Article

Pion–Argon Inclusive Cross-Section Measurement on ProtoDUNE-SP †

Yinrui Liu ⓘ **on behalf of the DUNE Collaboration**

University of Chicago, Chicago, IL 60637, USA; yinrui@uchicago.edu

† Presented at the 23rd International Workshop on Neutrinos from Accelerators, Salt Lake City, UT, USA, 30 July–6 August 2022.

Abstract: The pion–argon cross-section measurement is crucial to understanding effects such as final state interactions, which account for a large source of systematic uncertainty in neutrino oscillation experiments. ProtoDUNE-SP, with its beam of charged particles, can provide such experimental constraints. This paper elaborates on the methodology to measure the cross-section on large-scale liquid argon time projection chambers like ProtoDUNE-SP. We use the 1 GeV Monte-Carlo (MC) sample to demonstrate the analysis procedures. The cross-section measurements for pion kinetic energy ranging from 350 MeV to 950 MeV are performed on the MC sample. 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.

Keywords: ProtoDUNE-SP; hadron beam; liquid argon; cross-section

1. Introduction

Liquid argon time projection chamber (LArTPC) technology is promising, and is used in many modern neutrino experiments, including the future Deep Underground Neutrino Experiment (DUNE) [1]. Knowledge of 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 nuclei, which can produce hadronic particles, and these features will be used to identify the neutrino. For another, when a neutrino interacts with an argon nucleus, the initially produced pions can also interact within the nucleus, which will change the kinematics of particles emitted from the nucleus. This is called the hadronic final state interaction (FSI) effect. These effects serve as corrections for data and 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, while the experimental measurements are scarce. Only the LADS collaboration [2] and the LArIAT collaboration [3] have performed such measurements. ProtoDUNE-SP [4], 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 it 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 [3], we develop the energy slicing method [5]. 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 *nitial 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*.
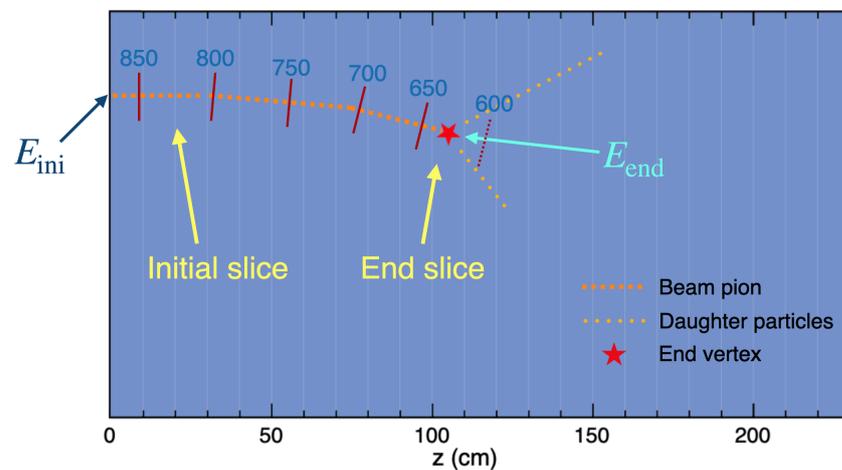
Figure 1. Demonstration of the energy slicing method. In this example, $E_{ini} = 866$ MeV, $E_{end} = 627$ MeV.

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 the slice ID 1, and [0, 350] has the 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 in each energy bin, where the energy of the beam pion reaches the upper boundary of the energy bin, forms the *incident histogram* N_{inc}, and it can be calculated as

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

Finally, from the equation in which the cross-section σ is defined,

$$\frac{N_{int}(i)}{N_{inc}(i)} = 1 - e^{\frac{\rho N_A}{M_{Ar}} \sigma \Delta x},$$ (2)

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

$$\sigma(E) = \frac{M_{\text{Ar}}}{\rho N_A \Delta E} \frac{dE}{dx}(E) \ln\left(\frac{N_{\text{inc}}(i(E))}{N_{\text{inc}}(i(E)) - N_{\text{int}}(i(E))}\right),$$ (3)

where M_{Ar} is the molar mass of argon, ρ is the density of LAr, N_A is the Avogadro constant. ΔE is the energy bin width, while $\Delta x = \frac{dE}{dx} \cdot \Delta E$ is the spatial bin width. The stopping power of the pion in LAr, $\frac{dE}{dx}(E)$, is evaluated from theory at the midpoint of each bin using the Bethe–Bloch formula [6] and is assumed to be a constant within each energy bin.

