Reference [42]. The analysis presented in this thesis utilized the CNN scores as an alternate track/shower discrimination technique. Scores for the full reconstructed particle were calculated by averaging over all hits in the particle. Cuts can be placed on these average scores to categorize full particles as tracks or showers. The use of this within this analysis is described further in Chapter 6

4.5 Detector Calibration

In order to conduct useful physics analyses such as the measurement presented here, the relationship between deposited energy in the detector and the response of the detector must be determined. Several effects that must be taken into account have already been described in Sections 4.1.3, 4.1.1, and 4.1.2. This section serves to describe the procedures taken to calibrate for these effects.

4.5.1 Space Charge Effect in ProtoDUNE-SP

As previously described, the steady flux of cosmic rays produces a buildup of charge from the slowly drifting Ar^{1+} ions produced by ionization. This so-called Space Charge Effect leads to a persistent distortions of the drift field. These alter the drift paths of ionization electrons and also affect the amount of prompt charge recombination, resulting in spacial distortions of reconstructed tracks and modified reconstructed dE/dx of tracks. The spatial distortions of reconstructed tracks is evident in Figure 4.9, where the end points of cathode-crossing cosmic rays are pulled inward from the edges of the detector (the dashed lines).

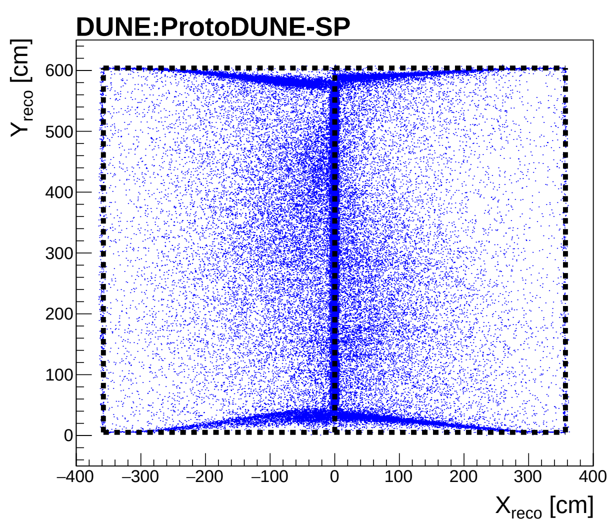


Figure 4.9: Projection of reconstructed track end points from cathode-crossing cosmic ray muons. The presence of SCE causes the end points to deviate from the boundaries of the TPC volumes represented by the dashed lines [30].

Figure 4.10 shows the distortions normal to four of the detector faces for events piercing the respective face within data events. These provide the magnitude of spatial distortions at these points of the detector in data events. A simulation of SCE was developed for ProtoDUNE-SP. This is shown in Figure 4.11, which is analogous to Figure 4.10. Data-MC discrepancies can be seen here that possibly stem from incorrect values of the Ar^{1+} drift velocity (amounting to a different amount of accumulated charge) and/or unsimulated flow of the liquid argon.

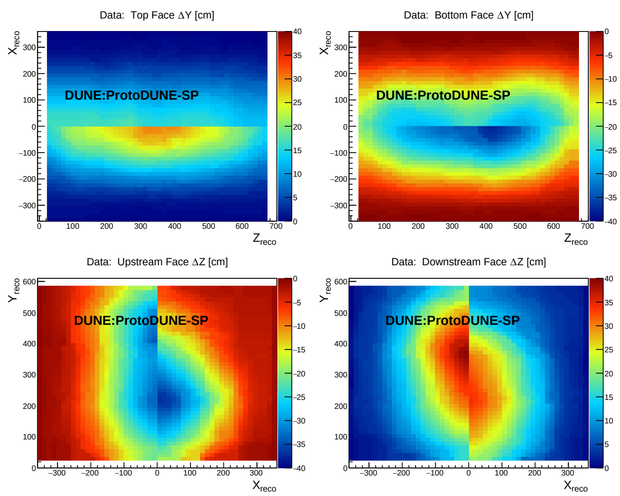


Figure 4.10: Spatial distortions normal to four detector faces from data events. Top: upstream & downstream relative to the z-direction. Bottom: Upper & lower faces relative to the y-direction. The reconstructed location of the end points of cathode-crossing tracks that pierce through the respective face show the distortions perpendicular to that face at the reconstructed 2D location [30].

In order to overcome the inability of the simulation to reproduce the SCE seen in data, a data-driven simulation of space charge was implemented. This consisted of creating a set of both spatial and electric field distortion maps to modify the nominal simulated distortions. These maps can also be used to correct for SCE in both data and MC by recovering the original positions and also accounting for the modified electric field. These are created as follows:

1. The ratio of the data to simulation map are taken for each of the relevant faces of the

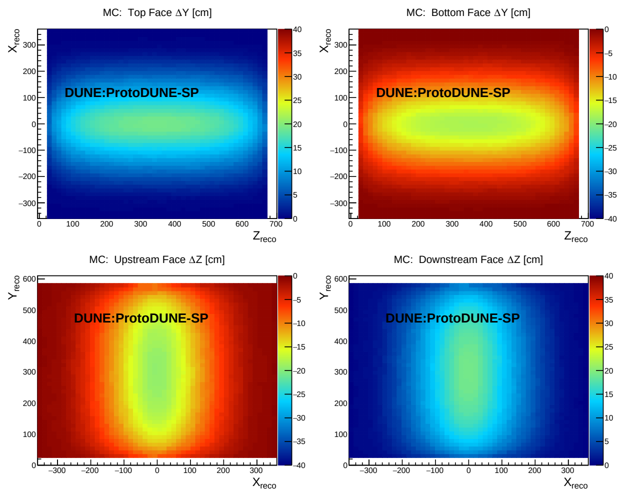


Figure 4.11: Spatial distortions normal to four detector faces from MC events. Top: upstream & downstream relative to the z-direction. Bottom: Upper & lower faces relative to the y-direction. The reconstructed location of the end points of cathode-crossing tracks that pierce through the respective face show the distortions perpendicular to that face at the reconstructed 2D location [30].

detector. This produces a 2D map of scale factors at each of these faces.

2. Spatial distortions in the y-direction are calculated by linearly interpolating the scale factor maps between the top and bottom faces. The same is done for z-direction distortions by interpolating the scale factor maps between the upstream and downstream faces. The x-direction distortions are then taken as the average between the distortions in y and z. This creates a 3D map with scale factors in all three directions. These are used to rescale the magnitudes of the spatial distortions maps.

3. The resulting distortions in each 3D map are then reversed in order to form maps that can be used to correct for the spatial distortions in both data and MC. The correction repositions the reconstructed ionization charge depositions to their original locations.

4. The gradient of the spatial distortion along the local drift direction (determined from the reversed maps) and the known drift velocity are used to form 3D electric field distortion maps.

These data-driven maps are used to modify the reconstructed position of ionization charge in simulation as well as to improve the prediction of prompt recombination effects [30].

4.5.2 Electron Lifetime

As discussed in Section 4.1.2, impurities in the liquid argon can capture drifting electrons before they reach the instrumentation wires. This reduces the amount of charge reaching the collection plane wires and is measured as an exponential decay as a function of drift time as in Equation 4.4, where τ is the drift electron lifetime. A larger τ corresponds to a higher liquid argon purity.

The electron lifetime can be measured by fitting the dQ/dx of cosmic ray collection plane hits as a function of drift time. To do this, cosmic rays that pass through the CRT and the front and back faces of the TPC were selected. The CRT was used to measure the initial time t_0 at which the track traveled through the TPC. The difference between the time the hit collected on the wire and t_0 was used as the drift time. The most probable value of dQ/dx for hits in slices of 100 μs drift times was fit according to Equation 4.5 to extract the lifetime. Two example fits, taken at the beginning and end of the beam data run are shown in Figure 4.12. The later data shows a higher lifetime resulting from higher purity [30].

$$\frac{dQ(t)_{\mathrm{MPV}}}{dx} = \frac{dQ_{0,\mathrm{MPV}}}{dx}\exp(-(t_{hit} - t_0)/\tau) \tag{4.5}$$

4.5.3 Energy Calibration

Reconstructed dQ/dx is affected by electronics gain variations, SCE, and attenuation. Previous sections describe the calibrations for these. Additional effects have also been calibrated out via a two-step process laid out here: first to equalize the detector response (using a sample of throughgoing cosmic rays), then with a determination of the absolute energy scale (using a sample of stopping cosmic rays).

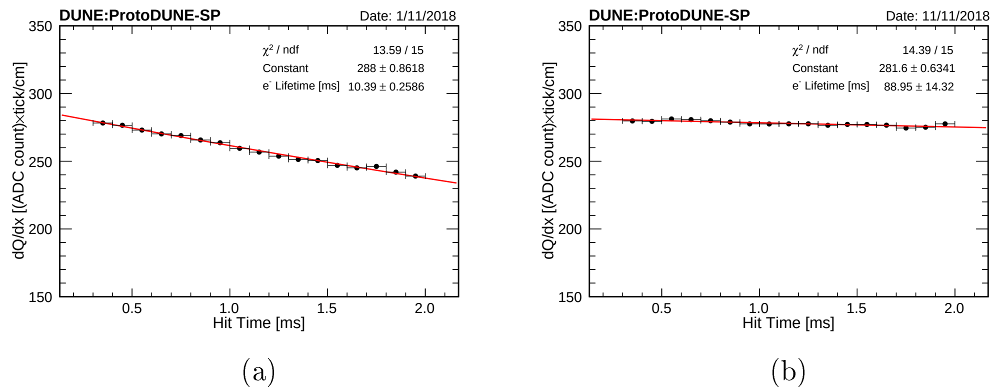


Figure 4.12: Fits to the drift electron lifetime τ for data collected at two different periods of time. Left is an earlier period with a lower period and shows a lower lifetime (10.39 ± 0.2586 ms) compared to the right (88.95 ± 14.32 ms) [30]

The equalization step accounts for nonuniformities from various effects that depend separately on x and $y - z$ positions of the hit. Effects that depend on $y - z$ position include non-uniform wire response from nearby dead channels, detector features such as the electron diverters, and transverse diffusion. This portion of the equalization step is done separately for each half of the detector on either side of the central cathode and as a function of y and z. The median dQ/dx value of hits in a given $y - z$ bin is determined and compared to the median dQ/dx value of the half of the detector wherein the hit lies ($x > 0$ or $x < 0$) to obtain a correction factor defined as such:

$$C(y, z) = \frac{(dQ/dx)^{\text{G}}_{\text{YZ}}}{(dQ/dx)^{\text{L}}_{\text{YZ}}} \tag{4.6}$$

where the numerator is the global median dQ/dx value on that side of the detector and the denominator is the median value on that $y - z$ bin.

Following this, effects that depend on x position such as longitudinal diffusion are equalized. Similar to the corrections in the $y - z$ plane, the median dQ/dx value of hits in an x bin are compared to the median value of all hits in the detector. This produces a correction factor that depends on x position as such defined in Equation 4.7.

$$C(y, z) = \frac{(dQ/dx)^{\text{G}}_{\text{YZ}}}{(dQ/dx)^{\text{L}}_{\text{YZ}}} \tag{4.7}$$

Finally, the dQ/dx values are normalized to the average value at the two anodes using the following factor:

$$N_Q = \frac{(dQ/dx)^{\mathrm{A}}}{(dQ/dx)^{\mathrm{G}}} \tag{4.8}$$

where the numerator is the average of the mean values at either anode, and the denominator is the mean value over the whole TPC. Thus, the dQ/dx of every hit in an event is equalized according to Equation 4.9.

$$(dQ/dx)_{\mathrm{C}} = N_Q C(y,z) C(x) (dQ/dx) \tag{4.9}$$

Next the measured dQ/dx must be translated to the energy loss of the particle per unit length dE/dx using a sample of cosmic muons that stop in the detector. dQ/dx values in the minimum ionizing region (120 to 200 cm from the end of the track) are converted to dE/dx using Equation 4.10 from the modified Box model [31] and fit to values predicted by Landau-Vavilon theory [43] as a function of residual range (the distance along the track from the hit to the end of the track).

$$\frac{dE}{dx} = \left(\exp\left(\frac{(dQ/dx)_{\mathrm{C}}}{C_{\mathrm{cal}}} \frac{\beta' W_{\mathrm{ion}}}{\rho \mathcal{E}} - \alpha\right)\right)\left(\frac{\rho \mathcal{E}}{\beta'}\right) \tag{4.10}$$

In Equation 4.10, W_{ion} is the amount of energy required to ionize an Argon atom (equal to 23.6 x $10^{-}6$ MeV/electron), ρ is the density of liquid argon at ProtoDUNE-SP operating temperature (equal to 1.38 g/cm^3), $\mathcal{E}$ is the local electric field at the location of the hit, α and β' are modified Box model parameters and were measured by ArgoNeuT with values of 0.93 and 0.212 (kV/cm)(g/cm^2)/MeV) respectively [31]. Finally, C_{cal} is a calibration constant that accounts for electronics gain and ADC conversion, and corrects for any residual effects not explicitly calibrated previously and is the parameter of interest in the fit [30].

The normalization factor N_Q, equalization maps $C(y,z)$ and $C(x)$, and calibration constant C_{cal} are measured separately for MC and each run of data, and are applied during analysis when extracting the values of dE/dx for each hit considered.

4.6 Monte Carlo Simulation

The simulation of test beam events in the TPC begins with the simulation of test beam particles generated within the beam line. A dedicated Geant4 [44] simulation of the beam line transports particles from their production point toward the face of the ProtoDUNE-SP TPC. More details can be found in Reference [45]. The rest of the PDSP simulation chain is based in the analysis framework LArSoft [46]. The beam line simulation results are passed to an event generator module that creates particles to be simulated by Geant4. The events are created when a "primary" particle (such as a π^+) travels through two triggering planes and reaches the outside of the PDSP cryostat structure. Additionally, checks are performed for when particles interact or decay (if applicable) in the beam line such that events are also created if some downstream particle (for example a μ^+ from a π^+ decay) reaches the cryostat. Without this check, the rate of test beam muons was severely underpredicted by the event generator in early simulation productions. The set of simulation used in this analysis included this hierarchy check. Each event created by this event generator is assigned a primary particle: either the original particle or the last-extant particle which reached the cryostat structure (i.e. the μ^+ described above). These primary particles serve as the main particles considered in the analysis. Other particles originating from the beam line which are "in time" with the beam are passed on to the next stage of the simulation. These additional particles are added in if they are within 4.5ms of the primary particle, similar to what can occur in events in data. Cosmic-ray particles as simulated by CORSIKA [47] are overlaid on the event and passed on to the Geant4 step as well.

All particles generated by the beam-based event generator and the overlaid cosmics are then passed to Geant4 to simulate their transport through the detector. It also facilitates the interaction of hadrons with the detector material (via the Bertini Cascade Model [48]), thus serving as the signal interaction model of the analysis. The full geometry of the detector is considered, allowing for particles to interact and lose energy within the uninstrumented portion of the detector geometry (i.e. the steel cryostat structure, insulation, etc.). As charged particles travel through the LAr portion of the detector, ionization is created which is then passed on to the drift simulation step of the simulation.

The drift simulation transports the ionization electrons produced during the Geant4 simulation stage along field lines toward the wire planes. The nominal electric field map used within the simulation is distorted according to the data-driven SCE maps discussed earlier in Section 4.5.1. The full electronics response to the ionization drift and collection onto the wires is simulated, creating waveforms which are then passed to the reconstruction chain described earlier.

CHAPTER 5

PROTODUNE-SP BEAM LINE

Test beam particles are delivered to ProtoDUNE-SP from an extension of the existing H4 beam line in the CERN North Area. This beam line is known as the H4-VLE (very low energy) beam line as it supplies particles (π^+, μ^+, e^+, K^+, and p) in the momentum range 0.3 – 7 GeV/c. Within the North Area Secondary Beam facility, protons form the CERN Super Proton Synchrotron (SPS) impinge on a beryllium target to create a beam of secondary particles. These particles are transported through the H4 beam line before impinging on a secondary target to create the test beam for ProtoDUNE-SP. These test beam particles are momentum-selected[1] and transported through the H4-VLE beam line toward ProtoDUNE-SP.

5.1 Beam Line Instrumentation

The H4-VLE beam line is instrumented with a set of various devices to aid in particle identification (PID), momentum reconstruction, and tracking the beam. The layout of the beam line is shown in Figure 5.1. The instrumentation consists of scintillating planes (XBTF) for triggering and time of flight (TOF) measurements; scintillating fiber monitors (XBPF) for profiling, tracking, and momentum reconstruction; and Cherenkov detectors (XCET) as part of the PID process. Throughout the beam line are bending magnets which direct the beam toward ProtoDUNE-SP – with one also being used as part of a momentum spectrometer.

5.1.1 Fiber Monitors

The XBPF profile monitors [49] are comprised of a set of 192 square scintillating fibers of width 1 mm set side-by-side to provide a measurement of a beam particle's position in one direction. Two can be placed in perpendicular orientations to provide a 2D measurement of

[1] Nominal momentum settings consist of 0.3, 0.5, 1, 2, 3, 6, and 7 GeV/c

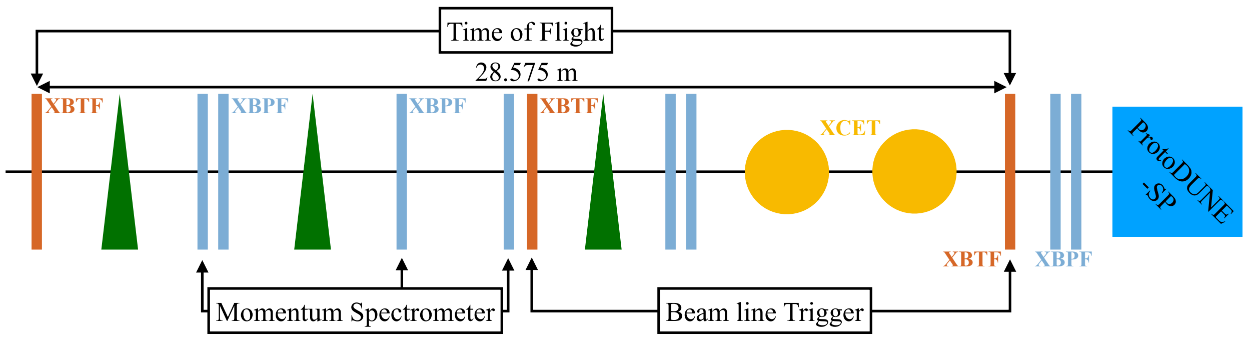


Figure 5.1: Diagram of the H4-VLE beam line instrumentation layout. XBTF (orange lines) are scintillating planes used for triggering and TOF measurement; XBPF (blue lines) are scintillating fiber monitors used for tracking and momentum reconstruction; XCET (orange circles) are Cherenkov detectors used for PID (sometimes in conjunction with the TOF); the green triangles are bending magnets throughout the beam line.

the particle's position. Each fiber is connected to an individual Hammamatsu S13360-130 silicon photomultiplier (SiPM) on one end[2]. Figure 5.2 shows a photograph of a prototype XBPF module taken from [49]. Further discussions of these devices and their readout are found in this reference as well as Reference [45]. The XBPF data was packaged such that, for each trigger in the beam line, the statuses (on/off) of the 192 fibers were separated into six 32-bit words. Two examples of this decoding is given in Figure 5.3.

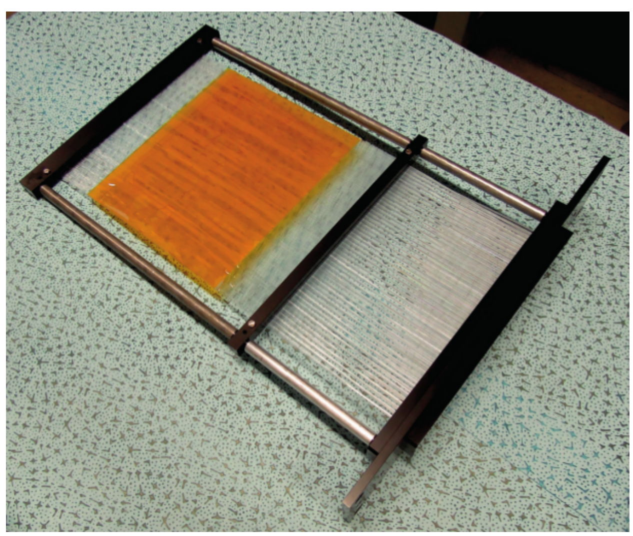
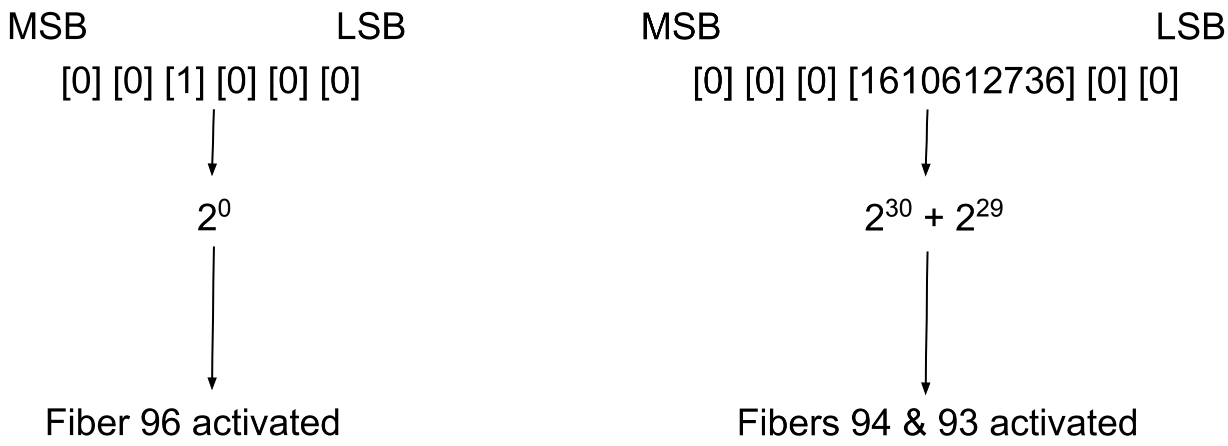


Figure 5.2: XBPF module. Taken from Reference [49].

[2]On the other end of the set of fibers is an aluminized mylar mirror.



(a) (b)

Figure 5.3: Two examples of XBPF data decoding. The most significant bit (MSB) and least significant bit (LSB) are labeled at the top of each example. a) The 0th fiber in the fourth 32-bit word is active. Thus the active fiber is $(0 + 3 * 32) = 96$. b) Two fibers are active: the 30th and 29th fibers in the third 32-bit word. Thus fibers $(30 + 2 * 32) = 94$ and $(29 + 2 * 32) = 93$ are active.

The last two sets of XBPF devices (shown immediately before the XCET devices and after the last XBTF plane in Figure 5.1) were used for tracking the particle as it entered into the TPC. 2D positions were reconstructed in both sets of XBPFs and used to create a trajectory between these points along the beam direction. This trajectory was further projected to the face of the active TPC to give the reconstructed position at the beam window.

These projected trajectories were used within analysis to cut out events considered as background to our pion sample. The difference in position between this reconstructed beam point and the start of the reconstructed TPC track, as well as the angle between the reconstructed beam trajectory and the starting angle of the reconstructed TPC track, were used to exclude various backgrounds (i.e. cosmics or particles from “upstream” interactions before the start of the active TPC volume).

5.1.1.1 Issues with XBPFs

In Winter 2019, two issues were identified within the data obtained by the XBPF during the initial beam run. In the first, the rate of fiber activations for the upper half of the fibers in the first XBPF was higher than the lower half. This can be seen in Figure 5.4 where the number of activations for each fiber for each selected event in the first XBPF are plotted. In talks with the device experts, this was determined to be caused by a configuration issue in the ASICs controlling the readout of this XBPF. This amounted to a higher efficiency in the upper half of fibers in this device. However, this was not an issue in ProtoDUNE-SP data analysis as this effect was suppressed by the lower trigger rate of the ProtoDUNE-SP detector compared to the trigger rate of the beam line (a subset of beam line particles triggered the detector).

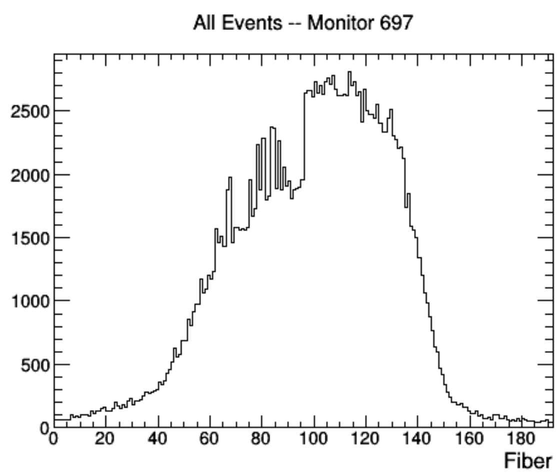


Figure 5.4: Active fibers in the first XBPF device from every event from a 1 GeV/c run. Note, multiple fibers can be active in any one event. The jump in rate at fiber 96 is due to a configuration problem in the readout electronics.

The second issue identified was due to a bug in the software controlling the data acquisition for the XBPFs, and occurred in all XBPF devices. In this issue, systematically repeated hits were being recorded in the last 64 fibers of each XBPF. This can be seen in Figure 5.5, where a bump is present starting near fiber 128 of the second XBPF device. Figure 5.6 highlights this issue, as it shows the number of times a fiber was activated in two subsequent events. A large spike in this rate can be seen starting at fiber 128. In discussions with the device expert, this was determined to be due to a software bug, in which the data in the last two words was not being cleared between events in the XBPF devices. This caused a "hangover" in the apparent activation of fibers in these two words. This resulted in extra reconstructed hits seen during analysis, which led to ambiguity in the reconstructed momentum and incident tracks. An attempt to mitigate this was implemented by simply scanning the last two words of each event for repeated fibers, and then masking the repeated fibers (in the second event). The results for this are shown in Figure 5.7

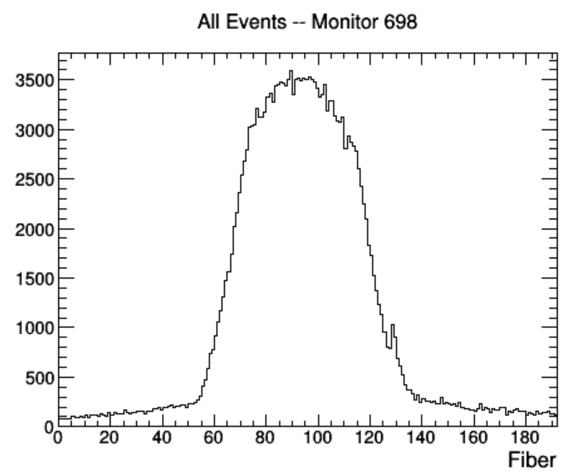


Figure 5.5: Active fibers in the second XBPF device from every event from a 1 GeV/c run. Note, multiple fibers can be active in any one event. A small bump can be seen starting around fiber 128.

