

Pion-argon inclusive cross-section measurement on ProtoDUNE-SP

Yinrui Liu^{†,*}

*Department of Physics, University of Chicago,
Chicago, IL 60637, USA*

E-mail: yinrui@uchicago.edu

The pion-argon cross-section measurement is crucial to understand effects such as final-state interactions, which account for a large source of systematic uncertainty in neutrino oscillation experiments. ProtoDUNE-SP, with its charged particle beam data, can provide such experimental constraints. This paper elaborates the methodology to measure the cross-section on large-scale LArTPCs like ProtoDUNE-SP. We use the 1 GeV MC sample to demonstrate the analysis procedures. The cross-sections for pion kinetic energy ranging from 350 MeV to 950 MeV are measured on the MC. The consistency of the MC results with its input values serves as validation of the method and the procedures, which we will later use to perform measurements on the data sample.

1. Introduction

The liquid Argon Time Projection Chamber (LArTPC) is a promising technology used by many modern neutrino experiments, including the future Deep Underground Neutrino Experiment (DUNE) [1]. Knowledge on the pion-argon cross-section is important for us to understand and explain the data collected in the LArTPC. For one thing, neutrinos are detected by their interaction with argon atoms, which can produce hadronic particles, and these features will be used to identify the neutrino. For another, when a neutrino interacts with an argon atom, the initially produced pions can also interact within the nucleus, which will change the kinematics of particles emitted from the atom. This is called the hadronic final-state interactions (FSI) effect. These effects account for a major source of systematic uncertainty in neutrino oscillation analyses, which need constraints from the pion-argon cross-section.

Currently, the pion-argon cross-section is predicted mainly by interpolating data from lighter and heavier nuclei [2], while the experimental measurements are scarce. Only the LADS collaboration [3] and the LArIAT collaboration [4] have performed such measurements. ProtoDUNE-SP [5], which contains 770 tons of LAr, is a prototype for one of DUNE's far detectors. Charged particles, including pions with various momentum modes ranging from 300 MeV to 7 GeV, are delivered into the TPC through a beam plug, making it a good place to measure the pion-argon cross-section.

In this paper, we first introduce the energy-slicing method in Section 2, which is used to extract the cross-section. After that, we describe the analysis procedures in Section 3, including selections, background subtraction, and unfolding. Finally, in Section 4, we use a subset of the 1 GeV MC as a fake data sample to exemplify using the method and the procedures to measure the cross-section and compare to the input cross-sections as validation. A few discussions towards measuring real data are given in Section 5.

2. Energy-slicing method

Based on the idea of the thin-slice method proposed by the LArIAT collaboration [4], we develop the energy-slicing method [6]. Figure 1 shows one TPC of ProtoDUNE-SP. A beam pion is shown as the dashed orange track with some possible daughter particles originating from the end vertex. The initial kinetic energy of the beam pion at the front-face of the TPC is denoted as E_{ini} , and the kinetic energy at the end vertex is denoted as E_{end} . Given these two energies, the track can be divided into several energy slices by pre-defined energy bins, determined according to the energy resolution and the statistics on a case-by-case basis. In this analysis, the energy bins are chosen to be [0, 350, 450, 500, 550, 600, 650, 700, 750, 800, 850, 950, ∞) MeV. The bin edges are indicated by red bars in Figure 1, where the last bar is dashed because the beam pion stops before it reaches that energy. The first complete energy slice is referred to as the *initial slice*; the slice where the track stops is called the *end slice*. If the signal interaction occurs at the end vertex, then the end slice is also called an *interaction slice*.