3. Analysis Procedures

The whole MC sample (the nominal test beam simulation of the ProtoDUNE-SP experiment [4] is employed) is divided into two halves: one is used as *fake data*, and the other is referred to as *true MC*. Using slice ID histograms derived from its true energy information, we can extract the true 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

$$E_{\text{end}} = E_{\text{ini}} - \int \frac{dE}{dx} dx_{\text{reco}}.$$ (4)

To conduct the energy slicing method, we need to select beam pion tracks. Based on machine learning, Pandora [7] 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 misidentification. Proton tracks have a very different dE/dx curve compared to pions and muons [4], so this feature can be used to reduce proton background. In order to mitigate muon background, a CNN-based Michel electron identifier is built [8] and gives the detected beam track a daughter Michel score that can be used to distinguish pions from stopping muons. After full selection, we have 80.1% of pion inelastic events, which are the signal events for the inclusive measurement. The remaining backgrounds, estimated using the true MC, are given in Table 1.

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

Total true MC	39,970	Pion inelastic	32,014	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 the true MC, from the reconstructed histogram after selections, N_{reco}, which represents N_{ini}, N_{end}, or N_{int}. However, to account for the difference in background fractions f between the (fake) data sample

and the true 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

$$N_{\text{reco}}^{\text{sig}} = N_{\text{reco}} \cdot \left(1 - \sum_i f_i^{\text{data}}\right) = N_{\text{reco}} \cdot \left(1 - \sum_i f_i^{\text{MC}} \cdot \alpha_i\right). \tag{5}$$

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 the true 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 are statistically the same as the true MC. However, this may not be the case for real data.

Finally, we perform unfolding [9] to transform reconstructed histograms into true histograms, expressed as

$$N_{\text{true}}^{\text{sig}} = M_{\text{unfolding}} \cdot N_{\text{reco}}^{\text{sig}}. \tag{6}$$

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 true MC sample. With R_{ij} as input, we use the d'Agostini method [10] implemented in the RooUnfold package [11] 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 $N^2 \cdot \text{ID}_{\text{ini}} + N \cdot \text{ID}_{\text{end}} + \text{ID}_{\text{int}}$ (the definition of the combined variable can be understood as flattening a 3D array into a 1D array), where N indicates the number of bins in each histogram. 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 (3) 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 true 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.
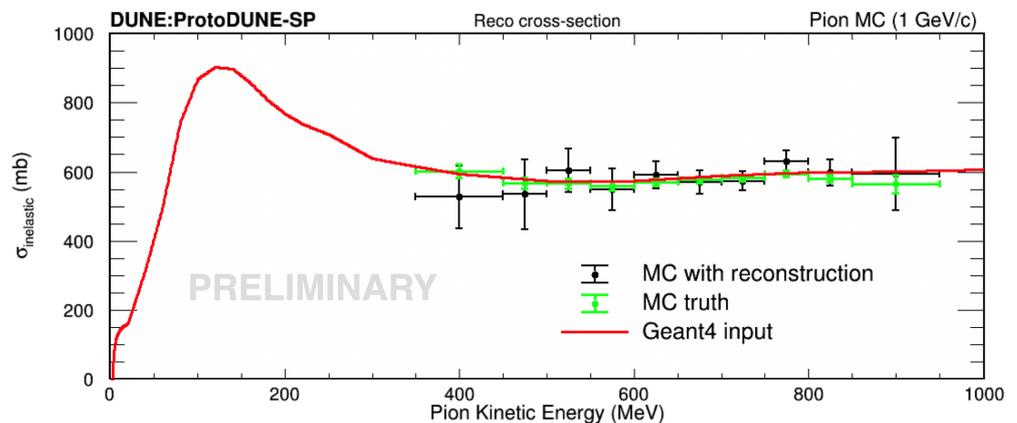

Figure 2. Pion–argon inclusive cross-section results for the 1 GeV fake data sample.