5.1.2 Momentum Reconstruction Using XBPFs

Within Figure 5.1, three XBPFs are labeled as "Momentum Spectrometer." Coincident signals in these three monitors were used to measure the deflection of the test beam particle by the bending magnet which the monitors surround. The angle of deflection is then used, along with the known magnetic field, to reconstruct the particle's momentum. A diagram of this measurement technique is shown in Figure 5.8. The lateral position (χ_1, χ_2, χ_3) of the particle – given by the activated fiber in each of the three XBPFs – is used with the known

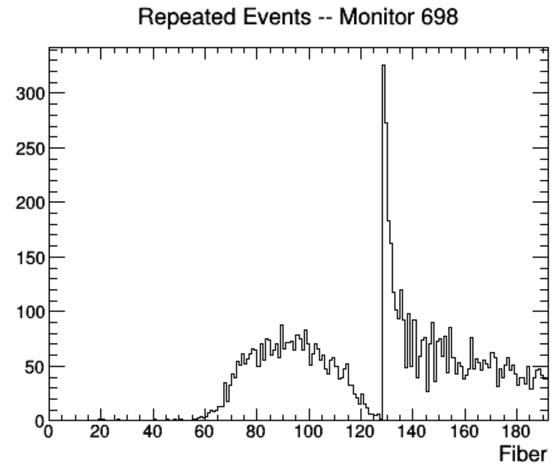


Figure 5.6: Rate of repeated fiber activations in the second XBPF device from every event from a 1 GeV/c run. The large jump at fiber 128 highlights the issue.

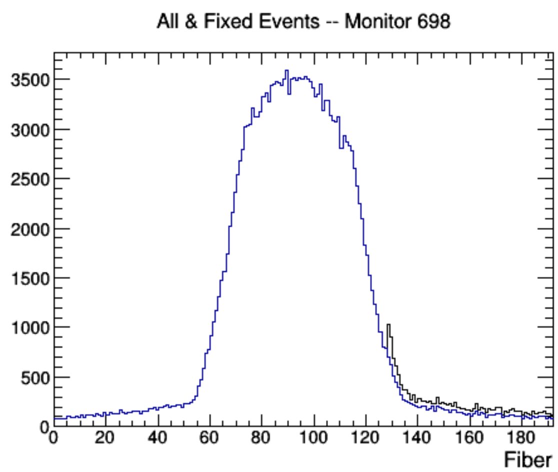


Figure 5.7: Active fibers in the second XBPF device from every event from a 1 GeV/c run before (black) and after (blue) the mitigation procedure.

distances between each monitor (L_1, L_2, L_3) in Equations 5.1 and 5.2 to reconstruct the momentum. In these, $M \equiv \alpha + \chi_1$, $\alpha = \frac{\chi_3 L_2 - \chi_2 L_3}{L_3 - L_2} \cos\theta_0$, $\Delta L \equiv L_3 - L_2$, and $\Delta\chi \equiv \chi_2 - \chi_3$. θ_0 is the nominal bending angle of the beam and is equal to 120.003 mrad [45]. This measurement has a nominal 2% resolution according to Monte Carlo studies [45].

$$\cos\theta = \frac{M[\Delta L \tan\theta_0 + \Delta\chi\cos\theta_0] + L_1 \Delta L}{\sqrt{[M^2 + L_1^2][(\Delta L \tan\theta_0 + \Delta\chi\cos\theta_0)^2 + \Delta L^2]}} \tag{5.1}$$

$$p = \frac{299.7924}{\theta} \times \int_0^{L_{\mathrm{mag}}} (Bdl) \tag{5.2}$$

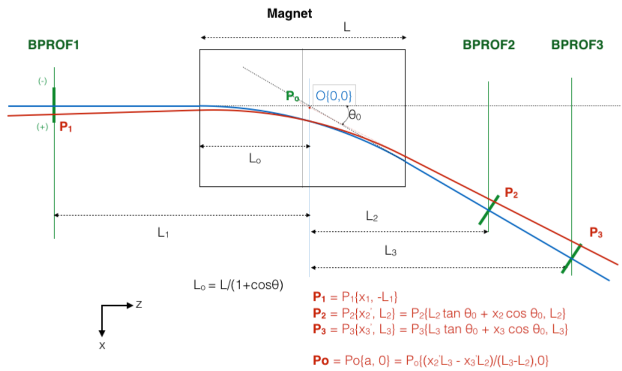


Figure 5.8: Momentum Spectrometer technique

Shortly after commissioning, a 5% offset in the reconstructed momentum was observed, and this was determined to originate from a bulk shift of the fibers in the third profiler of the spectrometer. Monte Carlo studies determined the fiber shift to be 1.45±0.18mm in the plane perpendicular to the beam. This was used as a systematic uncertainty within this analysis and will be discussed further in Chapter 8.

5.1.3 Scintillating Planes

The XBTF scintillating planes are of similar design to the XBPF – sets of 192 fibers arranged side-by-side and set perpendicular to the beam direction – but without individual readout of the fibers. Instead, the fibers are bundled into two groups, which are read out by two separate Hammamatsu H11934-200 photomultiplier tubes (PMTs). Figure 5.9 shows the bundled nature of the XBTF fibers.

The first and third (last) XBTFs – as shown in Figure 5.1 – are used for measure the TOF of the test beam particle over a distance of 28.575 m. The second (middle) and last XBTFs are used as a trigger for the rest of the beam line instrumentation as well as a prerequisite for triggering beam-type events within the ProtoDUNE-SP detector.

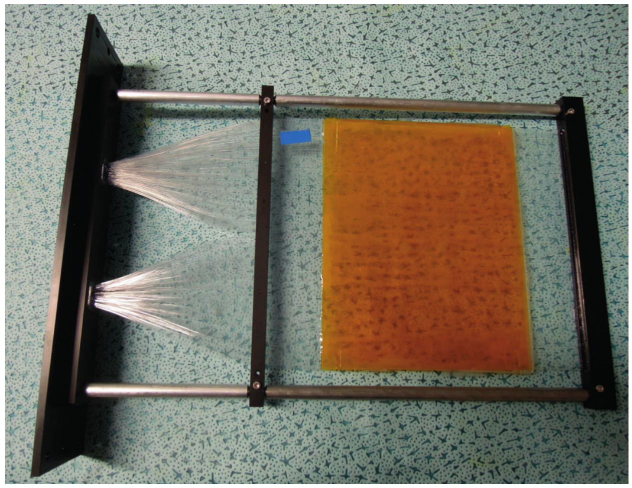


Figure 5.9: XBTF module. The bundling of the two sets of fibers can bee seen on the left [49].

5.1.3.1 Issue with XBTFs

A ∼4 ns “jitter” can be seen in the left plot of Figure 5.11b where a small bump in the TOF distribution exists around 100 ns. The cause for the issue was never identified. Furthermore, it has little effect on the analysis, due to the cuts used for PID (see below).

5.1.4 Cherenkov Devices

The two Cherenkov devices each consist of a 1.9 m long tube filled with the radiator gas (CO2) followed by a stainless steel enclosure. This enclosure houses a PMT at the bottom to collect the Cherenkov light and a curved mirror to guide Cherenkov light toward the PMT. The fill-pressure of the two devices were set to two different values to allow for discrimination between certain particle types. Figure 5.10 [45] shows the Cherenkov threshold pressure of CO2 at various momenta for different particle types, as well as the maximum possible pressure value for the two XCET devices. Consider an example setup at 3 GeV/c momentum. One device can be set above the electron threshold but below the μ/π thresholds, while the other can be set above the μ/π threshold but below the K/p threshold in order to distinguish positrons, muons/pions, and kaons/protons. Further discussion of Cherenkov devices within the beam line PID algorithm will be given in the following section.

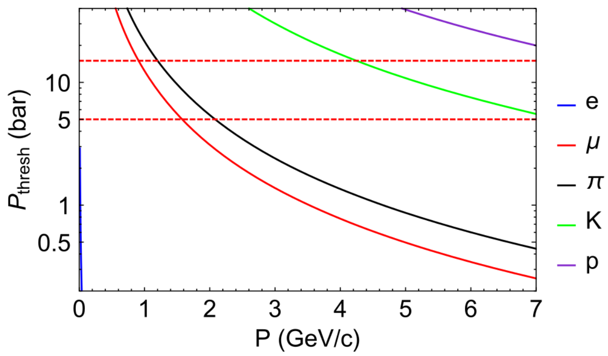


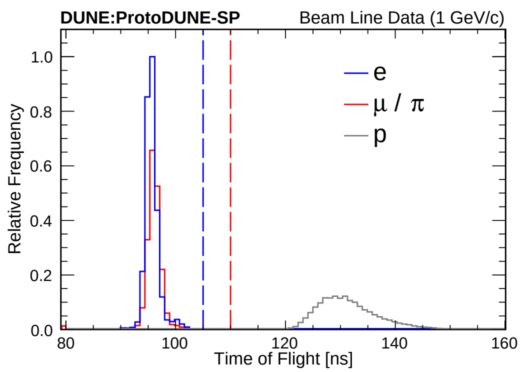
Figure 5.10: CO2 Cherenkov threshold pressures across ProtoDUNE's beam momentum range for the various particles present in the beam line. The dashed red lines show the maximum pressures for the two Cherenkov devices present in the beam line. Taken from Reference [45].

5.2 Beam Line PID

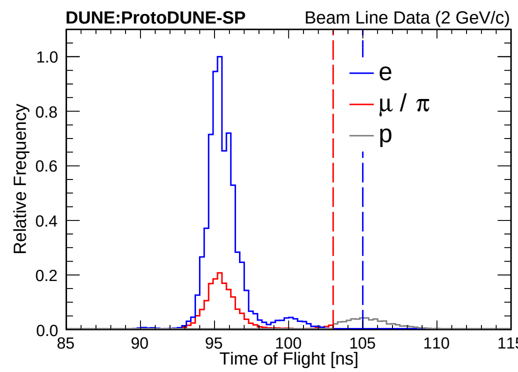
As mentioned above, the Cherenkov devices and TOF as measured by the XBTFs were used for PID of the beam line particles. Table 5.1 shows the conditions of the Cherenkov devices and TOF value used for the PID algorithm across the various nominal momentum settings. As shown in this table, for nominal momenta below 3 GeV/c, one Cherenkov device is used to distinguish e from the other particles, and the TOF is then used to distinguish μ/π from p. At 3 GeV/c, both Cherenkov devices are used to separate e, μ/π, and K/p. Finally, at 6 – 7 GeV/c, the two Cherenkov devices are used to separate $e/\mu/\pi$, K, and p. Figure 5.11 demonstrates this for the various beam momentum settings.

		Momentum (GeV/c)			
		1	**2**	**3**	**6 - 7**
e	TOF (ns)	0, 105	0, 105	–	–
	Low-p Status	1	1	1	1
	High-p Status	–	–	1	1
μ / π	TOF (ns)	0, 110	0, 103	–	–
	Low-p Status	0	0	0	1
	High-p Status	–	–	1	1
K	TOF (ns)	–	–	–	–
	Low-p Status	–	–	0	0
	High-p Status	–	–	0	1
p	TOF (ns)	110, 160	103, 160	–	–
	Low-p Status	0	0	0	0
	High-p Status	–	–	0	0

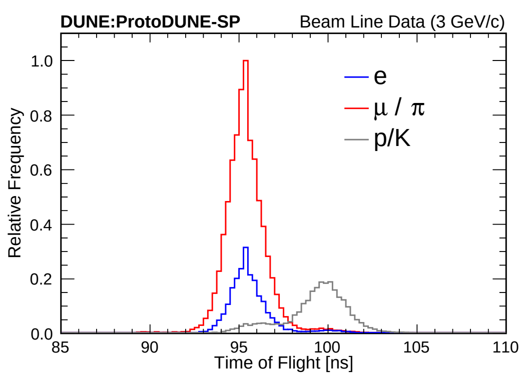
TABLE 5.1: A summary of beam line instrumentation logic used in the identification of particle types. Each cell reflects how a particular type of instrumentation is used at a given reference momentum. When time of flight is used, the values of the lower and upper cuts are given in nanoseconds. In the case of the high-pressure Cherenkov ("High-p Status") and the low-pressure Cherenkov ("Low-p Status"), zero and one represent the absence and presence of a signal respectively. When a given piece of instrumentation is not involved in a logic decision at a given momenta, a dash is used.



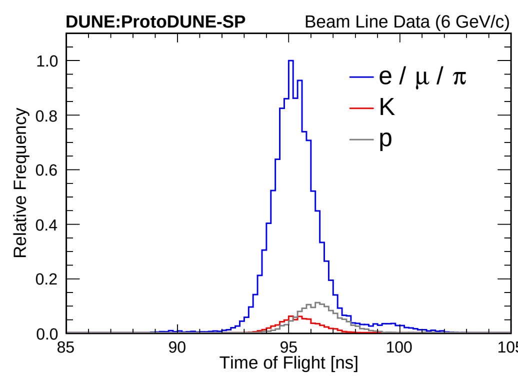
(a) Nominal beam momentum = 1 GeV/c. Vertical lines represent the time of flight cuts used for electrons (blue), and muons/pions (red).



(b) Nominal beam momentum = 2 GeV/c. Vertical lines represent the time of flight cuts used for electrons (blue), and muons/pions (red).



(c) Nominal beam momentum = 3 GeV/c.



(d) Nominal beam momentum = 6 GeV/c.

Figure 5.11: Time of flight distributions for different reference momenta, separated by particle using the PID techniques listed in table 5.1. The distributions are normalized such that the maximum height is equal to 1. Taken from Reference [30].

CHAPTER 6

EVENT SELECTION

This section describes the characterization of reconstructed data and MC events. Included is a set of data-MC comparisons detailing the cuts used in the selection. The data shown here is from Run 5387 of the initial ProtoDUNE-SP running period in the Fall of 2018.

For data, an event is included in the set if it passes the following criteria:

1. It is an event that was triggered by the beam line
2. It follows the π/μ beam line selection.
3. It has singular hits in each beam profile monitor. This is to eliminate ambiguity in the beam line momentum and tracking reconstruction.

For MC, due to the lack of fully simulated beam line instrumentation, the only requirement is that the simulated event was generated from a (primary) π^+ or μ^+ in the beam line simulation. Only μ^+ and π^+ are considered because at 1 GeV/c (the beam momentum used for this analysis), the beam line PID can distinguish π^+ and μ^+ from e^+ and p, but not from each other. The criteria for the beam line PID can seen in Table 5.1. Figure 5.11a shows that the protons are well separated by the TOF cut used to select μ^+/π^+. The MC events have been normalized to the number of data events that pass the aforementioned data criteria.

6.1 Truth Definitions

The MC events which pass the above criteria are separated into the following categories based on truth information of the primary beam particle:

1. Muons: The primary beam particle was a μ^+.

2. Upstream Interaction: The primary beam π^+ did not reach the TPC Fiducial Volume (FV).

3. Past FV: The primary beam π^+ extended past the FV in the z-direction[1].

4. Background Interaction: The primary beam π^+ interacted within the FV, and that interaction was a background (not Absorption or Charge Exchange) inelastic interaction. This includes any inelastic interaction between the primary π^+ and an Ar nucleus with an outgoing $\pi^\pm$ above a momentum threshold of 150 MeV/c. Note: this threshold is discussed further below.

5. Absorption: The primary beam π^+ interacted within the FV in an Absorption interaction. This is the first type of signal event and is defined as a π^+ which interacted with an Ar nucleus and resulted in no outgoing $\pi^\pm$ (above threshold) or π^0.

6. Charge Exchange: The primary beam π^+ interacted within the FV in a Charge Exchange interaction. This is the second type of signal event and is defined similarly to Absorption, but with any number of π^0 present.

7. Other: The primary beam π^+ ended within the FV, and did not interact inelastically (i.e. it decayed in flight or came to a stop and then decayed at rest).

The signal categories have been defined to occur within the FV (defined for primary particles ending before $z = 222$ cm). This is due to the fact that the grounded electron diverters created electric field distortions (as described in Section 4.2.2) which caused reconstructed tracks to break in their vicinity. A data-driven simulation of the electric field distortions was implemented, which attempted to account for this effect in MC. The track-breaking effect can be seen in the Figure 6.1, which shows the reconstructed endpoint of beam tracks in the

[1]Reminder: using a right-handed coordinate system, the z-direction is horizontal and follows the beam direction, the x-direction is horizontal and points away from the wires on the beam side TPCs, and the y-direction is vertical and points up.

TPC in the z direction. The legend shows the truth categories described in the previous section. Note that the exact effect is not perfectly modeled by the simulation, and a systematic

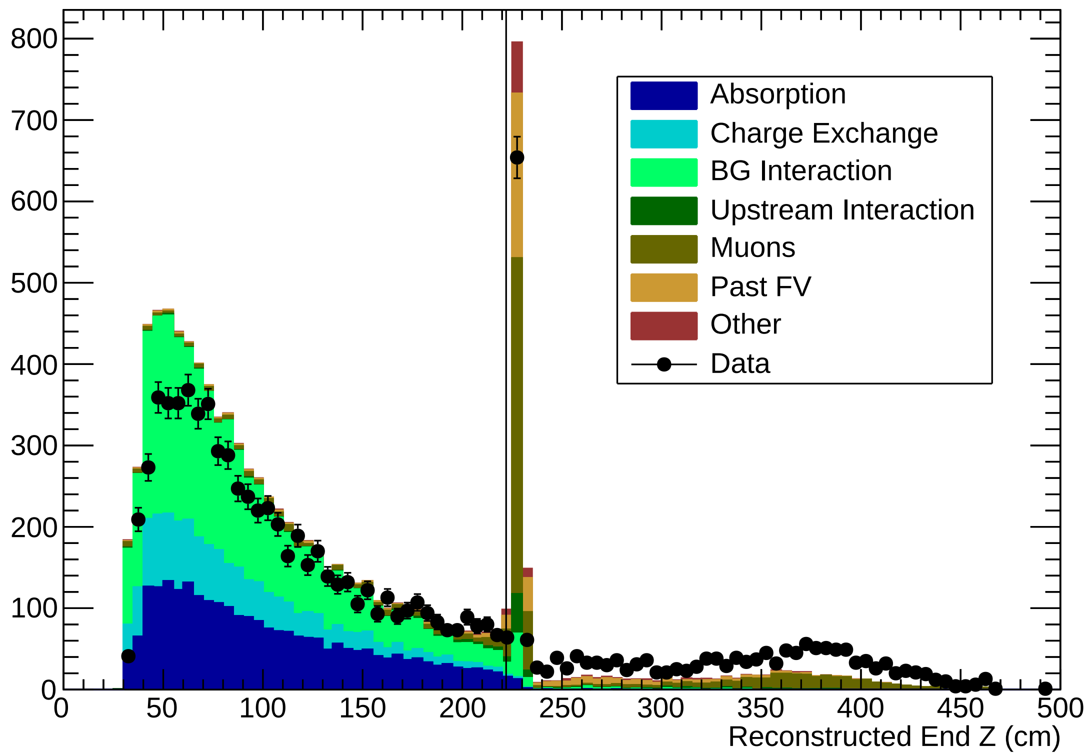


Figure 6.1: Reconstructed endpoint of beam tracks within the TPC. The vertical line represents the FV cut at 222 cm. The spike immediately after the FV cut is the track-breaking effect from the grounded electron diverters.

uncertainty on the strength of this effect in MC was implemented. This will be discussed further in Section 8.3.

Additionally, as mentioned above, the signal definitions are defined to have no charged pions above a momentum threshold of 150 MeV/c. This is due to the inefficiency to reconstruct charged pions exiting the primary interactions which are below this threshold.

6.2 Event Selection

The events that pass the previously stated criteria are then categorized according to the results of the TPC reconstruction. Ultimately, attempts are made to distinguish π^+ from μ^+ and then to distinguish absorption and charge exchange interactions from other π^+ interactions and stopping π^+. Every event is accounted for and characterized into one of the following categories:

1. The event contained no Pandora-reconstructed beam track *or* it did not leave enough hits on the collection plane wires. If, after the Pandora reconstruction described in Section 4.4 is performed, either no beam object was found or the beam object was reconstructed as a shower, the event is placed in this category. Also, events are placed here if there are not enough hits on the collection plane wires, as these are used in later cuts and in binning the events.

2. The event contained a reconstructed beam track, but it was not considered consistent with coming from the beam. This is done in order to pick out events in which the pion interacted upstream of the TPC FV or if Pandora erroneously reconstructed a cosmic particle as the beam track. This is described in Section 6.3.

3. The event was consistent with the incident track, but it extended past the FV cut in the z-dimension (222 cm).

4. The event remained in the FV, but it was rejected by the combined absorption and charge exchange selection. The selection criteria for this and the following two categories is presented in Section 6.4.

5. The event passed the combined absorption/charge exchange selection, and was distinguished as an absorption interaction.

6. The event passed the combined absorption/charge exchange selection, and was distinguished as a charge exchange interaction.

6.3 Beam Cuts

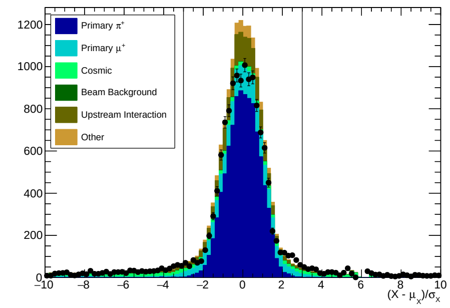
Sometimes, the wrong particle is identified as the beam particle by the Pandora reconstruction. As mentioned above, these could come from a cosmic muon or particle resulting from an interaction before the active volume of the detector. Information about the starting position and direction of the reconstructed track identified as beam is used to separate these out and place them in their own category. For the cuts in position, the mean (μ) and width (σ) of the beam track distribution in x, y, and z is found (using SCE-corrected information). Any track that is at least 3σ away from the mean in any direction is categorized as inconsistent with the beam. Additionally, the direction of the track is taken from the vector connecting the SCE-corrected start and end points of the track. The means of the angles relative to the Cartesian planes ($\bar{\theta}_x$, $\bar{\theta}_y$, $\bar{\theta}_z$) are found. Then, for each beam track, the cosine of the angle between its direction and the mean direction (defined by the mean angles) is found. This is defined in Equation 6.1. Any track which has $\cos(\theta) < 0.95$ is considered inconsistent with the beam.

$$\cos(\theta) = \cos(\theta_x) * \cos(\bar{\theta}_x) + \cos(\theta_y) * \cos(\bar{\theta}_y) + \cos(\theta_z) * \cos(\bar{\theta}_z) \tag{6.1}$$

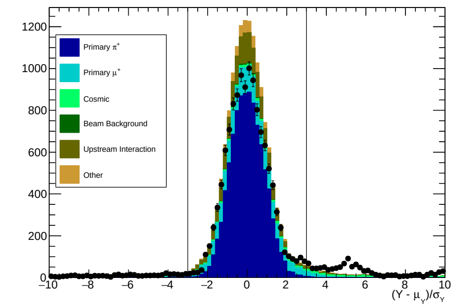
Figure 6.2 shows the distributions of the position (relative to the μ and σ) of the beam in each direction, as well as two views of the $\cos(\theta)$ distribution. The 3σ cuts in position and the $\cos(\theta) < 0.95$ cut in direction are shown as the vertical black lines. As can be seen in these plots, the cosmic particles and upstream interactions tend to have extreme angles and positions.

6.4 Absorption and Charge Exchange Selection

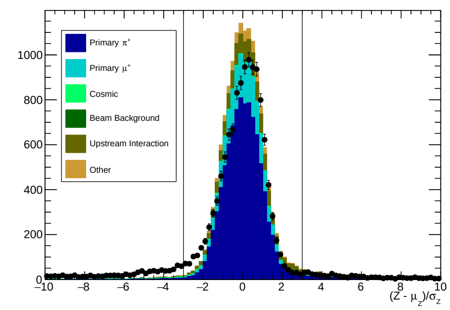
By analyzing reconstructed particles that have been associated to the TPC beam track as daughter particles, the tracks ending within the FV are separated between two categories: 1) absorption or charge exchange 2) other events. Because both absorption and charge exchange events contain no charged pion (above threshold) in the final state, the selection strategy is



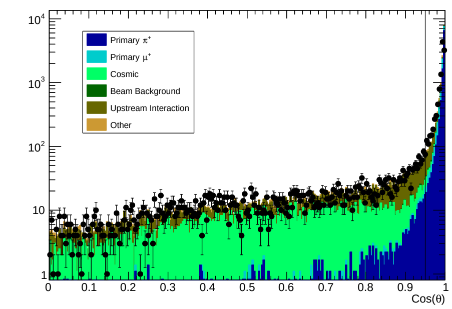
(a) SCE-corrected start x position reconstructed TPC track relative to mean and width of all beam tracks.



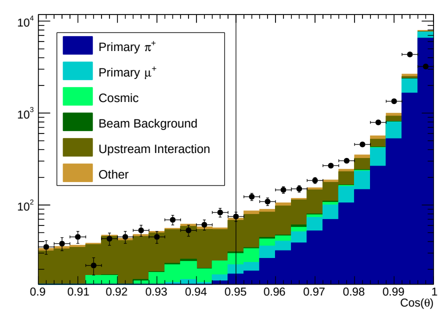
(b) SCE-corrected start y position reconstructed TPC track relative to mean and width of all beam tracks. Note that a bump in the distribution exists towards the right of the plot. This was recently identified as caused by a known detector effect present in data. Further investigation and possible treatments in MC simulation is ongoing.



(c) SCE-corrected start y position reconstructed TPC track relative to mean and width of all beam tracks.



(d) Cosine of the angle between reconstructed beam track and the mean track.



(e) Cosine of the angle between reconstructed beam track and the mean track. Zoomed-in to highlight cut region.

Figure 6.2: Distributions used to determine consistency with the beam line. The vertical black lines represent the cut values used.

to identify events with a charged pion daughter.

The daughter particles are separated between track-like (ideally from μ, π, p, etc.) and shower-like (ideally from e, γ) objects using the results of the CNN described in Section 4.4.1. For daughter particles of beam tracks, the CNN-based track/shower discrimination performed better than Pandora's native track/shower discrimination, and so was used for this analysis. Each hit in an event receives a set of scores produced by the CNN that encodes the degree to which it appears to be produced by a track-like particle or a shower-like particle. For each associated daughter, the scores from all of its hits are averaged to produce aggregated scores for the reconstructed particle. The Pandora reconstruction software was configured to reconstruct both a track-like and shower-like object for each reconstructed particle cluster in an event, so that analyzers could use alternate track/shower discrimination (such as the CNN method described here) and access the information accordingly.

