PROCEEDINGS OF SCIENCE

Automated Reconstruction, Signal Processing and Particle Identification in DUNE

Michael Wallbank[*†]
University of Sheffield
E-mail: wallbank@fnal.gov

Liquid Argon TPC (LArTPC) technology is increasingly prevalent in large-scale detectors designed to observe neutrino scattering events induced by accelerators or by natural sources. LArTPCs consist of a very high fraction of active detector material with spatial resolutions on the order of a few millimetres. Three-dimensional interactions are imaged in multiple two-dimensional views by the process of projection onto planes of wires. The goal of automated reconstruction is to correctly classify each neutrino scattering event by the flavor of the incoming neutrino, to separate charged-current events from neutral-current and other backgrounds, and to measure the energies of the incoming neutrinos. Detection of neutrinos from supernova bursts and also searching for nucleon decay are important uses of automated reconstruction algorithms. Because the amount of spatial detail is high and the amount of scattering is high, this reconstruction presents many challenges that are now being investigated with sophisticated techniques. Signal processing, track and shower identification, particle identification, and event classification by a variety of innovative algorithms are reviewed.

EPS-HEP 2017, European Physical Society conference on High Energy Physics
5-12 July 2017
Venice, Italy

[*]Speaker.
[†]On behalf of the DUNE collaboration.

PoS(EPS-HEP2017)815

1. Introduction

Liquid argon time projection chambers (LArTPCs) provide a robust and elegant method for measuring the properties of neutrino interactions above a few tens of MeV by providing 3D event imaging with excellent spatial and energy resolution. A lot of effort has been focussed on the challenge of reconstructing events in a LArTPC over the past few years. The development of a common software framework, LArSoft [1], by several LArTPC experiments has helped significantly with the development of new algorithms and the sharing of expertise. These proceedings review the latest reconstruction chain used by the Deep Underground Neutrino Experiment (DUNE) [2].

DUNE will make world-leading measurements and will take large amounts of high-precision, high-quality data. It is therefore imperative that all processing, including signal processing, reconstruction, particle identification, event selection and analysis, be automated.

2. Signal Processing

Raw data from the TPC is readout as a waveform of digitised charge (ADC) as a function of time ticks for each wire in the detector. If collected on an induction wire, the signal takes the form of a bipolar pulse and on the collection wires it is unipolar. The first step in the reconstruction is to convert the raw signal from each wire to a pulse of charge, of Gaussian shape, by removing the detector effects. This is achieved by passing the raw data through a calibrated deconvolution algorithm to ‘remove’ the impact of field and electronics responses [3]. The aim is to recover the number of drifting ionisation electrons and this technique has the advantages of being robust and fast.

3. Hit Finding

Following deconvolution, all signals take the form of Gaussian charge pulses. Hit finding proceeds by identifying these peaks as areas of charge above a given threshold. Once a pulse is found, a sum of n Gaussians is fit to the waveform, where n is defined by the number of peaks initially identified within the pulse. Each Gaussian is defined as a hit.

The wrapped induction wires utilised by the DUNE far detector readout gives rise to an inherent ambiguity in the location of the hit along the length of the channel. The process of determining which wire segment the charge was collected on is referred to as disambiguation. The ambiguity is broken using information from the collection planes, which are not wrapped. The induction wire angles are either slightly different, giving rise to unique ‘triple points’ of deposited charge, or chosen such that each intersects a collection wire no more than once; either method facilitates a complete disambiguation.

4. Pattern Recognition

There are numerous pattern recognition algorithms implemented within LArSoft, optimised for specific purposes, which will be briefly overviewed here. There are two general approaches, each utilising the same event information, which proceed from the output of the raw signal processing and result in fully reconstructed tracks, showers, vertices and energy. The reconstruction is either performed in each 2D plane separately before combining this information to create 3D objects, or 3D image reconstruction is conducted directly, with pattern recognition achieved by studying this 3D image. The majority of developments so far have utilised the former approach.

4.1 Line Cluster

In LArSoft, a cluster is a collection of hits on one wire plane representing a single particle. The intent of Line Cluster [4] is to construct 2D line-like clusters using local information. The algorithm constructs short line-like 'seeds' of close hits in areas of low hit density, where hit proximity is a good indication that the hits are associated with a common particle, and then 'crawls' along other nearby hits, adding them to the cluster if they are determined similar enough to current constituent hits. The conditions are related to the hit charges and impact parameters whilst allowing large dE/dx fluctuations towards the end of a stopping track and attempting to reject hits from δ-rays.

4.2 Blurred Cluster

The Blurred Cluster algorithm [5] attempts to construct 2D shower-like clusters from deposits left by showering particles. It specialises especially in the separation of nearby showering particles in the reconstruction of, e.g., π^0 decay. It proceeds by first applying a weighted 2D Gaussian smearing to the hit map in order to introduce 'fake hits' and more realistically distribute the charge deposited in the detector. Clustering then utilises a nearest-neighbour approach within the blurred region, before removing the fake hits to leave just the original shower hits.