It is worth noting that E_{ini} is larger than the energy in the initial slice, and the piece of track before the initial slice is referred to as an *incomplete slice*, which is not usable. On the contrary, E_{end} is inside the end slice. For convenience, we define the slice ID from 1 to the number of energy bins, N , starting with the bin with the highest energy. For example, energy bin [950, ∞) has slice

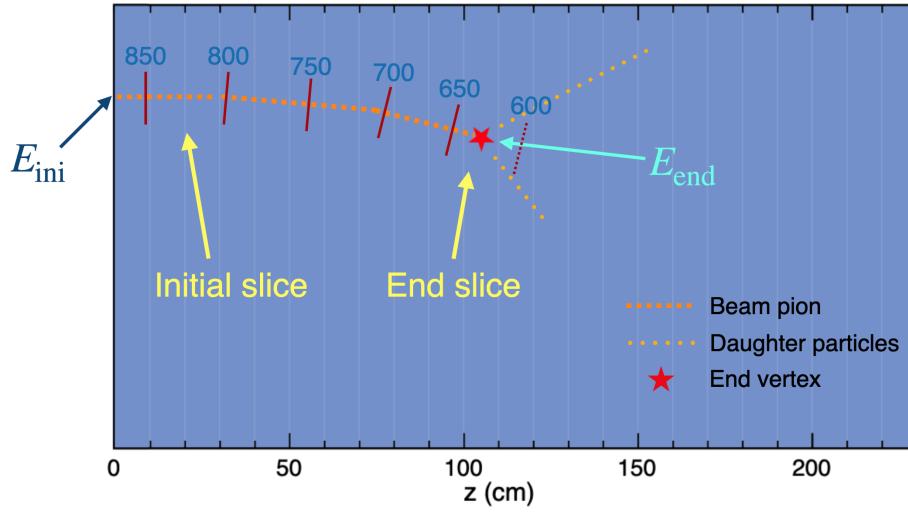


Figure 1: Demonstration of the energy-slicing method. In this example, $E_{\text{ini}} = 866$ MeV, $E_{\text{end}} = 627$ MeV.

ID 1, and $[0, 350]$ has slice ID 12. Therefore, for each event which has a beam pion track in the TPC, there is an initial slice ID, an end slice ID, as well as an interaction slice ID, which is assigned as null if the interaction occurring at the end vertex is not the signal interaction. However, if the beam pion stops inside the incomplete slice, the whole track is not usable, so all three slice IDs will be assigned as null.

For all events with a beam pion track in the TPC, the distribution of initial slice ID is called the *initial histogram* N_{ini} , and similarly we have the *end histogram* N_{end} and the *interaction histogram* N_{int} . The number of incident events, where the beam pion goes through the whole range of the energy bin, forms the *incident histogram* N_{inc} , and it can be calculated as

$$N_{\text{inc}}(i) = \sum_{j=i}^N N_{\text{end}}(j) - \sum_{j=i+1}^N N_{\text{ini}}(j). \quad (1)$$

Finally, the cross-section in each energy bin can be given by

$$\sigma(E) = \frac{M_{\text{Ar}}}{\rho N_A \Delta E} \frac{dE}{dx}(E) \ln \left(\frac{N_{\text{inc}}(i(E))}{N_{\text{inc}}(i(E)) - N_{\text{int}}(i(E))} \right), \quad (2)$$

where M_{Ar} is the molar mass of argon, ρ is the density of LAr, N_A is the Avogadro constant, and ΔE is the energy bin width. The stopping power of pion in LAr, $\frac{dE}{dx}(E)$, is evaluated from theory at the midpoint of each bin using the Bethe-Bloch formula [7] and is assumed to be a constant within each energy bin.

3. Analysis procedures

The whole MC sample is divided into two halves: one is used as *fake data*, and the other is referred to as *truth MC*. Using slice ID histograms derived from its truth energy information, we can extract the truth cross-section of the fake data sample, which can then be used to validate the slicing method. On the other hand, we can also treat fake data the same as real data and perform

the same procedures to measure the cross-section using reconstruction information. The results can be used to validate the procedures, which will later be applied to measuring real data. The reconstructed E_{ini} is derived from the measured beamline instrumented kinetic energy minus a constant upstream energy loss, the value of which is estimated using MC. E_{end} is calculated from E_{ini} and the reconstructed track length based on the theoretical Bethe-Bloch formula as

$$E_{\text{end}} = E_{\text{ini}} - \int \frac{dE}{dx} dx_{\text{reco}}. \quad (3)$$