At time of writing, only the calibration for collection plane hits was at a suitable state, and so only these hits were used to calculate the aggregated scores. A cut on the track-like score of the daughter particle at 0.3 (shown in Figure 6.3) was used to separate the daughters into shower-like and track-like. Here, the track score of every reconstructed particle associated as a daughter particle to the primary reconstructed TPC particle. The MC has been categorized by the true particle that created the reconstructed particle object. The fields "Daughter+" and "Daughter++" represent particles that are products of reinteractions of final state particles and so on. The field "Self" refers to segments of the true primary particle that were associated as a daughter (i.e. the track ended early). Finally, the field "γ" represents photons emitted by the nucleus following a primary interaction (i.e. from nuclear de-excitation), while "$\pi^0\gamma$" represents photons truly originating from the decay of a π^0 created within a primary interaction.

If the daughter is considered track-like, an attempt is made to identify charged pions by identifying particles that appear to be a Minimum Ionizing Particle (MIP). This MIP-like determination is done first by looking at the energy deposited per unit length by the

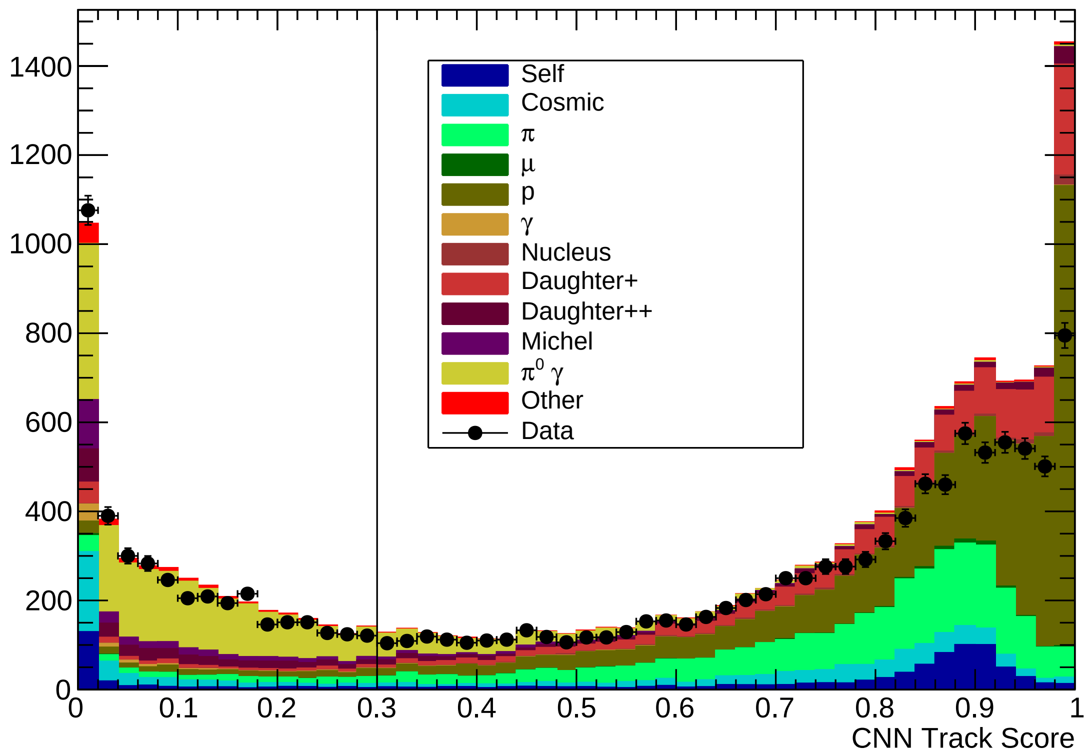
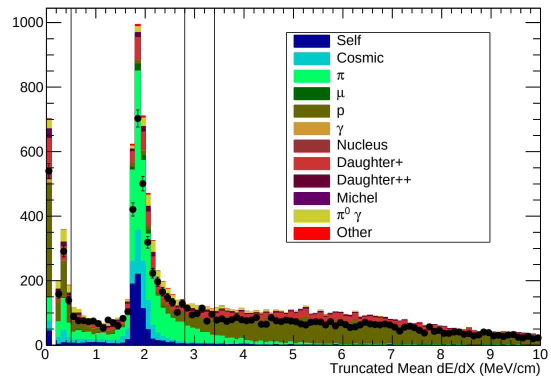
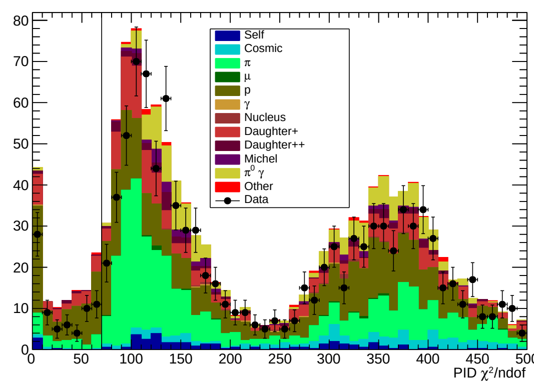


Figure 6.3: CNN Track scores of all reconstructed particles associated as daughters to the primary beam.

reconstructed track. For this, the truncated-mean dE/dx (defined to be the total energy deposited by a reconstructed hit divided by the track pitch of that hit) is used in order to exclude the large energy deposits from stopping particles. In its calculation, the lowest 16% and highest 16% of hits in a track are ignored. The distribution of the truncated mean dE/dx for all daughter tracks is shown in Figure 6.4a. Particles are immediately considered MIP-like if they fall between 0.5 and 2.8 MeV/cm, and are considered not MIP-like if they are above 3.4 MeV/cm. For other particles (those that fall below .5 MeV/cm or between 2.8 and 3.4 MeV/cm), another step is done in the selection. This step consists of comparing the dE/dx of each hit in the track to the expectation value for protons and producing a χ^2 value. Ideally, protons should have a low χ^2 and pions should have a high χ^2. This is shown in Figure 6.4b. These particles are considered MIP-like if they have a χ^2 above 70.



(a) Truncated mean dE/dX of all reconstructed tracks associated as daughters to the primary beam.

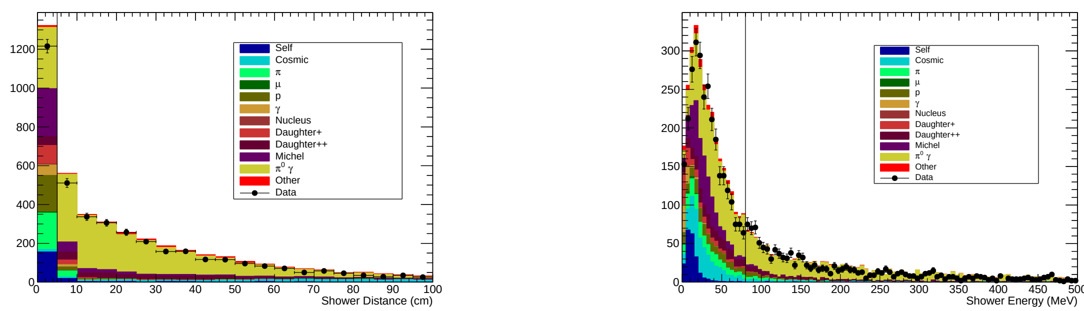


(b) PID χ^2 value of all reconstructed tracks associated as daughters to the primary beam.

Figure 6.4: Distributions used for combined absorption and charge and exchange selection. Vertical black lines represent the cuts used. Note that the events were separated into multiple regions of truncated mean dE/dX, as indicated by the multiple black lines on the left plot.

If any of the daughter particles appears MIP-like, it is considered to be a charged pion originating from the primary interaction, and the event is rejected from the absorption and charge exchange selection.

Following the combined absorption and charge exchange selection, these interactions are separated by attempting to identify showers originating from the decay of π^0 daughters. A daughter shower is considered as coming from a π^0 decay if it is at least 5 cm away from the end of the primary track and has at least 80 MeV of energy. These cuts are chosen to exclude any activity around the interaction vertex which originated from lower-energy pions and protons or nuclear de-excitation photons from both the primary and downstream interactions. This can be seen in Figure 6.5.



(a) Distance between the end of the reconstructed beam track and start of showers associated as daughters to the primary track.

(b) Total deposited energy of all showers associated as daughters to the primary track.

Figure 6.5: Distributions used to separate absorption from charge exchange. The black vertical lines represent the cuts used.

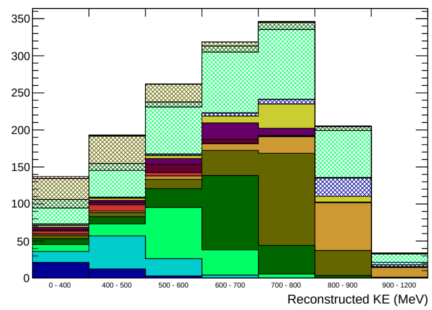
Channel	Efficiency	Purity
Absorption	0.53	0.52
Charge Exchange	0.23	0.80

TABLE 6.1: Effiency and purity of the signal categories.

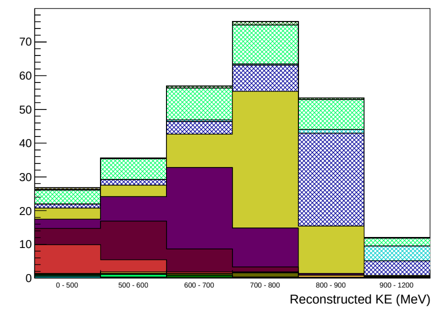
6.5 Binning

The bin variables used for the event selection categories described in the previous section are as follows. Events that fall in the first two categories (no beam track and events that fail the beam cuts) are each placed in single, unitless bins. Events that end past the FV cut are binned according to their SCE-uncorrected ending position in z. Finally, the three "interaction" categories (absorption, charge exchange, and other) are binned according to their ending kinetic energy. This is determined by first calculating their reconstructed kinetic energy using the reconstructed beam line momentum, assuming they are pions, and then subtracting the energy of each collection plane hit up to but not including the last. Occasionally, large hits from large amounts of vertex activity or from crossing cosmic tracks saturate the cold electronics, resulting in seemingly enormous reconstructed energy deposits on the order of a few hundred to a thousand MeV. Thus, any hit above 80 MeV is ignored in the calculation. This value was chosen such that the saturated electronics hits are skipped, but truly large energy deposits like from overlapping hits are kept.

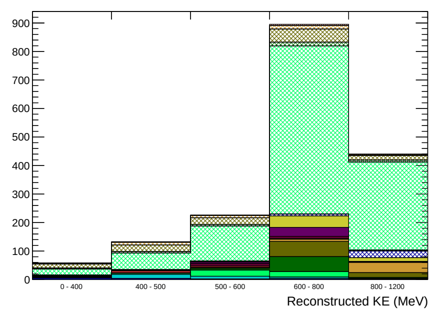
This binning is shown in Figure 6.6, where the reconstructed distributions from the nominal MC are shown. Additionally, these distributions are broken down by their true category (i.e. a signal interaction in a given energy bin or a muon). These are displayed as the stacked histograms, where the different colors represent the specific true category. The bin edges for the interaction distributions were chosen based on the smearing between true and reconstructed kinetic energy shown as the spread in the different colored portions of the stacks. The purity and efficiency for the absorption and charge exchange selections is also shown in Table 6.1.



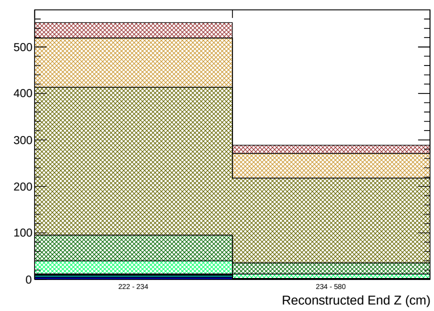
(a) Selected Absorption



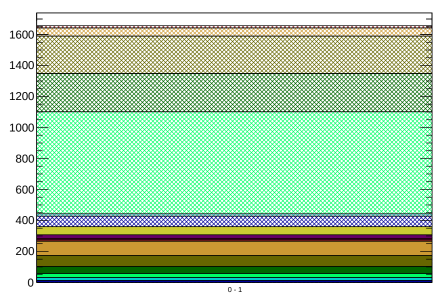
(b) Selected Charge Exchange



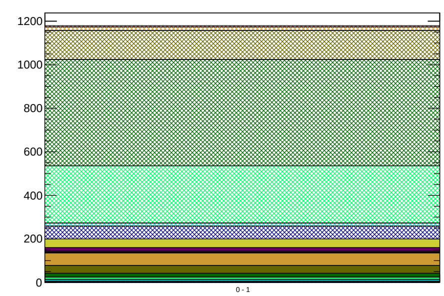
(c) Selected other interactions.



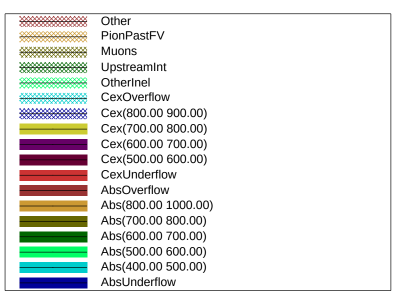
(d) Selected as ending past FV



(e) Failed beam cuts



(f) No reconstructed beam track



(g)

Figure 6.6: Reconstructed distributions of events from the nominal MC. The distributions are broken down by true categories, shown in 6.6g.

6.6 Selected MC Event Displays

This section provides some examples of successes and failures in the event selection within the MC sample used in the analysis. Shown in the following figures are reconstructed event displays in the view of the collection plane near the beam entrance. The x-axis is the wire number which is equivalent to the position in z. The y-axis is the time (or tick) at which the charge reached the wire plane. This is equivalent to the horizontal position away from the wire plane.

The first example, shown in Figure 6.7, is a true absorption interaction correctly selected as absorption. The pion is shown as the yellow track entering from the left, and it interacts with a nucleus. The interaction produces two protons. These are correctly identified as proton-like tracks by the event selection and are shown as the light blue and pink tracks in the display.

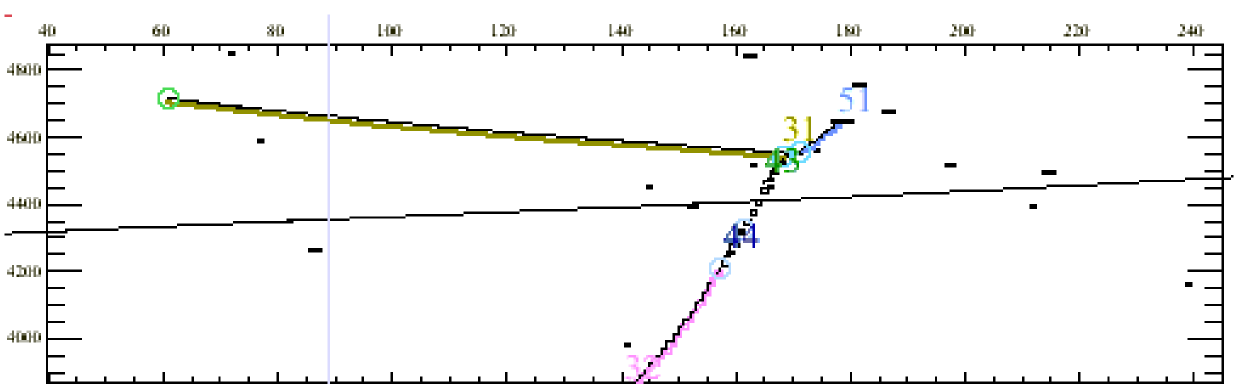


Figure 6.7: MC absorption event correctly identified as absorption.

Next, Figure 6.8 shows a true charge exchange event incorrectly identified as an absorption event. Again, the beam enters from the left shown as the tan track, and interacts with a nucleus. A very energetic proton exits the interaction and travels toward the lower right of the display shown as the green track. This proton is correctly identified as a proton. However, a π^0 also exits the interaction. Near the vertex, one of the γs produced by the decay of this is identified as a small shower (represented by the black rectangles near the vertex). Its reconstructed energy is too low to be identified as resulting from a π^0. The other

γ is not identified. The pink track extending from the top to bottom of the plot is a cosmic muon.

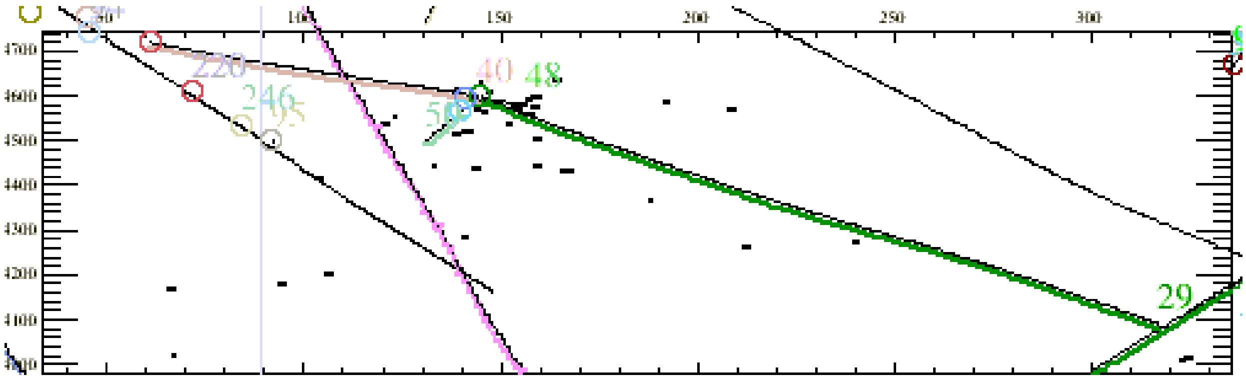


Figure 6.8: MC charge exchange event incorrectly identified as absorption.

The third example in Figure 6.9 is a background inelastic event (one a charged pion in the final state) selected as absorption. The beam pion enters from the left (shown as the red track), and strikes a nucleus. Both a π^+ and π^0 exit the interaction. The π^0 promptly decays, and the resulting showers are not associated to this event as daughters. The π^+ is reconstructed as the tan track exiting the interaction, but it does not appear to be a pion when its calorimetry information is checked in the event selection procedure. The light blue track extending from the top right to the bottom left is a cosmic muon.

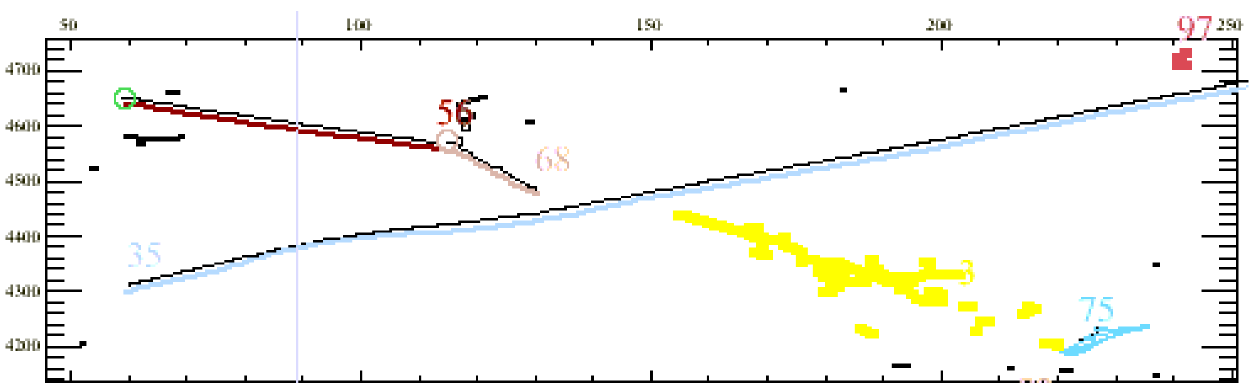


Figure 6.9: MC background inelastic event incorrectly identified as absorption.

The next example in Figure 6.10 shows the pion as a tan track entering from the top left before interacting with a nucleus. A resulting proton is reconstructed as the light blue track heading toward the bottom of the figure. A π^0 exits the interaction and promptly decays.

The resulting photons are reconstructed as the red and yellow showers, and identified as such.

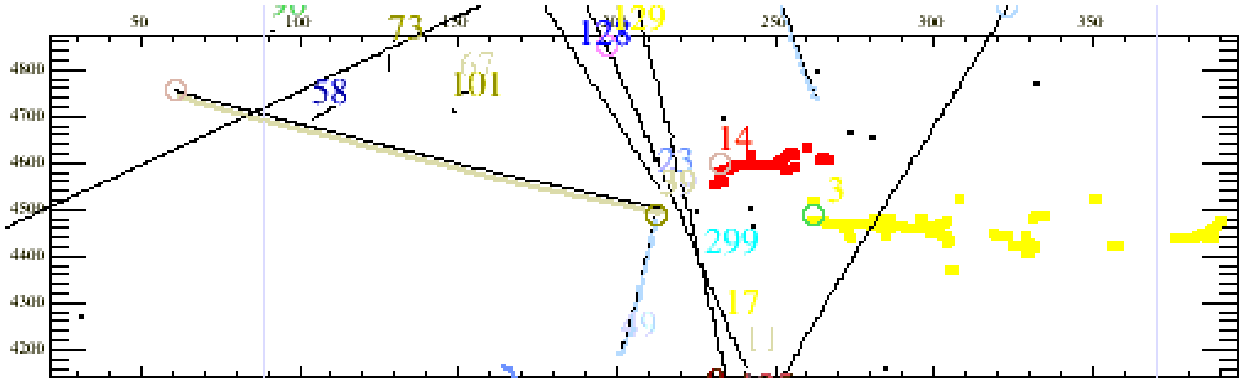


Figure 6.10: MC charge exchange event correctly identified as charge exchange.

Figure 6.11 shows an absorption event misidentified as charge exchange. The beam pion is reconstructed as the pink track and interacts with a nucleus. A neutron exits the interaction before itself interacting and resulting in a proton track (the tan track toward the right). A proton also exits the interaction, but was reconstructed as a shower (represented by the black boxes at the end of the track). This proton appears as a π^0 shower and so the event is selected as charge exchange.

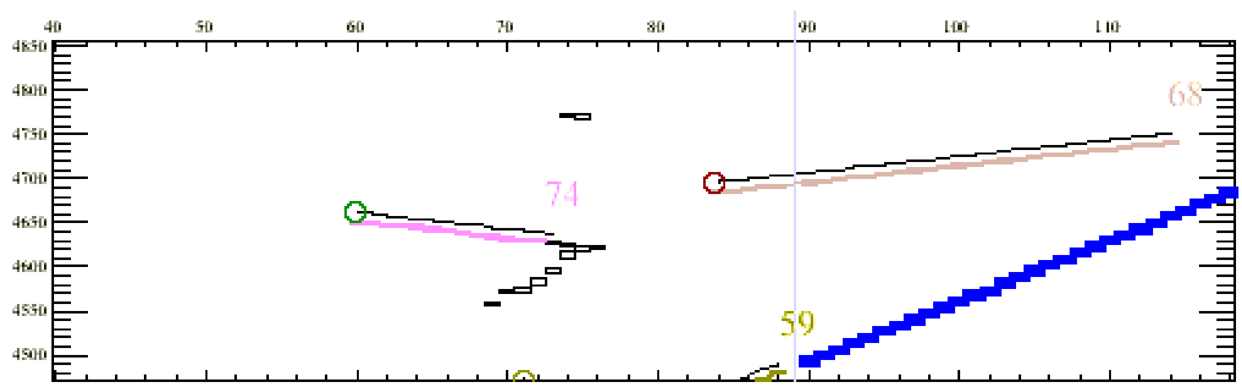


Figure 6.11: MC absorption event incorrectly identified as charge exchange.

Figure 6.12 shows a background inelastic misidentified as charge exchange. The pion, reconstructed as the tan track, enters from the left and ends in an inelastic interaction. A high energy π^+ exits the interaction and reinteracts nearby the primary interaction, resulting in a charge exchange event. The π^0 from the secondary interaction decays, and a shower is reconstructed and associated to the primary interaction.

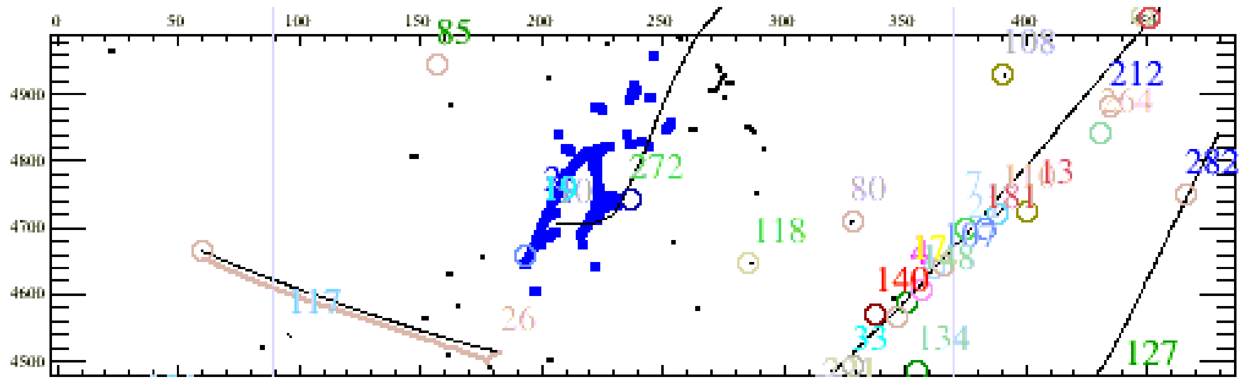


Figure 6.12: MC background inelastic event incorrectly identified as charge exchange.

Finally, Figure 6.13 shows a muon that is misidentified as a background inelastic interaction. The muon, reconstructed as the yellow track, enters from the top left of the plot and its track is ended prematurely. The remainder of the muon is reconstructed as a MIP-like track (the red track) and associated as a daughter to the primary track. The blue and yellow tracks toward the left of the figure extending from top to bottom are cosmic muons.

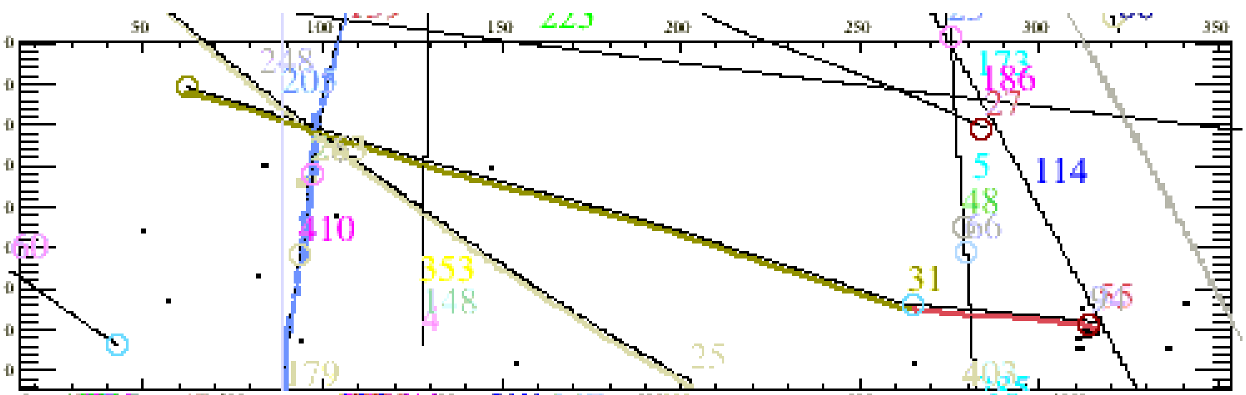


Figure 6.13: MC muon incorrectly identified as a background inelastic interaction.