4.3 Projection Matching Algorithm

The Projection Matching Algorithm (PMA) [6] was primarily developed as a technique of 3D reconstruction of individual particle trajectories and was designed to address the challenging issue of transformation from a set of independently reconstructed 2D projections of objects into a 3D representation. PMA uses clusters of hits (from, e.g., Line Cluster) as input and also attempts to correct cluster assignment of hits using properties of 3D reconstructed objects.

The concept is to build and optimise objects in 3D space, formed as polygonal lines with iteratively increasing number of segments, by minimising a 'cost function' in all available 2D projections into other planes. This cost function represents the 2D distance of hits to the optimised object in each projection, a penalty for track curvature and the 3D distance of various feature points to the optimised objects. Once optimised, the output tracks also contain basic physics information, such as particle directionality and dE/dx evolution along the trajectories.

4.4 Pandora

The Pandora Software Development Kit [7] was created to address the problem of identifying energy deposits from individual particles in fine-grained detectors. It promotes the idea of a multi-algorithm approach to pattern-recognition and utilises large numbers of decoupled algorithms to characterise the the input hits. There or over 70 algorithms, designed to target specific event topologies and controls operations such as collecting hits into clusters, merging or splitting clusters or collecting clusters in order to build particles.

4.5 EMShower

The EMShower algorithm [8] aims to find final 3D reconstructed showers and all associated properties of the showering particles. It is intentionally high-level by design and relies heavily on previous reconstruction, particularly BlurredCluster (Section 4.2), PMA (Section 4.5) and Pandora (Section 4.4).

The reconstruction proceeds in two general steps: first the shower objects, including all associated hits in each of the views, are found; following this, the properties of these showers, including the start point, direction, initial dE/dx and total deposited energy, are determined using multiple pattern recognition and calorimetric reconstruction algorithms. The particle objects are either formed using the associations between 2D Blurred Clusters and 3D tracks provided by PMA, or by utilising hits associated with showering objects identified using Pandora reconstruction.

4.6 WireCell

WireCell [9] adopts the latter approach discussed previously and constructs a 3D event image prior to pattern-recognition and complete reconstruction. This is achieved using time, geometry and charge information and has the advantage of using the full TPC information. The strong requirement of the time/geometry/charge information provides a natural way to suppress electronic noise at the cost of being more sensitive to hit inefficiencies, and, since track and shower hypotheses are not used, the 3D imaging works for any event topology. Once the 3D images are reconstruction, pattern-recognition proceeds to identify the content of the image.

5. Calorimetric Energy Reconstruction and Particle Identification

As charged particles traverse a LAr volume, they deposit energy through ionisation and scintillation. The ionising power provides information on the particle energy and species and is reconstructed by taking all relevant hits, for example along a track, and converting the hit area, in ADC counts, to charge, in units of fC, using information from crossing muons or test stands. Two corrections are made to convert dQ/dx to dE/dx: a lifetime correction to account for attenuation of drift electrons by electronegative impurities and a recombination correction.

The dE/dx may be used to distinguish between showering particles and, if a track stops within the detector volume, the energy loss as a function of residual range may be utilised as a powerful method for particle identification. This technique is known as PIDA [10].

6. Summary

The reconstruction of events in a LArTPC is being actively developed and much encouraging recent progress has been made. A fully automated reconstruction chain is in place in DUNE to reconstruct tracks, showers and calorimetric information for various physics analyses.

References

[1] *larsoft.org* .

[2] DUNE Collaboration, *Long-Baseline Neutrino Facility (LBNF) and Deep Underground Neutrino Experiment (DUNE): The LBNF and DUNE Projects*, *arXiv:1601.05471 [physics.ins-det]* (2016) , [1601.05471].

[3] The MicroBooNE Collaboration, *A Method to Extract the Charge Distribution Arriving at the TPC Wire Planes in MicroBooNE*, *MicroBooNE Public Note* **1017** (2016) .

[4] *https://cdcvs.fnal.gov/redmine/documents/727* .

[5] *http://larsoft.org/single-record/?pdb=110* .

[6] M. Antonello and Others, *Precise 3D track reconstruction algorithm for the ICARUS T600 liquid argon time projection chamber detector*, *Adv. High Energy Phys.* **2013** (2013) 260820, [1210.5089].

[7] J. S. Marshall and M. A. Thomson, *The Pandora software development kit for pattern recognition*, *The European Physical Journal C* **75** (2015) 439.

[8] *http://larsoft.org/single-record/?pdb=113* .

[9] *http://lar.bnl.gov/wire-cell/* .

[10] R. Acciarri, C. Adams, J. Asaadi, B. Baller, T. Bolton, C. Bromberg et al., *A study of electron recombination using highly ionizing particles in the ArgoNeuT Liquid Argon TPC*, *Journal of Instrumentation* **8** (2013) P08005.