

Study of Kaon Identification in LArIAT

Thesis submitted under the recommendation of
Ed Kearns, Ph.D, for the degree of Bachelor of Arts

Daniel L. Smith

April 2018

Contents

1	Introduction	3
1.1	Proton Decay Theory	3
1.2	Liquid Argon Time Projection Chamber (LArTPC)	4
2	LArIAT	6
2.1	Experiment Introduction and Physics Goals	6
2.2	LArIAT's LArTPC	7
2.3	Beam & Beam Line Instrumentation	9
2.3.1	Beam	9
2.3.2	TOF system	9
2.3.3	Wire Chamber System	10
2.3.4	Other Detectors	10
2.4	LArIAT's Role in Proton Decay	11
3	Beam Line Selection of Kaon Sample	12
3.1	TOF Calibration and Selection	12
3.2	Wire Chamber Calibration	13
3.3	Wire Chamber to TPC Track Matching	21
3.4	Mass Selection	22
3.5	Kaon Data Sample	24
4	Proton Decay Signals and Backgrounds	26
4.1	dE/dx Calibration	27
4.2	Bragg Peak Fit	29
4.3	PIDA Performance on Simulation	32
4.4	Data-MC Comparison of PIDA	35
4.5	PIDA Application to Proton Decay	36
4.6	Proton Decay MC	36
5	PIDAN, a proposed improvement	40
6	Conclusion	45
7	Acknowledgements	46

1 Introduction

1.1 Proton Decay Theory

For as long as science has been a practice, there has been an attempt to understand all of nature in a single theory of everything. A single defining theory would be the end of our understanding of the universe: with it, all physical reality can be simulated and understood. Democritus predicted that forces between all atoms described all of nature. The contemporary understanding is that our universe is governed by the Standard Model of particle physics, a mating of two more fundamental theories - quantum chromodynamics and electroweak theory. However, this Standard Model is not the coveted theory of everything, but our closest approximation to it and the best-working theoretical framework that can be studied to approach the desired end goal.

Electricity and magnetism were discovered to be a single force, electromagnetism, by scientists in the 19th century. Many decades later, electroweak theory was created from the unification of electromagnetism and the weak forces, a behavior only observed at energy scales above 100 GeV. The next theoretical step is to combine the strong nuclear and electroweak forces in what are called Grand Unified Theories (GUTs). The difficulty of such theories is that the unification of the forces can only be observed at energy scales much above what is currently possible in particle accelerators: approximately 10^{16} GeV, compared to the 10^5 available at the LHC. To find evidence for GUTs, we must look for evidence outside of accelerators.

Many GUTs unify the forces with the addition of the theory of supersymmetry (SUSY). SUSY introduces a superpartner for each known particle that has a half-integer difference in spin. The consequence of SUSY GUTs is that the global gauge symmetry of baryon number is broken. Under our current understanding of the universe, baryon number conservation prevents protons from decaying because they are the least massive baryon. However, evidence of proton decay would show that baryon number is not conserved and would be very suggestive of the naturalness of SUSY GUTs.

For the channel $p \rightarrow K^+\bar{\nu}$, the existence of the kaon as a decay product is a result of a colored Higgsino, the superpartner of the Higgs, that creates a baryon-number violating dimension-five operator. In minimal supersymmetric gauge SU(5) theories, the decay rate for the kaon has already been effectively ruled out. However, using gauge symmetry SU(10) or a non-minimal SU(5) theory, the predicted lifetime is above the current experimental limit of 5.9×10^{33} years [1][2].

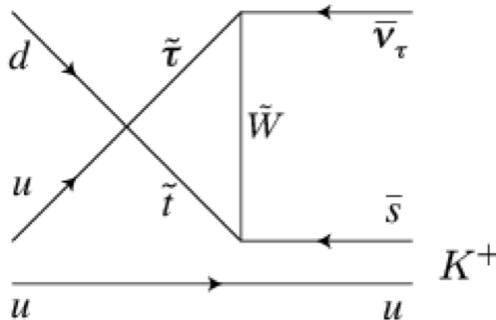


Figure 1: Feynman diagram of the theoretical decay $p \rightarrow K^+\bar{\nu}$. [1]

Due to the extremely long theoretical lifetime (greater than 10^{34} years) of the proton, a detector created to measure proton decay must be able to observe an extremely large number of protons (on the order of tens of tons of active volume) for long time periods (approximately years), all with a well-understood and low background. The choice detector has historically been water Cherenkov detectors. However, in the theoretical decay channel $p \rightarrow K^+\bar{\nu}$, the produced kaon travels slower

than the speed of light in water and thus has a very low event identification efficiency ($< 10\%$) in this detector type[2]. The next generation of liquid argon time projection chambers (LArTPCs) will be able to make competing measurements of this channel, even with a thirty year time disadvantage to water Cherenkov detectors, due the LArTPC's millimeter-level tracking and high energy resolution. Under optimal detector performance, LArTPCs have been theorized to be able to detect this channel with 98.5% efficiency and a pseudo-zero background [3].

1.2 Liquid Argon Time Projection Chamber (LArTPC)

A time projection chamber (TPC) is a detector that is capable of reconstructing 3D particle trajectories and interactions. This is done by using an electric field to collect ionization electrons, created from a charged particle, on instrumented wire planes. A LArTPC uses liquid argon as its active material because it is density, cheap and has high scintillation light yield, useful for triggering.

The detector is composed of a volume of material that becomes ionized when a charged particle passes through the medium. An electric field is applied to the volume via a high-voltage cathode plane on one side and several lower-potential anode wire planes on the opposite side of the volume. The ionized electrons drift towards the anode wires and then induce charge and collect on the wires. This charge is saved using charge-to-time (QTC) digitizers. An event can be triggered by the the ionization and scintillation light collected by photo-multiplier tubes (PMTs) situated just outside of the TPC. The light is orders of magnitude faster than the electron drift velocity ($\sim 1 \times 10^3$ m/s), making it useful as an effective event trigger. Timing information is essential for 3D reconstruction as time is transformed into position via the known electron drift velocity.

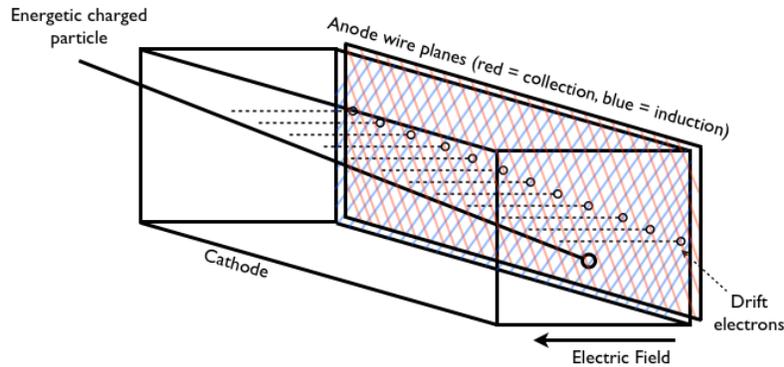


Figure 2: Pictorial diagram of a particle entering a TPC and how the electrons are collected. [4]

The detector is composed of a volume of a material with desirable ionization properties. The original TPC design by David R. Nygren in the late 1970s [6] called for a gaseous medium to allow for stronger electric fields and thus higher event rates. For future LArTPCs, a denser material is optimal as it allows for more protons in the active volume for proton decay searches and allows more active target material for a neutrino detection experiment. As well, a denser material allows for the TPC to act as a more effective lossless calorimeter. A calorimeter reconstructs particle energy by stopping the particle in the volume and summing the energy that was deposited. A denser material is more likely to completely contain events, allowing for better energy reconstruction. Out of available materials that fit the density requirement, liquid argon was chosen for its abundance, low price and high scintillation properties [7].

The sensing anode wire planes are typically composed of three planes: one shield plane (used to optimize electric field uniformity) and two sensing planes of differing voltage and differing angle in

regards to each other. The differing voltage on each plane allows for ionization electrons to move past the outside most plane(s) (induction planes) and collect on the inside most plane (collection plane). This is done to allow multiple planes to observe the same charge, giving multiple non-orthogonal views of an event. The planes are not typically perpendicular to one-another to lower reconstruction ambiguity [7]. Due to electrons moving towards and then moving away from the induction planes, the signal is bipolar compared to the unipolar signal on the collection plane.

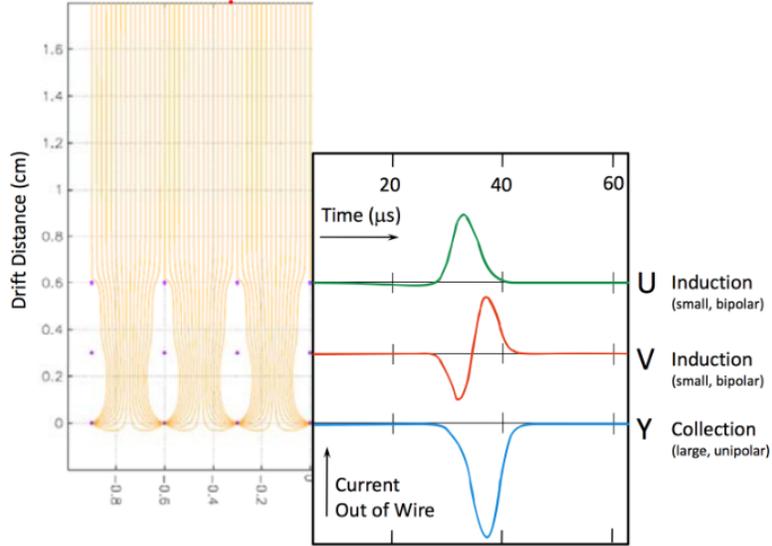


Figure 3: Electric field and signals on each wire plane. [8]

The properties of electrons drifting in a material are well understood and can be used both to convert induced signals into reconstructed energy and to reconstruct 3D particle trajectories. Electron energy is reconstructed by correcting for electron lifetime (electron attenuation from impurities in the material, a function of time spent drifting), electron recombination (electrons recombining with the argon before beginning to drift, a function of electric field and charge-space density) and ultimately ADC-to-Charge calibration done using the known energy deposition of minimum ionizing particles, typically cosmic muons. Trajectory reconstruction is more complicated. After an event is triggered by the PMTs and data begins to be collected, the electrons that reach the sensing wires first in time are closest to the wire planes in space. These signals are given a position in the drift dimension by using their arrival times and the known electron drift velocity. This extra dimension, combined with the two dimensions from the minimum two wire planes, allows for full three dimensional reconstruction.

2 LArIAT

2.1 Experiment Introduction and Physics Goals

Liquid Argon in a Test Beam (LArIAT) is an experiment at the Fermi National Accelerator Laboratory composed of a small LArTPC deployed in an instrumented beam of electrons, muons, pions, protons, kaons and, in small number, antiprotons and deuterons. The experiment's instrumented beam line allows for the selection of particle species and, thus, a varied physics agenda for the experiment.

LArIAT originated as a low-cost calibration for the LArTPC detector technology on known particle species as an input for on-going and future neutrino and proton decay experiments. The experiment is composed of the decommissioned ArgoNeut cryostat, the refurbished ArgoNeut TPC, an argon filtration system developed on-site by the LAPD experiment, and beam line detectors made available by the Fermilab Test Beam Facility.

LArIAT's most developed analysis is a total pion interaction cross section on liquid argon. This measurement, besides being the first such measurement on argon, is used to verify and tune simulation models while also verifying completely automatic event reconstruction. All previously published results using LArTPCs have been either partially or completely done manually, making LArIAT one of the first to be able to produce a physics results using the automatic software alone.

Besides the pion analysis and this thesis on kaon identification for proton decay searches, ongoing and future projects in LArIAT include electron-photon separation efficiencies, proton-antiproton annihilation as an input for nucleon decay searches, muon charge-sign determination, general particle identification efficiencies, and general simulation-verification studies for the reconstruction software. LArIAT is also a test bed for several different light-collection systems with the aim of improving energy resolution via scintillation light collection.

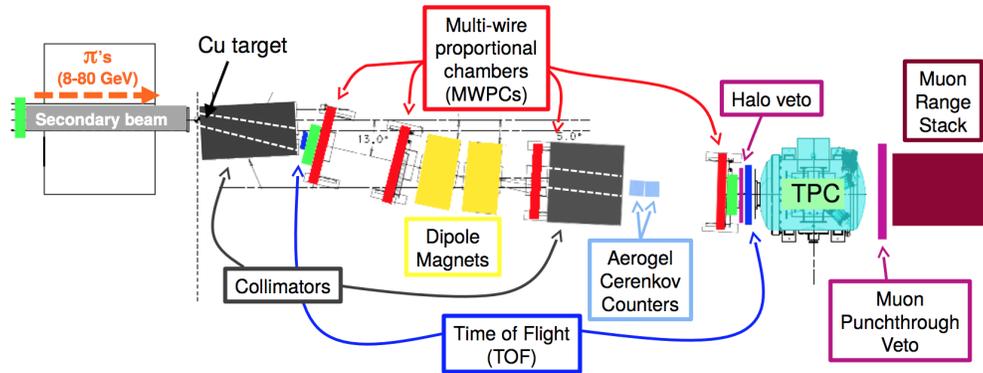


Figure 4: LArIAT's beam line instrumentation and LArTPC. The beam is from left to right.

LArIAT has gone through several runs over a three year period punctuated by Fermilab's seasonal beam shutdown. The first run was primarily a calibration run to develop reconstruction techniques, improve purity and tune beam settings. The second run was used to collect large amounts of statistics with the tuned system for the various physics goals of the experiments. The third and final run was used primarily to test different wire pitches and electron collection techniques, including PixLAR [9]. For this analysis, only data from the second run, called Run 2, is used.

2.2 LArIAT's LArTPC

The TPC is a $47\text{w} \times 40\text{h} \times 90\text{l cm}^3$, 170 L liquid argon time projection chamber. It sits after the majority of the beam line instrumentation to allow for the traditional identification and momentum reconstruction of the particles entering the TPC.

LArIAT is equipped with three anode planes. For Run II, the planes are composed of parallel conducting wires separated by 4 mm with each plane separated from its neighboring plane by 4 mm. The planes are called, from inner to outer most, shield, induction and collection planes and are set at a bias voltage of -298.8 V , -18.5 V and $+338.5\text{ V}$, respectively. The planes are at 0° , $+60^\circ$ and -60° from the beam direction, respectively. The cathode plane is a sheet of copper located on the opposite side of the TPC from the anodes planes. It set at a bias voltage -23.17 kV , creating a $468 \pm 21\text{ V/cm}$ electric field throughout the active volume of the TPC.

The uniformity of the electric field is maintained by the combination of field shapers, a stable voltage applied to the cathode plane, and the shield plane. The field shapers are 23, 1 cm rectangular copper bands that line the walls of the TPC perpendicular to the cathode and anode planes. Each spacer is held at a regularly decreasing potential as to reduce the fringe fields on the edges of the TPC. The stability of the cathode plane is maintained via two pots between the high voltage supply and TPC feed-through which act as two low-pass filters. The shield plane acts to satisfy the boundary conditions of a static, uniform field set up by the cathode and field shapers. By not allowing a discontinuity, the shield plane maintains the uniform field until deposited charge drift past the plane.

The induction and collection plane are instrumented and their signals are recorded, each with 240 wires and thus 240 read-out channels. The signals are first amplified by cold preamp motherboards that are inside of the cryostat and in the liquid argon to lower noise. The signals then leave the cryostat via a feed through and are sent to CAEN V1740 64-Channel Waveform Digitizers to be sampled at 3.90625 MHz, or a 2^8 ns period.

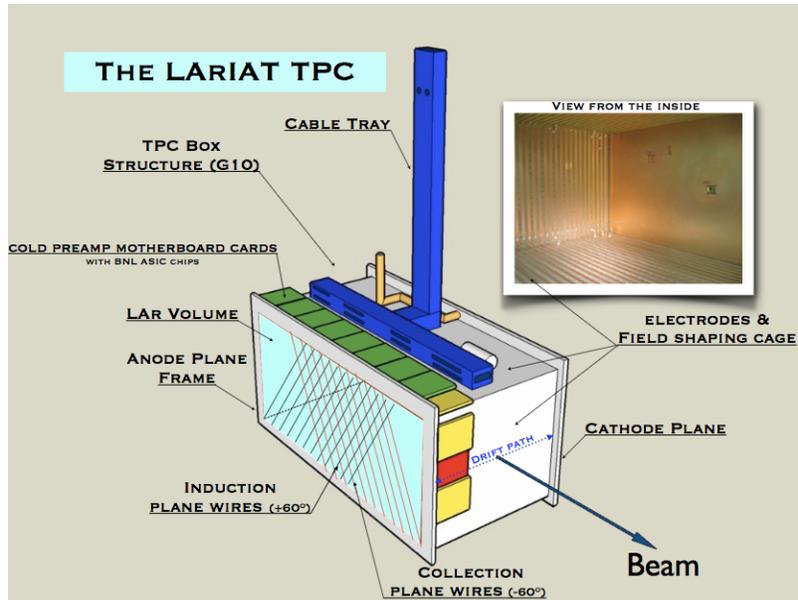


Figure 5: Diagram of LArIAT's TPC, including the anode and cathode planes, the read out electronics and the cable feed through out of the cryostat.