6.7 Selected Data Event Displays

This section provides example events in 5 of the 6 selection categories (all but the category "no-track" category) used in the fit to data. The dataset containing these events, Run 5809, is different from the one used to display the event selection cuts.

The first example in Figure 6.14 is a selected absorption event. The pion candidate enters from the left, and appears to interact with a nucleus. The reconstruction does not associate any tracks as daughters to this primary particle. Despite this, there appear to be a pair of heavily-ionizing protons exiting the interaction. A cosmic muon crosses the primary track in a nearly-vertical trajectory, and a pair of cosmic muons appear toward the right.

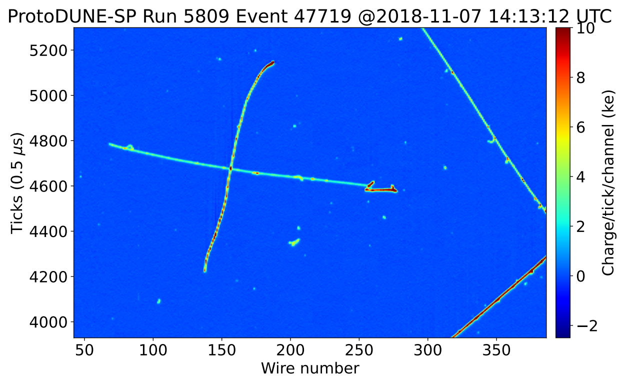


Figure 6.14: Selected absorption event.

The second example shown in Figure 6.15 is a selected charge exchange event. The pion candidate enters from the left, and appears to interact with a nucleus. Clearly seen after the interaction is an apparent shower structure resulting from the decay of a π^0.

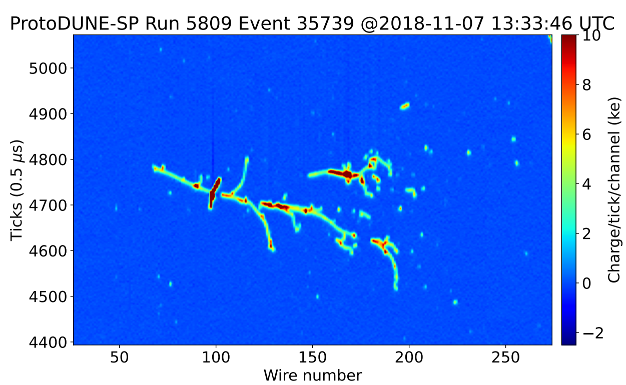


Figure 6.15: Selected charge exchange event.

The third example in Figure 6.16 is a selected background inelastic interaction. The pion candidate enters from the left, and results in an interaction with multiple particles exiting. A daughter pion candidate travels from the interaction toward the top right of the plot.

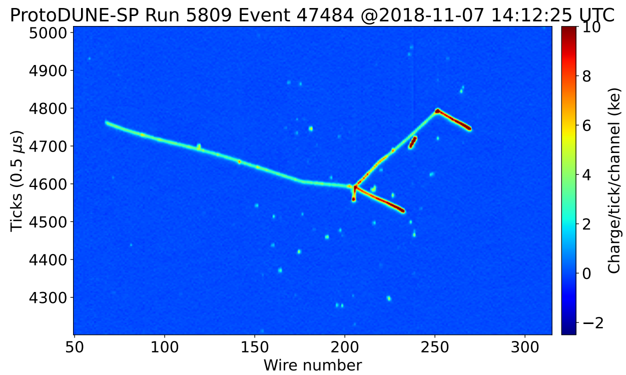


Figure 6.16: Selected background inelastic interaction event.

The fourth example in Figure 6.17 is a π^+/μ^+ candidate extending past the fiducial volume. The primary particle appears to come to a stop near wire number 700. A break in the particle's ionization track is seen near wire number 500. This is the dead region caused by the grounded electron diverters. Additionally, a cosmic muon is seen crossing the primary track.

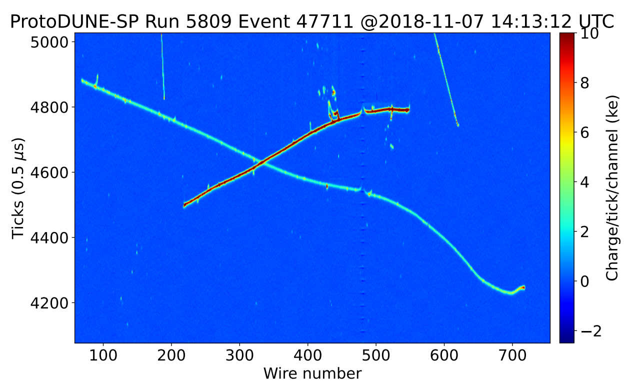


Figure 6.17: Event selected as extending past the fiducial volume.

Finally, in Figure 6.18 is another π^+/μ^+ candidate that extends past the fiducial volume. This time, however, the reconstruction (not shown) ends near the grounded electron diverters. The remainder of the primary particle's ionization to the right of the grounded electron diverters is reconstructed as a separate track, and is associated as a daughter to the primary track.

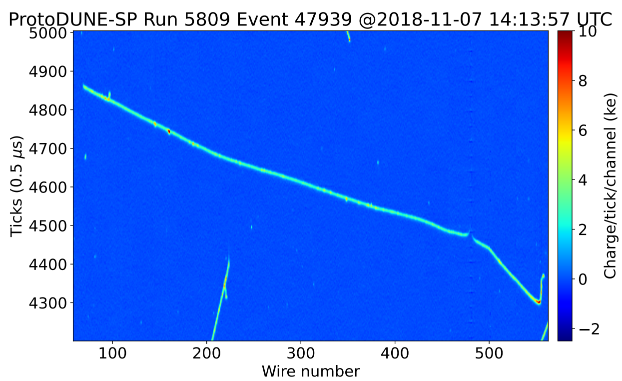


Figure 6.18: Event selected as extending past the fiducial volume, and specifically ending near the electron diverter region.

CHAPTER 7

CROSS SECTION MEASUREMENT TECHNIQUE

This analysis measures π^+ - Ar absorption and charge exchange cross sections using beam-triggered events in ProtoDUNE-SP. The measurement employs a fit which extracts the number of signal (absorption and charge exchange) interactions as well as the number of background events (incident muons, non-signal interactions, stopping pions) from this data. Truth-level information (information representing the exact results of the simulation, rather than reconstructed information taken from a simulated detector response) is then used to extract the cross section according to a technique derived from the Liquid Argon in A Test Beam experiment (LArIAT) [50]. That technique, known as the "Thin Slice Method" (described in Section 7.2) was used to measure hadron cross sections using a LArTPC wherein the detection medium (LAr) also serves as the target. This method is distinct from measurements using thin targets. This chapter first describes these thin target cross section measurements, as well as the Thin Slice Method. It then specifies how the Thin Slice Method is used on truth information to extract the cross section from simulation. It then describes the statistical fit used to interpret the data.

7.1 Thin Target Cross Section Experiment

Historically, hadron scattering experiments have been performed by firing a beam of particles onto a thin piece of material as a target. By counting the number of interactions, the cross section for an interaction can be measured as a function of the incident energy (since the target is thin, a negligible amount of energy is lost before an interaction, and the cross section is measured at the incident beam energy). A simple cartoon of the experimental setup can be seen in Figure 7.1. Here, a beam of pions of width A and flux Φ impinges on a target of thickness t. After passing through the target, N_{Inter} pions have interacted, while N_{Surv} have passed through without interacting. The cross section can be extracted from

Equation 7.1.

$$\frac{N_{Inter}}{\Phi A} = \frac{N_{Inter}}{N_{Inc}} = 1 - e^{-nt\sigma} \tag{7.1}$$

Here, σ is the cross section for the relevant interaction, n is the number density of atoms in the target material, and N_{Inc} is the number of incident pions as given by $\Phi \times A$. This can be slightly simplified by expanding the exponential term around t as such:

$$\frac{N_{Inter}}{\Phi A} = \frac{N_{Inter}}{N_{Inc}} \approx 1 - (1 - nt\sigma + \mathcal{O}(t^2)) = nt\sigma \tag{7.2}$$

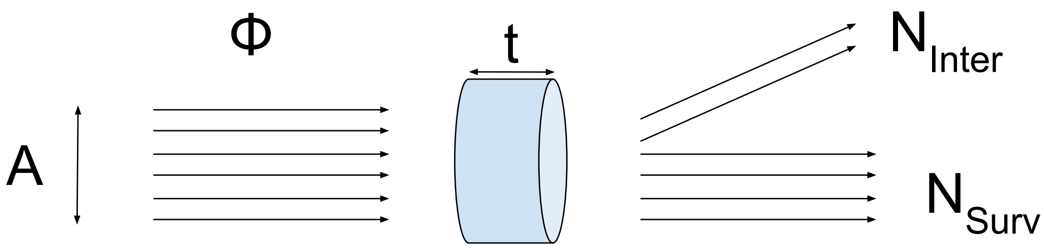


Figure 7.1: Cartoon of a thin target scattering experiment.

7.2 The Thin Slice Method

By virtue of being a LArTPC, ProtoDUNE-SP is not thin, and thus cannot be used for the simple thin target experiment as described above. However, LArIAT [50] used a method they called the Thin Slice Method to mock-up a series of multiple thin target experiments in an extended volume of LAr in order to measure hadronic cross sections in an extended LAr volume. The segmentation created by the collection plane wires allows analyzers to treat an extended volume of LAr as if it were multiple thin targets stacked in front of one another.

This can be seen in Figure 7.2, where a cartoon of a pion track in a LArTPC is shown. The vertical dashed lines represent the collection wires of the TPC, and the red dot represents the point at which the pion interacts. One can treat every slice the pion passes through (up to and including the slice which contains the interaction) as a separate thin target experiment. In each of these, the pion enters the slice and either interacts, or decays. From this, one can count the number of incident pions (N_{Inc} as described above) by counting the number of times a pion enters a slice (it passes by a new wire) and the number of interactions (N_{Inter}) to extract the cross section.

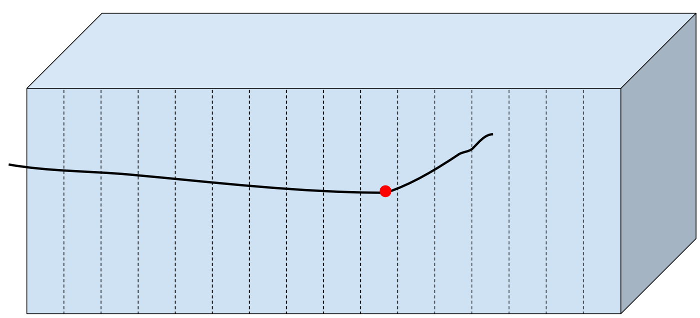


Figure 7.2: Cartoon of the thin slice method applied to a pion track within a LArTPC. The red point represents a hadronic interaction.

If the energy of the pion is known as it enters each slice, then energy-dependence is added to Equation 7.2, as reflected in Equation 7.3. Here, the thickness t is the width of the wire spacings.

$$\frac{N_{Inter}(E)}{N_{Inc}(E)} = nt\sigma(E) \tag{7.3}$$

Mechanically, this calculation is achieved by using two histograms thus called “Incident” and “Interacting” which respectively represent the denominator and numerator of Equation

7.3. As the pion enters into a new slice, the Incident histogram is filled at the corresponding energy. This is done for the entire pion track up to the end, meaning a track can contribute multiple entries in the histogram. For example, in Figure 7.2, the pion track will contribute an entry for every section of Ar up to and including the interaction point (represented by the red dot). If the pion undergoes an interaction of interest, the Interacting histogram is filled according to the energy of the pion as it entered the final slice (this will be the same energy for the final entry into the Incident histogram). A demonstration of this is shown in Figure 7.3.

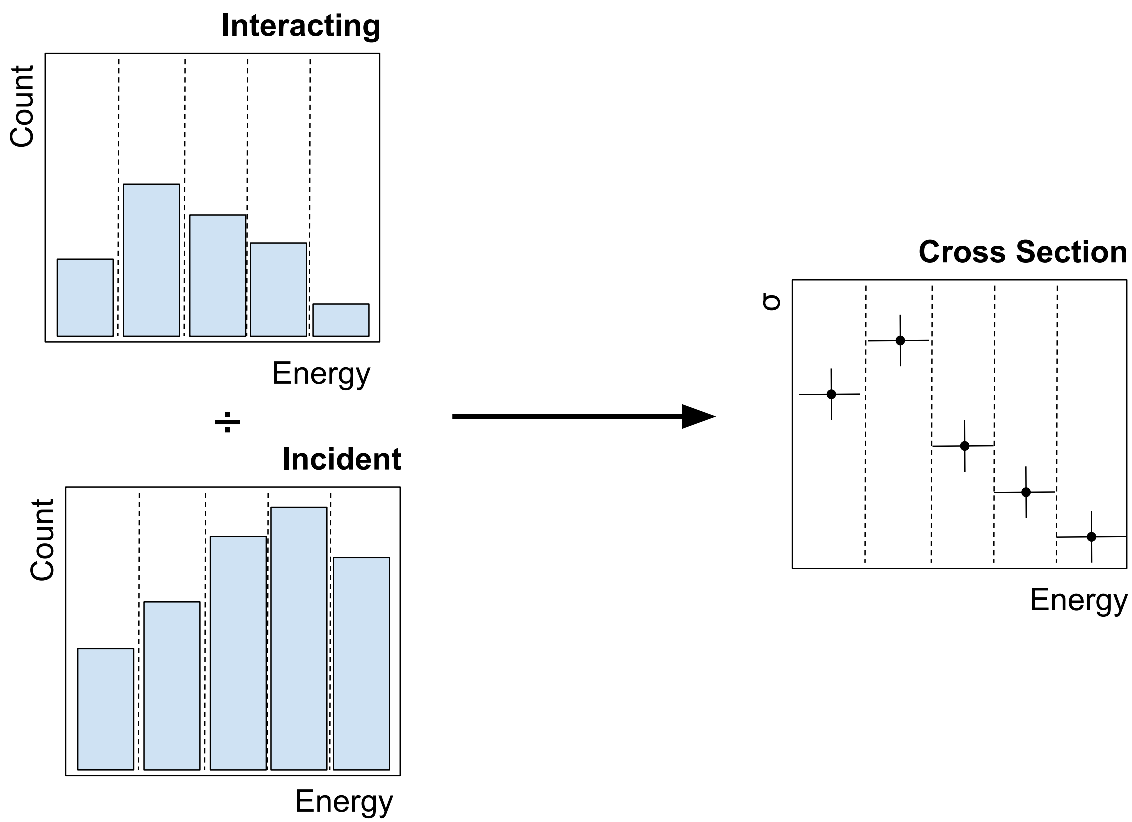


Figure 7.3: Demonstration of the cross section calculation using Equation 7.3.

7.3 Thin Slice Method on Truth Information

The previous section described how the Thin Slice Method could be used on reconstructed information to determine hadronic cross sections. The measurement presented in this thesis is slightly different, but is generally based on this method. Rather than using reconstructed information to determine the Incident histogram, it is taken directly from truth information from ProtoDUNE-SP Monte Carlo simulation. This simulation will be modified by performing a fit to collected data. This fit, known as a "template fit" which is described later in Section 7.4, will vary the number of signal and background interactions (binned in true ending kinetic energy) within the MC. This will of course change any true Interacting histogram created from this information. It will, in turn, also change the true Incident histogram, as the number of slices (equivalently, the distance traveled by the pion) depend on the pion's starting and ending energy. In this way, the varied MC which best describes the data can be used to extract a varied cross section. This section describes the procedure used to extract the cross section from truth information.

The ProtoDUNE-SP MC simulation contains a set of π^+ and μ^+ created by the beam impinging on the detector. Pions that interact before the start of the LAr are ignored and do not contribute to the Incident distribution. For all other pions (those that enter into the TPC), their energy at the initial TPC point (E_0) is used as an entry in the Incident distribution. Using a uniform spacing[1], the energy deposited by the pion as it was simulated by Geant4 is separated into "slices". The energy at each slice boundary crossed by the pion is calculated by summing the energy deposited in the previous slice and subtracting that from the previous incident energy. Thus, the energy as the pion crosses slice boundary i is equal to $E_{i-1} - \delta E_{i-1,i}$ where $\delta E_{i-1,i}$ is the energy deposited between slice boundaries $i-1$ and i. This is demonstrated in Figure 7.4, where the labels E_i represent the energy of the pion as it crosses each slice boundary. All of the energies after E_0 are then given an entry

[1]Note, the width of the spacing to extract the cross section from truth info is arbitrary. For this analysis, the wire spacing (.47974cm) was used.

in the Incident distribution as well. This occurs for every pion that reaches the TPC, and along each pion up to some fiducial volume edge. Then, for each pion ending in a signal interaction within the fiducial volume, the energy of the pion at its interaction point is used as an entry in the Interacting histogram. The resulting Interacting and Incident histograms are used as in Equation 7.3 and Figure 7.3 to compute the cross section.

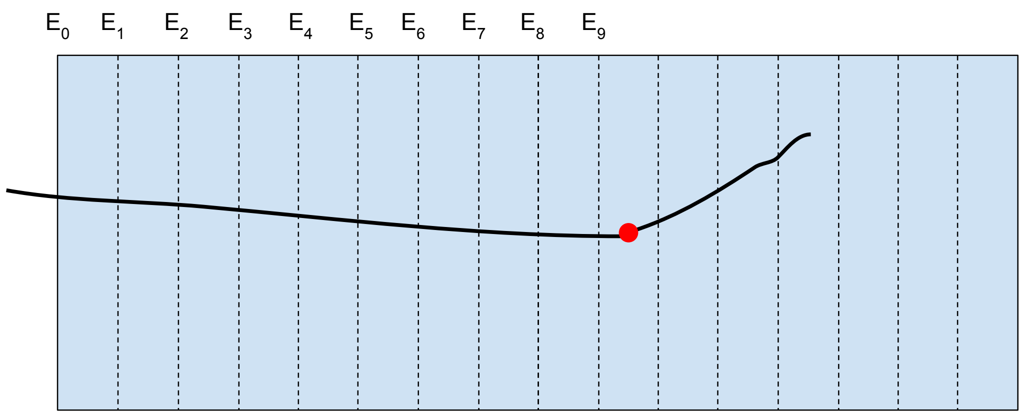


Figure 7.4: Cartoon diagram showing a pion track split up into multiple slices and the energy denoted at each slice boundary.

It is instructive to consider this measurement technique under a varied cross section model. If the cross section is higher over the momentum range of the simulated pions, more interactions will occur (N_{Inter} will be higher). The pions will (on average) travel through less slices before they interact, and thus contribute less events to the Incident histogram. For an overall lower cross section, the inverse is true: less interactions occur, and the pions travel further on average (creating more entries in the Incident histogram). This line of thought can be extended to more complicated variations in shape as well. The number of interactions at a given energy will change, and so too will the entries in the Incident distribution. This serves as the guiding principle used in this measurement: if one is able to measure the number of interacting pions at a given energy (and equally importantly, the number of pions that do not interact), the Thin Slice Method can extract cross sections using truth information from a varied Monte Carlo simulation that best describes the data. The following section describes the fit strategy used to interpret the data in terms of a varied Monte Carlo sample in order to extract the absorption and charge exchange cross sections in this manner.

7.4 Fit Strategy

This analysis uses a binned maximum likelihood fit to 1 GeV/c momentum ProtoDUNE-SP beam line events to estimate the number of signal and background interactions in the data set. The fit results in a set of varied MC which best matches this data and from which the signal cross sections are extracted. A set of signal parameters ($\vec{\theta}$) and nuisance (also called systematic) parameters ($\vec{p}$) controlled by the fit vary simulated π^+ and μ^+ events. The fit attempts to find the set of parameters that best describe the data by maximizing the likelihood $L(\vec{\theta}, \vec{p}; \vec{n})$ to observe a set of events $\vec{n}$ given the model parameters $\vec{\theta}$ and $\vec{p}$. Additionally, we include constraints to the nuisance parameters represented by predictions of their central values $\vec{q}$ and the uncertainties on these predictions represented by a covariance matrix V_{Cov}. As such, the likelihood L is made of two components: a statistical term and a systematic term:

$$L(\vec{\theta}, \vec{p}; \vec{n}) = L_{\text{Stat}}(\vec{\theta}, \vec{p}; \vec{n}) L_{Syst}(\vec{p}; \vec{q}, V_{\text{Cov}}) \tag{7.4}$$

For compatibility with the fitting routines (discussed later) in finding the best fit parameters and their uncertainties, the minimum of twice the negative log-likelihood ($-2 \ln L$) is found instead of the maximum likelihood[2]. Additionally, minimizing this value is equivalent to minimizing twice the negative of the natural logarithm of the likelihood ratio λ [29]. The likelihood ratio is defined as

$$\lambda = L(\vec{\theta}, \vec{p}; \vec{n}) / L(\vec{\theta_T}, \vec{p_T}; \vec{n}) \tag{7.5}$$

[2]The fitting routines implemented in ROOT work by minimizing rather than maximizing some value.

where $\vec{\theta_T}, \vec{p_T}$ represents the true, unknown underlying model. Plugging Equation 7.4 into this results in the following.

$$\lambda = L_{\text{Stat}}(\vec{\theta}, \vec{p}; \vec{n}) L_{\text{Syst}}(\vec{p}; \vec{q}, V_{\text{Cov}}) / L_{\text{Stat}}(\vec{\theta_T}, \vec{p_T}; \vec{n}) \tag{7.6}$$

Where there is no true value of L_{Syst} shown in the denominator, as it is trivially equal to one. $-2\ln\lambda$ is thus defined as

$$-2\ln\lambda = -2\ln\left(L_{\text{Stat}}(\vec{\theta}, \vec{p}; \vec{n}) / L_{\text{Stat}}(\vec{\theta_T}; \vec{n})\right) - 2\ln L_{\text{Syst}}(\vec{p}; \vec{q}, V_{\text{Cov}}). \tag{7.7}$$

In this fit, we are seeking to categorize a fixed number of events (a set of beam line-triggered events) based on the results of ProtoDUNE-SP reconstruction described in Section 4.4. As such, the likelihood L_{Stat} is the multinomial likelihood as defined in Equation 7.8.

$$L_{\text{Stat}}(\vec{\theta}, \vec{p}; \vec{n}) = N! N^N \prod_j y_j(\vec{\theta}, \vec{p})^{n_j} / n_j! \tag{7.8}$$

Here, $y_j(\vec{\theta}, \vec{p})$ and n_j are the number of predicted and measured events in reconstructed bin j, and $N = \sum_j n_j = \sum_j y_j(\vec{\theta}, \vec{p})$ is the total number of beam line events. As stated before, $L_{\text{Stat}}(\vec{\theta_T}, \vec{p_T}; \vec{n})$ depends on some true underlying model denoted by $\vec{\theta_T}, \vec{p_T}$. This model is unknown, but $L_{\text{Stat}}(\vec{\theta_T}, \vec{p_T}; \vec{n})$ is estimated using the measured events as shown in Equation 7.9.

$$L_{\text{Stat}}(\vec{\theta_T}, \vec{p_T}; \vec{n}) = N! N^N \prod_j n_j^{n_j} / n_j! \tag{7.9}$$

From this, the statistical portion of $-2\ln\lambda$ is defined as follows.

$$-2\ln\lambda_{\text{Stat}} = 2\sum_j n_j \ln\frac{n_j}{y_j} \tag{7.10}$$

The systematic term $-2\ln\lambda_{\text{Syst}}$ is a constraint term that assumes the systematic parameters $\vec{p}$ are Gaussian distributed around their central values $\vec{q}$ and whose uncertainties are described by a covariance V_{Cov}:

$$-2\ln\lambda_{\text{Syst}} = \sum_{i,j}(p_i - q_i)(V^{-1}_{\text{Cov}})_{ij}(p_j - q_j). \tag{7.11}$$

With this, the full statistic minimized by the fit is given by Equation 7.12.

$$-2\ln\lambda = 2\sum_j n_j \ln\frac{n_j}{y_j} + \sum_{i,j}(p_i - q_i)(V^{-1}_{\text{Cov}})_{ij}(p_j - q_j) \tag{7.12}$$

A crucial step in the analysis is the extraction of true information (the set of true events from which the number of signal interactions and slices which form the cross section calculation as in Section 7.3) from reconstructed quantities. In general, this is known as "unfolding" and is a common problem within High Energy Physics [51]. Several unfolding techniques exist, each with their own benefits and drawbacks (typically, a balance is made between biased results, bin-to-bin correlations, uncertainty, and smoothness) [52]. The fit done within this analysis, known as a template fit, performs the role of unfolding. Included in the set of parameters are a set of "template weights" assigned to the MC signal events which vary the normalization of signal events in a given true energy bin, and which also have a subsequent effect on the predicted reconstructed distributions. The fit simultaneously varies the template weights and the other parameters, then compares the resulting predicted reconstructed distributions to the measured distributions until it converges at a minimum $-2\ln\lambda$ value.

The role of the template parameters is highlighted in Equation 7.13, which shows the relationship between the true and reconstructed events as predicted by MC. $\hat{y}_i$ represents the number of events in true bin i for the indicated true category (absorption, charge exchange, muon background, or pion backgrounds). The events have a chance ϵ^k to be selected as some selection category k when the reconstructed information is passed through the event selection (described in Section 6). Reconstruction effects smear the events from some true bin i to some reconstructed bin j in selection category k. This is represented by $t^k_{i,j}$ which can be thought of as a "smearing matrix." In general, ϵ^k and $t^k_{i,j}$ depend on the true category they act on. $\hat{y}_i$, ϵ^k and $t^k_{i,j}$ all depend on some subset of the fit parameters $\vec{\theta}$ and can be modified at each step in the fit. A parameter f^μ is used to vary the normalization of muons in the sample, as this is uncertain. Lastly, c_i^{Abs}, c_i^{Cex} are the template parameters that control the normalization of absorption and charge exchange events in true bin i. The sums extend over the number of true bins n_T for the different true categories. Since the number of impinging π^+/μ^+ is known and static, the fit is constrained as in Equation 7.14.

$$y_j^k = \sum_i^{n_T} c_i^{\text{Abs}} \hat{y}_i^{\text{Abs}} \epsilon^k t_{i,j}^k + \sum_i^{n_T} c_i^{\text{Cex}} \hat{y}_i^{\text{Cex}} \epsilon^k t_{i,j}^k + \sum_i^{n_T} f^\mu \hat{y}_i^\mu \epsilon^k t_{i,j}^k + \sum_l^{n_{\pi\text{BG}}} \sum_i^{n_T} \hat{y}_i^l \epsilon^k t_{i,j}^k \quad (7.13)$$

$$N = \sum_j y_j = \sum_j n_j \quad (7.14)$$

In addition to the constraint on the overall number of incident particles, the number of incident particles in bins of true initial momentum (where it was generated by the beam event generator module), is also held constant. This has been omitted from Equation 7.14 for clarity.