To conduct the energy-slicing method, we need to select beam pion tracks. Based on machine learning, Pandora [8] selects one track in each event which is most likely to be the beam track. The information of this track and its daughters will then be used for further selections. In addition to technical cuts ensuring useful information from the track, we also perform some specific cuts to veto certain types of backgrounds. The beam particle enters the TPC through the beam plug, and we know approximately where the particle enters the TPC as well as its direction. Thus, we set cuts on the start position as well as the angle of the detected beam track in order to reduce non-beam tracks caused by mis-identification. Proton tracks have a very different dE/dx curve compared to pions and muons [5], so this feature can be used to reduce proton background. In order to mitigate muon background, a CNN-based Michel electron identifier is built [9] and gives the detected beam track a daughter Michel score that can be used to distinguish pions from stopping muons. After full selections, we have 80.1% pion inelastic events, which are the signal events for the inclusive measurement. The remaining backgrounds, estimated using the truth MC, are given in Table 1.

Table 1: The number of events of each type after full selections.

Total truth MC	39970	Pion inelastic	32014	Pion decay	298	Muon	3237	Cosmic origin	12
Non-beam proton	1590	Non-beam pion	1520	Non-beam muon	892	Shower	212	Other non-beam	195

Next, we subtract the background histograms, estimated by truth MC, from the reconstructed histogram after selections, N_{reco} , which represents N_{ini} , N_{end} , or N_{int} . However, to account for difference of background fractions f between the (fake) data sample and the truth MC, we perform a data-driven method by introducing a scale factor α_i for each background type i . Thus, the reconstructed signal histogram is

$$N_{\text{reco}}^{\text{sig}} = N_{\text{reco}} \cdot \left(1 - \sum_i f_i^{\text{data}} \right) = N_{\text{reco}} \cdot \left(1 - \sum_i f_i^{\text{MC}} \cdot \alpha_i \right). \quad (4)$$

In practice, we only consider $\alpha_i \neq 1$ for the three major backgrounds, which are muon, non-beam proton, and non-beam pion. α_i is fitted for the best agreement between data and truth MC in the sideband of a variable distribution used in data selection, where background i dominates. For the fake data sample, the fitted results of the three α_i are all consistent with unity, which is reasonable since the fake data is statistically the same with the truth MC. However, this may not be the case for real data.

Finally, we perform unfolding [10] to transform reconstructed histograms into true histograms, expressed as

$$N_{\text{true}}^{\text{sig}} = M_{\text{unfolding}} \cdot N_{\text{reco}}^{\text{sig}}. \quad (5)$$

The response matrix $R_{ij} = P(x \in \text{reco bin } i | y \in \text{true bin } j)$ is estimated using true beam pion tracks in the truth MC sample. We use the d'Agostini method [11] implemented in the RooUnfold package [12] to model the unfolding matrix $M_{\text{unfolding}}$. To take into account the correlations among the three histograms, N_{ini} , N_{end} , and N_{int} , we combine them as one variable $(N_{\text{ini}}, N_{\text{end}}, N_{\text{int}})$ and perform 3-D unfolding. The index of the combined variable is $N^2 \cdot \text{ID}_{\text{ini}} + N \cdot \text{ID}_{\text{end}} + \text{ID}_{\text{int}}$. To be conservative, the number of iterations in the d'Agostini method is chosen to be as large as 20 to obtain the results in Section 4, but this parameter should be optimized based on further studies.

4. Validation results

The true histograms derived after unfolding can be plugged into Equation 1 and then Equation 2 to calculate the cross-section. Figure 2 shows the 1 GeV fake data cross-section results. The red curve is the Geant4 input, which is used to generate the MC. The green points are derived using the truth information from the fake data sample, and their consistency with the red curve suggests the feasibility of the energy-slicing method. The black points are the measured results for the fake data sample. They also agree with the red curve, which validates the procedures described in Section 3. Cross-sections in the underflow and the overflow energy bins are not calculated. The horizontal error bars indicate the bin width, and the vertical error bars are purely statistical.