For the effective operation of the LArTPC, the argon is kept at cryogenic temperatures using a cryostat. In LArIAT, the cryostat was in large part recycled from the ArgoNeut experiment, a LArTPC neutrino experiment in the NuMI beam at Fermilab that finished data-taking in March 2010 [5]. The cryostat is composed of two stainless steel vessels separated by an insulating vacuum. The inner volume of the cryostat is $76r \times 130l$ cm and contains the liquid argon, the TPC, the cold readout electronics and the scintillation light collection system. There exists three shafts connecting the inner and outer vessels: on top of the vessel to refill argon, supply high voltage and readout the TPC signals; on the side for the scintillation light collection system; and on the bottom for the argon purification system.

With the original ArgoNeut cryostat, an entering beam particle would need to traverse two 4.8 mm stainless steel ports and 15 cm of liquid argon before reaching the active TPC volume. To reduce the amount of dead material, a beam window was added to the outer vessel and a so-called excluder was added to the inner vessel. The outer vessel's window is 22.9 cm in diameter and is blanked with a 0.4 mm titanium sheet. The inner vessel's excluder is a concavity that reduces the amount of argon separating the inner vessel wall and the active TPC volume. These modifications amount to $0.4 X_0$ of material as opposed to the original $1.6 X_0$.

ArgoNeut's purification system had a flow rate deemed too long to achieve the desired level of purity for LArIAT. Instead, LArIAT uses a system developed by LAPD, a test stand experiment at Fermilab studying argon purity [10]. The purification system receives argon from boiled-off gases from the top of the cryostat and from the draining pipe on the bottom of the cryostat. The boil-off gases are filtered for nitrogen, re-condensed, and mixed with the liquid drained from the bottom of the TPC. The mixture is then pumped through a cartridge bedded with molecular sieves and activated copper material in order to remove the water oxygen impurities. The resulting purified argon returns to the TPC after being re-cooled in a condenser.

Measurement of the argon purity is done via cosmic muons that trigger on two scintillator paddles on opposite corners of the TPC. Cosmic muons offer an ample source of minimally ionizing particles and the orientation of the paddles forces a known and varying drift distance. The combination of a known amount of deposited charge and drift distance can be used to find electron lifetime, a function of argon purity. Purity changes with time and needs to be measured on a sub-run basis. For this analysis, the electron lifetime is approximately $800 \mu\text{s}$ on average.

2.3 Beam & Beam Line Instrumentation

2.3.1 Beam

LArIAT's particle beam is low energy (up to 2 GeV) and low intensity because of the experiment's physics goals, small volume and the inherently slow data acquisition rate of the TPC. To meet the desired beam specifications, the Fermilab Test Beam Facility and Accelerator Division supplies LArIAT with a "tertiary beam," or a beam created from two separate targets. The primary beam is composed of 120 GeV protons from the lab's Main Injector. These protons are directed to the test beam facility where they are made incident upon an aluminum target. The resulting produced beam have a tunable energy between 8 to 68 GeV. This secondary beam is incident upon a copper target inside of LArIAT's enclosure, creating a tertiary beam of particles. It is this tertiary beam that the experiment studies. To lower beam intensity and background halo, the target exists within a collimator and the beam line instrumentation is at 13° from the secondary beam.

The tertiary beam arrives once a minute in 4.2 second-long spills, corresponding to the rate that protons are removed from the Main Injector. The spill has an RF structure, coined "RF buckets", of 53 MHz or 19 ns. The tertiary beam was designed to have a 25 Hz event rate, or events arriving at the TPC every ~ 10 ms. The TPC has a drift window of $O(100 \mu s)$ so multiple particles entering the TPC within the same drift window should be kept to a minimum. For this analysis, the data was taken with a slightly higher beam energy (64 GeV) than originally planned to optimize kaon production, leading to more pile up than expected. This is discussed in further detail in the WC-TPC matching section, Section 3.3, of this thesis.

2.3.2 TOF system

LArIAT measures particle velocity for particle identification and mass reconstruction. This is done by finding the time difference between signals create by a particle between two scintillator paddles in the beam line. The two scintillator paddles, each with two PMTs, are positioned in the most downstream portion of the beam and immediately in front of the cryostat in order to maximize the distance traveled, and thus time difference, of a given particle. For Run II, the upstream and downstream paddles have an active area of approximately 10×6 cm and 14×14 cm, respectively. To maximize light collection, each PMT is mounted to the scintillator paddle via a light guide. PMT signals are digitized by a CAEN V1751 digitizer at a sampling rate of 1 ns.

The process of reconstruction involves:

- Find the rising edge of signals for each PMT
- Match coincidence signals on each paddle. This process is done to reduce noise.
- For each set of upstream and downstream hits within 100 ns of each other, take the time difference and save the result.

Edge finding is done by performing a cut on the numerical derivative of the digitized PMT signal. The rising edge of a pulse above the numerical derivative of the pedestal is labeled as a hit. The 5th order numerical derivative is used and is equal to

$$f'(x) \approx \frac{-f(x+2h) + 8f(x+h) - 8f(x-h) + f(x-2h)}{12h} \quad (1)$$

where $f(x)$ is the digitized PMT signal and h is the granularity of the numerical derivative, or 1 ns as determined by the digitizer.

2.3.3 Wire Chamber System

The wire chamber system in LArIAT tracks particles through the beam line instrumentation and into the TPC. It is composed of four multi-wire proportional chambers that were developed in-house by the Test Beam Facility team [11].

The wire chambers are filled with 85% argon gas and 15% isobutane. Each has an active region of 12.8×12.8 cm, with 128 vertical and 128 horizontal wires spaced at 1 mm. The wires are held at a bias voltage between 2400 and 2500 V. In a test beam setting, they performed at an event trigger efficiency of approximately 98%.

As a charged particle passes through wire chamber, it ionizes the gas mixture. The ionization electrons drift towards the nearest wire due to the electric field created by the wire. As the electrons approach a wire, they enter a region of very high electric field, accelerate and begin an electromagnetic avalanche. The avalanche creates a signal on a wire, which is then further amplified, discriminated and sent to TDCs to be saved at a resolution of 1.18 ns. The TDCs can accept hits on the order of megahertz, far greater than the intensity of LArIAT's beam.

The reconstruction of the wire chamber tracks is a large part of the analysis for this thesis. For this reason, it is discussed in detail in Section 3.2.

Wire chamber tracking is used for momentum determination. The amount a particle is deflected in a magnetic field is a function of the particle's momentum. Two dipole bending electromagnets are deployed between the two upstream and the two downstream wire chambers. The relationship between magnet current and magnetic field is well understood, so momentum can be reconstructed with the amount of particle deflection as determined by the tracking system. This can be expressed as the simple equation,

$$p = \frac{q}{\theta_{def}} \int_0^L B_y dz, \quad (2)$$

where θ_{def} , the amount of particle deflection, is retrieved from the tracking system and the integral can be performed from geometric considerations.

LArIAT's tracking system is also used to make a geometric connection between a track reconstructed in the beam line and a track reconstructed in the TPC. Due to the large difference in time scales of each detector, timing can not be used to make a correlation between beam line and TPC objects. By projecting the beam line track forward from the last wire chamber to the front of the TPC, tracks can be spatially matched, assuring that the particle that triggered the beam line is the same that created a track within the TPC. This is especially important because of track multiplicity at higher energy beam settings.

2.3.4 Other Detectors

Not every system in LArIAT's beam line instrumentation was used in this analysis. The muon range stack is a detector composed of several layers of steel with scintillator paddles and PMTs between the steel layers. The stack is positioned after the TPC and is used primarily to separate relativistic pions from muons. An aerogel threshold Cherenkov detector is deployed in LArIAT to also separate pions and muons. The detector has two blocks aerogel, each with a slightly different index-of-refraction and each with two PMTs. Since a particle only produces Cherenkov radiation after a threshold velocity and this velocity is dependent on mass for a given momentum and index of refraction, the detector can be used to separate pions and muons in the energy range from 200 to 400 MeV/c. Also in the enclosure are various scintillator paddles positioned around the beam line to act as triggers or vetos.

2.4 LArIAT's Role in Proton Decay

The LArIAT experiment has collected a set of real data K^+ events in liquid argon, the first such set to exist. Besides the fundamental cross section measurements, LArIAT can begin the first steps in identifying kaons using LArTPC information as an input to future proton decay searches.

LArIAT is unable to perform an entire data-driven proton decay sensitivity study for the desired channel. With the low volume of the detector, all pions produced in kaon decay leave the detector before stopping or interacting, making their identification extremely difficult. Future experiments will not suffer from this as they will have the volume and the fiducial volume cuts so that charged particle tracks never leave the detector before stopping. For this reasons, this analysis and LArIAT does not claim to perform an entire proton decay study to its fullest. What the experiment can offer, however, is a data driven analysis of one element of the particle identification scheme to be used in a full analysis: particle identification via track energy deposition.

This thesis is dedicated to demonstrating the current identification power of energy deposition particle identification algorithms on K^+ as an input to future proton decay studies. Roughly, it is split into four sections:

- Work done to create a relatively pure sample of K^+ events using the beam line system, resulting in 1998 K^+ candidate events with an approximate background of 10%.
- Discussion of current energy deposition PID algorithms, how they perform on LArIAT beam-like simulated events, where they fail and ultimately how they compare to LArIAT data. I then discuss potential backgrounds to proton decay events that such algorithms must be able to suppress.
- Current energy deposition PID performance on an approximate proton decay MC sample.
- Proposed improvement to current energy deposition PID algorithms, called PIDAN.

3 Beam Line Selection of Kaon Sample

Due to the relative rarity of the kaon in the beam entering LArIAT - approximately 0.5% of the total - it was necessary to improve and verify beam line particle identification. This section includes analysis of the time-of-flight system and the tracking system to ultimately select particle species using reconstructed mass.

3.1 TOF Calibration and Selection

The calibration and efficiency of the time-of-flight (TOF) system is important for the effective selection of kaon events.

For correct mass determination, measured velocity must be accurate. Due to issues arising from using different clocks, cables, PMTs or other spurious effects, there may exist a timing offset in the TOF system which would cause incorrect mass reconstruction. To remove this bias, an absolute calibration of the TOF system is done using relativistic particles. These particles are traveling at a known velocity, approximately the speed of light and, since the TOF paddles are separated by a known distance determined by surveying, these particles have an absolute TOF that can be used to determine an absolute offset in the system.

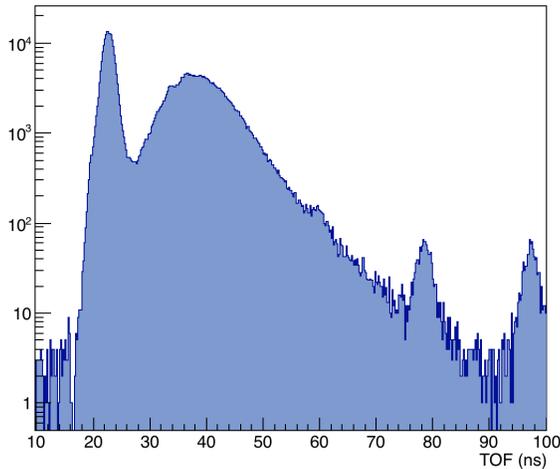


Figure 6: Reconstructed TOF distribution from Run 2 Data.

The distribution of reconstructed TOF from LArIAT Run 2 data, as observed in Figure 6, exhibits a sharp peak at low values of TOF. This peak is from relativistic pions and muons in the beam. The mean of this peak is used to find the absolute offset while the standard deviation determines the resolution of the measurement. The Gaussian fit over this peak is

$$TOF_{rel} = 22.29 \pm 0.77 \text{ ns.} \quad (3)$$

The distance between the TOF paddles is equal to 6.652 m, making the expected time difference for relativistic particles 22.19 ns. The measured TOF value is consistent with the expected value, within error bars.

In the relativistic peak, the beam line system has no particle selection power. In this region and at the energies of LArIAT, the timing resolution from TOF is too poor to separate electrons, pions and muons. For the effective selection of a pure sample of kaons, this peak is cut from the sample. Using the above fit, the cut removes events that are below four standard deviations from the mean, or $TOF_{reco} < 25.37 \text{ ns}$.

The efficiency of the TOF reconstruction cannot yet be determined from simulation because LArIAT’s beam line simulation is still being developed. However, a data-driven method can be devised using the TOF coincidence trigger that is saved along with the digitized TOF signals. This trigger means that signals arrived within a time window at all four PMTs on both upstream and downstream paddles as determined by a discriminator in hardware. For the full analysis, a TOF trigger is not required because the coincidence rate is very low, potentially due to high discriminator levels and timing slew within the scintillation paddles. However, these events can be used to determine efficiency because, if the event has a TOF trigger on 4 PMTs, it effectively must have a reconstructible TOF.

On a subsample of 6500 events from Run 2, 245 have a TOF trigger and, of these, 240 have a reconstructed TOF. This is an efficiency of $98 \pm 2 \%$.

This approach for determining efficiency could be flawed because the the trigger may preferentially pick large signals, making the triggered subset unrepresentative of the entire data set. To investigate this, I plot in Figure 7 the area-normalized distributions of signal amplitudes for triggered TOF events (in blue) and all non-cosmic events (in purple) from the same 6500 events. The agreement between the two sets affirms that the triggered set is representative of the whole.

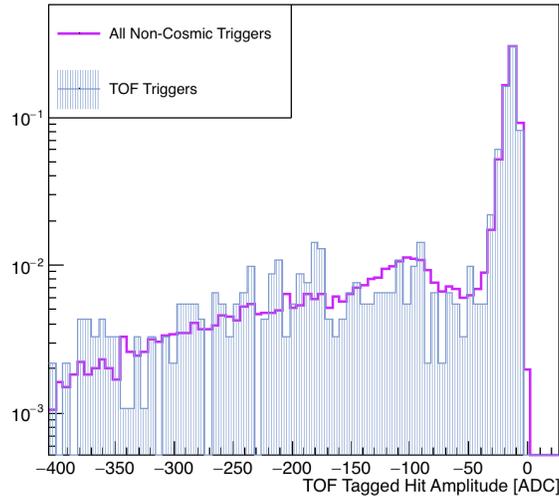


Figure 7: Comparison between hit amplitudes of TOF Triggered and All Non-Cosmic Events. Area normalized.

The rate of misreconstructed TOF is very low due to the coincidence requirements of hits on both PMTs in a paddle. This rate can be approximated by looking at unrealistically low TOF values, below ~ 18 ns, in Figure 6. This background is three orders of magnitude below the signal. An additional background is present from the RF beam structure, seen in the plot as peaks visible at 40 ns, 60 ns, 80 ns and 100 ns. These events are suppressed by wire chamber reconstruction and exist in a area of the TOF-momentum phase space outside of the area of interest for kaon selection.

3.2 Wire Chamber Calibration

LArIAT’s wire chamber system is used for momentum determination and for matching beam line tracks with tracks reconstructed within the TPC. It is a relatively complex system and, due to requirements created by the rarity of the kaon in the beam, roughly half of the work for this thesis is dedicated to the improvement and calibration of the wire chamber system.

The algorithm previously used for tracking was developed to meet the requirements of the pion

analysis in LArIAT, an analysis that is not statistics limited and thus does not require high precision in tracking nor background suppression. Simplified for brevity, the original algorithm reconstructs a track by the steps enumerated below.

1. Cluster hits in each wire chamber using the DBScan Algorithm [15]
2. Create a wire chamber object of each possible permutation of a four-hit track
3. Remove objects that do not meet straightness requirements in the YZ plane, the plane perpendicular to the bend direction in the magnets
4. In the case of multiple potential tracks, save only the track that is the straightest in the YZ plane

This algorithm does not perform full trajectory tracking of the particle bending in the magnets. It instead takes advantage of the approximation that the magnetic fields in the bending magnets are uniform and identical. With this assumption, one can measure momentum using only the current of the magnets and the angle the incident particle makes with the z-axis entering and leaving the magnetic fields, according to the formula

$$p = \frac{\int \vec{B} \cdot d\vec{l}}{\sin(\theta_{in}) - \sin(\theta_{out})}. \quad (4)$$

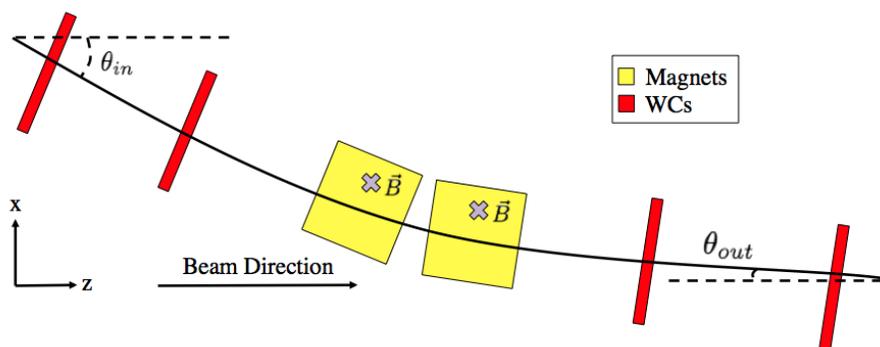


Figure 8: Diagram highlighting different quantities used in momentum reconstruction.