Thus, the fit changes $\vec{\theta}$ and $\vec{p}$ until the measured and predicted reconstruction distributions best match. The result of the fit is a set of best-fit parameters $\vec{\theta}_0$ and $\vec{p}_0$ and their covariance which will be used for error propagation as described in the Section 7.5. The best-fit parameter values produce a set of modified MC events that can be used as in Section 7.3 to extract cross sections.

The fit uses the MIGRAD [53] routine of the Minuit2 [54] minimizer library within ROOT [55] to find the maximum likelihood ratio. The MIGRAD routine estimates the gradient of the likelihood ratio surface at each fit point and follows the gradient until it reaches the best-fit point. After finding the best-fit point, the HESSE routine within Minuit2 is called. This computes the Hessian matrix: the second derivative of the $-2\ln\lambda$ surface around the best fit point. The Hessian matrix is inverted to create the covariance matrix which describes the post-fit uncertainties and correlations of the fit parameters.

7.5 Error Propagation

The output of the fit – the best-fit parameters $\vec{\phi}_0$[3] and their associated covariance matrix Σ – can be used to propagate the post-fit errors to the extracted cross sections. First, the Cholesky decomposition [56] of the post-fit covariance matrix is computed. This representation of the covariance matrix (shown in Equation 7.15) is the product of an upper triangular matrix R with positive diagonal elements and its transpose R^T.

$$\Sigma = R^T R \tag{7.15}$$

A random set of fit parameters $\vec{\phi}_t$ (also known as a "throw") can be generated by multiplying a random unit Gaussian vector $\vec{r}_t$ by R and adding this to the best-fit parameter values $\vec{\phi}_0$, as shown in Equation 7.16. $\vec{\theta}_t$ will be randomly distributed with the same covariances of the post-fit covariance matrix [56].

$$\vec{\theta}_t = \vec{\phi}_0 + R\vec{r}_t \tag{7.16}$$

This procedure is repeated on the order of 1000 times to generate an ensemble of throws. Each set of thrown parameters is used to calculate the cross section as described in Section 7.3. The cross section covariance matrix V is computed as in Equation 7.17, where V_{ij} is the covariance between bins i and j, σ_{it} is the cross section in bin i for throw t and σ_{i0} is the best-fit cross section in bin i. Note: the bins i, j include both absorption and charge exchange to account for the covariances between these channels.

$$V_{ij} = \frac{1}{N}\sum_{t}^{N}(\sigma_{it} - \sigma_{i0})(\sigma_{jt} - \sigma_{j0}) \tag{7.17}$$

[3]The set of parameters $\vec{\phi}$ includes both the signal parameters $\vec{\theta}$ and systematic parameters $\vec{p}$.

If any parameter is thrown into an unphysical region (i.e. for the template parameters, below zero), the throw is repeated until all parameters are within their allowed regions. This may results in truncated Gaussian distributions for any parameters that experience this issue. If this truncated area is small, the distribution is considered valid and has a negligible effect on the cross section covariance.

This throwing procedure makes an assumption that the likelihood surface around the best fit point is distributed according to a multivariate Gaussian. If this assumption holds, the covariance matrix from the fit describes a multidimensional contour with constant χ^2 around the best-fit point which represents the probable spread of fit parameters. Additionally, the cross section covariance created by this propagation procedure describes a constant-χ^2 contour centered around the best-fit cross section point [57][58].

CHAPTER 8

SYSTEMATIC UNCERTAINTIES

This chapter describes the systematic uncertainties and their implementation within the analysis. The uncertainties discussed stem from the dE/dx calibration, the reconstructed beam line momentum,the modeling (via Geant4) of the hadrons as they pass through the detector, the effect of the electron diverters on reconstructed track, and differences in the rate of both events without a reconstructed track and those failing the beam cuts. These uncertainties are parameterized within the fit and are constrained by a covariance within the systematic term given in Equation 7.11.

8.1 dE/dX Calibration

Section 4.5.3 describes how the measured charge per unit distance dQ/dx is translated into the energy deposited per unit distance dE/dx (which is used for the energy measurements of particles in this analysis). Part of this dE/dx extraction is the determination of a calibration constant C_{cal} by analyzing stopping muons. There is some uncertainty in what this calibration constant is, and as such, it has been implemented as a systematic parameter in the fit.

As C_{cal} is varied within the fit, it has two large effects. The first is to change the MIP-like separation of daughter tracks during the event selection as described in Section 6.4, and the second is to migrate events between bins since more apparent energy will be accounted for in the energy reconstruction. This parameter was first implemented within the fit by rescaling the dE/dx in each step of the fit where the prediction histograms are refilled before comparing to data. This caused instability within the fit, as events would fail to migrate bins until the parameter was turned enough. This "threshold" behavior caused discontinuities in the $-2\ln\lambda$ surface, and so a different approach was opted for. Instead of implementing this effect directly on the events, a weighting scheme was implemented, where varied MC samples were created for various values of C_{cal}. In each bin of the prediction histograms, the ratio to nominal was taken to form a weight for that bin and C_{cal}. These weights were then interpolated between in order to form a smoothly-varying surface that could be used within the fit. Each step of the fit, the events are given a weight which depends on the bin the event falls into and the value of C_{cal} for that fit step.

8.2 Beam Momentum

Section 5.1.2 describes how the beam line instrumentation reconstructs momentum using sets of fiber monitors surrounding a bending magnet in the beam line. This section also mentions a bulk shift to the fibers in one of the monitors that affected the reconstructed momentum. This shift was found to be 1.45±0.18mm. In addition to the uncertainty in the shift is an estimated 1% uncertainty on the magnetic field. Recalling Equation 5.2 (repeated here), these systematic uncertainties affect the reconstructed momentum as such: the fiber shift varies θ, which can then cancel out a variation in B.

$$p = \frac{299.7924}{\theta} \times \int_0^{L_{\text{mag}}} (Bdl) \tag{5.2}$$

These parameters would then be degenerate within the fit, and so these effects were combined into a single momentum rescaling parameter c_p. The prior uncertainty on c_p is given by the shifts to p due to variations in both parameters added in quadrature. The effect of the variation to B is trivially 1%. For the effect of the fiber shift, the nominal beam line MC simulation was ran with the fibers in the third monitor shifted by its 1σ uncertainty (0.18mm). This results in an average 0.7% shift in the reconstructed momentum. The uncertainty on c_p is thus given in Equation 8.1.

$$\sigma_{c_p} = \sqrt{.007^2 + .01^2} = .012 \tag{8.1}$$

Within the analysis, the effect of this scaling parameter is to change the difference between true and reconstructed momentum r (defined in Equation 8.2).

$$r = \frac{p_{\text{Reco}} - p_{\text{True}}}{p_{\text{True}}} \tag{8.2}$$

The beam simulation show this is Gaussian distributed with mean μ and width σ. Some variation to c_p will then result in a distribution with varied μ' and σ'. An event can then be given a weight according to its value of r and the μ' and σ' resulting from the value of c_p within one step of the fit. This weight is given in Equation 8.3, which is the ratio of two Gaussian distributions.

$$w = \frac{\sigma}{\sigma'} \exp\left(\frac{(r-\mu)^2}{2\sigma^2} - \frac{(r-\mu')^2}{2\sigma'^2}\right) \tag{8.3}$$

The dependences of μ' and σ' on c_p were found from studies of the beam line MC simulation and used within the fit to form the weights as defined in 8.3.

During fit validation, it was found that this beam momentum parameter created instability in the fit due to its tendency to create extremely large weights for certain events at large parameter variations. This made it difficult to properly assess the post-fit error of the other parameters. As such, this parameter was chosen to be fixed during fits. Its pre-fit uncertainty was propagated to the cross section uncertainties by adding it in quadrature to post-fit parameter covariance matrix.

8.3 Electron Diverter Effect

As shown in Figures 6.1 and 8.1, the simulation of the grounded electron diverters (which causes tracks to prematurely break), differs from data. To account for the uncertainty in the strength of the track-breaking effect, a simple weighting scheme was developed to artificially vary the track-breaking strength.

The weighting scheme varies the fraction of tracks ending above 222 cm, which end in the "track-breaking" region of 222–234 cm. This fraction, f, is defined as

$$f = \frac{N_{\text{Break}}}{N_{>222}} = \frac{N_{\text{Break}}}{N_{>234} + N_{\text{Break}}} \tag{8.4}$$

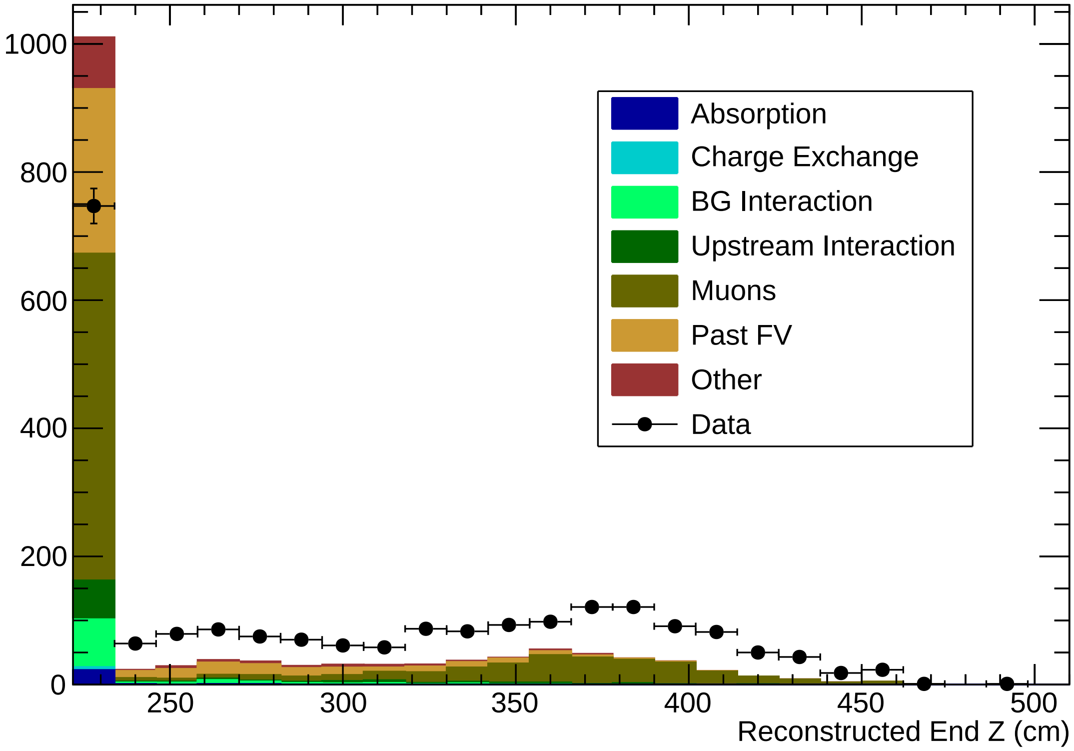


Figure 8.1: Enhanced view of the reconstructed endpoint of beam tracks within the TPC.

where N_{Break} is the number of broken tracks (ending between 222 and 234 cm), and $N_{>222}$ and $N_{>234}$ are the number of tracks above 222 and 234 cm respectively. The probability for a track ending above 222 cm to break is thus f, while the probability for a track to not break is $1 - f$.

Consider some variation as such: $f \rightarrow f' = cf$. Each track ending above 222cm is thus given a weight as follows, depending on if was or was not broken.

$$W_{\text{Break}} = \frac{f'}{f} = \frac{cf}{f} = c \tag{8.5}$$

$$W_{>234} = \frac{1 - f'}{1 - f} = \frac{1 - cf}{1 - f} \tag{8.6}$$

The nominal value of f, the fraction of events ending in the electron diverter region, in MC is 0.6133. The central value of the scale factor c was set to 0.50 as taken from comparisons between MC and data shown here. The uncertainty on this was set naively to 20%.

8.4 Beam Efficiencies

It was found that the Pandora had an apparent difference between data and MC in the efficiency for identifying beam particles in the TPC. Additionally, shape differences in the position and direction of reconstructed beam tracks (possibly due to inaccuracies in mapping SCE as described in Section 4.5.1) created a difference in the fraction of events passing the beam cuts. These two uncertainties were parameterized as efficiency-like effects by varying the numbers of events in the following categories: 1) no reconstructed beam track, 2) reconstructed beam track that fails the beam cuts, 3) reconstructed beam track that passes the beam cuts. Let the fraction of events categorized as such be represented by f_1, f_2, and f_3 respectively. These fractions sum to one ($f_3 = 1 - f_2 - f_1$) and can be varied as follows.

Consider some variation to these fractions (these are, in effect, variations to the two efficiency-like effects):

$$\begin{aligned} f_1 &\rightarrow f_1' = c_1 f_1 \\ f_2 &\rightarrow f_2' = c_2 f_2 \\ f_3 &\rightarrow f_3' = 1 - c_2 f_2 - c_1 f_1 \end{aligned} \tag{8.7}$$

Similar to the previous section, the events are given weights according to how they are categorized:

$$\begin{aligned} W_1 &= \frac{f_1'}{f_1} = c_1 \\ W_2 &= \frac{f_2'}{f_2} = c_2 \\ W_3 &= \frac{f_3'}{f_3} = \frac{1 - c_2 f_2 - c_1 f_1}{1 - f_2 - f_1} \end{aligned} \tag{8.8}$$

The nominal values for the fraction of events with no track or failing the beam cuts in MC are 0.164 and 0.2305 respectively. The central value of the no-track parameter was set to 1.62 taken from comparisons to data and MC, and its uncertainty was naively set to 20%. The central value and uncertainty for the beam cut parameter were naively set to 1.00 and 10% respectively.

8.5 Hadronic Interaction Modelling

In addition to uncertainties in the modeling of the detector systems described in the previous few sections, there are uncertainties in the hadronic interaction model. While the π^+ absorption and charge exchange interactions are measured by this analysis, the rate of background interactions (quasielastic, double charge exchange, production) can differ between data and MC as well. This can lead to wrongly estimated rates of categorization errors within the fit, and cause biased results of the signal interactions. The same is true of the rate of proton interactions as well. Protons are often emitted into the detector as a result of the primary π^+-Ar interactions, and can go on to interact in the nearby argon, producing their own interaction products. These products can influence the event selection and produce categorization errors. Thus, differing rates of proton-argon interactions within data and MC can also bias the cross section results.

To facilitate the propagation of hadronic modeling uncertainties related to the Geant4 stage of the MC simulation (as discussed in Section 4.6), the Geant4Reweight [59] framework was used. This framework is able to create weights for events based on some variation applied to a cross section model in GEANT4. The weights created from this framework work by determining how likely the event was to occur given the nominal cross sections and the set of steps taken by a particle, and then comparing this to how likely the same event was to occur under some variation. The weights are generated under some flat scale factor applied over a user-defined region of momentum. The momentum regions and prior uncertainties for each variation were determined by a crude examination of the spread of models studied within Reference [27]. The description of the systematic parameters are given in Table 8.1.

Geant4Reweight creates a weight for each parameter by running over each π^+ and proton

Channel	Momentum Range	Prior Uncertainty
π^+ Quasielastic	0–500 MeV/c	±36%
π^+ Quasielastic	500–2000 MeV/c	±33%
π^+ Pion Production	0–2000 MeV/c	±33%
π^+ Double Charge Exchange	0–2000 MeV/c	±33%
Proton Reaction	0–2000 MeV/c	±33%

TABLE 8.1: Description of the Geant4Reweight parameters used within the fit.

created within the event and calculating a weight for that particle. These are all multiplied together to create full event weights. For each parameter, a weight is created at intervals of 10% from -90% to +100%. In order to create a smoothly varying effect within the fit, the variations must be interpolated between. Prior to the fit, sets of MC are produced at each variation for each parameter (note: only one parameter is varied at a time). For each truth category and reconstructed bin, the ratio between the varied and nominal MC are calculated and interpolated between using a spline. Then, when running the fit, each parameter contributes a weight to the event corresponding to the value of the spline at the parameter's value. All weights from all Geant4Reweight parameters are multiplied together when creating the predicted distributions for each step of the fit.

8.6 Systematic Covariance Matrix

Table 8.2 summarizes the size of the prior uncertainties that comprise the systematic covariance matrix. Note that all uncertainties described in this section are treated as uncorrelated before the fit.

Parameter	Nominal Value	Prior Uncertainty
$dE/dX\ C_{cal}$	$1.011 \times 10^{-}3$	$\pm 10\%$
Beam Momentum	1.00	$\pm 1.2\%$
Electron Diverter Fraction	0.5	± 0.20
No Track Fraction	1.62	± 0.20
Failed Beam Cuts Fraction	1.00	± 0.10
π^+ Quasielastic Low	1.00	$\pm 36\%$
π^+ Quasielastic High	1.00	$\pm 33\%$
π^+ Pion Production	1.00	$\pm 33\%$
π^+ Double Charge Exchange	1.00	$\pm 33\%$
Proton Reaction	1.00	$\pm 33\%$

TABLE 8.2: Description of the Geant4Reweight parameters used within the fit.

CHAPTER 9

FIT VALIDATION

This chapter demonstrates validation of the fit framework described in Section 7.4. It includes the systematic uncertainties detailed in Chapter 8. In all tests, a set of MC simulation produced according to the 1 GeV/c beam setting is fit to various fake data inputs also produced from MC simulation. These inputs could be the nominal MC or a set of varied MC. The specifics of the fake data will be described in each section.

To evaluate how the fit performed, several quantities will be examined including the post-fit values of the parameters, the extracted cross sections, and a goodness of fit metric. Particular attention will be paid toward the post-fit values of the systematic parameters as they compare to their prior uncertainties. The goodness of fit will be investigated by comparing the minimum $-2\ln\lambda$ (defined in Section 7.4) found by the fit in question to the distribution of minimum $-2\ln\lambda$ found in a set of fits to systematically and statistically varied fake data. This comparison will take the form of a p-value, defined to be the probability of a fit resulting in a $-2\ln\lambda_{\mathrm{Min}}$ at least as extreme as the one in question. This is defined in Equation 9.1 where t_{Fit} represents the $-2\ln\lambda_{\mathrm{Min}}$ of the fit in question, and $f(t)$ is the distribution of $-2\ln\lambda_{\mathrm{Min}}$ found from the set of systematically and statistically varied fake data.

$$p = \int_{t_{\mathrm{Fit}}}^{\infty} f(t)dt \tag{9.1}$$

Figure 9.1 shows the distribution of $-2\ln\lambda_{\mathrm{Min}}$ from 1000 toy fits to systematically and statistically varied fake data. The systematic variations were created with the systematic parameters chosen according to the input covariance matrix (in a manner similar to the post-fit throws described in Section 7.5). Then, each set of systematically-varied fake data was statistically fluctuated. This distribution will be used throughout the following sections to determine p-values for each fit.

Finally, the cross sections extracted from the post-fit MC will be compared to the cross

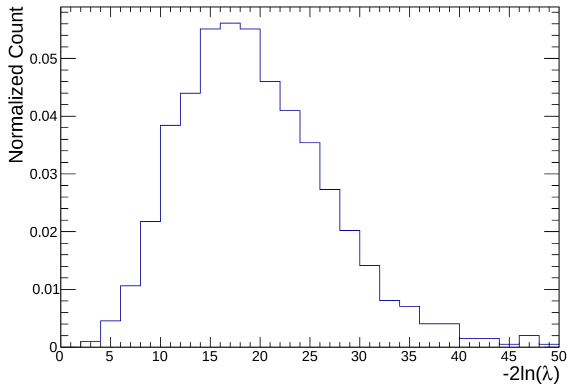


Figure 9.1: Post-fit $-2\ln\lambda$ distribution of the toy experiments described above.

sections as produced by the fake data input using the χ^2 defined in Equation 9.2.

$$\chi^2_\sigma = \sum_{i,j} (\sigma_i - \bar{\sigma}_i)(V^\sigma)^{-1}_{i,j}(\sigma_j - \bar{\sigma}_j) \tag{9.2}$$

Here, σ_i represents the measured cross section in bin i, $\bar{\sigma}_i$ represents the cross section from either the nominal MC or fake data input (this will be specified), and $(V^\sigma)^{-1}_{i,j}$ is the value of bin i, j of the inverted cross section covariance matrix as computed in the error propagation procedure described in Section 7.5. This will be used similar to the minimum fit statistic distribution discussed above to determine a p-value for the cross section results. The distribution of χ^2_σ from the set of 1000 toy fits is shown in Figure 9.2.

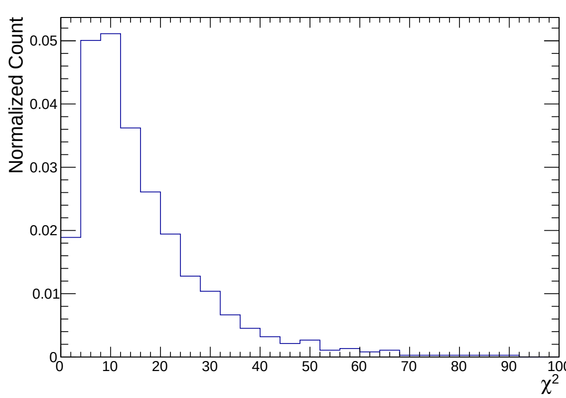


Figure 9.2: χ^2_σ distribution of the toy experiments described above.

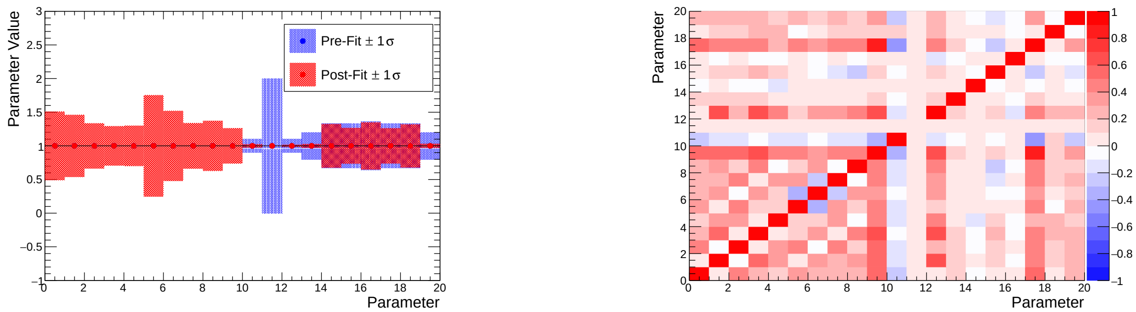
9.1 Asimov Fit

The first validation test is a simple "Asimov" fit. In this fit, the input fake data is the same as the nominal MC within the fit. This tests the base functionality of the fit and whether or not the fit can correctly identify the minimum (the starting point of the fit). It also shows the level of sensitivity the fit has for the signal and nuisance parameters. The results are shown in Figures 9.3, 9.4, and 9.5. The first shows that the best-fit parameters are at the starting point, as expected. The parameters in these plots are enumerated as in Table 9.1. This will be the same for the rest of the chapter.

0	Absorption factor 400–500 MeV/c	10	Beam cut efficiency
1	Absorption factor 500–600 MeV/c	11	Beam momentum resolution
2	Absorption factor 600–700 MeV/c	12	dE/dX calibration constant
3	Absorption factor 700–800 MeV/c	13	Electron diverter effect strength
4	Absorption factor 800–1000 MeV/c	14	Geant4Reweight Double Charge Exchange
5	Charge Exchange factor 500–600 MeV/c	15	Geant4Reweight Pion Production
6	Charge Exchange factor 600–700 MeV/c	16	Geant4Reweight Quasielastic Low
7	Charge Exchange factor 700–800 MeV/c	17	Geant4Reweight Quasielastic High
8	Charge Exchange factor 800–900 MeV/c	18	Geant4Reweight Proton
9	Muon factor	19	No-track efficiency

TABLE 9.1: The parameters used within the fit. The numbers correspond to the bins shown in the figures throughout the chapter.

The post-fit and nominal MC reconstructed distributions in Figure 9.4 are identical to the Asimov fake data, and both distributions have a $-2\ln\lambda$ of 0 with respect to the fake data as expected from this closure test.



(a) Pre-fit and post-fit parameters.

(b) Post-fit correlation matrix of the fit parameters.

Figure 9.3: Asimov fit results.

Pre-fit $-2\ln\lambda_{\text{Stat}}$	0.00
Post-fit $-2\ln\lambda_{\text{Stat}}$	0.00
Post-fit $-2\ln\lambda_{\text{Syst}}$	0.00
Fit p-value	1.00
Nominal χ^2_σ	0.00
Fake Data χ^2_σ	0.00
Nominal σ p-value	1.00
Fake Data σ p-value	1.00

TABLE 9.2: Numerical results of the fit to Asimov fake data.