5. Conclusions and Outlook

In this paper, we describe the energy slicing method and use a 1 GeV fake data sample to validate the procedures of measuring pion–argon-inclusive cross-sections. 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 collecting data next year. It will offer opportunities to measure the pion–argon cross-section in a wider momentum range with higher statistics.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The author declares no conflict of interest.

References

1. Abi, B.; Acciarri, R.; Acero, M.A.; Adamov, G.; Adams, D.; Adinolfi, M.; Ahmad, Z.; Ahmed, J.; Alion, T.; Monsalve, S.A.; et al. Deep Underground Neutrino Experiment (DUNE), Far Detector Technical Design Report, Volume II: DUNE Physics. *arXiv* **2020**, arXiv:2002.03005.
2. Rowntree, D.; Androic, D.; Backenstoss, G.; Bosnar, D.; Breuer, H.; Dooling, T.; Furic, M.; Gram, P.A.M.; Gregory, N.K.; Hoffart, A.; et al. π^+ absorption on N and Ar. *Phys. Rev. C* **1999**, *60*, 054610. [CrossRef]
3. Gramellini, E.; Ho, J.; Acciarri, R.; Adams, C.; Asaadi, J.; Backfish, M.; Badgett, W.; Baller, B.; Basque, V.; Rodrigues, O.B.; et al. Measurement of the π–Ar total hadronic cross-section at the LArIAT experiment. *Phys. Rev. D* **2022**, *106*, 052009. [CrossRef]
4. Abi, B.; Abud, A.A.; Acciarri, R.; Acero, M.A.; Adamov, G.; Adamowski, M.; Adams, D.; Adrien, P.; Adinolfi, M.; Ahmad, Z.; et al. First results on ProtoDUNE-SP liquid argon time projection chamber performance from a beam test at the CERN Neutrino Platform. *JINST* **2020**, *15*, P12004. [CrossRef]
5. Stocker, F. Measurement of the Pion Absorption Cross-Section with the ProtoDUNE Experiment. Ph.D. Thesis, University of Bern, Bern, Switzerland, 2021.
6. Workman, R.L.; Burkert, V.D.; Crede, V.; Klempt, E.; Thoma, U.; Tiator, L.; Agashe, K.; Aielli, G.; Allanach, B.C.; Amsler, C.; et al. Review of Particle Physics. *PTEP* **2022**, *2022*, 083C01. [CrossRef]
7. Marshall, J.S.; Thomson, M.A. The Pandora Software Development Kit for Pattern Recognition. *Eur. Phys. J. C* **2015**, *75*, 439. [CrossRef]
8. Abud, A.A.; Abi, B.; Acciarri, R.; Acero, M.A.; Adames, M.R.; Adamov, G.; Adamowski, M.; Adams, D.; Adinolfi, M.; Aduszkiewicz, A.; et al. Separation of track- and shower-like energy deposits in ProtoDUNE-SP using a convolutional neural network. *Eur. Phys. J. C* **2022**, *82*, 903. [CrossRef]
9. Cowan, G. A survey of unfolding methods for particle physics. *Conf. Proc. C* **2002**, *0203181*, 248–257.

10. D'Agostini, G. A Multidimensional unfolding method based on Bayes' theorem. *Nucl. Instrum. Method Phys. Res. Sect. A Accel. Spectrom. Detect. Assoc. Equip.* **1995**, *362*, 487–498. [CrossRef]
11. Brenner, L.; Verschuuren, P.; Balasubramanian, R.; Burgard, C.; Croft, V.; Cowan, G.; Verkerke, W. Comparison of unfolding methods using RooFitUnfold. *Int. J. Mod. Phys. A* **2020**, *35*, 2050145. [CrossRef]