Using a preliminary beam line simulation, it was estimated that, of events that have a reconstructed track using this method, roughly 0.5% have a catastrophic mis-reconstruction [12]. The source of this error is not well understood: it may arise from matching noise hits with real particle tracks or from the misreconstruction of hits at the wire chamber closest to the target because it experiences a markable higher flux of particles than all other wire chambers. Either way, the error results in a reconstructed momentum arbitrarily far from the actual value. For the pion sample, this level of error is reasonable but, for the kaon sample, this error is on the order of the kaon beam composition, resulting in an insufficiently high background.

I improve upon this original algorithm by introducing full, probabilistic trajectory tracking of particles through the beam line system. For this analysis, I am using Geant4, a software suite that uses Geant4's particle propagation and simulation framework [14]. Geant4 is powerful and includes much of the software necessary to create an optimized tracking system. For this thesis and to avoid radical changes to wire chamber reconstruction, my new algorithm does not use all of Geant4's potential: it only uses Geant4 to measure the difference between the WC data hits and projected particle location for the WC objects created from routines in the original algorithm. However, I have incorporated

Geane into the LArIAT software framework, lowering the overhead to apply more sophisticated tracking in the future [13].

The fundamentals of the Geane-based algorithm are enumerated below.

1. Load LArIAT's Geometry into Geane / Geant4
2. Load a simulated, non-uniform magnetic field map of LArIAT's magnets
3. Using the original algorithm, reconstruct WC hits in 3D space, with the additional requirement that hits occur in time.
4. Using the original algorithm, create a wire chamber object of each possible permutation of a four track hit that meet straightness requirements in the YZ plane
5. Using the original algorithm, reconstruct momentum
6. Using the two upstream wire chambers as a directional vector and the reconstructed momentum, propagate a generic charged particle from the most-upstream WC, through the magnets and past the two downstream WCs.
7. Save all tracks that meet requirements on the residual between the propagated and measured hits at the two downstream WCs.

This full tracking method requires a strong understand of LArIAT's geometry and magnetic fields. As well, requirements of straightness, time coincidence and downstream WC residuals are tunable parameters and need to be optimized for the kaon analysis.

The determination of LArIAT's geometry is one of larger challenges for the full tracking approach. For the millimeter-level tracking that the wire chamber system is capable of, the geometry of the beam line needs to be known to this same level of precision. LArIAT's geometry was determined by Fermilab surveyors and, for Run 2, the survey was done twice: at the start and end of the run. Unfortunately, the error on these measurements is not well understood and the measurements are not self-consistent within several millimeters, as can be seen in Table 1.

	Feb. 2016			Oct. 2016		
	x [cm]	y [cm]	z [cm]	x [cm]	y [cm]	z [cm]
WC1	106.86	2.31	-682.92	106.87	2.35	-682.91
WC2	73.36	2.31	-537.76	73.38	2.31	-537.80
WC3	40.70	2.01	-339.20	40.81	2.02	-339.12
WC4	27.66	0.26	-95.31	27.68	0.31	-95.33

Table 1: Surveyed positions of the four wire chambers at the start and end of Run 2.

Typically, this discrepancy would be resolved by using the surveying values as starting points in a calibration of the system using the data itself: the tracking from the wire chambers should be able to measure wire chamber location more accurately than the surveyors. However, this is difficult because LArIAT is equipped with only four wire chambers and thus does not have a very constrained system. Since the minimum number of wire chambers required to measure a particle's momentum is three, it would appear that, without information from other detectors, the system can only calibrate the location of one wire chamber relative to the other three.

LArIAT's LArTPC cannot be used for tracking calibration. The amount of material in between the most downstream WC and the active TPC volume causes multiple scattering of the particle. When projecting a TPC track back to the most downstream WC, the multiple scattering causes a variance in position larger than the errors in the survey, making this method unusable.

The only other detector that can aid in calibration is the TOF system. The TOF can be used to measure the momentum of slow proton (TOF > 35 ns) because, in this region of reconstructed TOF-Momentum space, protons are the majority of the signal and slow enough to be separated by the TOF. The TOF-derived momentum measurement can be used to determine how much the proton should bend in the magnet, allowing for the calibration of the position of two WCs in relation to the other two.

The final WC positions used in this analysis can be seen in Table 2. The average of the two surveys was taken and then the slow-proton method was used to calibrate the furthest WCs from the magnets in the bending plane.

	x [cm]	y [cm]	z [cm]
WC1	$106.86 - 0.41$	2.31	-682.92
WC2	73.36	2.31	-537.76
WC3	40.70	2.01	-339.20
WC4	$27.66 + 0.50$	0.26	-95.31

Table 2: The tuned WC positions used in this analysis.

There is room for improvement in this position calibration. For example, a straight line fit can be used to calibrate the non-bending plane or the three-point calibration can be used in tandem with the slow protons to calibrate three out of the four wire chambers. Moreover, the survey measurements can be improved simply by verifying their correctness with a professional lab technician, something that is necessary due to the survey document complexity.

The magnetic fields used in tracking come from a simulated magnetic field map. The map is a file of five million points that have the vector direction of the magnetic field within a volume that encompasses one of LArIAT's dipole magnets and its fringe fields. The field map was created by a magnetic field simulator developed in-house by Dr. Douglas Jensen [16]. A pictorial representation of the map can be seen in Figure 9.

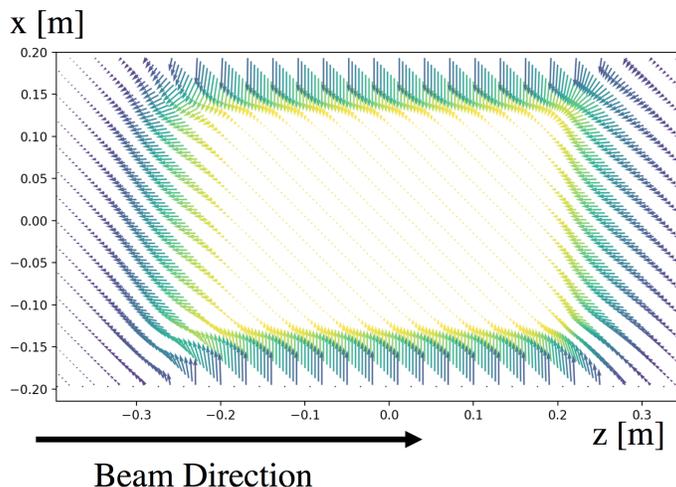


Figure 9: A slice of the magnetic field map. The color represents the magnitude of the B-field in the y direction.

After the implementation of the Geane software and magnetic field map, it becomes necessary to determine various parameters in reconstruction to optimize for efficiency and low background. Without the beam line simulation, this optimization is done using two sets of data events. The first, so-called "picky tracks", requires that each wire chamber has exactly one unique hit in each plane. By requiring one hit in each wire chamber plane, each hit most likely originated from the same incident particle. Without the use of simulation, this calibration set is the best available to use to test the efficiency of the reconstruction. This set is too statistics limited for the kaon analysis: events that meet this requirement are rare due to the target halo at the first wire chamber. The second set of events, the set of events used for this analysis, is defined to be events with one or more hits per wire chamber plane.

I do not consider a third possible set of events: events with one or more hits per wire chamber plane for three or more of the wire chambers. Since wire chamber momentum can be reconstructed with three wire chambers, we can boost statistics by accepting events with hits on only three of the WCs. In theory, such events should be rare because of the high efficiency of the wire chambers. However, in practice, such events could represent as much as a 30% increase in events with a reconstructed wire chamber track. For the purposes of the kaon analysis, the added uncertainty in momentum determination and increase in track misreconstruction makes these events unusable.

The first new parameter in the new algorithm to determine is the timing requirement for matching TDC hits on different planes within a given wire chamber. At a resolution of 1.18 ns/TDC, the TDC should have strong separation power for out-of-time hits, especially with LArIAT's low beam luminosity. Plotted in Figure 10 is the time difference between each hit in both planes in the first wire chamber. The calibration set, in blue, has one entry per event, and all other events, in purple, have between 1 and 30 entries per event. Both sets are created from the same subset of 6529 events, and the distribution is unnormalized to show relative hit rate between the sets.

Due to the beam structure, WC hits are reconstructed only if $\Delta\text{TDC} < 20.0$ for the hits within a given WC. This should remove many out-of-time particles since 20 ns separates each bucket in a beam spill. It is important to note that the distribution observed in Figure 10 suggests a timing resolution far worse than the prescribed 1.18 ns/TDC. In the calibration set, the majority of the hits originated from the same particle traversing the wire chamber but it still produces hits with a standard deviation at approximately 7 TDCs, or 8.25 ns. The reason for this apparent under-performance is not well understood and contributes to the difficulty in reconstructing wire chamber tracks.

After WC hit reconstruction, the next step is to create each possible permutation of a four-hit track. Within this set, there should exist the best and 'correct' single beam line track that created the trigger. It becomes the job of the rest of the wire chamber algorithm to find which track is the most probable.

Before full track propagation through Geane, a basic straightness requirement is created to eliminate a large percentage of obviously improbable tracks. The YZ plane is perpendicular to the bending of particles through the magnets, meaning that tracks should be approximately straight in this plane. To determine a track's straightness, the algorithm performs a linear fit on the points and then returns an average residual of the points to this fit. This average residual is plotted in Figure 11. Note that the calibration events has one entry per event while all other events have between one and 30 entries per event.

The scope of the challenge of wire chamber reconstruction is clear in this plot. The number of possible tracks grows with each addition hit in a wire chamber and, with the first wire chamber's proximity to the target, the number of potential tracks to an event can become large. It is clear that this parameter alone can not separate the signal, the true track, from noise.

The algorithm rejects tracks with a average residual greater than 7.0 mm. This relatively low value was chosen to lower the computational burden of running a large number of potential tracks through the full Geane tracker in the next step of reconstruction. Also, events removed in the calibration run, about 30% of the set, are difficult to label as true events that should be saved or

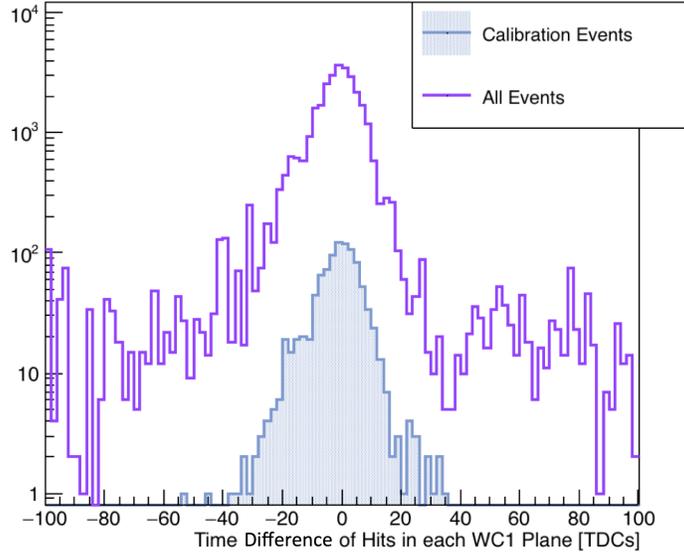


Figure 10: Δ TDC of hits inside the first wire chamber.

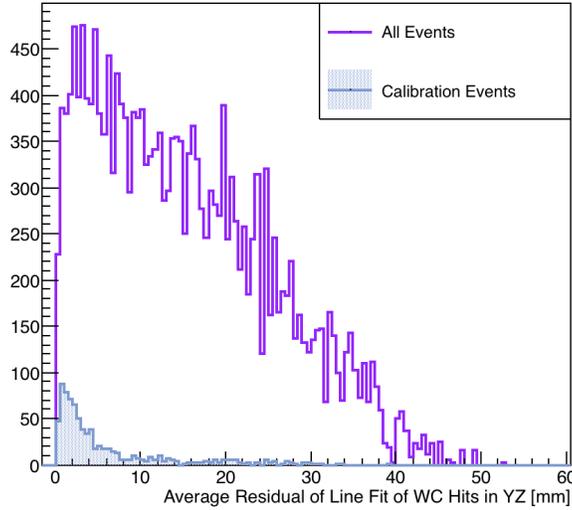


Figure 11: Average Residual of linear fit to hits in wire chambers in the YZ plane. Entry per potential track, so one for the calibration run and between 1 and 30 for all other events.

noise that is entering the calibration set. For these reasons, this parameter's value may warrant further optimization in the future.

After the straightness requirement, the next step in reconstruction is to do full track propagation and select only tracks that meet tracking error requirements. A generic charged particle is launched from the most upstream wire chamber with the momentum reconstructed using the original algorithm's routine. The direction of this launched particle is determined by the vector created from the hits on the two upstream WCs. The propagation software steps the generic particle through LArIAT geometry, bends the particle through the two dipole magnets with their fully simulated field map, and then past the two downstream WCs. The algorithm finds the residual between the propagated track and the measured hit at both downstream WCs. It is this residual that is used to determine

how realistic a track is.

An example event display for a data event can be seen in Figure 12. This event shows the Geane propagated track in red, from the calibrated geometry, and green, from the original geometry. The magnets and wire chambers are shown in light green and purple. The reconstructed wire chamber hits are in black. It is clear from this event display that, with the tuned geometry and updated magnetic field maps, the Geane propagation is doing a reasonable job at propagating the particle through the wire chamber system. An accuracy of this propagation can be estimated from the residual plots in Figure 13.

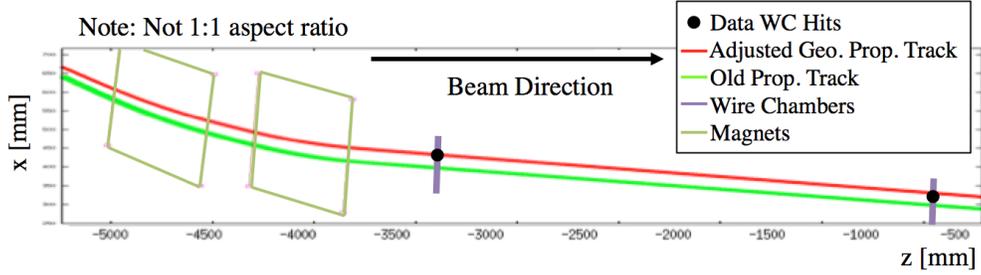


Figure 12: Event display of wire chamber data event overlaid with the Geane propagated track from the original and updated geometry.

Plotted in Figure 13 is the distribution of the residuals from the two most downstream WCs. The distributions are not entirely correlated, making a cut on both residuals productive. The cut values chosen are to reject a track with a residual greater than 2.0 cm at WC3 or 4.0 cm at WC4.

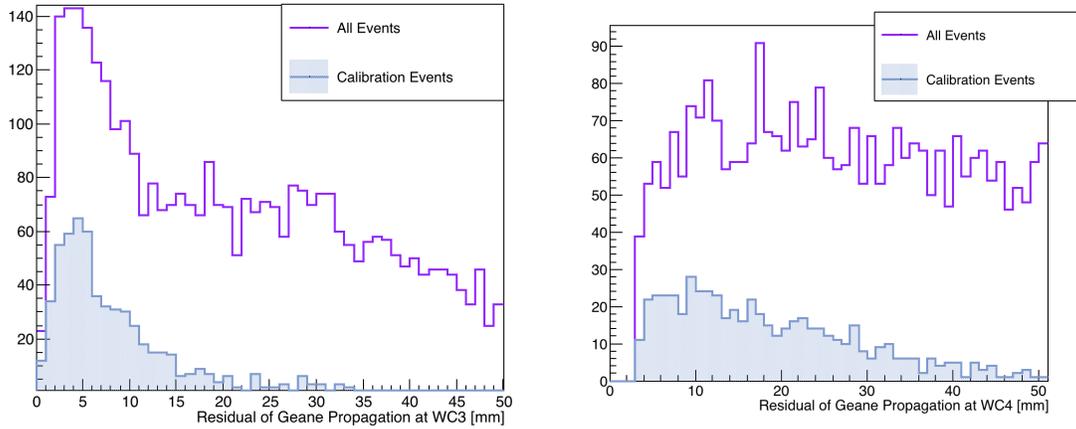


Figure 13: Geane Propagation Residuals at the two most downstream WCs, WC3 (left) and WC4 (right).