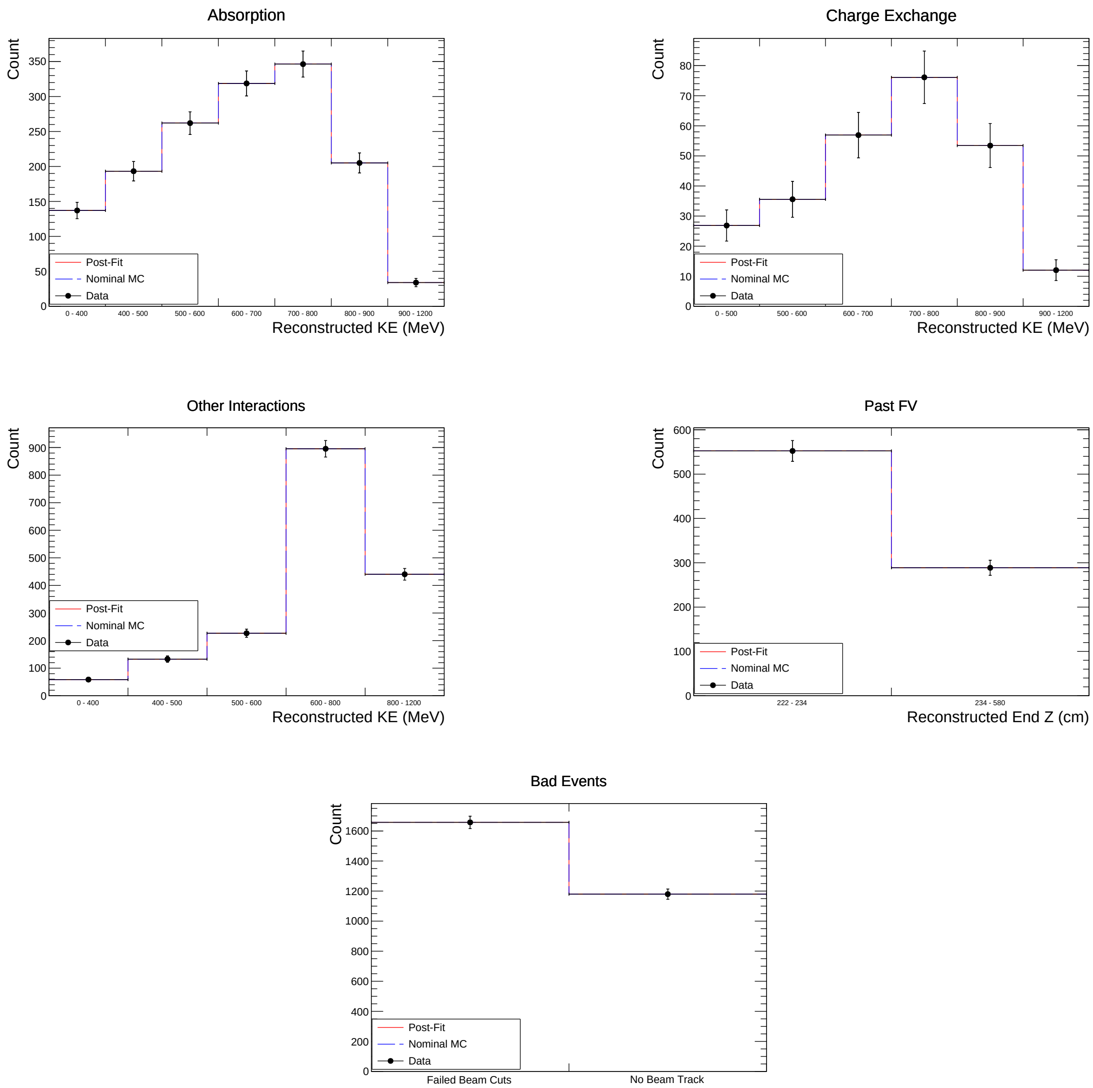
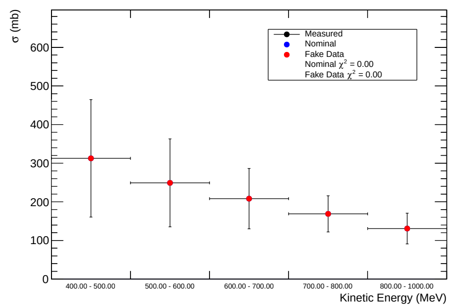
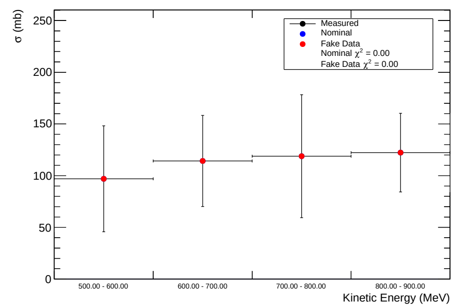


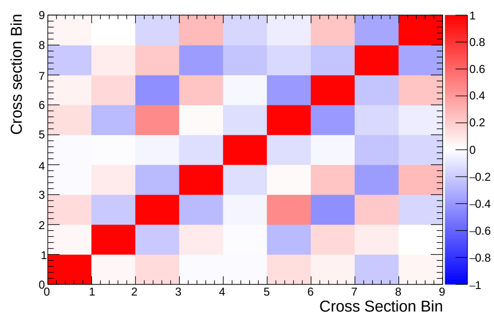
Figure 9.4: Reconstructed distributions of events in data (black points), Nominal MC (blue histogram), and post-fit results (red histogram) for the Asimov fit.



(a) Absorption



(b) Charge Exchange

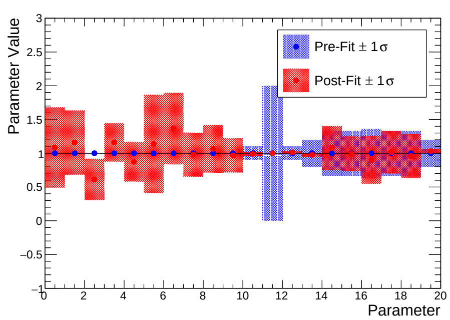


(c) Cross Section Correlations

Figure 9.5: Cross sections extracted from truth information taken from the post-fit MC ("Measured", black points), Nominal (blue points), and Asimov Fake Data (red points). 9.5c is the correlation between the cross sections. The first five rows are the absorption, and the last four rows are charge exchange. Note that the correlations between the two cross section types are included.

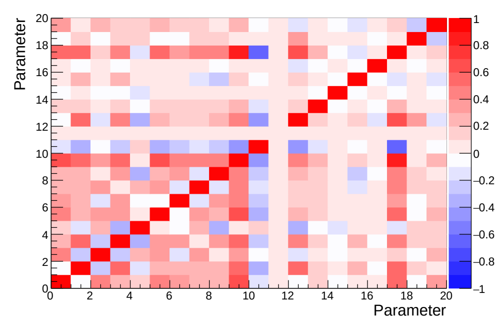
9.2 Statistically Independent Nominal MC

This test is similar to the previous Asimov fit where the input fake data is the nominal MC. However, half of the nominal MC was used as the input fake data, and the other half was used as the input MC. This is to test the performance of the fit to a statistically-independent set of nominal MC. The input fake data is expected to deviate from the input MC by a normal statistical fluctuation. This can be seen in Figure 9.7, where the fake data points no longer lay directly on top of the input MC. As can also be seen in Table 9.3, the post-fit $-2\ln\lambda$ between the post-fit and fake data reconstructed distributions is less than that between the pre-fit and fake data distributions, as expected. In Figure 9.6b, the systematic parameters can be seen to vary from nominal, but within the set of prior uncertainties presented within the plot (as the blue bands). Finally, the χ^2_σ between the measured and fake data cross sections as shown in Figure 9.8 shows the measured cross section is statistically consistent with the cross section extracted from the fake data set.



(a) Pre-fit and post-fit parameters.

(b)



(c) Post-fit correlation matrix of the fit parameters.

(d)

Figure 9.6: Fit results for the statistically independent nominal MC fit.

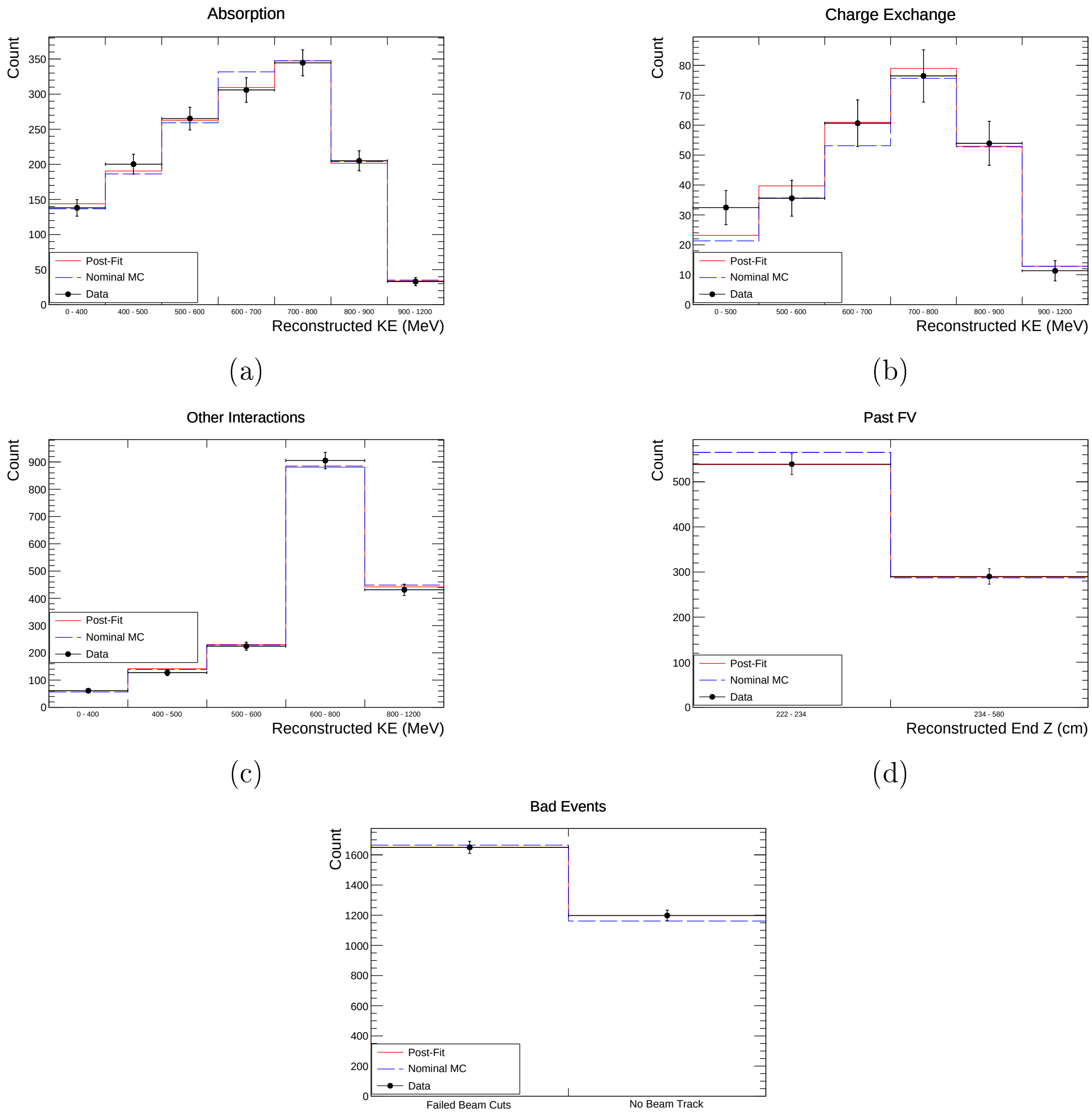
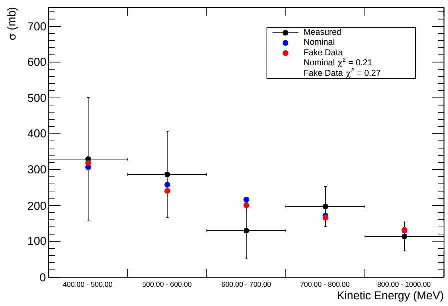
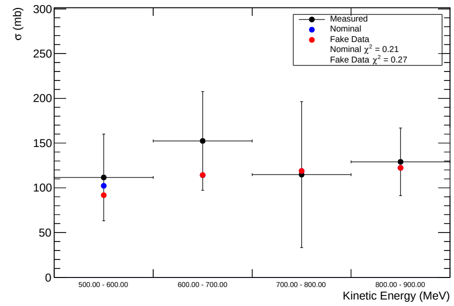


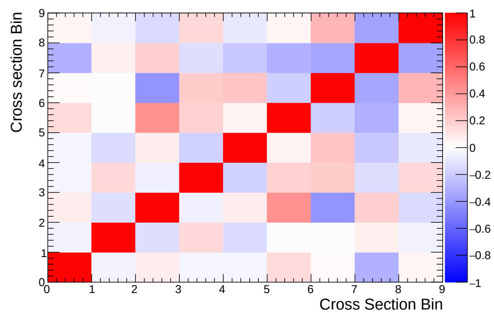
Figure 9.7: Reconstructed distributions of events in data (black points), Nominal MC (blue histogram), and post-fit results (red histogram) for the statistically independent nominal MC fit.



(a) Absorption



(b) Charge Exchange



(c) Cross Section Correlations

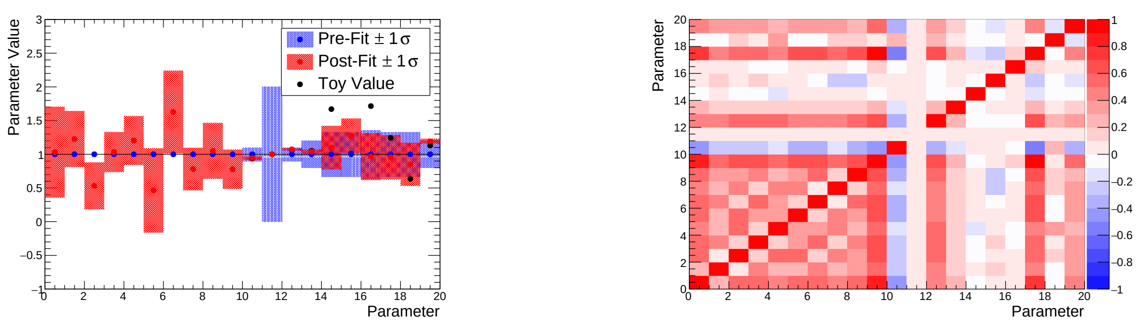
Figure 9.8: Cross sections extracted from truth information taken from the post-fit MC ("Measured", black points), Nominal (blue points), and Fake Data produced from statistically independent nominal MC (red points). 9.8c is the correlation between the cross sections. The first five rows are absorption, and the last four rows are charge exchange. Note that the correlations between the two cross section types are included.

Pre-fit $-2\ln\lambda_{\text{Stat}}$	14.72
Post-fit $-2\ln\lambda_{\text{Stat}}$	7.54
Post-fit $-2\ln\lambda_{\text{Syst}}$	0.21
Fit p-value	0.97
Nominal χ^2_σ	0.21
Fake Data χ^2_σ	0.27
Nominal σ p-value	1.00
Fake Dataσ p-value	1.00

TABLE 9.3: Numerical results of the fit to statistically independent fake data.

9.3 Systematic Variation

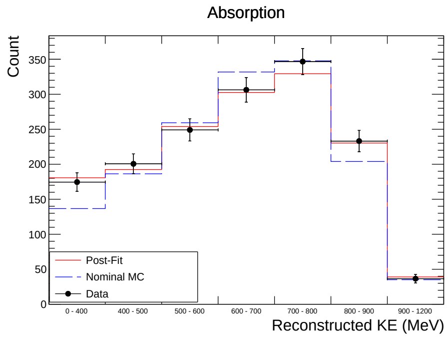
In this test, the fake data has been generated by using a statistically independent set of MC produced with varied systematic parameters. This is generated by first creating a random set of systematic parameter values. To create these values, a vector of random, unit-Gaussian distributed values is produced and then multiplied by the lower triangle of the Cholesky decomposition of the prior covariance matrix of the systematic parameters. This produces a set of values for the parameters with all correlations encoded. These systematic parameter values are then applied to half of the MC sample. Both the fake data reconstructed distributions and the cross sections are extracted from this varied MC sample. The other half of the MC is used as the input MC to be varied within the fit. The results are shown in Figures 9.9, 9.10, and 9.11. Figure 9.9a now includes the input systematic parameters used to create the variation (labeled "Toy Values"). As can be seen in this figure, the post-fit systematic parameters approach the input values. Shown in Table 9.4, the $-2\ln\lambda$ between MC and fake data show a large reduction as a result of the fit. Finally, in Figure 9.11 one can see that the χ^2_σ between the measured and fake data cross sections shows a consistent fit result.



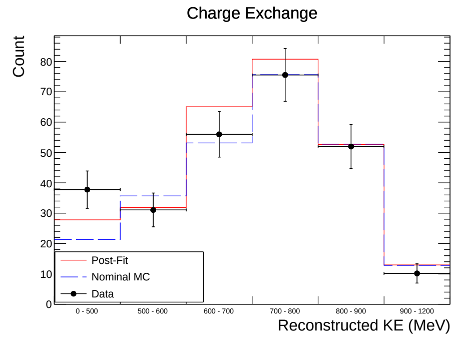
(a) Pre-fit and post-fit parameters.

(b) Post-fit correlation matrix of the fit parameters.

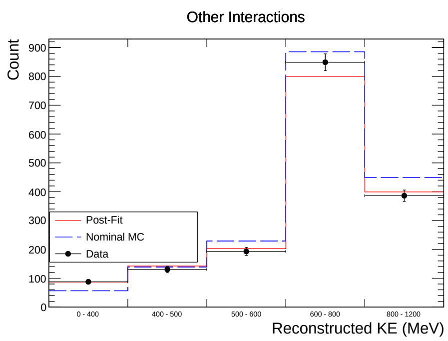
Figure 9.9: Fit results for the systematically varied fit.



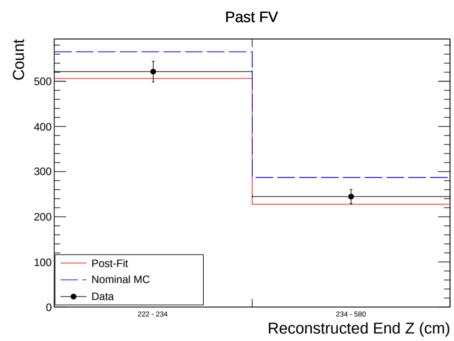
(a) Reconstructed distribution of events selected as Absorption.



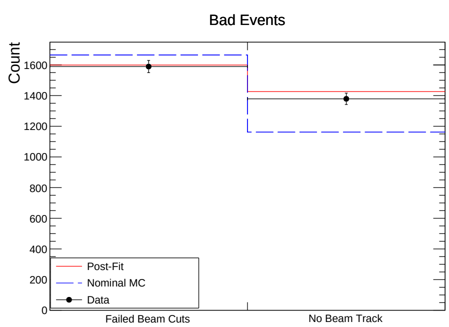
(b) Reconstructed distribution of events selected as Charge Exchange.



(c) Reconstructed distribution of events selected as neither Absorption or Charge Exchange.

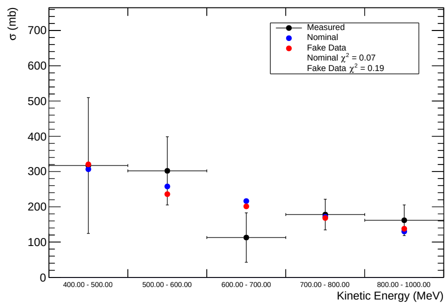


(d) Reconstructed distribution of events which extend past the Fiducial Volume

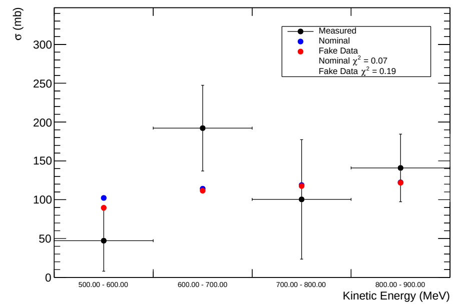


(e) The number of events which fail beam cuts (left bin) or lack a reconstructed beam track (right bin).

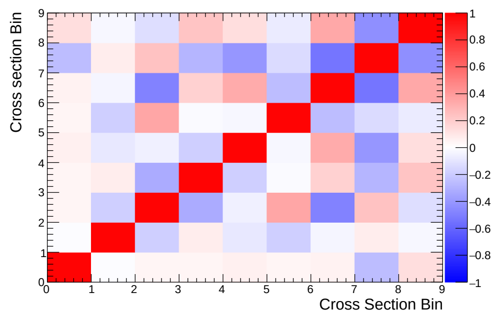
Figure 9.10: Reconstructed distributions of events in data (black points), Nominal MC (blue histogram), and post-fit results (red histogram) for the systematically varied fit.



(a) Absorption



(b) Charge Exchange



(c) Cross Section Correlations

Figure 9.11: Cross sections extracted from truth information taken from the post-fit MC ("Measured", black points), Nominal (blue points), and systematically varied Fake Data (red points). 9.11c is the correlation between the cross sections. The first five rows are absorption, and the last four rows are charge exchange. Note that the correlations between the two cross section types are included.

Pre-fit $-2\ln\lambda_{\text{Stat}}$	155.87
Post-fit $-2\ln\lambda_{\text{Stat}}$	12.69
Post-fit $-2\ln\lambda_{\text{Syst}}$	2.84
Fit p-value	0.69
Nominal χ^2_σ	0.07
Fake Data χ^2_σ	0.19
Nominal σ p-value	1.00
Fake Data σ p-value	1.00

TABLE 9.4: Numerical results of the fit to systematically and statistically varied fake data.

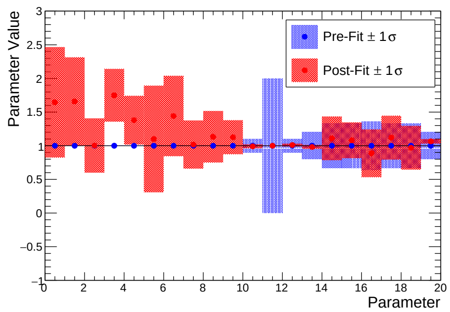
9.4 Geant4Reweight Fake Data

For this test, fake data is produced by reweighting[1] half of the nominal MC according to some set of π^+-Ar and p-Ar cross section variations using Geant4Reweight. Three sets of fake data were created. The first set was created varying the signal cross sections by some "reasonable" amount (i.e. similar to the level of the prior uncertainties of the Geant4Reweight parameters). The second set was created by varying the signal cross sections by an amount larger than the prior uncertainties on the Geant4Reweight parameters. The final set was created by varying both the signal and background cross sections. The background cross sections were varied in a different parameterization than those used in the fit: the bins of the Geant4Reweight variations in the fake data did not align with the bins in the fit parameters.

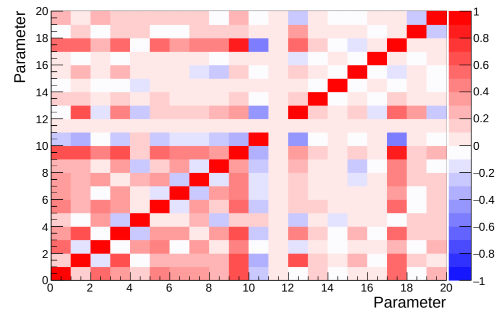
9.4.1 Reasonable Variations

The first set of fake data was created with the absorption cross section increased by 30% and the charge exchange cross section reduced by 10% across the full MC momentum range. Shown in Figure 9.12a, the systematic parameters are kept within their prior uncertainties. The reconstructed distributions in Figure 9.13 shows the fit ends in good agreement with the fake data distributions as can be seen in the post-fit $-2\ln\lambda$. Finally, the cross section extracted from the fit agree quite well with the cross sections extracted from the fake data set as can be seen in the "Fake Data χ^2" in Figure 9.14.

[1] A process to produce varied Monte Carlo samples assuming alternate cross section models, described in Section 8.5.



(a) Pre-fit and post-fit parameters.



(b) Post-fit correlation matrix of the fit parameters.

Figure 9.12: Fit results for the reasonable-variation Geant4Reweight fake data fit.

Pre-fit $-2\ln\lambda_{\text{Stat}}$	30.34
Post-fit $-2\ln\lambda_{\text{Stat}}$	7.65
Post-fit $-2\ln\lambda_{\text{Syst}}$	0.54
Fit p-value	0.97
Nominal χ^2_σ	3.58
Fake Data χ^2_σ	0.45
Nominal σ p-value	0.94
Fake Data σ p-value	1.00

TABLE 9.5: Numerical results of the fit to reasonable-variation Geant4Reweight fake data.

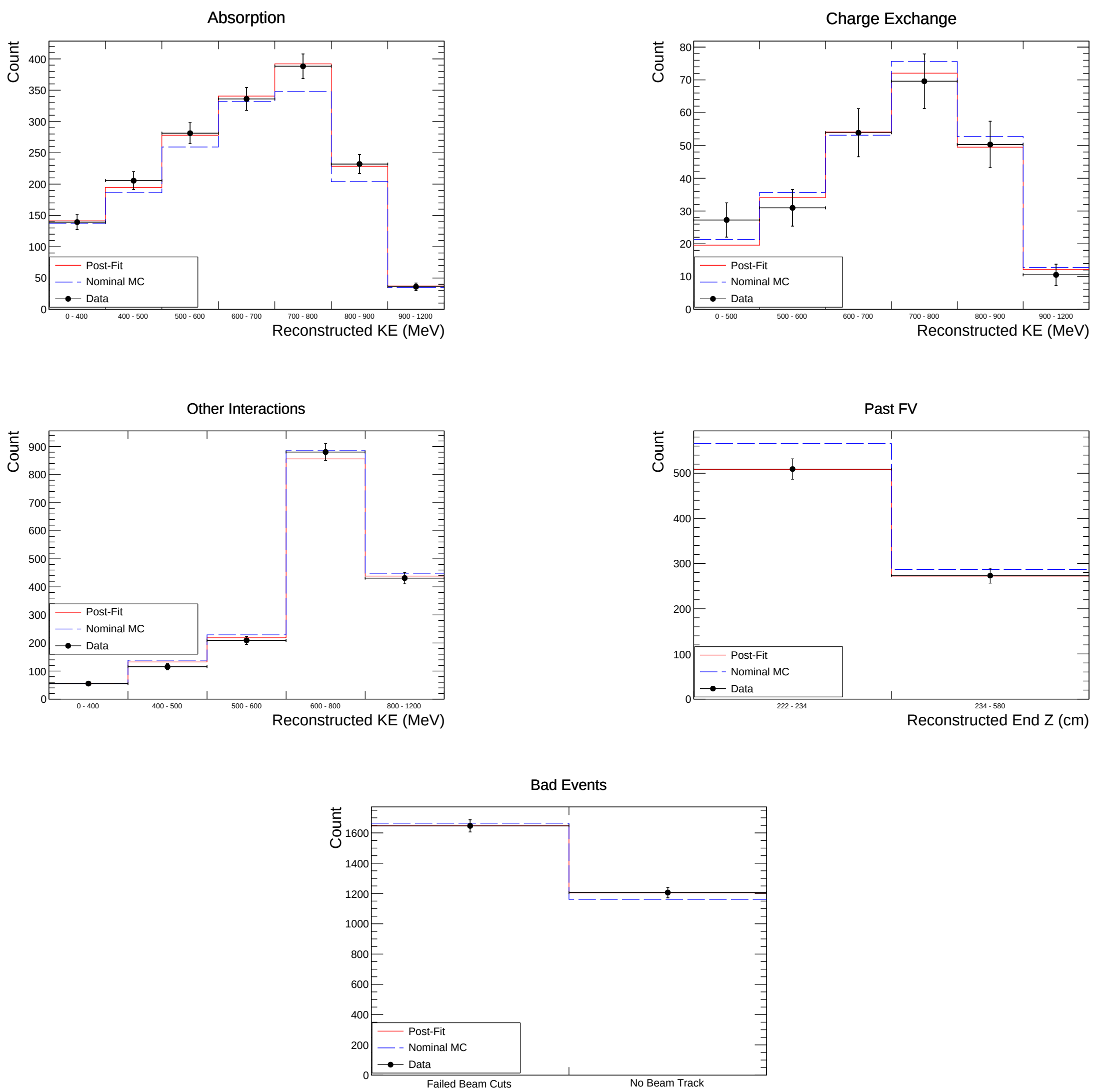
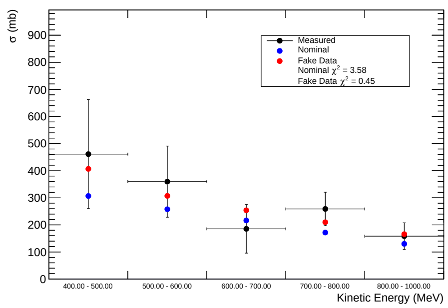
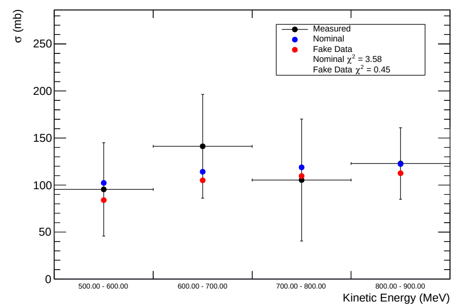


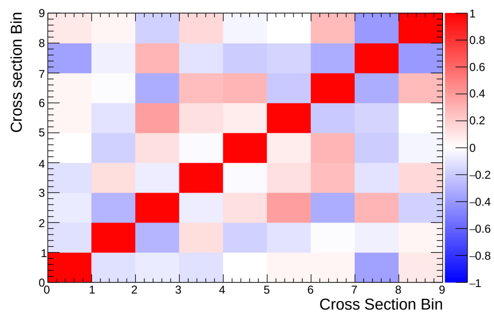
Figure 9.13: Reconstructed distributions of events in data (black points), Nominal MC (blue histogram), and post-fit results (red histogram) for the reasonable Geant4Reweight fake data fit.