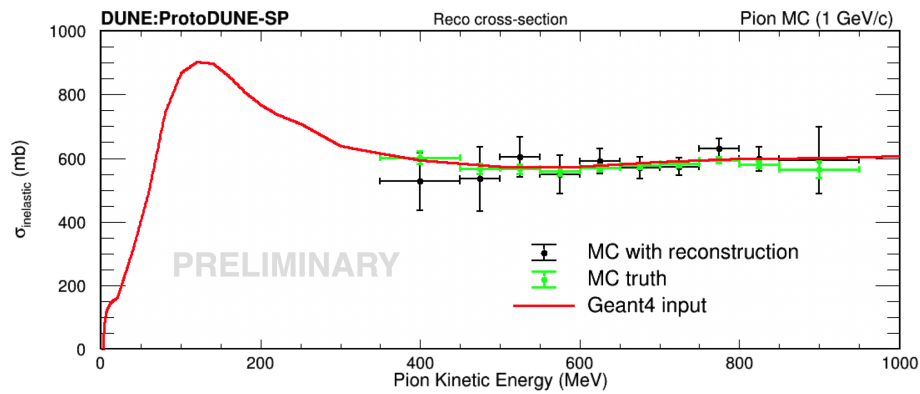


Figure 2: Pion-argon inclusive cross-section results for the 1 GeV fake data sample.

5. Conclusions and outlook

In this paper, we introduce the energy-slicing method and measure the pion-argon inclusive cross-section using a 1 GeV fake data sample for validation. When applying these procedures to measuring real data, we first need to study the differences between data and MC. We will reweight MC accordingly in order to make it more consistent with real data, since both the background estimations and the response matrix rely on MC to model. The differences between real data and MC will also be included in studies of systematic uncertainties.

The energy-slicing method can be applied to other momentum modes in ProtoDUNE-SP as well. It can also be used to measure the exclusive cross-section with further cuts on daughter particles. Moreover, ProtoDUNE Run 2, with a slightly different detector configuration, is currently

being assembled and is scheduled to start data-taking next year. It will offer opportunities to measure the pion-argon cross-section in a wider momentum range with higher statistics.

References

- [1] DUNE collaboration, *Deep Underground Neutrino Experiment (DUNE), Far Detector Technical Design Report, Volume II: DUNE Physics*, [2002.03005](#).
- [2] A.S. Carroll, I.H. Chiang, C.B. Dover, T.F. Kycia, K.K. Li, P.O. Mazur et al., *Pion-Nucleus Total Cross-Sections in the (3,3) Resonance Region*, *Phys. Rev. C* **14** (1976) 635.
- [3] LADS collaboration, *π^+ absorption on N and Ar*, *Phys. Rev. C* **60** (1999) 054610.
- [4] LArIAT collaboration, *Measurement of the π -Ar total hadronic cross section at the LArIAT experiment*, *Phys. Rev. D* **106** (2022) 052009 [[2108.00040](#)].
- [5] DUNE collaboration, *First results on ProtoDUNE-SP liquid argon time projection chamber performance from a beam test at the CERN Neutrino Platform*, *JINST* **15** (2020) P12004 [[2007.06722](#)].
- [6] F. Stocker, *Measurement of the Pion Absorption Cross-Section with the ProtoDUNE Experiment*, Ph.D. thesis, University of Bern, 2021.
- [7] PARTICLE DATA GROUP collaboration, *Review of Particle Physics*, *PTEP* **2022** (2022) 083C01.
- [8] J.S. Marshall and M.A. Thomson, *The Pandora Software Development Kit for Pattern Recognition*, *Eur. Phys. J. C* **75** (2015) 439 [[1506.05348](#)].
- [9] DUNE collaboration, *Separation of track- and shower-like energy deposits in ProtoDUNE-SP using a convolutional neural network*, *Eur. Phys. J. C* **82** (2022) 903 [[2203.17053](#)].
- [10] G. Cowan, *A survey of unfolding methods for particle physics*, *Conf. Proc. C* **0203181** (2002) 248.
- [11] G. D'Agostini, *A Multidimensional unfolding method based on Bayes' theorem*, *Nucl. Instrum. Meth. A* **362** (1995) 487.
- [12] L. Brenner, R. Balasubramanian, C. Burgard, W. Verkerke, G. Cowan, P. Verschuur et al., *Comparison of unfolding methods using RooFitUnfold*, *Int. J. Mod. Phys. A* **35** (2020) 2050145 [[1910.14654](#)].