If a given track passes all criterion described, it is saved. It is important to note that this method returns every track reconstructed during a trigger window, while the previous method only returned the single, straightest track. The high efficient TOF system shows us that it is very rare for an event to have more than one beam track for a given trigger, due to the low luminosity of the beam. For this reason, there should be only one, 'best' reconstructed track. This is determined later when matching beam line tracks to TPC tracks.

To determine the efficiency of this cut scheme, I use the same data-driven method as in the TOF system. A trigger is saved alongside the data for events that have coincidence signals on all WCs.

This trigger can be used to determine efficiency as all triggered events should have a reconstructible track. However, again as in the TOF system, this trigger is not used for the kaon analysis as the coincidence rate is very low, possibly due to poorly tuned discriminators. On the subset of 6529 beam events, 218 have a wire chamber trigger and, of these, 64 have at least one reconstructed wire chamber track. This represents an efficiency of $29.4 \pm 0.4\%$. This low value is similar to that seen in the original wire chamber system and it is not well known why it is far from the expect 95% efficiency of the system. I was unable to find the origin of the inefficiency of the wire chamber, but most likely comes from higher-than-optimal voltage applied to the wire chambers, causing a bad signal-to-noise.

Slow protons can be used to determine both the absolute accuracy and resolution of the wire chamber reconstruction algorithm. At a TOF greater than 35 ns, the TOF system can be used to determine the momentum of protons at an uncertainty of approximately 25 MeV/c over the momentum range of the data sample. This was determined by propagating the uncertainty of the TOF measurement through the equation for momentum,

$$p = \frac{mc}{\sqrt{\left(\frac{ct}{l}\right)^2 - 1}}, \quad (5)$$

where m is the mass of the proton, t is the TOF and l is the distance between the TOF paddles, equal to 6.652 m. The difference between the momentum determined by the wire chamber system and momentum determined by the TOF system for slow protons is plotted in Figure 14. The events plotted are from the "picky track" calibration set.

The produced distribution is centered at zero, signaling that the wire chamber system is successfully reconstructing momentum without any biases. The spread of the distribution, equal to 50 ± 1 MeV/c, can be used to find momentum reconstruction uncertainty. Assuming that the uncertainty in the TOF system and wire chamber systems are uncorrelated, the momentum uncertainty is approximately $\sigma_{WC} = \sqrt{\sigma_{Diff}^2 - \sigma_{TOF}^2} = \sqrt{(50)^2 - (25)^2} \approx 43$ MeV/c.

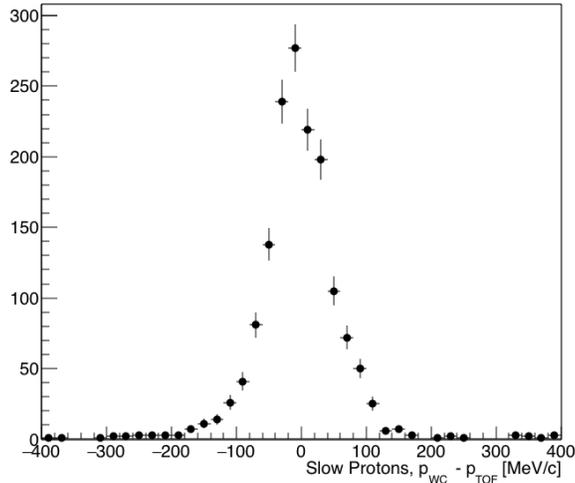


Figure 14: Difference of momentum as determined by TOF and from the wire chamber system for a data sample of slow (> 35 ns) protons.

An estimation of mis-reconstruction rate and ultimate background from the tracking system is determined in the next section discussing matching beam line tracks with TPC tracks.

3.3 Wire Chamber to TPC Track Matching

To take full advantage of the momentum and mass reconstruction from the beam line for a LArTPC study, it is necessary to match the beam line track reconstructed by the WC system with a particle's associated reconstructed TPC track. This is a non-trivial task due to several reasons: the different time scales of the TPC and beam electronics makes temporal associations besides PMT triggering impossible, the luminosity of the beam gives rise to multiple particles entering the TPC during a given drift window and there exists a substantial amount of material between the active TPC volume and the last WC.

The track matching, with no timing information, can only use spatial information. For a given beam line track, the track is projected to the face of the TPC. Then, for each reconstructed TPC track that begins within the first 2 cm of the active volume, an absolute angular difference ($\alpha \equiv \arccos\left(\frac{\vec{r}_{TPC} \cdot \vec{r}_{WC}}{\|\vec{r}_{TPC}\| \|\vec{r}_{WC}\|}\right)$) and spacial distance (Δr) is measured between the end of the beam line track and the start of the TPC track at the TPC face.

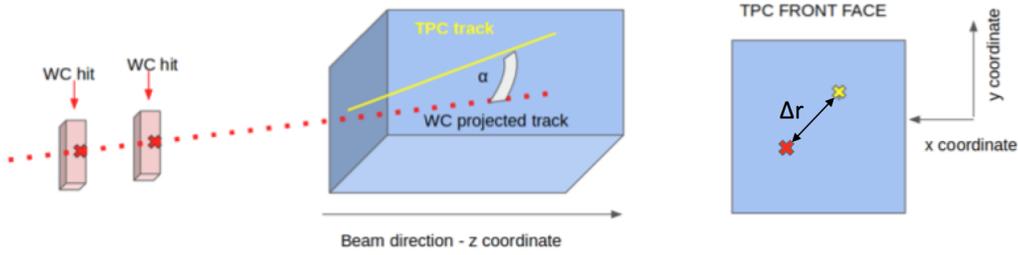


Figure 15: Diagram of the WC track and TPC matching algorithm.

Tracking matching inefficiencies are due, in part, to the substantiation amount of material between the last wire chamber and the active volume of the TPC. This material, besides requiring an energy correction, causes multiple scattering and thus positional uncertainty in the matching. The effect of multiple scattering manifests itself as a smearing in angle and space for the two parameters used in matching.

Background in this matching arises from high-occupancy events, where multiple tracks enter the TPC during a single TPC drift window. This background was reduced in LArIAT Run 2 by lowering the luminosity of the incident beam. In data in the Δr plot, background manifests itself as a broad Gaussian distribution underneath the sharper peak caused by particles that created the trigger.

To understand the matching background, a single particle gun is used to simulate a triggered event. The gun is fired from the location of the last wire chamber with an angular and positional spread determined from data. The momentum distribution is also weighted to match the data. Multiple-track events are not yet incorporated into the full simulation, but a reasonable approximation can be made by plotting the broader background gaussian, derived from data, onto the simulated set. In the Δr parameter space, this added background is approximately correct, as can be seen in the MC-Data agreement in Figure 16.

Figure 16 shows the Δr distribution after the α cut. The α cut is determined by the multiple scattering from the material before the TPC and is equal to $\alpha < 15^\circ$. This value represents the approximate three sigma range of the spread cause from multiple scattering alone and assumes that the angular uncertainty from the WC system is significantly better than that caused from multiple scattering.

From Figure 16, a cut value for the parameter Δr is selected to optimize background and efficiencies. The final cut values for matching are equal to,

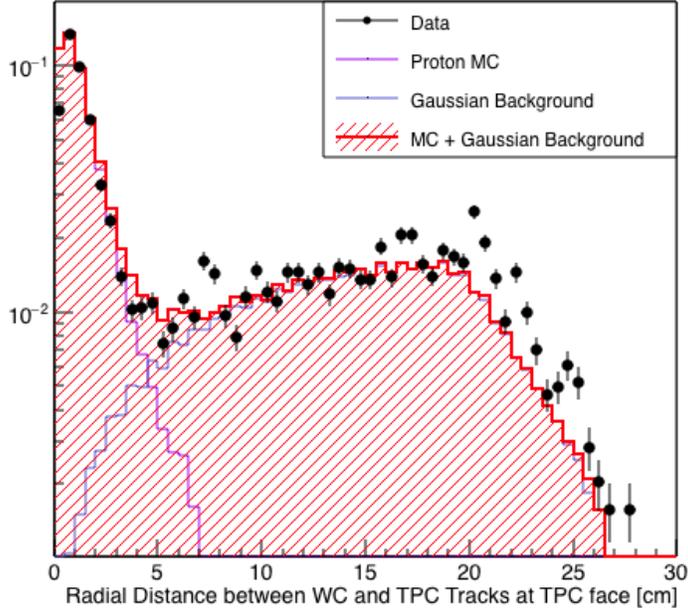


Figure 16: The matching parameter of Δr plotted against MC after the alpha cut.

$$\alpha < 15^\circ \quad \text{and} \quad \Delta r < 4.4\text{cm}. \quad (6)$$

Since these are the last cuts to be made on the final data sample to be used in this analysis, the background found here represents the final background in data caused by mismatching and misreconstructing WC tracking. At the determined value, there is a $\sim 10\%$ background. The background percentage is a function of event multiplicity and thus a function of beam intensity. Due to this, it is subject to change from run to run. The number quoted here and used is an average background for the kaon intensities used for the kaon sample.

The beam matching requirement removes a significant number of events, approximately 40%, due to particles stopping, interacting or decaying in the dead material before the TPC active volume. Plotted in Figure 17 is the reconstructed TOF vs. Momentum for all reconstructed momentum tracks for a large subset ($\sim 70\%$) of Run 2 data. The four distinct bands presents four different particles: from bottom to top in TOF, relativistic particles (composed of electrons, pions and muons), kaons, protons and then deuterons.

The matching has an understandable effect on the distributions. Below 600 MeV/c, highly ionizing protons can not penetrate the material to reach the TPC. Almost all deuterons at all energies meet a similar fate. The out-of-time particles seen as bands in TOF at 40, 60 and 80 ns are almost entirely suppressed in matching. As well, separation is improved between the kaon and proton bands.

3.4 Mass Selection

Using reconstructed TOF and momentum, finally mass can be determined to select for various particle species. Mass is reconstructed using the equation

$$m = \frac{p}{c} \sqrt{\left(\frac{ct}{l}\right)^2 - 1}, \quad (7)$$

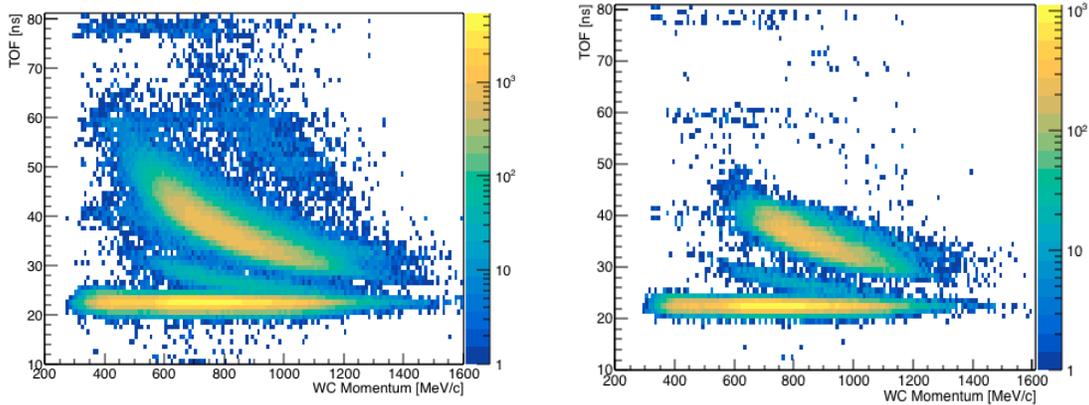


Figure 17: Reconstructed TOF vs. Momentum for a large subset of LArIAT Run 2 Data. Left is before WC-TPC matching and right is after matching.

where l is the length between the TOF paddles (6.675 m) and t is the TOF.

In a plot of TOF vs. Momentum, this mass equation results in approximate hyperbolas that are linear at low energy (classical limit) and decay into a straight line at the relativistic TOF. This is shown in the expectation curves in Figure 18. This figure shows the reconstructed TOF vs. Momentum for all of Run 2 data with positive bending magnet polarity that, events of interest to the kaon analysis. Only events with a single reconstructed wire chamber track are plotted.

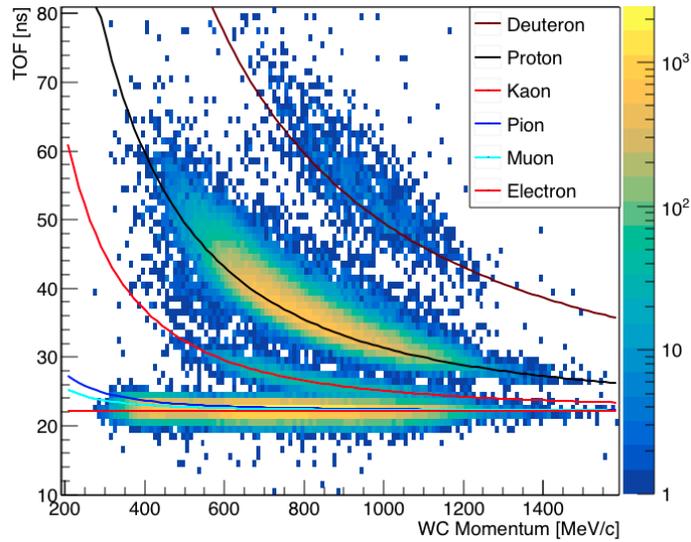


Figure 18: TOF Vs. Momentum for all of Run 2, positive bending magnet polarity events. Requirement that only a single wire chamber track is reconstructed.

A non-negligible number of events have multiple wire chamber tracks that both pass Geane track propagation requirements and TPC track matching. Since the TOF system shows us that is very rare for multiple tracks to occur during the same trigger, it is assumed that only one of these tracks represent the correct track. To determine which is the best, the track with the mass reconstructed closest to a known particle mass is kept.

3.5 Kaon Data Sample

Reconstructed mass is the primary cut used to create the final kaon data sample. After reconstructing TOF and a TPC-track-matched wire chamber track, mass is calculated by Equation 7 and a cut of 320 to 650 MeV/c^2 is applied to select for particles with the kaon mass. A final relativistic TOF cut is applied to arrive upon the final results.

A reduction table of events can be seen in Table 3. For the purposes of computational time, a mass cut is applied before the TPC-WC tracking matching requirement so that a significantly lower number of events have to have full TPC reconstruction. A plot of the mass as a function of cuts applied can be seen in Figure 19. Finally, six example event displays can be seen in Figure 20, with each taken from the collection plane.

As discussed in Section 3.3, the sample will have an approximate 10% background of pions, muons, electrons and protons. This background can be further suppressed with a range cut, removing particles with an improbably long primary track. However, for the purposed of creating the sample, this range cut is not applied.

Cut	Surviving Events
Number TOF Object > 0	1687288
Number WC Object > 0	388793
Mass > 320.0 and Mass < 650 MeV/c^2	6577
Unique WC-TPC Match	2072
TOF > 25.0 ns	1992

Table 3: Reduction table of events to create the K^+ sample.

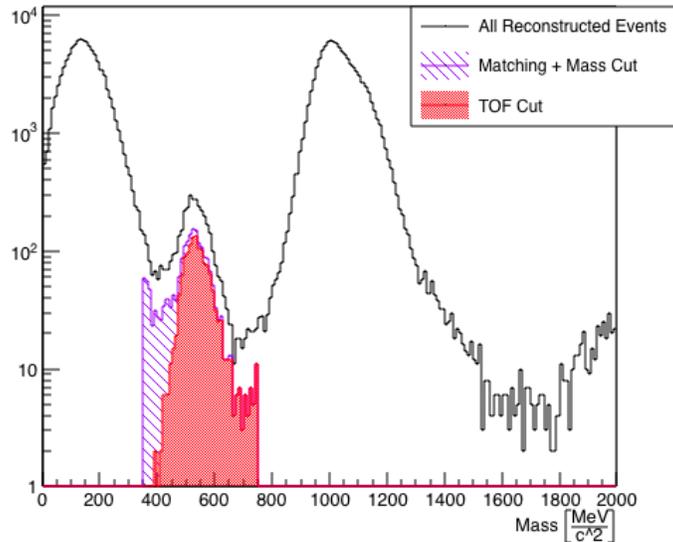


Figure 19: Reconstructed Mass as a function of cuts applied to select for kaons.