(a) Absorption



(b) Charge Exchange

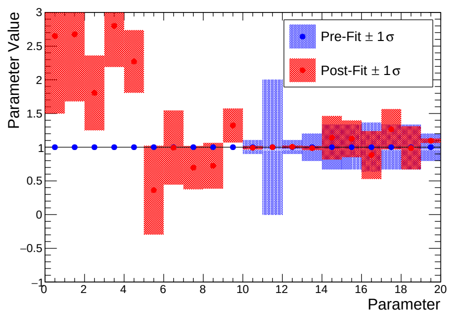


(c) Cross Section Correlations

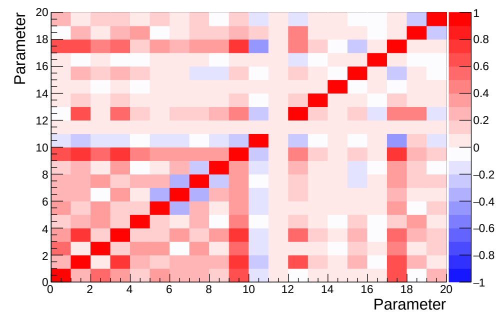
Figure 9.14: Cross sections extracted from truth information taken from the post-fit MC ("Measured", black points), Nominal (blue points), and less extreme Geant4Reweight Fake Data (red points). 9.14c is the correlation between the cross sections. The first five rows are absorption, and the last four rows are charge exchange. Note that the correlations between the two cross section types are included.

9.4.2 Plausible Variations

The second set of fake data created with Geant4Reweight contained an increase to the absorption cross section by 80% and a reduction of the charge exchange cross section across the full momentum range by 60%. Shown in Table 9.6, the fit ends with a consistent $-2\ln\lambda$ the post-fit nuisance parameters are within their prior uncertainties showing a successful fit to the fake data. In Figure 9.17, a drastic reduction in χ^2_σ is shown, indicating the fit can successfully pick out a variation to the signal cross sections at this level.



(a) Pre-fit and post-fit parameters.



(b) Post-fit correlation matrix of the fit parameters.

Figure 9.15: Fit results for the plausible-variation Geant4Reweight fake data fit.

Pre-fit $-2\ln\lambda_{\text{Stat}}$	151.75
Post-fit $-2\ln\lambda_{\text{Stat}}$	8.01
Post-fit $-2\ln\lambda_{\text{Syst}}$	1.33
Fit p-value	0.94
Nominal χ^2_σ	20.60
Fake Data χ^2_σ	0.92
Nominal σ p-value	0.29
Fake Data σ p-value	1.00

TABLE 9.6: Numerical results of the fit to plausible-variation Geant4Reweight fake data.

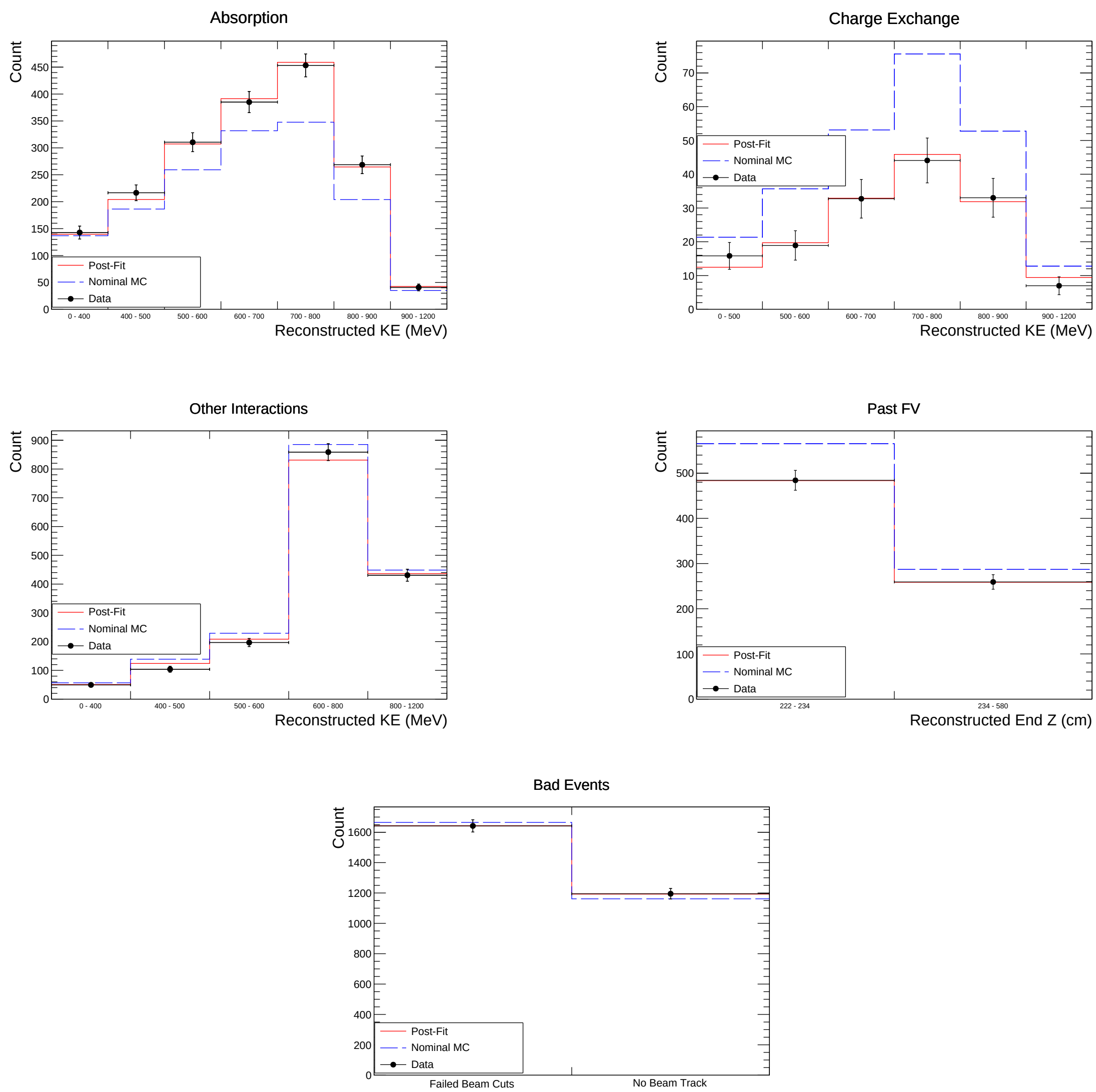
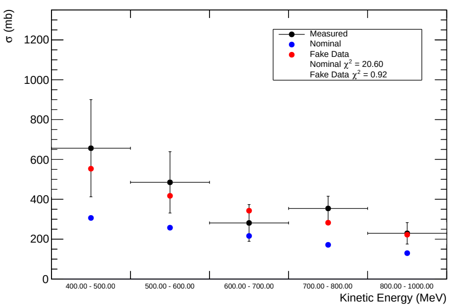
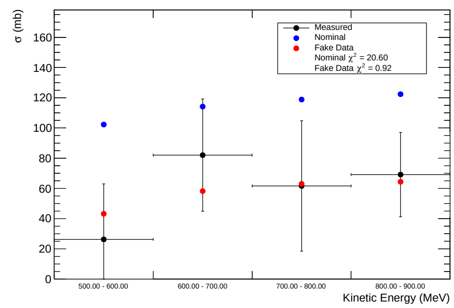


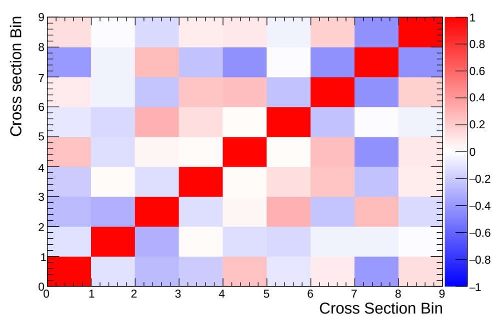
Figure 9.16: Reconstructed distributions of events in data (black points), Nominal MC (blue histogram), and post-fit results (red histogram) for the plausible-variation Geant4Reweight fake data fit.



(a) Absorption



(b) Charge Exchange

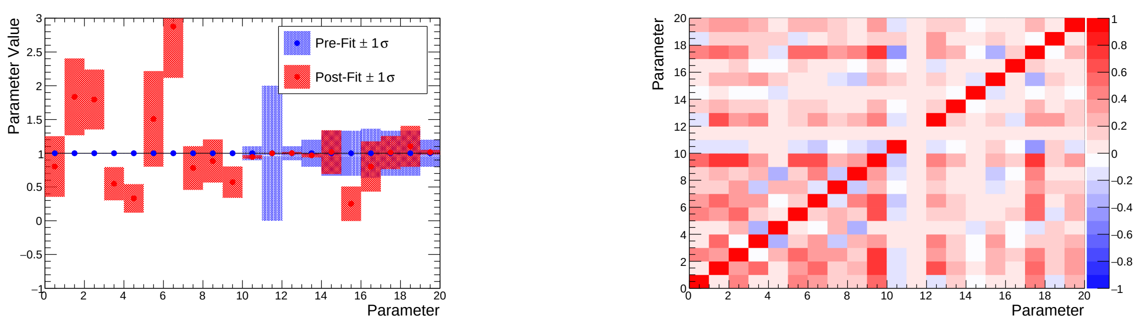


(c) Cross Section Correlations

Figure 9.17: Cross sections extracted from truth information taken from the post-fit MC ("Measured", black points), Nominal (blue points), and plausible-variation Geant4Reweight Fake Data (red points). 9.17c is the correlation between the cross sections. The first five rows are absorption, and the last four rows are charge exchange. Note that the correlations between the two cross section types are included.

9.4.3 Extreme Variations

The final set of fake data created with Geant4Reweight weights was intended to represent "extreme" variations. The variations applied to the MC are to increase the total inelastic cross section by 80% up to 800 MeV/c momentum and to reduce the total inelastic cross section by 60% above 800 MeV/c. As can be seen in the Figures 9.18 - 9.20, the fit finds a minimum, but the results indicate a poor result. A few of the systematic parameters shown in Figure 9.18a are pulled outside of their prior uncertainties. The post-fit $-2 \ln \lambda$ in Table 9.7 is large, indicating a bad goodness-of-fit. Similarly, the χ^2 between the measured cross sections and those extracted from fake data is actually *higher* than that to the nominal MC. This suggests that the parameterization used within the fit is not suitable for this fake data. As described in Section 8.5, the systematic parameter for pion production is a single bin from 0 to 2 GeV/c, whereas the variation used to create the fake data has a more complicated shape to it. This is an example of where the fit on data could fail to find a useful result. If this was seen in a fit to data, we would reconsider the parameterization of the Geant4Reweight parameters (i.e. the coarseness of the variation bins).



(a) Pre-fit and post-fit parameters.

(b) Post-fit correlation matrix of the fit parameters.

Figure 9.18: Fit results for the extreme Geant4Reweight fake data fit.

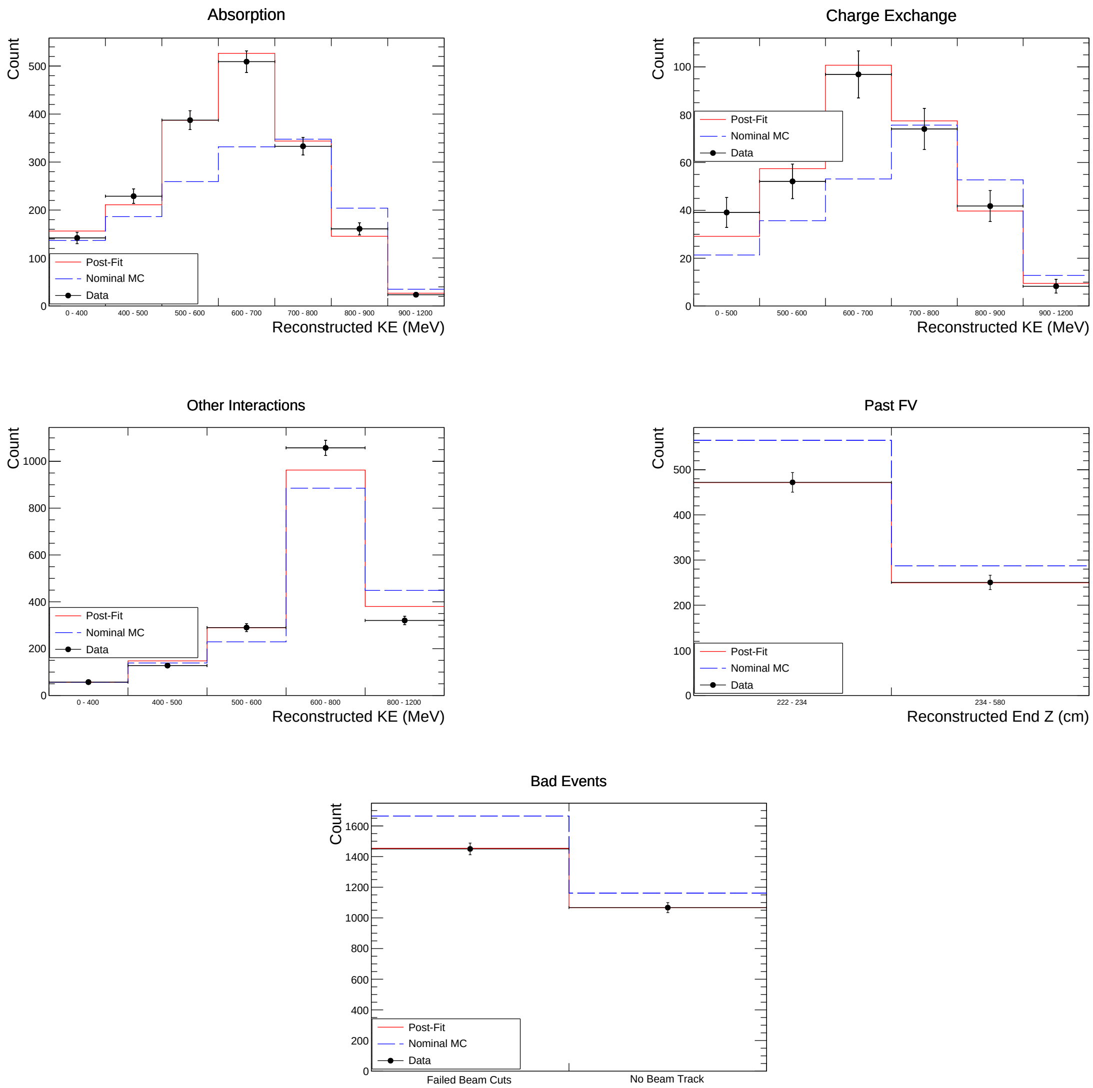
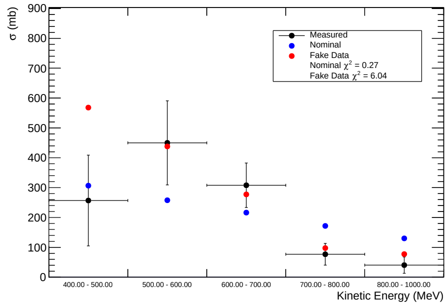
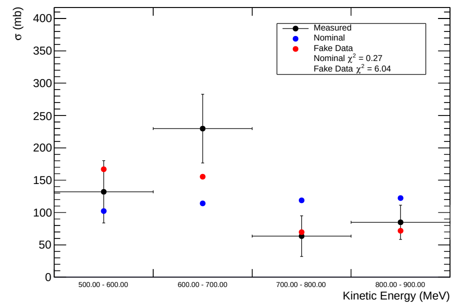


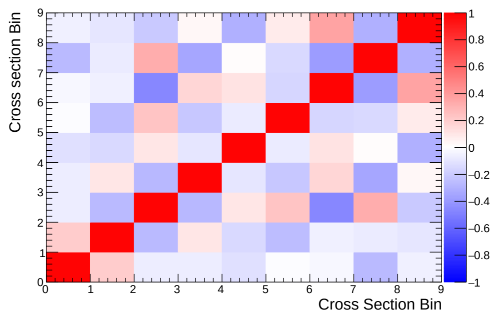
Figure 9.19: Reconstructed distributions of events in data (black points), Nominal MC (blue histogram), and post-fit results (red histogram) for the extreme Geant4Reweight fake data fit.



(a) Absorption



(b) Charge Exchange



(c) Cross Section Correlations

Figure 9.20: Cross sections extracted from truth information taken from the post-fit MC ("Measured", black points), Nominal (blue points), and extreme Geant4Reweight Fake Data (red points). 9.20c is the correlation between the cross sections. The first five rows are absorption, and the last four rows are charge exchange. Note that the correlations between the two cross section types are included.

Pre-fit $-2\ln\lambda_{\text{Stat}}$	358.02
Post-fit $-2\ln\lambda_{\text{Stat}}$	31.62
Post-fit $-2\ln\lambda_{\text{Syst}}$	5.87
Fit p-value	0.03
Nominal χ^2_σ	0.27
Fake Data χ^2_σ	6.04
Nominal σ p-value	1.00
Fake Data σ p-value	0.84

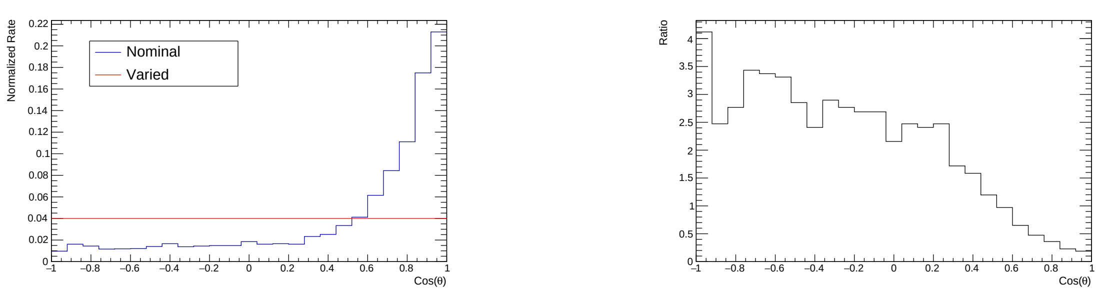
TABLE 9.7: Numerical results of the fit to unreasonable-variation Geant4Reweight fake data.

9.5 Pion Angle Variation

The next set of fake data used to validate the fit considers a variation to the outgoing direction of pions produced in quasielastic (QE) events. Note that this is defined according to the signal definition, and the outgoing pion is required to have above 150 MeV/c for the event to be considered QE. This is done to test whether the fit is resilient to mismodeling of outgoing pion kinematics in the Geant4 model used. The angular distribution of outgoing pions in QE events was modified by-hand to create a varied MC sample. This is done in bins of true momentum at interaction in order to prevent some variation to be imparted on the distribution of pion momentum at interaction. The bin edges for these are (0, 400, 600, 800, 1000, 2000) MeV/c. In each final momentum bin, the ratio of the varied distribution to the nominal distribution is used as an event weight to create a set of fake data. Cross sections from the set of varied MC used to create the fake data are extracted and compared to the post-fit MC.

9.5.1 Flat Distribution

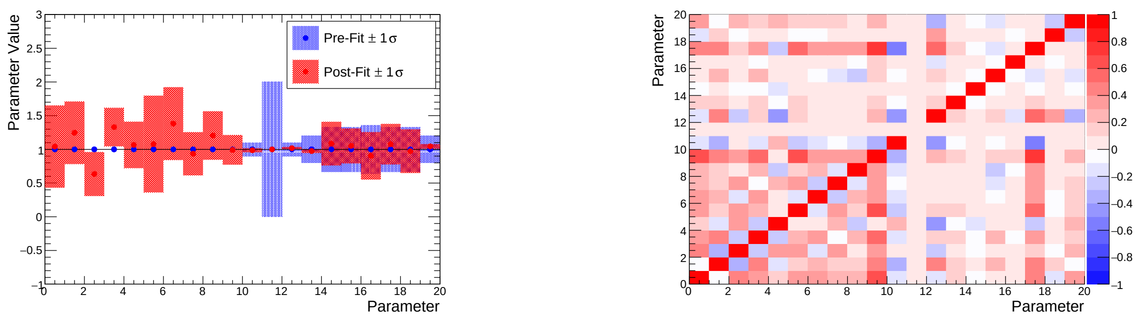
In this set of fake data, the angular distribution of final state pions is flattened. The nominal distribution, varied distribution, and ratio used for weighting from the 600–800 MeV/c bin are shown in Figure 9.21.



(a) Outgoing pion angular distributions. (b) Ratio used for weighting QE events.

Figure 9.21: Inputs for this fake data test.

As can be seen in Figure 9.23, small variations occur throughout the distributions. The most interesting bin is the lowest bin in Figure 9.23d, as it can be explained from the variation applied. This bin contains QE events with a forward-going pion. Since these forward-going events have been suppressed in the fake data, this bin has been lowered. Still, the results of this fit are promising. The post-fit $-2\ln\lambda$ in Table 9.8 shows that the fit is insensitive to variation, indicating a robustness against data-MC disagreements of this type. The small variations in the bins mentioned previously show low sensitivity to this type of physical variation. Furthermore, the χ^2 between the Measured and Nominal & Fake Data cross sections in Figure 9.24 show a consistent measurement.



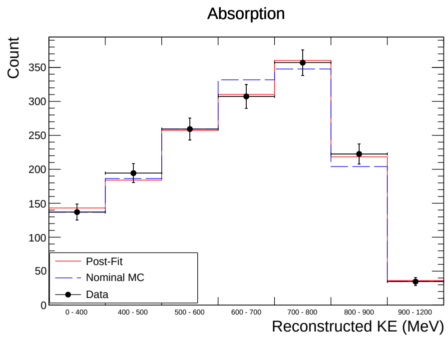
(a) Pre-fit and post-fit parameters.

(b) Post-fit correlation matrix of the fit parameters.

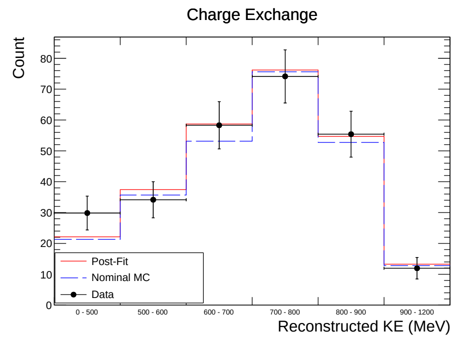
Figure 9.22: Fit results for the flat pion fit.

Pre-fit $-2\ln\lambda_{\text{Stat}}$	18.56
Post-fit $-2\ln\lambda_{\text{Stat}}$	6.61
Post-fit $-2\ln\lambda_{\text{Syst}}$	0.33
Fit p-value	0.98
Nominal χ^2_σ	0.31
Fake Data χ^2_σ	0.55
Nominal σ p-value	1.00
Fake Data σ p-value	1.00

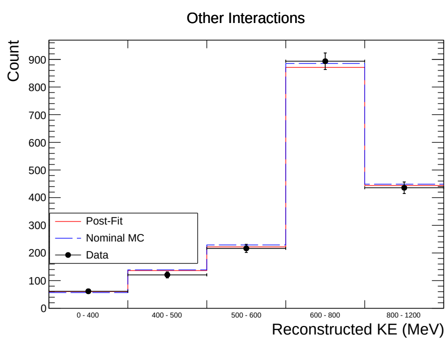
TABLE 9.8: Numerical results of the fit to varied pion angular distribution fake data.



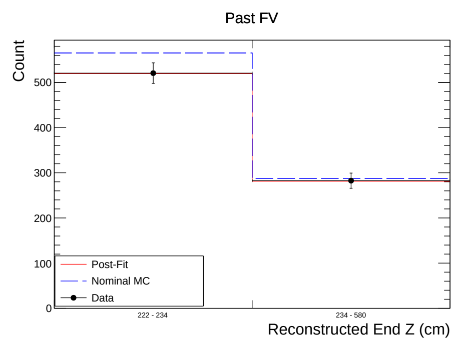
(a)



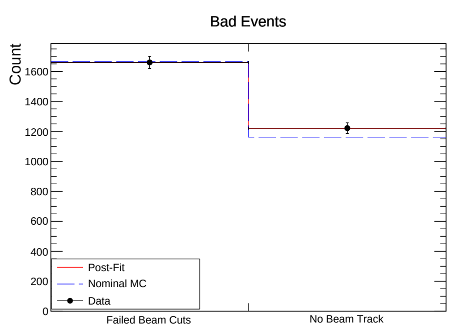
(b)



(c)

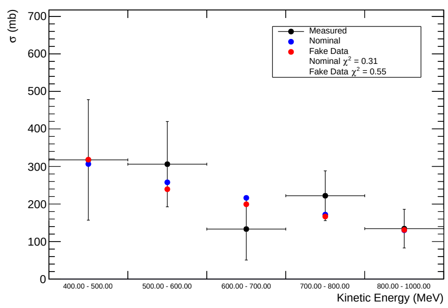


(d)

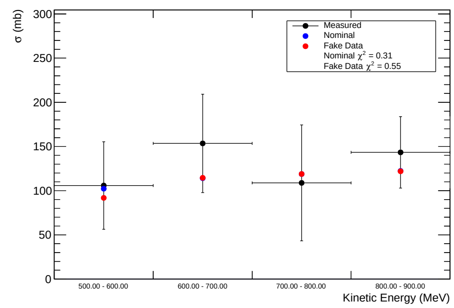


(e)

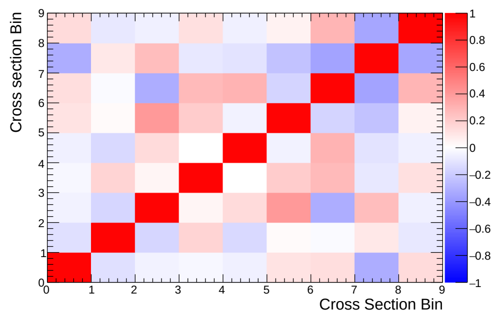
Figure 9.23: Reconstructed distributions of events in data (black points), Nominal MC (blue histogram), and post-fit results (red histogram) for the flat pion fit.