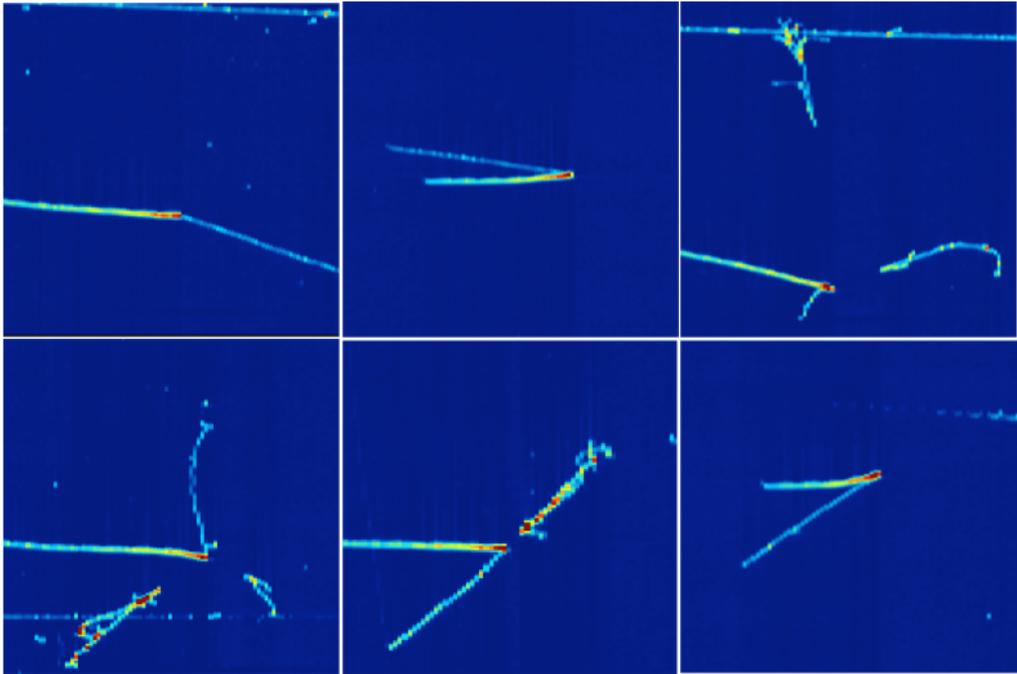


Figure 20: Six Kaon Candidate Data events from the set.

4 Proton Decay Signals and Backgrounds

A complete proton decay search includes extensive study of both signal identification efficiency and background contamination. According to a theory paper by A. Bueno[3], an underground LArTPC should have the energy precision and background rejection necessary to identify a kaon with ‘99% efficiency’ and with ‘less than 1% misidentification’ with other particles, allowing the detector to make competing measurements in the channel $p \rightarrow \bar{\nu}K^+$. The job of the experimentalist is to verify these claims via simulations of both the signal and potential backgrounds, show that the simulation agree with data ‘outside’ of the signal region, and then measuring the signal region for proton decay events or to set a lifetime limit.

The produced kaon and its decay products is the signal for proton decay in this channel. In the energy range of interest, 250 MeV/c to 450 MeV/c, the kaon decays at rest $\sim 90\%$ of the time, leaving a distinctive Beta-Bloch energy deposition curve as it slows: its Bragg peak. With sufficient energy and tracking resolution, this Bragg peak can be a powerful tool for particle identification. The kaon has a lifetime of $(1.2380 \pm 0.0020) \times 10^{-8}$ s and decays into the modes seen in Table 4.

Decay	Branching Ratio
$\mu^+\nu_\mu$	$63.55 \pm 0.11\%$
$\pi^+\pi^0$	$20.66 \pm 0.08\%$
$\pi^+\pi^+\pi^-$	$5.59 \pm 0.04\%$
$\pi^0e^+\nu_e$	$5.07 \pm 0.04\%$
$\pi^0\mu^+\nu_\mu$	$3.353 \pm 0.034\%$

Table 4: Decay modes and branching ratio for the K^+ . All remaining channels have branching ratios less than 2%.

The primary backgrounds for this channel come from atmospheric neutrinos and cosmic ray interactions. A common atmospheric neutrino interaction is the charged-current reaction $\nu_\mu + n \rightarrow \mu^- + p$. This signal is topologically identical to a proton decay event for $K^+ \rightarrow \mu\bar{\nu}$ if a vertex is reconstructed not at the neutrino interaction point but instead the proton stopping point. To suppress this background, there is a need to create an algorithm that, from energy deposition, can separate a forward-going kaon from a backward-going proton. Due to the abundance of these neutrino interactions, this algorithm would need to have near perfect efficiency to reach the proposed pseudo-background-less measurement.

Another source of background is from NC neutrino-proton elastic scattering. This reaction can create free, highly-ionizing protons in the detector, a signal similar to that of the kaon produced from proton decay. Topologically, these two reactions are very different: the K^+ decays into minimally-ionizing particles while the proton typically stops and does not interact. However, even this may be spoofed because an energetic proton can collide with an argon nucleus and create mesons.

A perfectly identified kaon still does not mean a zero background as kaons are produced, in a very low number, from neutrino interactions in the reaction $\nu_\mu N \rightarrow \mu^- K^+ N$. In such reactions, topology needs to be used to reject the background: a kaon produced from proton decay must not have any other particles at its point of creation because the kaon exits the nucleus with only an unseen neutrino. This type of event classification that uses topological information is not studied in this analysis due to LArIAT’s small size.

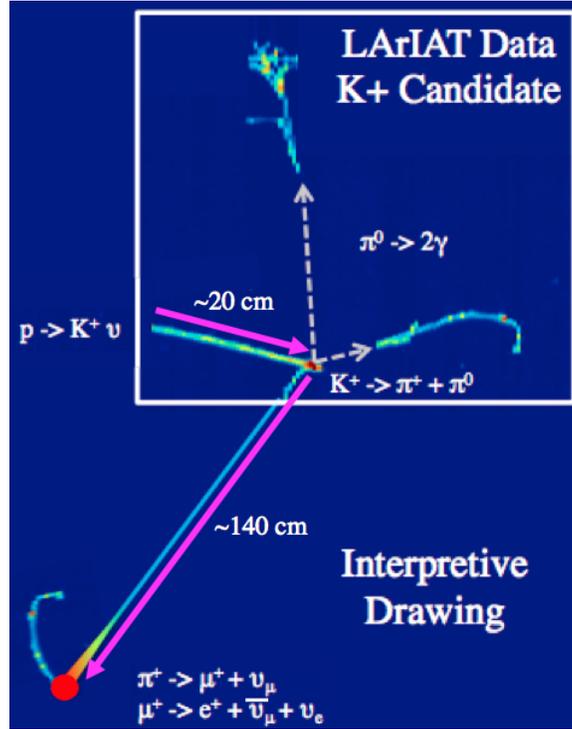


Figure 21: An interpretive drawing as to what a proton decay event would look like, with LArIAT data to act as the produced kaon. Neutral particles are labeled with dashed arrows. The escaped pion represents topological difficulties faced by LArIAT’s small size.

4.1 dE/dx Calibration

To make effective PID algorithms using LArTPC information, the detector must be well calibrated. A particle’s energy deposition as a function of its momentum is well understood and follows the Bethe-Bloche equation [17],

$$-\left\langle \frac{dE}{dx} \right\rangle = \frac{4\pi}{m_e c^2} \frac{n z^2}{\beta^2} \left(\frac{e^2}{4\pi\epsilon_0} \right)^2 \left(\log \left[\frac{2m_e c^2 \beta^2}{I(1-\beta^2)} \right] - \beta^2 \right), \quad (8)$$

where n is electron number density, I is mean excitation energy and z is the charge of the incident particle. To be able to take advantage of this equation, the relationship between signals recorded on wires in the TPC and the energy deposited by an incident particle must be well known.

When a charged particle enters the TPC, it ionizes the argon. The released ionization electrons then begin to drift toward the wire planes of the TPC. Various processes effect the location and decrease the number of electrons that make the entire journey to the wires.

The first effect is electron recombination, or the ionization electrons recombining with the ionized argon atoms before drifting to the wire planes. This is a function of the electric field and the charge space-density because the coulomb force between an electron and ionized argon atom is the primary cause for recombination. In a LArTPC, both space charge and electric field is known so one can correct for recombination. This correction is done via a “box” method that parameterizes a model for electron recombination via two parameters. This algorithm was developed and tested both by the ICARUS and Argoneut experiments [8]. This correction should be agnostic of experiment if the next two corrections to charge are done properly.

The second effect that reduces electron number is from electro-negative impurities in the argon, specifically oxygen and nitrogen. These contaminants are kept at a minimum using the filtration systems. Nevertheless, a correction is applied to the number of observed electrons to account for electrons lost from impurities. Due to the statistical mechanics of this process, the lose of electrons can be written as an electron lifetime, equal to

$$Q_{obs} = Q_{dep}e^{-t/\tau}. \quad (9)$$

Since drift time is dependent only on location from the wires and the applied electric field, the amount of time an electron spends drifting can be measured and the number of electrons lost can be corrected for. Determination of the τ is done via cosmic rays that are triggered on two PMT paddles on opposite diagonals of the TPC. These cosmic muons are used for lifetime determination because they deposit a known amount of charge at a known distance. Impurities in the argon change over time so the τ is determined on a run-by-run basis (every \sim hour).

After both impurity and recombination corrections, there remains one last parameter of the calorimetry: electrons-to-ADCs. When an electron comes to the wires, it induces some voltage and this voltage is read-out as a number of ADCs. This response should be linear so that the relationship between ADCs and electrons is described by one scaling factor. To find this scaling factor, we must use particles that deposit a known amount of charge. The particles used are beam pion and muons in the momentum range 500 to 1000 MeV/c because they are minimally ionizing particles (MIPs). As seen in Figure 22, these particles are expected to deposit in the TPC ~ 1.64 MeV/cm along a particles path.

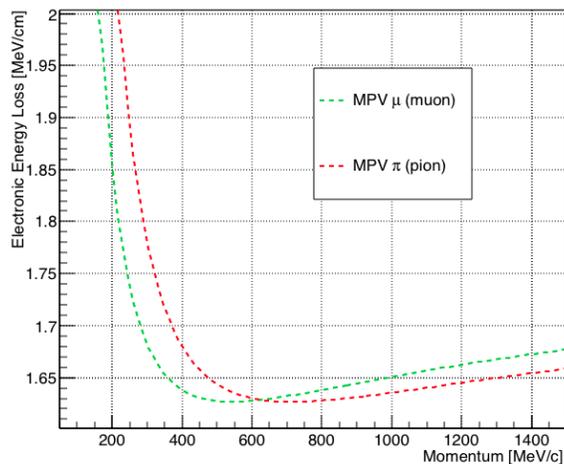


Figure 22: The most-probably-value of the energy deposition for pions and muons as determined by the Bethe-Bloch equation in liquid argon. The kaon MPV is greater than 2 MeV/cm in this momentum range.

To measure the most-probable-value (MPV) of the Landau distribution for the energy deposition created by a particle of a given momentum, one must take into consideration the smearing of dE/dx from electron dispersion in the argon or other detector effects. All fits are thus done via a Landau convoluted with a Gaussian, with the MPV value from the Landau used in calibration and the variance of the Gaussian used to smear the simulation to match data. A result of the calibration can be seen in Figure 23[17].

The agreement seen in Figure 23 between the data and simulation is not perfect. It is good for pions, but progressively worse for the most highly ionizing protons. This can be caused by an incorrect momentum distribution used to produce the MC plot: the simulation is weighted to

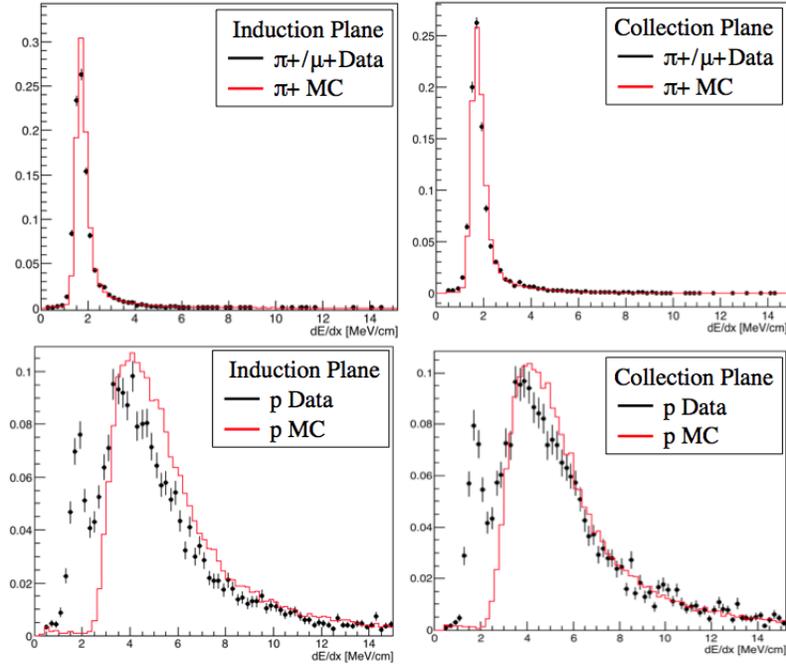


Figure 23: Data vs. MC comparison of dE/dx for MIPs (pions / muons) and protons in both planes after calibration. Note: There is notable MIP contamination in the proton data sample.

have the same momentum distribution as data, but it was assumed that pions and protons have identical momentum profiles, which is may not be true because fewer protons may be produced at lower momentum. The disagreement might also be caused by poorly tuned electron recombination or electron lifetime since higher charge events are most sensitive to these corrections. For the time being, this data-MC agreement is deemed sufficient to move onto the next step of this analysis, using energy deposition and tracking for particle identification.

4.2 Bragg Peak Fit

A well-tuned LArTPC should have fractional energy resolution on the order of percents for a well-behaved, non-showering charged track like those seen in proton decay events. This, combined with the fact that a LArTPC is a loss-less calorimeter (all deposited energy is collected and measured) means that such a detector is capable of measuring the energy deposition of an particle along its entire track to a high level of accuracy. The energy deposition information is powerful because a particle's energy deposition is a function of its momentum, as described by the Bethe-Bloch equation, and thus can be used for particle identification

The momentum range of interest for the kaon produced in a decaying proton is ~ 200 to 500 MeV/c as determined from two body decay of a proton smeared by Fermi energy and final state interactions. In this momentum range in liquid argon, kaons have a total track range of 2 to 30 cm before stopping. For a proton background of the same momentum, a range cut can easily reduce background: the range is a function of particle mass so any track with a range lower than 2 cm in this momentum range must be a proton. For this reason, to study backgrounds, I am not looking at protons in the momentum range ~ 200 to 500 MeV/c, but instead at protons in the momentum range 300 to 650 MeV/c because protons of this momentum have the same range as the kaons signal, 2 to 30 cm.

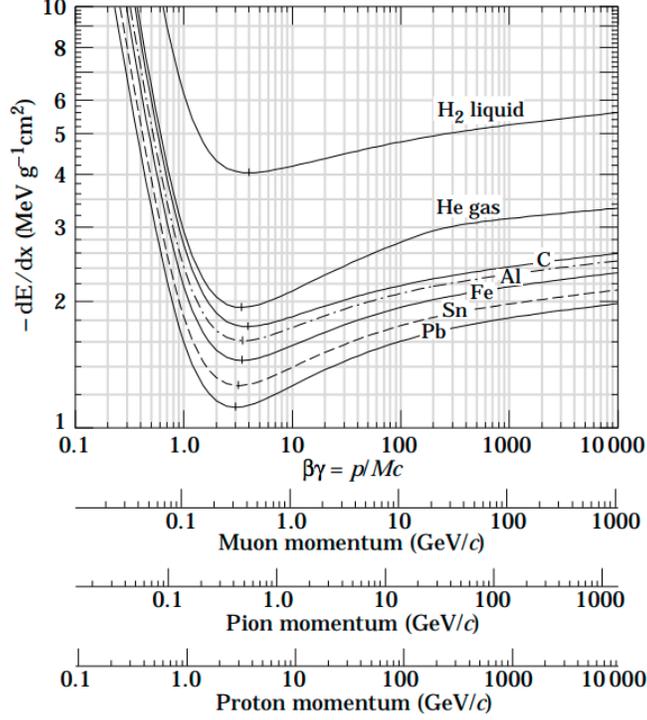


Figure 24: The energy deposition of various particles at various momentums and in various materials. Argon is closes in atomic mass to Aluminum on this plot [17].

With a wire pitch of 0.40 cm, LArIAT would reconstruction a track of this range using hits from 5 to 75 wires in a single plane. The objective is to demonstrate that, for a given track of length 5 to 75 wires, the energy and tracking resolution available from a LArTPC is adequately high to be able to separate protons and kaons using their energy deposition curves.

The Bethe-Bloch equation is complex - freshmen year, it was a weekend project to plot it with correct units. As seen in Figure 24, it describes the energy deposition of a particle over orders of magnitudes of momentum. For the low momentums of interest in this study, the relativistic corrections in the equation can be neglected, giving rise to the formula,

$$-\left\langle \frac{dE}{dx} \right\rangle = \frac{4\pi}{m_e c^2} \frac{nz^2}{\beta^2} \left(\frac{e^2}{4\pi\epsilon_0} \right)^2 \log \left[\frac{2m_e c^2 \beta^2}{I} \right]. \quad (10)$$

In order to perform a fit on the energy deposition curve, we move from momentum space into range space. Since range is a non-linear function of momentum, it can be used as a low-level proxy for the quantity. Thus, the curve we are interested in is not dE/dx vs. momentum, but dE/dx vs. residual range, where residual range is the amount of track left between a measured point and a tracks end. In this space, the simplified Beth-Bloch equation can be simplified further by parameterizing the equation in the form,

$$\frac{dE}{dx} = AR^b, \quad (11)$$

where R is the residual range and A and b are parameterization functions. Each particle type has a unique value for both A and b that described their low-energy Bethe-Bloch equation behavior within

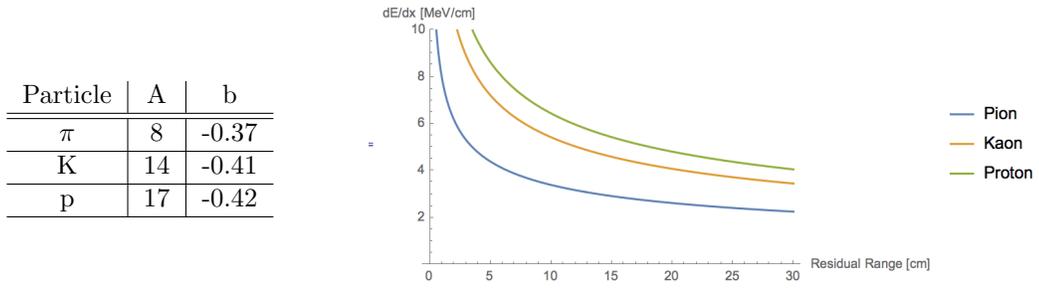


Figure 25: Left: Various parameterization values for each particle type. Right: dE/dx vs. residual range for each particle according to the parameterized equation.

$\sim 3\%$ error, in the case of the proton. The values of these parameters for each particle type, as well as the plot of each, is shown in Figure 25.

This parameterization was introduced by the Argoneut experiment and is used in their PID algorithm called PIDA [18]. To determine a particle type from the parameterized equation, the parameter b is held constant at -0.42 and the equation is flipped to solve for the average value of the parameter A , according to

$$\langle A \rangle = \frac{1}{N} \sum_i^N \left(\frac{dE}{dx} \right)_{obs,i} R_i^{0.42}, \quad (12)$$

where N is the number of reconstructed hits per track and i is a given hit in the track. The number of hits is a function of the number of wires traversed by a given particle trajectory and this value $\langle A \rangle$ can be measured for each plane. A particle species is determined by this measured $\langle A \rangle$ and a data-driven method can be used to determine the best values for selecting the various particle species.

It is important to note that this method is purposely simple. More powerful fits can be done using χ^2 test for various particle hypotheses. However, this method is to be used because it gives an easy-to-understand window into how well energy calibration is capable of particle determination. Such effects as poorly reconstructed tracks or bad calorimetric tuning have effects on the PIDA distribution that gives us both insight on how to improve reconstruction and a first-order approximation of detector performance.

4.3 PIDA Performance on Simulation

With perfect track reconstruction and energy resolution, PIDA results in extremely sharp, separated normal distributions around each of the expected value of $\langle A \rangle$ for the different particle species. However, tracking inefficiencies and energy smearing change the performance of PIDA. In this section, I step through how various detector effects change the performance of the PID algorithm.

I use a LArG4 single particle gun to produce simulated sets of pions, kaons and protons. These sets are made to be as similar to LArIAT data as possible to act as a fair comparison with data. The simulated particles are fired from one meter before the active volume to simulate the last WC location, are propagated through the simulated cryostat and nonactive argon to take into consideration energy uncertainty at the beginning of the track, and enter into a active volume of argon with simulated impurities, recombination effects and detector smearing. The simulated sets then go through the same track reconstruction as used in data. I simulate 5000 K^+ , pions and protons with a uniform momentum distribution from 500 to 1000 MeV/c

The PIDA values for the simulated sets can be seen in Figure 26. The left plot is all particles that come to a complete stop in the active volume while the right are all particles in the set. For the time being, the simulation here has a reconstruction quality cut that requires a 3D reconstructed track end within 2.0 cm of the MC Truth information. This requirement is made to first understand, under optimal reconstruction, what the PIDA distribution would look like.

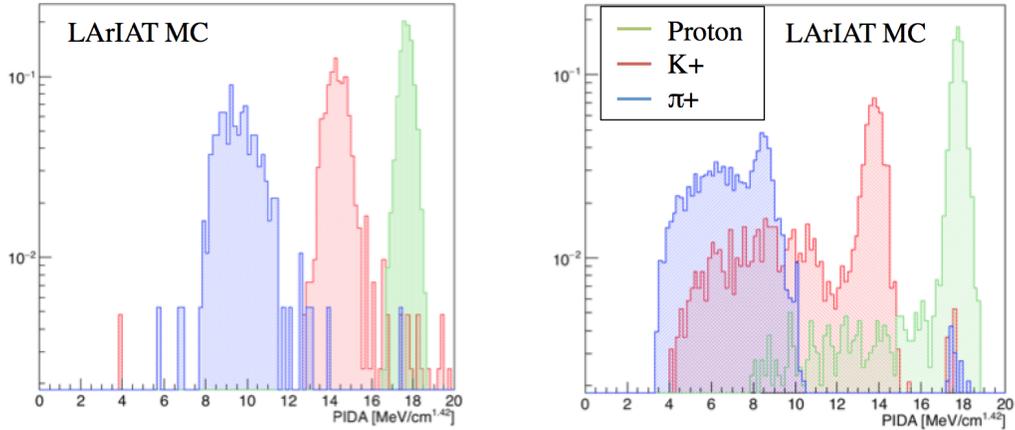


Figure 26: PIDA distribution for a set of stopping and all particles for pions, kaons and protons. Normalized so relative abundance is 1:1:1.

Looking first at the left plot in Figure 26, the separation for stopping particles with near-optimal reconstruction is extremely good. Protons have no downward smearing towards what would be the signal region for a kaon, or $12.0 < \langle A \rangle < 15.0$ MeV/c. A number of kaon events do exist within the proton band. These events are from the mis-reconstruction of energy deposited from the track of a decay product being added into the energy deposited from the primary track. This is made clear in Figure 27, which plots the dE/dx vs. Residual range for all the stopping kaon sample. The second band of anomalously high dE/dx points are from the addition of charge deposited from a minimally ionizing particles produced as decay products. The second band is 5 MeV/cm instead of only the expected 2 MeV/cm above the signal band because of recombination and lifetime over-corrections due to the apparently high charge-space density at these points.

When the kaon decays, the products leave the vertex isotropically. A small number of the time, the decay product track is immediately on top of the primary kaon, increasing the apparent energy deposited by the primary. This is rare. However, it is less rare for the decay tracks to be at a close angle with the primary from the perspective of a wire plane. This effect can be suppressed by taking

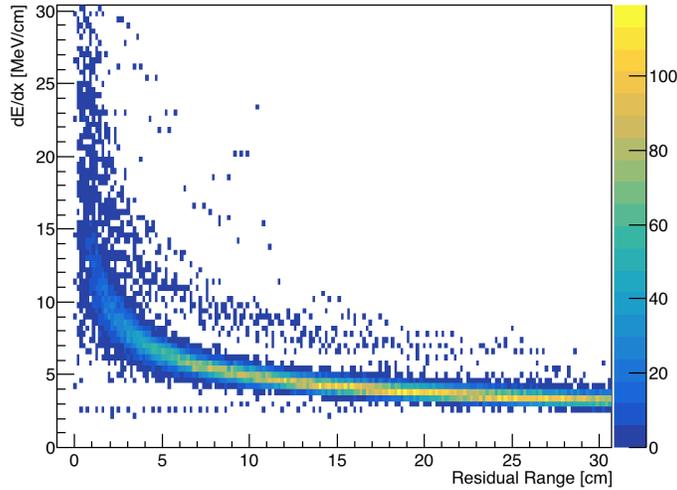


Figure 27: $\frac{dE}{dx}$ vs. Residual range for a stopping K^+ MC sample selected to have a well reconstructed track end point. Higher-than realistic dE/dx points are from decay-product tracks overlaid on top of the primary and misreconstructed together.

into consideration information from each wire plane: every plane has a unique “view” of the event, and thus may have better energy reconstruction than another plane.

Looking now at the right plot in Figure 26, every particle species exhibits a peak at their expected value of $\langle A \rangle$ but they also have a noticeable smearing of events approximately one order of magnitude below the peak. This is not due to reconstruction faults because this plot has track reconstruction quality cuts: a reconstructed track end point must be within 2 cm of an MC truth track. Instead, the smearing is due to particles not stopping inside of the TPC before depositing all of their energy, either due to exiting the active volume, decaying in flight, or interacting before stopping.

This smearing effect is mostly unimportant for proton decay studies. Non-stopping pions do not enter the kaon signal region. The smeared kaons are primarily from tracks leaving the detector before stopping or kaons decaying in flight, but, in DUNE, fiducial volume cuts will be applied so that kaons will never leave the detector and, at momentums below 500 MeV/c, $\sim 95\%$ of kaons decay at rest.

However, the most important element of Figure 26 is that a non-negligible number of protons enter the kaon signal region. These proton events, the majority of the time, are a result of protons interacting with the medium before stopping. This is a background that needs to be suppressed to almost zero to meet expectations and requirements of detector performance in order to be able to make proton decay measurements.

The final observation of this plot is that a small number of pions and kaons have events with a PIDA value compatible with protons. These events are from protons that are ejected from an argon nucleus from a hard scatter and become attached to end of a reconstructed track. These attached protons are short - less than 2 cm from the requirements of tracking quality - showing that a small number of hits, about 4 wires, can skew PIDA perform.

To make the reconstructed simulation as similar to data as possible, I now lift the restriction on tracking quality. This gives insight on reconstruction performance and how it changes PIDA performance at a level of reconstruction similar to data. The resulting PIDA distributions are plotted in Figure 28.

This plot represents the current state of LArSoft reconstruction as of April 2018. The additional errors from tracking uncertainties and misreconstruction of ends points amounts to further spreading

of all the distributions. Background due to protons in the kaon signal region are approximately the same as before. However, an additional background of pions has entered the kaon signal region. This is mostly due to protons that are ejected from an argon nucleus by the pion but reconstructed at the end of the track, artificially raising the PIDA value. Kaon identification efficiency, the fraction of kaons with PIDA values within the signal region, also goes down due to the additional smearing.

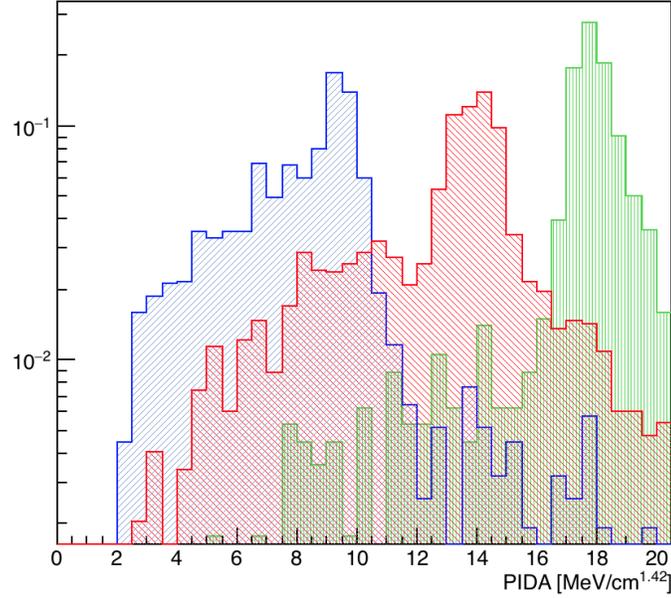


Figure 28: PIDA distribution for all particles entering TPC without quality cuts on tracking.

Stepping through how PIDA changes with various degrees of restrictions on purity has revealed three primary modes of failure. These failures can be seen in Figure 29, where the light blue represents particle tracks, the red triangles are Bragg peaks and the red arrows represent reconstructed tracks. From left to right, the problems in dE/dx energy deposition are interaction / decay before rest, misreconstruction of energy from one track to another near end points, and poorly reconstructed end points with tracks attached to the end of a primary.

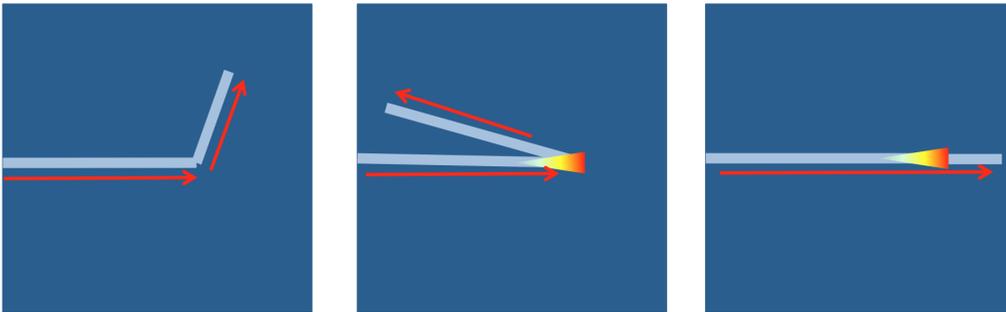


Figure 29: Failure modes of tracking that lower PIDA performance.

4.4 Data-MC Comparison of PIDA

With the PIDA for each simulated particle species under data-like conditions (no cuts made on MC truth information), we can compare these distributions to data. For the kaon data sample, the PIDA distribution plotted against simulation can be seen in Figure 30. The data points are the PIDA value measured for the kaon data sample that have a momentum sufficiently low such that they stop in the active volume of the TPC, or $p_{reco} < 750$ MeV/c. The momentum of the simulated samples are weighted to match the momentum spectrum of the data, scaled to a relative abundance of 90% K^+ , 5% proton and 5% π^+ and then area scaled to match the data sample.

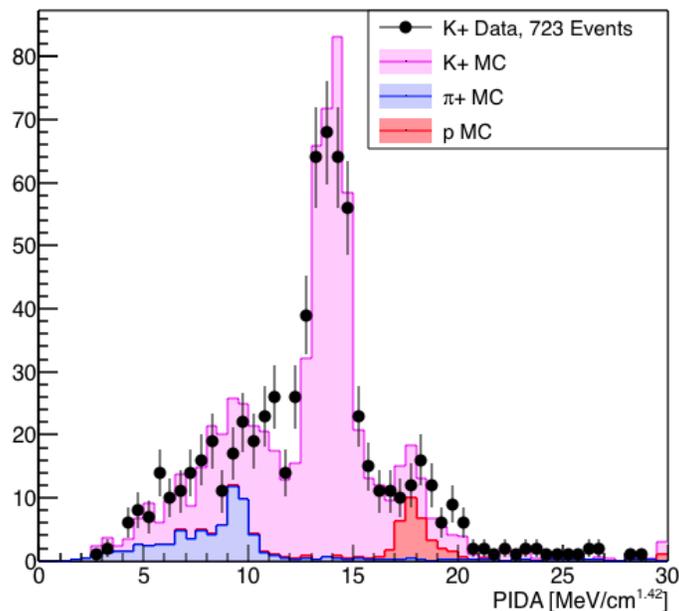


Figure 30: PIDA value for K^+ data sample, compared against stack histogram of 90% K^+ MC, 5% p MC and 5% π MC.

The agreement seen in this plot is a promising sight and one of the final results of this thesis. While the statement “our simulation agrees with the data” is not remarkable, this plot represents the first time that a data sample of kaons is proven to agree with simulation, reconstruction software and, ultimately, the particle identification techniques that will be deployed in a proton decay search in future LArTPCs. This plot helps confirm that all past and future analyses for proton decay searches that will have to be, by necessity, computer simulated studies connects back to the data.

It is important to note that there are other ways to compare both reconstruction and simulation performance against data. Other methods include the energy calibrations seen in Figure 23 or quantities like reconstructed track lengths for particles of a given momentum. However, by making the comparison of PIDA - a relatively high-level reconstruction PID technique that requires high tracking efficiencies, track direction determination and energy calibration - I demonstrate that, through the large number of errors that could propagate into this final quantity, the reconstruction in both data and simulation is performing sufficiently well to agree with one another - signaling an error in one is an error in another - and to show notable peaks and troughs at expect values of PIDA.

Finally, I note that this plot was produced using automatic reconstruction that is not near the performance expected for the LArTPC. This plot serves both as a glimpse into the progress made since manual reconstruction and as a representation of the work left before reconstruction meets expected performance. I present one proposed improvement to reconstruction in Section 5.

4.5 PIDA Application to Proton Decay

As discussed in previous sections, two of the largest backgrounds for proton decay in the kaon channel come from neutrino scatters that produce proton and protons produced in NN neutrino interactions that are reconstructed backwards as to have an identical topological appearance to kaon decay.