(a) Absorption



(b) Charge Exchange



(c) Cross Section Correlations

Figure 9.24: Cross sections extracted from truth information taken from the post-fit MC ("Measured", black points), Nominal (blue points), and flat pion Fake Data (red points). 9.24c is the correlation between the cross sections. The first five rows are absorption, and the last four rows are charge exchange. Note that the correlations between the two cross section types are included.

CHAPTER 10

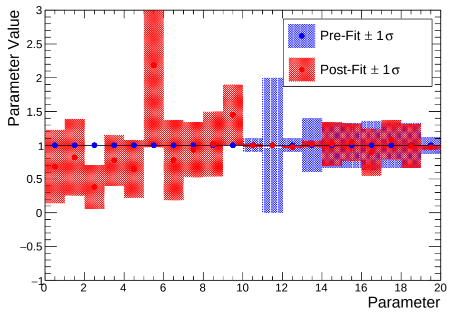
RESULTS

This chapter shows the results of the fit to ProtoDUNE-SP beam data from Run 5809 taken during its initial running period in the Fall of 2018. The data is comprised of beam-triggered events with a PID corresponding to π/μ (as defined in Section 5.2). The pre-fit and post-fit parameters and their corresponding uncertainties are shown in Figure 10.1a along with their post-fit correlation matrix. The pre-fit (referred to as "Nominal") MC, post-fit MC, and data event distributions are shown in Figure 10.2. Lastly, the cross sections extracted from the fit to data are shown in Figure 10.3. This figure contains the correlation matrix and 1-D error bars (taken from the diagonal of the covariance matrix) as computed from the error propagation procedure described in Section 7.5.

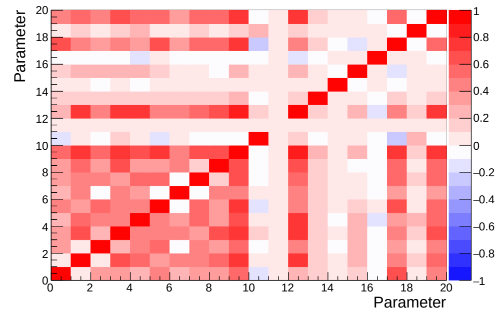
Figure 10.1a shows the effect on the systematic parameters as a result of the fit. The dE/dX calibration and beam cut efficiency parameters are tightly constrained within the fit, but remain within their prior uncertainties. The Geant4Reweight parameters remain unconstrained by the fit.

The event distributions shown in Figure 10.2 show good agreement between the post-fit MC and data distributions. The resulting p-value of the fit is 0.998, pointing to a successful parameterization of the fit. However, it suggests that the fit could be "too good". This could be due to the choice of parameterization of the fit (perhaps the effects of the efficiency-like systematic parameters are too strong), or it could be due to the size of the post-fit uncertainties being too large.

Finally, the extracted cross sections shown in Figure 10.3 remain consistent with the nominal-MC cross section, as indicated by the χ^2_σ of 6.70 shown in these plots. The σ p-value of this is 0.81, indicating the post-fit cross sections remain consistent with the nominal cross sections.



(a) Pre-fit and post-fit parameters.



(b) Post-fit correlation matrix of the fit parameters.

Figure 10.1: Fit results for the fit to ProtoDUNE-SP data.

Pre-fit $-2\ln\lambda_{\text{Stat}}$	874.37
Post-fit $-2\ln\lambda_{\text{Stat}}$	4.14
Post-fit $-2\ln\lambda_{\text{Syst}}$	0.31
Fit p-value	0.998
Nominal χ^2_σ	6.70
Nominal σ p-value	0.81

TABLE 10.1: Numerical results of the fit to ProtoDUNE-SP run 5809 data.

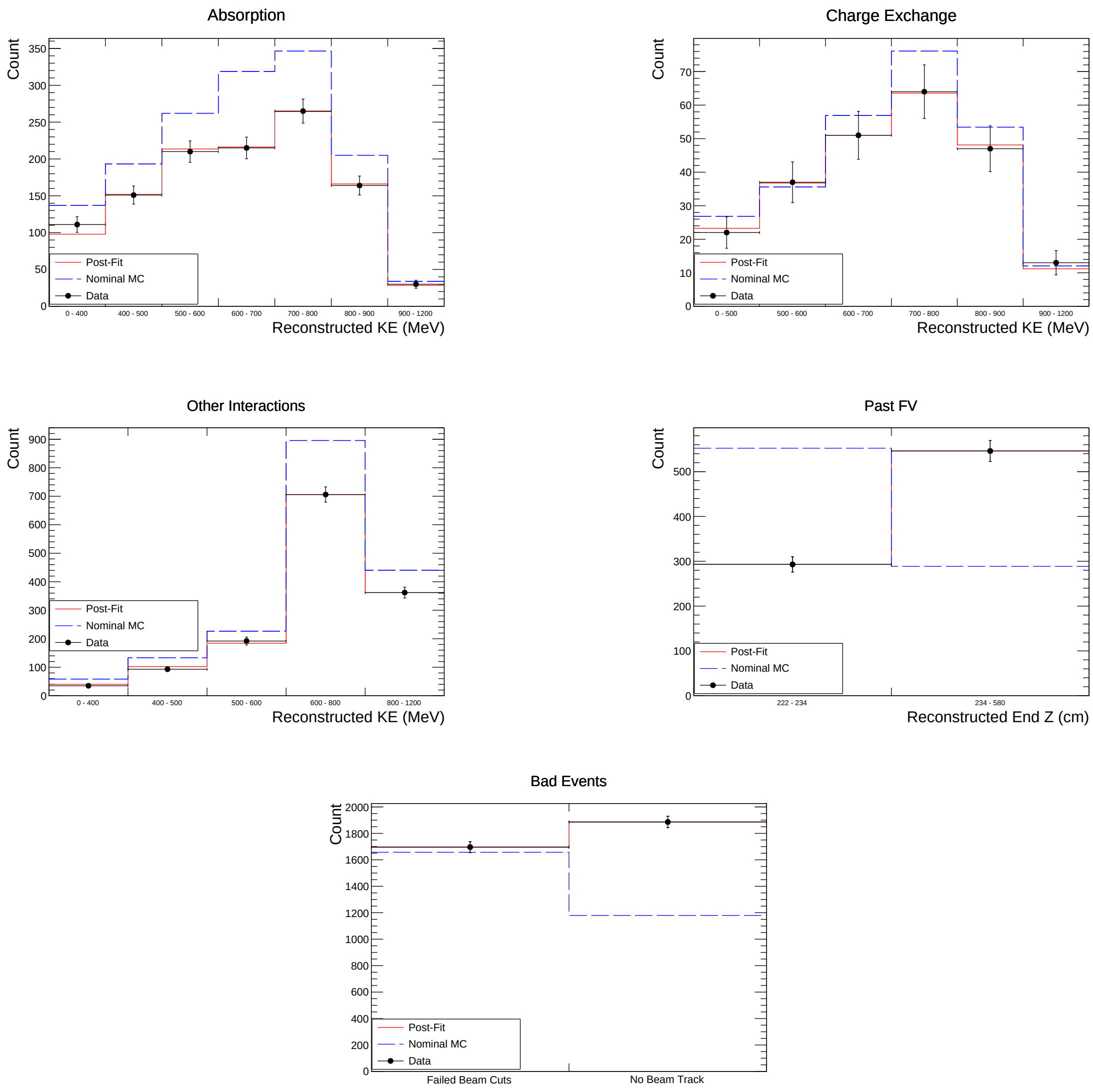
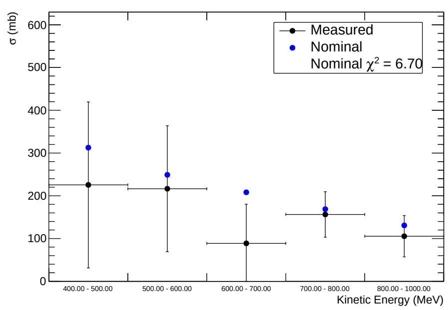
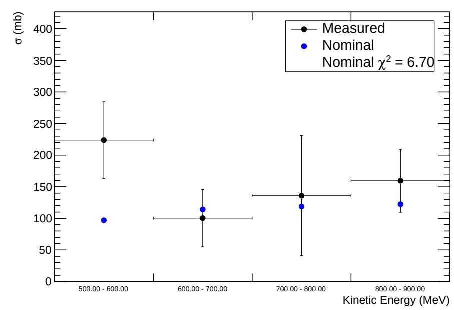


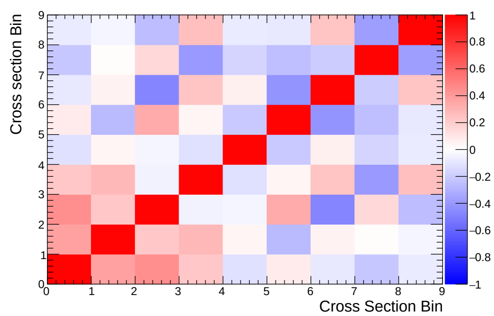
Figure 10.2: Reconstructed distributions of events in data (black points), Nominal MC (blue histogram), and post-fit results (red histogram) for the fit to real data.



(a) Absorption



(b) Charge Exchange



(c) Cross Section Correlations

Figure 10.3: Cross sections extracted from truth information taken from the post-fit MC ("Measured", black points) and Nominal (blue points). 10.3c is the correlation between the cross sections. The first four rows are the absorption, and the last three rows are charge exchange. Note that the correlations between the two cross section types are included.

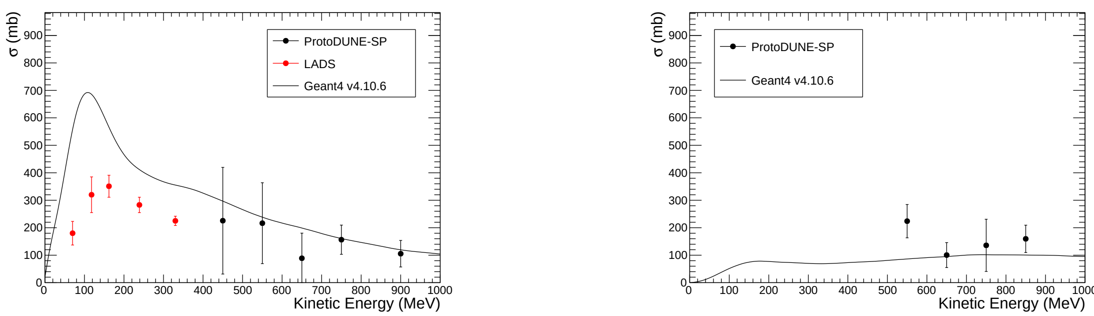
10.1 Future Work

Some shortcomings within the analysis are worth addressing here. The cross sections shown in this chapter have relatively large error bars, and remain compatible with the nominal MC at this level of uncertainty. Though this compatibility is not an issue, a reduction in the measurement's uncertainty would allow us to determine if there is a significant difference to the cross section models used in DUNE's simulation. A larger data set (such as more runs taken during Fall 2018) would of course reduce the statistical uncertainties, while a better understanding of the underlying cause of the Pandora reconstruction efficiency would improve the systematic uncertainties. The improved understanding of what is causing the data–MC discrepancies regarding the reconstruction efficiency could provide a more suitable uncertainty parameterization than the ad-hoc efficiency factors currently used in the fits. We also chose to neglect SCE uncertainties in this analysis, though these are in development and will be added in the future. Following the implementation of the SCE uncertainties, the other runs taken in Fall 2018 will be added. This will increase the size of the data set in the fit by a factor of 9. These issues will all be iterated upon in future work as this analysis moves toward publication by the DUNE collaboration. Additionally, a planned second run of ProtoDUNE-SP will provide even more data for this and future measurements.

10.2 Conclusion

Presented in this thesis is one of the first measurements of π^+ interactions on Argon using ProtoDUNE-SP data. This measurement would have been impossible without the large amount of work undertaken within the DUNE collaboration to construct and commission this detector, currently the largest single-phase LArTPC to have operated. The rapid data-taking in the Fall of 2018 was followed by an immense effort to carefully categorize and calibrate the data and also to produce an accurate simulation of the detector. An exciting future awaits ProtoDUNE-SP, as additional configurations of the detector and more beam data are planned in the coming years.

This analysis shows the first ever measurement of π^+–Ar charge exchange and the first measurement of π^+–Ar absorption in this energy range. The LADS collaboration measured[22] π^+–Ar absorption at an energy range below that shown in this thesis. The LADS measurement is compared to the measurement from this analysis as well as to the prediction from Geant4[1] in Figure 10.4a, while the this measurement of charge exchange is compared to Geant4 in Figure 10.4b. Of interest is the disagreement between the Geant4 model and the LADS data in the resonance region. A similar analysis using data from ProtoDUNE-SP at lower momentum will provide a chance to explore this region further. At higher momentum, the results of this analysis show close agreement with the Geant4 model for both channels.



(a) Absorption

(b) Charge Exchange

Figure 10.4: The measured π^+–Ar absorption (left) and charge exchange (right) cross sections compared to the nominal Geant4 model. Additionally, the left plot contains an earlier measurement of absorption from the LADS experiment [22].

The fit used to perform this analysis is based on analyses from the T2K experiment to measure ν_μ and $\bar{\nu}_\mu$ interactions in their near detector ND280 [60][61][62][63][64]. It includes systematic uncertainties due to the detector and signal model. The benefit of doing such an analysis is that it produces robust estimates of the uncertainties and correlations on the extracted cross sections. This allows the data to be properly compared to interaction models used within detector simulations for upcoming experiments including the Short-Baseline

[1]These curves were generated by an application within the Geant4Reweight framework. A 150 MeV/c pion momentum threshold was applied in order to match the signal definition for these measurements. The LADS data does not include this threshold.

Neutrino Program at Fermilab and, ultimately, DUNE. By doing this, important systematic uncertainties regarding the rate of secondary interactions of neutrino interaction products can be constrained, and will help allow DUNE to achieve the experimental precision required for its physics goals.

BIBLIOGRAPHY

[1] G. Aad et al. Observation of a new particle in the search for the standard model higgs boson with the atlas detector at the lhc. *Physics Letters B*, 716(1):129, Sep 2012.

[2] S. Chatrchyan et al. Observation of a new boson at a mass of 125 gev with the cms experiment at the lhc. *Physics Letters B*, 716(1):3061, Sep 2012.

[3] Mandl, F. and Shaw, Graham. *Quantum Field Theory; Second Edition*. Wiley, 2010.

[4] Martin, Brian R. and Shaw, Graham. *Particle Physics; Third Edition*. The Manchester Physics Series. Wiley, New York, NY, 2008.

[5] R. Aaij et al. Observation of $j/\psi p$ resonances consistent with pentaquark states in $\Lambda_b^0 \to j/\psi K^- p$ decays. *Phys. Rev. Lett.*, 115:072001, Aug 2015.

[6] Raymond Davis, Don S. Harmer, and Kenneth C. Hoffman. Search for neutrinos from the sun. *Phys. Rev. Lett.*, 20:1205–1209, May 1968.

[7] David J Griffiths. *Introduction to elementary particles; 2nd rev. version*. Physics textbook. Wiley, New York, NY, 2008.

[8] Y. Fukuda et al. Evidence for oscillation of atmospheric neutrinos. *Phys. Rev. Lett.*, 81:1562–1567, Aug 1998.

[9] Q. R. Ahmad et al. Measurement of the rate of $\nu_e + d \to p + p + e^-$ interactions produced by ^8b solar neutrinos at the sudbury neutrino observatory. *Phys. Rev. Lett.*, 87:071301, Jul 2001.

[10] https://neutrinos.fnal.gov/mysteries/mass-ordering/#moreinfo.

[11] K. Abe et. al. Constraint on the matterantimatter symmetry-violating phase in neutrino oscillations. *Nature*, 580(7803):339344, Apr 2020.

[12] B. Abi et al. Deep underground neutrino experiment (dune), far detector technical design report, volume ii: Dune physics, 2020.

[13] K. Abe et al. (Hyper-Kamiokande Proto-Collaboration). Physics potential of a long-baseline neutrino oscillation experiment using a J-PARC neutrino beam and Hyper-Kamiokande. *Progress of Theoretical and Experimental Physics*, 2015(5), 05 2015. 053C02.

[14] Carlo Giunti and Kim Chung Wook. *Fundamentals of Neutrino Physics and Astrophysics*. Oxford Univ., Oxford, 2007.

[15] X. Qian and P. Vogel. Neutrino mass hierarchy. *Progress in Particle and Nuclear Physics*, 83:130, Jul 2015.

[16] Ulrich Mosel. Neutrino interactions with nucleons and nuclei: Importance for long-baseline experiments. *Annual Review of Nuclear and Particle Science*, 66(1):171–195, 2016.

[17] J. A. Formaggio and G. P. Zeller. From ev to eev: Neutrino cross sections across energy scales. *Reviews of Modern Physics*, 84(3):13071341, Sep 2012.

[18] L. Alvarez-Ruso et al. Nustec white paper: Status and challenges of neutrino-nucleus scattering. *Progress in Particle and Nuclear Physics*, 100:1–68, 2018.

[19] T.-S. H. Lee and R. P. Redwine. Pion-nucleus interactions. *Annual Review of Nuclear and Particle Science*, 52(1):23–63, 2002.

[20] Torleif Erik Oskar Ericson and W. Weise. *Pions and Nuclei*, volume 74. Clarendon Press, Oxford, UK, 1988.

[21] Jörg Hüfner. Pions interact with nuclei. *Physics Reports*, 21(1):1 – 79, 1975.

[22] B. Kotlinski et al. (The LADS Collaboration). Pion absorption reactions on n, ar, and xe. *The European Physical Journal A*, 9:537–552, 12 2000.

[23] I. Esteban et al. Global analysis of three-flavour neutrino oscillations: synergies and tensions in the determination of θ_{23}, δ_{CP}, and the mass ordering. *Journal of High Energy Physics*, 2019(1), Jan 2019.

[24] B. Abi et al. Deep underground neutrino experiment (dune), far detector technical design report, volume ii: Dune physics, 2020.

[25] https://www.dunescience.org/.

[26] A. Ashkenazi. Connections between neutrino and electron scattering. Neutrino 2020 – Virtual Meeting, 2020.

[27] E. S. Pinzon Guerra, C. Wilkinson, S. Bhadra, S. Bolognesi, J. Calcutt, P. de Perio, S. Dolan, T. Feusels, G. A. Fiorentini, Y. Hayato, K. Ieki, K. Mahn, K. S. McFarland, V. Paolone, L. Pickering, R. Tacik, H. A. Tanaka, R. Terri, M. O. Wascko, M. J. Wilking, C. Wret, and M. Yu. Using world charged pion–nucleus scattering data to constrain an intranuclear cascade model. *arXiv e-prints*, page arXiv:1812.06912, December 2018.

[28] R. Acciarri, C. Adams, R. An, A. Aparicio, S. Aponte, J. Asaadi, M. Auger, N. Ayoub, L. Bagby, and B. Baller et al. Design and construction of the microboone detector. *Journal of Instrumentation*, 12(02):P02017P02017, Feb 2017.

[29] P. A. Zyla et al. (Particle Data Group). Review of Particle Physics. *Progress of Theoretical and Experimental Physics*, 2020(8), 08 2020. 083C01.

[30] B. Abi et al. First results on ProtoDUNE-SP liquid argon time projection chamber performance from a beam test at the CERN neutrino platform. *Journal of Instrumentation*, 15(12):P12004–P12004, dec 2020.

[31] R Acciarri, C Adams, J Asaadi, B Baller, T Bolton, C Bromberg, F Cavanna, E Church, D Edmunds, and A Ereditato et al. A study of electron recombination using highly ionizing particles in the argoneut liquid argon tpc. *Journal of Instrumentation*, 8(08):P08005P08005, Aug 2013.

[32] S. Amoruso et al. Study of electron recombination in liquid argon with the icarus tpc. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 523(3):275 – 286, 2004.

[33] D Montanari, J Bremer, A Gendotti, M Geynisman, S Hentschel, T Loew, D Mladenov, C Montanari, S Murphy, M Nessi, B Norris, F Noto, A Rubbia, R Sharma, D Smargianaki, J Stewart, C Vignoli, P Wilson, and S Wu. Development of membrane cryostats for large liquid argon neutrino detectors. *IOP Conference Series: Materials Science and Engineering*, 101:012049, dec 2015.

[34] B. Abi et al. The Single-Phase ProtoDUNE Technical Design Report. 6 2017.

[35] B. Abi et al. Deep underground neutrino experiment (dune), far detector technical design report, volume i: Introduction to dune, 2020.

[36] D. Adams, M. Bass, M. Bishai, C. Bromberg, J. Calcutt, H. Chen, J. Fried, I. Furic, S. Gao, D. Gastler, J. Hugon, J. Joshi, B. Kirby, F. Liu, K. Mahn, M. Mooney, C. Morris, C. Pereyra, X. Pons, V. Radeka, E. Raguzin, D. Shooltz, M. Spanu, A. Timilsina, S. Tufanli, M. Tzanov, B. Viren, W. Gu, Z. Williams, K. Wood, E. Worcester, M. Worcester, G. Yang, and J. Zhang. The ProtoDUNE-SP LArTPC electronics production, commissioning, and performance. *Journal of Instrumentation*, 15(06):P06017–P06017, jun 2020.

[37] R. Herbst, R. Claus, M. Freytag, G. Haller, M. Huffer, S. Maldonado, K. Nishimura, C. O'Grady, J. Panetta, A. Perazzo, B. Reese, L. Ruckman, J. G. Thayer, and M. Weaver. Design of the slac rce platform: A general purpose atca based data acquisition system. In *2014 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC)*, pages 1–4, 2014.

[38] J. Anderson, K. Bauer, A. Borga, H. Boterenbrood, H. Chen, K. Chen, G. Drake, M. Dönszelmann, D. Francis, D. Guest, B. Gorini, M. Joos, F. Lanni, G. Lehmann

Miotto, L. Levinson, J. Narevicius, W. Panduro Vazquez, A. Roich, S. Ryu, F. Schreuder, J. Schumacher, W. Vandelli, J. Vermeulen, D. Whiteson, W. Wu, and J. Zhang. FELIX: a PCIe based high-throughput approach for interfacing front-end and trigger electronics in the ATLAS upgrade framework. *Journal of Instrumentation*, 11(12):C12023–C12023, dec 2016.

[39] K. Biery et al. artdaq: DAQ software development made simple. *Journal of Physics: Conference Series*, 898:032013, oct 2017.

[40] J. S. Marshall and M. A. Thomson. The pandora software development kit for pattern recognition. *The European Physical Journal C*, 75(9), Sep 2015.

[41] R. Acciarri et al. (MicroBooNE Collaboration). The pandora multi-algorithm approach to automated pattern recognition of cosmic-ray muon and neutrino events in the microboone detector, 2017.

[42] A Reynolds. *Evaluating the low-energy response of the ProtoDUNE-SP detector using Michel electrons*. PhD thesis, University of Oxford, 2020.

[43] M. Tanabashi et al. Review of particle physics. *Phys. Rev. D*, 98:030001, Aug 2018.

[44] S. Agostinelli et al. Geant4a simulation toolkit. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 506(3):250–303, 2003.

[45] A. C. Booth, N. Charitonidis, P. Chatzidaki, Y. Karyotakis, E. Nowak, I. Ortega-Ruiz, M. Rosenthal, and P. Sala. Particle production, transport, and identification in the regime of $1-7$ GeV/c. *Phys. Rev. Accel. Beams*, 22:061003, Jun 2019.

[46] Eric D. Church. Larsoft: A software package for liquid argon time projection drift chambers, 2014.

[47] D. Heck, J. Knapp, J. N. Capdevielle, G. Schatz, and T. Thouw. CORSIKA: A Monte Carlo code to simulate extensive air showers. 2 1998.

[48] Geant4 Collaboration. Physics reference manual *Release 10.6*, 2020.

[49] Inaki Ortega Ruiz. *Accurate Profile Measurement of the low Intensity Secondary Beams in the CERN Experimental Areas*. PhD thesis, Ecole Polytechnique, Lausanne, 2018.

[50] Roberto Acciarri et al. The Liquid Argon In A Testbeam (LArIAT) Experiment. *JINST*, 15(04):P04026, 2020.

[51] Stefan Schmitt. Data unfolding methods in high energy physics. *EPJ Web of Conferences*, 137:11008, 2017.

[52] Stephen Dolan. What we measure when we measure σ *Cross-section extraction techniques at T2K*.

[53] F. James and M. Roos. Minuit - a system for function minimization and analysis of the parameter errors and correlations. *Computer Physics Communications*, 10(6):343–367, 1975.

[54] Fred James and Matthias Winkler. Minuit2.

[55] Rene Brun and Fons Rademakers. Root an object oriented data analysis framework. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 389(1):81–86, 1997. New Computing Techniques in Physics Research V.

[56] J. E. Gentle. *Matrix Algebra: Theory, Computations and Applications in Statistics; Second Edition*. Springer, 2017.

[57] Andrew Cudd. *Measurement of the charged current muon neutrino differential cross section on scintillator with zero pions in the final state with the T2K on/off-axis near detectors*. PhD thesis, Michigan State University, 2020.

[58] Stephen Dolan. *Probing nuclear effects in neutrino-nucleus scattering at the T2K off-axis near detector using transverse kinematic imbalances.* PhD thesis, University of Oxford, 2017.

[59] J. Calcutt, C. Thorpe, K. Mahn, and Laura Fields. Geant4reweight: a framework for evaluating and propagating hadronic interaction uncertainties in geant4, 2021.

[60] K. Abe et al. Simultaneous measurement of the muon neutrino charged-current cross section on oxygen and carbon without pions in the final state at t2k. *Phys. Rev. D*, 101:112004, Jun 2020.

[61] K. Abe et al. First combined measurement of the muon neutrino and antineutrino charged-current cross section without pions in the final state at t2k. *Phys. Rev. D*, 101:112001, Jun 2020.

[62] K. Abe et al. First t2k measurement of transverse kinematic imbalance in the muon-neutrino charged-current single-π^+ production channel containing at least one proton. *Phys. Rev. D*, 103:112009, Jun 2021.

[63] K. Abe et al. Characterization of nuclear effects in muon-neutrino scattering on hydrocarbon with a measurement of final-state kinematics and correlations in charged-current pionless interactions at t2k. *Phys. Rev. D*, 98:032003, Aug 2018.

[64] K. Abe et al. Measurement of double-differential muon neutrino charged-current interactions on c_8h_8 without pions in the final state using the t2k off-axis beam. *Phys. Rev. D*, 93:112012, Jun 2016.