To discuss how well PIDA currently performs at suppressing each of these backgrounds, I plot in Figure 31 the PIDA values for a sample of kaons plotted against a sample of protons reconstructed forward (on the left) and backward (on the right). Each set has a flat momentum spectrum of 500 to 1000 MeV /c. Each sample has the quality requirement that a reconstructed track falls within 2 cm of a truth track and that the kaon or proton stops in the active volume. The backward reconstruction means that the vertex and end-point for the proton are flipped so that the Bragg peak of the track exists at what is now labeled the vertex. Due to their high values of dE/dx at high values of residual range, this theoretically raises the value of PIDA up and further from the kaon signal region.

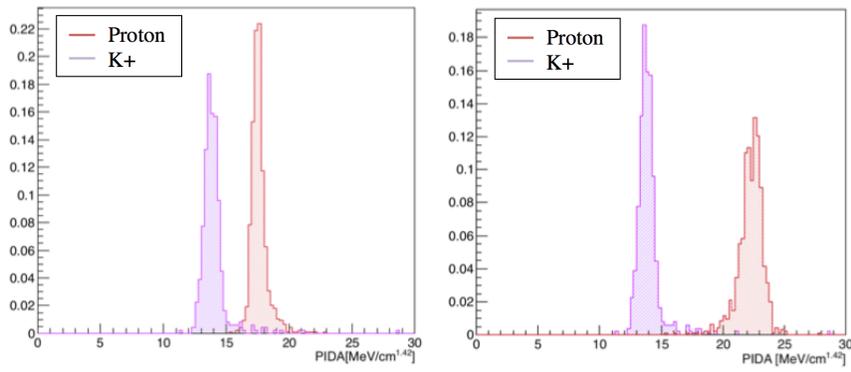


Figure 31: PIDA distributions for simulated kaons and protons reconstructed in the correct direction (left) and backwards (right). Area normalized so relative abundance is 1:1.

From the few number of simulated events, we can begin to get estimates of backgrounds. Of the 949 normally reconstructed protons, $O(10)$ fall into the signal region of $PIDA < 15.0 \text{ MeV/cm}^{1.42}$, giving an approximate false identification rate of 1%, while, for backwards reconstruction, zero events fall within the same region, resulting in a $< 0.1\%$ background.

This too can be verified against LArIAT data, but with a number of caveats. LArIAT beam line selection results in an approximate 7.5% mismatch rate with other tracks. Since protons encompass about half of the signal, this rate turns into about a 5% background of MIPS in the proton sample. I apply a range cut to lower this background, requiring that the events do not travel further than the theoretical maximum of a proton or kaon with the momentum determined by the beam line system. However, this cut still allows for a non-negligible number of pions in the proton and kaon sample. Plotted in Figure 32 is the kaon data sample, proton data sample and flipped proton data sample for events that pass beam line selection and TPC range cut. Note the result is much broader because the MC plots use a MC truth quality requirement.

4.6 Proton Decay MC

Previous to this point in this thesis, all simulated sets were created to emulate the beam K^+ data events available in LArIAT. These particles are of a higher momentum than those produced in proton decay and are mono-directional, instead of the expected isotropic. Due to time limitations, instead of creating a full simulated proton decay sample using Genie and then importing it into LArIAT's simulation software, I created an approximate proton decay sample of 1000 K^+ events uniformly distributed in momentum from 250 to 450 MeV/c and with isotropic direction. Due to similar time

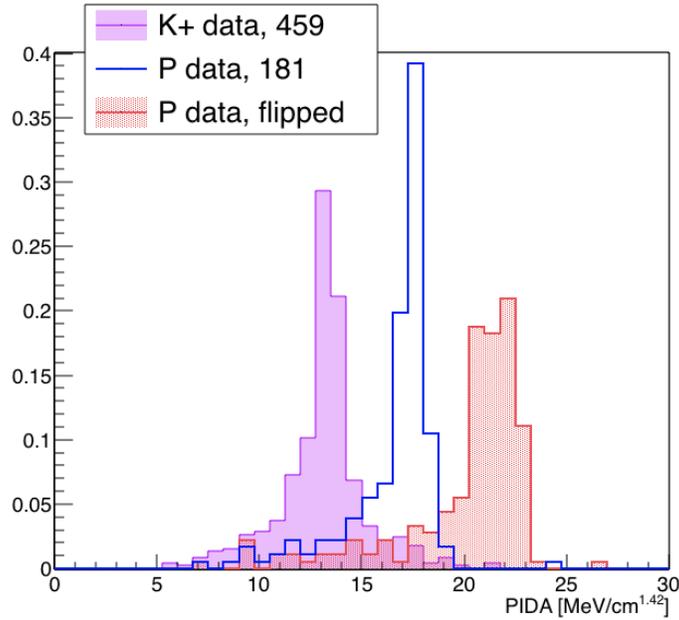


Figure 32: PIDA distributions for data kaons and protons reconstructed in the correct direction (left) and backwards (right). Area normalized so relative abundance is 1:1.

restrictions, I also created an isotropic sample of protons that is uniform in momentum from 300 to 650 MeV/c to emulate a background.

These events have worse PIDA performance than beam events, as can be seen in Figure 33. Due to the lower track lengths seen in these sets, fewer number of hits are available for reconstruction into tracks. This results in a significantly lower track reconstruction efficiency. As well, the isotropic distribution makes the events much more susceptible to reconstruction errors due to a poor ‘view’ of the event from the wire planes. In LArIAT, the wires are at an angle that optimizes the probability of a well reconstructed beam track event. The singular direction of beam particles minimizes tracks that, for example, run parallel or perpendicular to a wire, depositing all charge into a single wire or small number of wires. To mitigate this effect, most LArTPCs have three or more wire planes. However, LArIAT and thus this simulation set, only has two planes, leading to potentially worse tracking and PID performance.

Tracking inefficiencies arises from the bipolar signals on induction planes. The bipolar shape is an extra challenge in reconstruction as it is more difficult to determine the amount of charge that was deposited in a given hit. Complicated topologies where a large amounts of charge might approach a single wire for a large amount of time leads to asymmetric and difficult to reconstruct pulse shapes. A deconvolution process is performed on the wire waveforms in order to remove detector and magnetic field response. However, the process can be imperfect, especially with the complicated pulse shapes on the induction plane. Seen in Figure 34 is an example LArIAT simulated event from the K^+ isotropic sample. The raw and deconvolution waveforms of the highlighted wire is plotted below the event display. The deconvolution signal should be a series of Gaussians or a long, smoothed step function. It is clear that the deconvolution algorithm attempted to reconstruct several hits in the large negative induction area seen in the raw wave function, but did not succeed. This failure results in misreconstructed track and poor energy reconstruction.

Poor energy reconstruction causes a failure of the PIDA algorithm as it relies upon dE/dx information. Plotted in Figure 35 is the PIDA distribution for the isotropic set of protons and kaons,

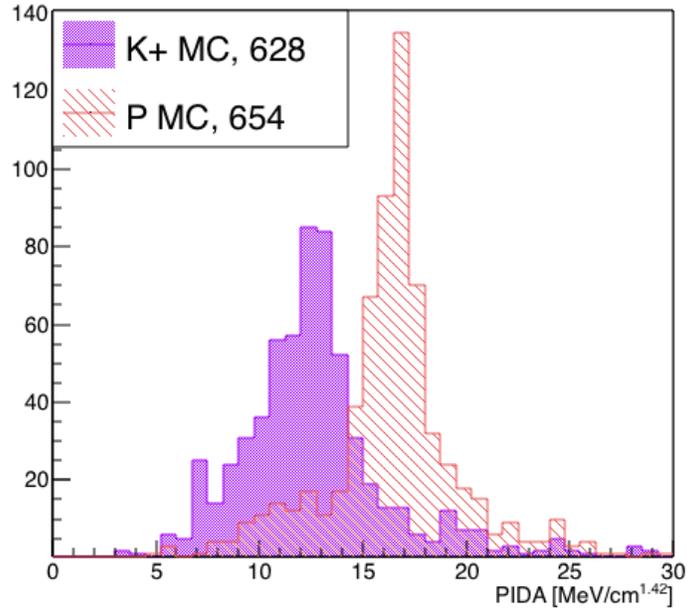


Figure 33: PIDA distributions for the isotropic set of simulated 1000 kaon and 1000 proton events. No quality cuts are made using truth information. No normalization is done to show relative efficiencies and separation.

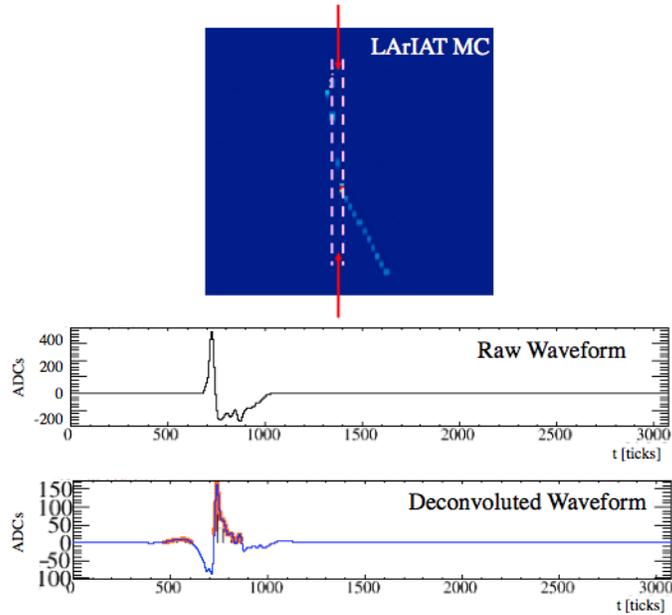


Figure 34: Example waveform from the induction plane before and after the deconvolution process. Taken from a simulated K^+ event.

but taking reconstructed energy information from the induction plane instead of the collection plane. For both protons and kaons, the PIDA distribution falls off at high values of $\langle A \rangle$ where expected,

but there is a large smearing towards lower PIDA values due to errors caused by the deconvolution process. Other experiments [8] have had much better success with deconvolution, implying that the algorithm is not properly tuned for the specific signal shapes collected in LArIAT.

Poor induction plane deconvolution impacts more than just energy reconstruction: with fewer hits reconstructed and with lower hit reconstruction on the induction plane, off-horizontal tracks have a much lower tracking reconstruction efficiency. It is for this reason that a large chunk, $\sim 30\%$, of the relatively high energy kaon and proton events in the isotropic sample do not have a reconstructed track at the vertex.

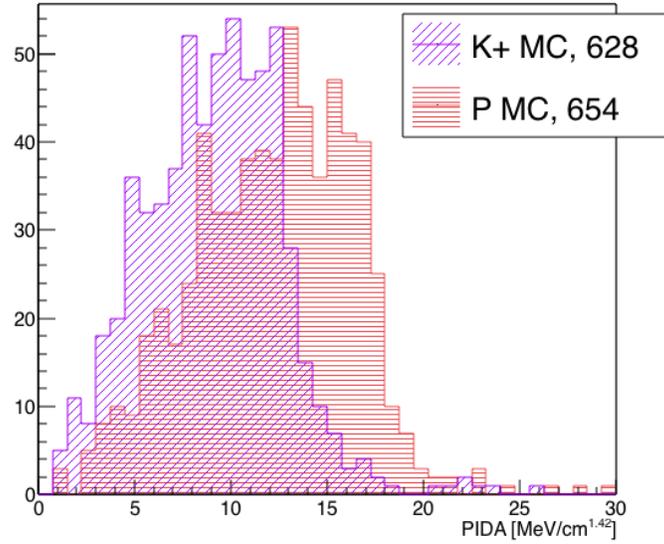


Figure 35: Induction plane PIDA distributions for the isotropic set of simulated 1000 kaon and 1000 proton events. No quality cuts are made using truth information. No normalization is done to show relative efficiencies and separation.

5 PIDAN, a proposed improvement

I propose an improvement to the PIDA algorithm by finding an optimal track end point via the energy deposition curve instead of relying entirely on tracking. Before finding $\langle A \rangle$, the algorithm would first find, based on the energy deposition curve, the most likely end point for the track. With this newly found end point, a new residual range is calculated for each point and then the value for $\langle A \rangle$ is measured. I call this algorithm PIDAN.

This method, at the bare minimum, would verify that the end point of a track is well reconstructed. If a Bragg peak is found with high confidence at the end of the track, this can be used to verify that a track is well reconstructed. However, I propose this addition to PIDA to combat the issues with tracking that were discussed in previous sections that cause lower than expected values of $\langle A \rangle$. These issues include,

- Improve end-point resolution. Instead of using tracking alone, this method would use the known relationship between energy deposition and track range to find an optimal track end.
- Find misreconstructed tracks attached to other tracks. Instead of assuming that a single Bragg peak exists at the end of a track, PIDAN seeks out Bragg peaks that may be in the middle of a track. In the case of a proton decay event, this can help find kaons that have an attached pion or muon.
- Determine particle direction. The Bragg peak is heavily directional, potentially helping to reduce backgrounds from backward-reconstructed protons.
- Find a projected stopping point of a track if a particle's track ends before it comes to a completely stop. This can only be done if an amount of a particle's Bragg peak is visible, typically < 30 cm range. This allows PIDAN to identify a number of decay-in-flight events.

To find the most probable location of the endpoint, I first create dE/dx vs. residual range plots using simulated stopping particles. From these plots, I find the most probable value dE/dx and the asymmetric variance per given residual range point. I then fit both the asymmetric errors and dE/dx points to the function $f(R) = aR^b + c$ where a, b and c are fit parameters. This results in three fits per particle species. Each particle species has its own distribution, and thus the fit process needs to be done for each desired particle species. This analysis calls for the separation of pions, protons and kaons, leading to three separate plots and 27 fit parameters. I use this method instead of the full Bethe Bloch equation to capture detector effects.

In seeking for the end point of a track, the fitted functions are used to calculate the χ^2 of the track assuming a given endpoint. The algorithm iterates over a range of endpoints within a given designated window (typically -30 cm to 30 cm from the reconstructed end point) and the best fit represents the most probable endpoint of the track. This is done for each particle species and the endpoint from the particle species with the most probable endpoint, as determined from the χ^2 , is used as the new track end point.

Plotted in Figure 37 is the difference between the endpoint that the PIDAN algorithm found and the reconstructed track endpoint. The points labeled MC Truth represents the difference of the MC truth endpoint and the reconstructed track endpoint. It is clear that the data agrees well with the reconstructed simulation which itself agrees well with MC truth information. The only exception in agreement is two peaks in data found at approximately -0.5 and 0.5 cm that is not seen in the MC reconstruction. These values correspond to the track pitch of the wire planes and may be from unstimulated physics when tracks end close to a wire because they are not present in the simulation.

There exists a correlation between the endpoint error determined using the MC truth information and the endpoint error determined by the PIDAN algorithm. This shows that the method definitively improves endpoint resolution. Using a sample of protons, the original endpoint resolution is found to be 1.49 ± 0.14 mm using just tracking and 1.17 ± 0.07 mm from the PIDAN method.

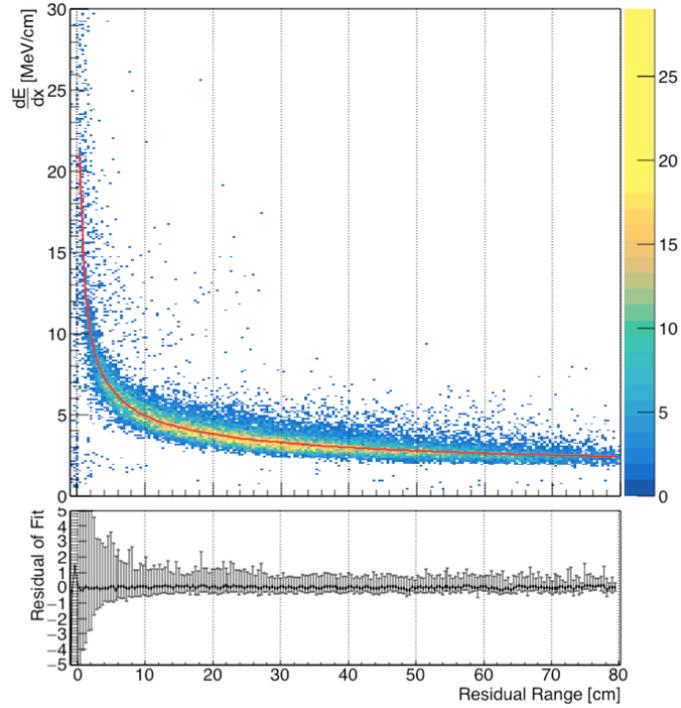


Figure 36: The template plot and residuals of the MPV fit for K^+ simulation.

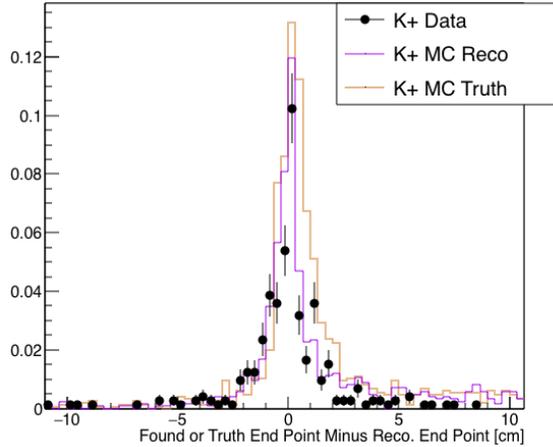


Figure 37: Correction factors as found by PIDAN algorithm compared against truth information.

In order to test how well the PIDAN algorithm is able to determine a projected endpoint from a track which does not stop before interacting or leaving the TPC, I use a set of simulated kaons with a momentum range that, at the lowest momentums, stops within the active volume but, at the highest momentums, leaves the active TPC volume before stopping. Plotted in Figure 39 is the identification efficiency of kaons in this simulated set using PIDA and PIDAN. K^+ efficiency represents the number of events with a primary track with measured PIDA between 10 and 15 $\text{MeV/cm}^{1.42}$ consistent with the expected value for kaons. The second x-axis in Figure 39 is the range determined by $z_{Reco} - z_{MC}$. For particles that stop within the active volume, the identification

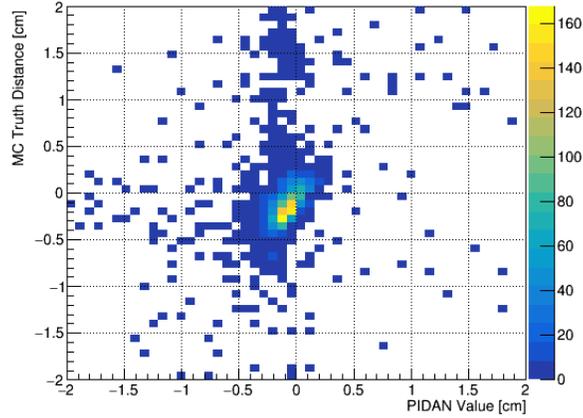


Figure 38: PIDAN correlation with MC Truth information.

efficiency for both PIDA and PIDAN is relatively high. At a threshold momentum, kaons begin to leave the active TPC volume and the K^+ identification falls to effectively zero for PIDA while the PIDAN method continues to be able to deduce the particle type as far as 50 cm from the endpoint. It is important to note that, at 50 cm, the ability of PIDAN to separate protons from kaons is very limited as their energy depositions at that range look too similar. This is why the kaon identification efficiency for PIDAN drops, as it begins to fail to separate kaons and protons but still finds a reasonable projected endpoint.

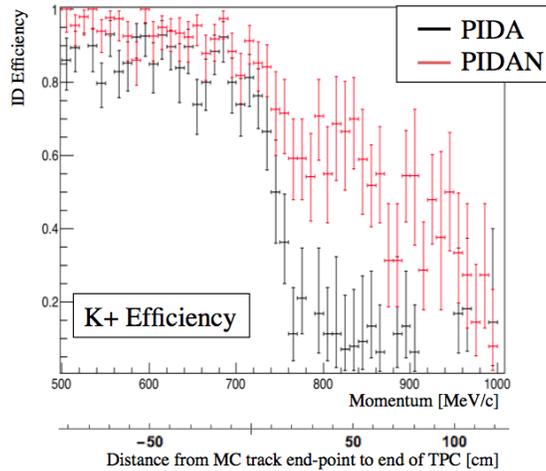


Figure 39: K^+ Identification efficiency as a function of momentum and amount of range outside of TPC detector for PIDA and PIDAN.

Since PIDAN finds the most probable location of a Bragg peak, it can be run on a set of hits forward and then backwards in order to check for a backwards-reconstructed track. To determine how well this method can separate forward kaons and backwards protons, I run the algorithm over the primary reconstructed track of each set forward and backwards for the proton decay-like simulated samples. If the algorithm finds that the most probable Bragg peak is backwards, it flips the dE/dx set to allow for normal measurement of PIDA. In plotting, I denote a track determined to be backwards as a negative PIDA value, as can be seen in Figure 40. The simulated samples used in this plot do not have reconstruction quality cuts applied.

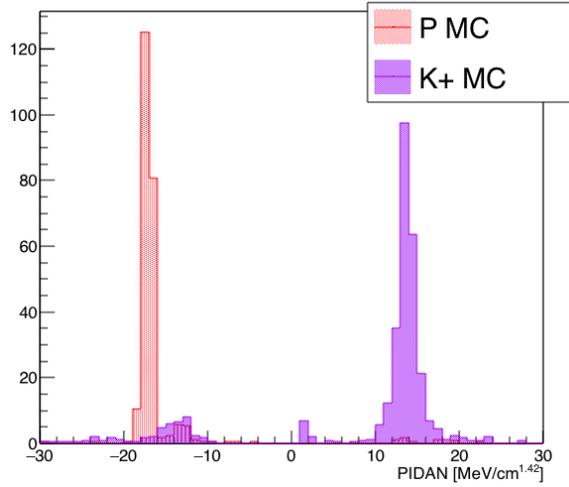


Figure 40: PIDAN ability to separate MC forward-going kaons and backwards-going protons. Protons are reconstructed backwards. If PIDAN determines a track is backwards, it saves the PIDAN value as negative. Both sets isotropic.

This plot shows that, for a backwards-reconstructed proton, PIDAN identifies it as backwards $\sim 98\%$ of the time while $\sim 15\%$ of forward-going kaons are labeled as backwards. This separation is even more impressive because there is no restriction on the performance of reconstruction for these events, meaning that this is the closest representation of how well PIDAN would perform at separating the two using current reconstruction efficiencies.

Plotted in Figure 41 are the PIDAN distributions for the isotropic simulated sets for forward-going protons and forward-going kaons. In comparing with the pure PIDA distribution for the same set seen in Figure 33, it is clear that PIDAN is sharpening the signal peaks for both protons and kaons and improving separation to some degree.

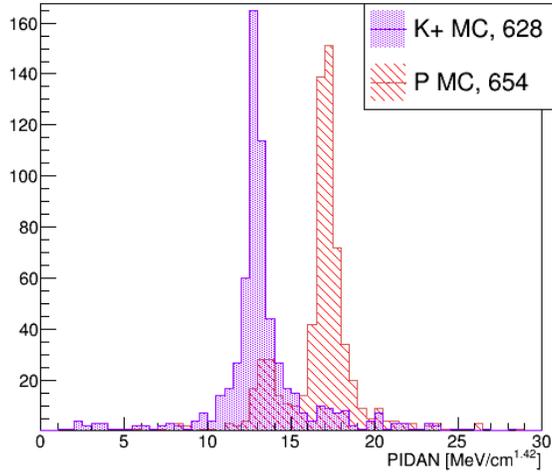


Figure 41: PIDAN for the isotropic simulated set of 100 kaons and 1000 protons.

The level of separation seen in Figure 41 is far from that proposed in literature for LArTPCs. The PIDAN method, while improving separation as compared to PIDA, is more a solution to the

lower than optimal tracking that is currently available in reconstruction. It is important to note that, every improvement in reconstruction over the coming years will improve the separation seen here. It is also important to note that energy deposition is only one element in particle identification: track topologies and invariant mass reconstruction are two other inputs to particle identification not used in this analysis due to the small size of LArIAT. Each additional input will only increase detector performance.

As seen in Figure 42, the PIDAN method performs well on the real data sample of kaons produced in LArIAT. This plot shows that the simulation, currently the only basis to LArTPC proton decay searches before larger detectors are build, is doing its job at representing reality and agreeing well with the K^+ data.

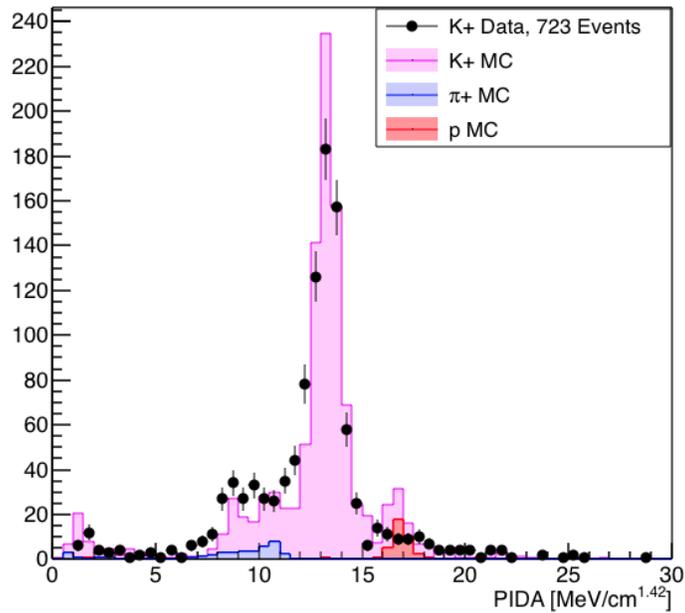


Figure 42: PIDAN stacked histogram for the kaon data sample.

6 Conclusion

As an input for future proton decay searches due to take place in large-scale LArTPCs, such as DUNE, I used the LArIAT detector, an experiment in the meson test beam at Fermilab, to select for a real data sample of 1992 K^+ candidate events. Half of this thesis was dedicated to developing the beam line reconstruction software necessary to select for the rare particle species as the K^+ composes only $\sim 0.5\%$ of the beam.

With these data events, I identify LArIAT's usefulness to proton decay searches for the channel $p \rightarrow \bar{\nu}K^+$ as a test bed for energy deposition particle identification algorithms. I then investigate the performance of one such algorithm, PIDA, on a simulated set of K^+ events and then show the agreement between the simulation and data. This agreement alone is a strong statement as this is the first time in a LArTPC that the simulation of kaon events is shown to agree with data, affirming the reality of proton decay studies that, out of necessity, are developed using simulation.

I then move on to create an approximate sample of proton decay events with corresponding proton decay background. Without the full simulation, these sets are weighted to 1:1 so only qualitative studies of background rejection are possible. With these samples, I again study the performance of PIDA and investigate new issues seen in the lower momentum set.

I finish the thesis by introducing PIDAN, an improvement to PIDA that is meant to lower the dependency of the algorithm on tracking resolution. I show mostly qualitatively promising improvements with this method over PIDA.

The ultimate take-away from this analysis is that the use of automatically reconstructed LArTPC data for proton decay searches is in development and, at the current date of April 2018, already shows promising performance that allows the detector to make better measurements of the channel $p \rightarrow K^+\bar{\nu}$ than larger and longer-lived water Cherenkov detectors.

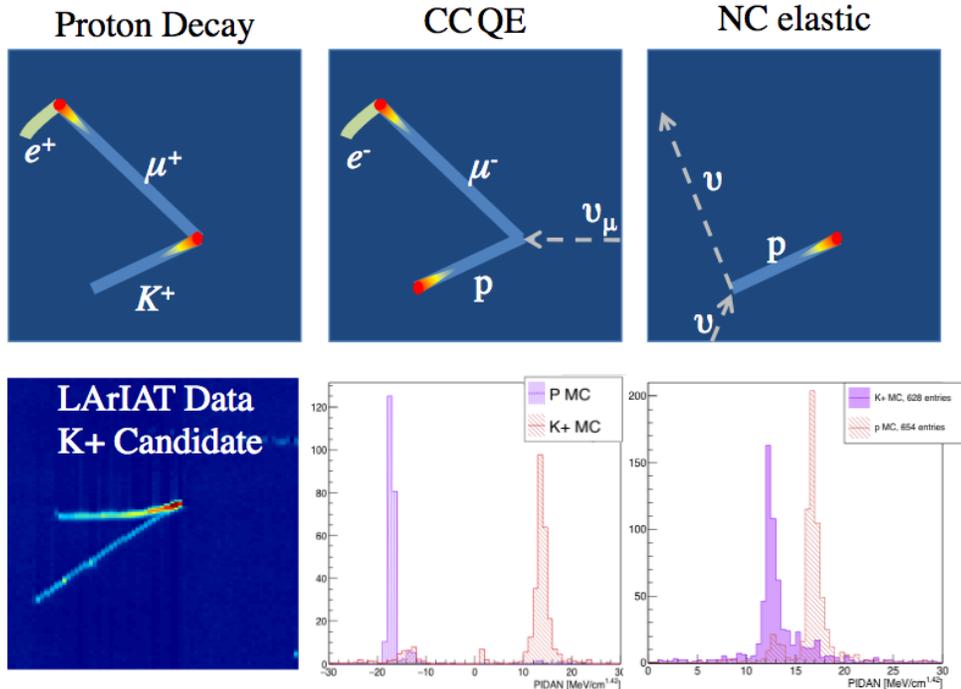


Figure 43: Final take-away results. Top: Drawings of proton decay signals and backgrounds. Bottom: Corresponding LArIAT data and current results of rejecting backgrounds using PIDAN.

7 Acknowledgements

The last four years have been the beginnings of my physics career and my success so far is thanks to a list of individuals too long to write out. However, I would like to take this opportunity to reflect and thank those who particularly stand out.

First, I thank the members of my defense committee. I would like to thank Prof. Ed Kearns who has, since the second week of my freshmen year, supported me as a mentor in my aspirations to become a physicist. I thank Prof. Rob Carey for his support as a mentor and as an example of a healthy mentality to approach physics. I thank Prof. Claudio Rebbi for his unceasing kindness starting from office hours in computational physics until today.

Next, I thank members of the LArIAT collaboration. This rag-tag group of physicists have taught me what it is like to be a professional scientist and have shown me that it is the right path for me. I could not have spent my undergraduate time with a better group.

To Ryan Linehan and Elena Gramellini, thank you for the countless time you've spent both as a mentor and as a friend.

To Jacob Levy, Madeleine O'Keefe, Nathan Ulberg, Nate Avish and Deepak Sathyan, thank you for the friendship. If nothing else, you guys have made my time at B.U. worth it.

Finally, I would like to thank Amina Li.

References

- [1] J.L. Hewett, *et al.*, “Planning the Future of U.S. Particle Physics (Snowmass 2013): Chapter 2: Intensity Frontier,” FERMILAB-CONF-14-019-CH02, 2013.
- [2] K. Abe, *et al.*, “Search for Proton Decay via $p \rightarrow \bar{\nu}K^+$ using 260 kiloton-year data of Super-Kamiokande,” *Phys. Rev. D* 90, 072005 (2014).
- [3] A. Bueno *et al.*, “Nucleon Decay Searches with large Liquid Argon TPC Detectors at Shallow Depths: atmospheric neutrinos and cosmogenic backgrounds.” *JHEP* 0704:041,2007.
- [4] Ryan Linehan, “A diagram of LArTPC design and basic operating principles,” Wikimedia, 2015.
- [5] ArgoNeut Collaboration, “ArgoNeut: About,” <http://t962.fnal.gov/About.html>, 2018.
- [6] D. Nygren, *et al.*, “1974 PEP Summer Study”, PEP-137, 1974.
- [7] C. Rubbia, “The Liquid-Argon Time projection Chamber: a new concept for Neutrino Detector,” CERN-EP/77-08, (1977).
- [8] MicroBooNE collaboration, “Michel Electron Reconstruction Using Cosmic-Ray Data from the MicroBooNE LArTPC”, *JINST* 12 P09014 (2017).
- [9] J. Assadi, “PixLAr (LArIAT with Pixels)”, Internal Report, LArTPC-doc-2828-v1, 2018.
- [10] M. Adamowski, *et al.*, “The Liquid Argon Purity Demonstrator,” FERMILAB-PUB-14-059-E, 2014.
- [11] Fermilab Test Beam Facility, “MWPC Tracking System”, <http://ftbf.fnal.gov/mwpc-tracking-system/>, 2018.
- [12] Greg Pulliam, “Beamline Composition Studies,” Internal Report, LArTPC-doc-2756-v1, 2017.
- [13] Daniel Smith, “Beam Line Tracking with Geane,” Internal Report, LArTPC-doc-3038-v1, 2018.
- [14] GEANT4 Collaboration, “GEANT4: A Simulation toolkit,” *Nucl. Instrum. Meth. A* 506 (2003) 250-303 SLAC-PUB-9350, FERMILAB-PUB-03-339
- [15] Martin Ester, *et al.*, “A density-based algorithm for discovering clusters in large spatial databases with noise.” *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining (KDD-96)*, 1996.
- [16] Douglas Jensen, “NDB fields,” Internal Report, LArTPC-doc-1782-v1, 2016.
- [17] Amsler *et al.*, “Passage of Particles Through Matter.” *Phys. Letters B*, vol 667, 1. 2008.
- [18] R. Acciarri, *et al.*, “A study of electron recombination using highly ionizing particles in the ArgoNeuT Liquid Argon TPC”, *JINST* 8 (2013) P08005.
- [19] R. Acciarri, *et al.*, “LArIAT: Liquid Argon In A Testbeam”, FERMILAB-PUB-14-268-E, 2014.